

Manual de
CÁLCULO

Artemio González López

Madrid, junio de 2018

Autor:

Artemio González López

Departamento de Física Teórica

Facultad de Ciencias Físicas

Avenida Complutense s/n

Ciudad Universitaria

28040 Madrid

© El autor

Se prohíbe la reproducción total o parcial de esta obra —incluido el diseño de portada—, sea cual fuere el medio, electrónico o mecánico, sin el consentimiento del autor.

ISBN-13: 978-84-608-8859-8

Índice general

1	Funciones de \mathbb{R}^n en \mathbb{R}^m. Límites y continuidad	1
1.1	Vectores cartesianos en \mathbb{R}^3	1
1.2	El espacio \mathbb{R}^n . Funciones de \mathbb{R}^n en \mathbb{R}^m	4
1.3	Espacios métricos, normados y euclidianos	8
1.4	Interior, exterior, frontera	12
1.5	Conjuntos abiertos, cerrados y compactos	15
1.6	Límites	18
1.7	Continuidad	23
1.8	Límite a lo largo de una curva	26
1.9	Notación o y O de Landau	29
1.10	Propiedades topológicas de las funciones continuas	32
2	Diferenciabilidad	37
2.1	Derivada de una función vectorial	37
2.2	Derivada de una función de \mathbb{R}^n en \mathbb{R}^m	37
2.3	Derivadas direccionales y parciales	43
2.4	Funciones de clase C^1	48
2.5	Regla de la cadena	50
2.6	Derivadas parciales de orden superior	55
2.7	Lema de Schwarz	57
2.8	Interpretación geométrica del gradiente y la derivada	61
2.9	Cálculo diferencial vectorial	65
3	Fórmula de Taylor. Extremos. Funciones inversas e implícitas	77
3.1	Fórmula de Taylor	77
3.2	Extremos	82
3.3	Formas cuadráticas	85
3.4	Criterio del hessiano	86
3.5	Teorema de la función inversa	92
3.5.1	Sistemas de coordenadas	96
3.6	Teorema de la función implícita	98
3.6.1	Derivación implícita	101
3.7	Extremos condicionados	103
4	Integrales múltiples	109
4.1	Integral doble en un rectángulo	109
4.1.1	Definición de integral	109
4.1.2	Criterio de integrabilidad de Lebesgue	112
4.1.3	Teorema de Fubini	115
4.2	Integral doble en recintos más generales	118
4.3	Integrales múltiples en \mathbb{R}^n	124

4.3.1	Integral múltiple en un n -paralelepípedo	124
4.3.2	Integral múltiple en un conjunto medible Jordan	126
4.4	Integrales triples	128
4.5	Cambios de variables en integrales múltiples	132
5	Integrales de línea y superficie	143
5.1	Integral de una función escalar a lo largo de una curva	143
5.1.1	Curvas parametrizadas	143
5.1.2	Longitud de una curva	145
5.1.3	Reparametrizaciones	146
5.1.4	Integral de trayectoria de una función escalar	148
5.2	Integral de línea de un campo vectorial	149
5.2.1	Integrales de línea	149
5.2.2	Campos conservativos	152
5.3	Integral de una función escalar sobre una superficie de \mathbb{R}^3	157
5.3.1	Superficies parametrizadas en \mathbb{R}^3	157
5.3.2	Producto vectorial fundamental	160
5.3.3	Área de una superficie parametrizada	162
5.3.4	Integral de superficie de una función escalar	168
5.4	Integral de superficie de un campo vectorial	170
6	Teoremas de Green, Stokes y Gauss	177
6.1	Teorema de Green	177
6.2	Teorema de Stokes	182
6.3	Teorema de Ostrogradsky–Gauss	186

Capítulo 1

Funciones de \mathbb{R}^n en \mathbb{R}^m . Límites y continuidad

1.1 Vectores cartesianos en \mathbb{R}^3

Intuitivamente, el espacio \mathbb{R}^3 es el conjunto de los vectores en el espacio ordinario (tridimensional). Como dichos vectores tienen tres *componentes* x_i ($i = 1, 2, 3$) según las direcciones de tres ejes perpendiculares (*ejes coordenados*), podemos asimilar \mathbb{R}^3 al conjunto

$$\mathbb{R}^3 = \{(x_1, x_2, x_3) \mid x_i \in \mathbb{R}, i = 1, 2, 3\}.$$

En particular, los vectores unitarios según las direcciones de los ejes coordenados (*vectores unitarios coordenados*) están dados por

$$\mathbf{e}_1 \equiv \mathbf{i} = (1, 0, 0), \quad \mathbf{e}_2 \equiv \mathbf{j} = (0, 1, 0), \quad \mathbf{e}_3 \equiv \mathbf{k} = (0, 0, 1).$$

Dos vectores $\mathbf{x} = (x_1, x_2, x_3), \mathbf{y} = (y_1, y_2, y_3) \in \mathbb{R}^3$ son iguales si y solo si tienen las mismas componentes, es decir si y solo si $x_i = y_i$ para todo $i = 1, 2, 3$. La suma $\mathbf{x} + \mathbf{y}$ de dichos vectores se define mediante la *regla del paralelogramo* (cf. Fig. 1.1), que es equivalente a la relación

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, x_3 + y_3).$$

Dado un número real λ (*escalar*), se define el vector $\lambda\mathbf{x} \in \mathbb{R}^3$ mediante la igualdad

$$\lambda\mathbf{x} = (\lambda x_1, \lambda x_2, \lambda x_3).$$

Los vectores \mathbf{x} y $\lambda\mathbf{x}$ son siempre paralelos (si $\lambda \neq 0$), y tienen la misma dirección (resp. direcciones opuestas) si $\lambda > 0$ (resp. $\lambda < 0$). Con las definiciones dadas de suma y producto por escalares, es inmediato comprobar que cualquier vector $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$ se puede escribir como la *combinación lineal*

$$\mathbf{x} = x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k} \equiv \sum_{i=1}^3 x_i\mathbf{e}_i.$$

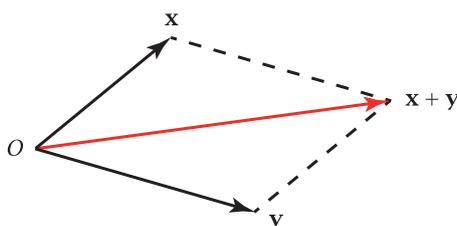


Figura 1.1: regla del paralelogramo ($O = (0, 0, 0)$ es el origen de coordenadas).

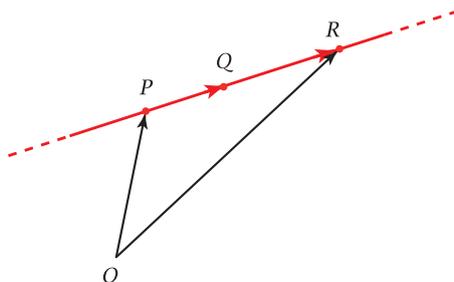


Figura 1.2: recta que pasa por el punto $\mathbf{a} = \overrightarrow{OP}$ en la dirección del vector $\mathbf{v} = \overrightarrow{PQ}$ ($\overrightarrow{OR} = \mathbf{x}$, $\overrightarrow{PR} = \mathbf{x} - \mathbf{a}$).

El *producto escalar* de dos vectores $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ es el número real definido por

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + x_3 y_3 \equiv \sum_{i=1}^3 x_i y_i.$$

Geoméricamente,

$$\mathbf{x} \cdot \mathbf{y} = \|\mathbf{x}\| \|\mathbf{y}\| \cos \theta(\mathbf{x}, \mathbf{y}), \quad (1.1)$$

donde $\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}}$ es la longitud (o *norma*) del vector \mathbf{x} (y análogamente $\|\mathbf{y}\|$) y $\theta(\mathbf{x}, \mathbf{y}) \in [0, \pi]$ es el *ángulo* formado por los vectores \mathbf{x}, \mathbf{y} . En particular, los vectores $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ son perpendiculares (u *ortogonales*) si $\mathbf{x} \cdot \mathbf{y} = 0$. Análogamente, se define el *producto vectorial* $\mathbf{x} \times \mathbf{y}$ de los vectores \mathbf{x}, \mathbf{y} mediante

$$\mathbf{x} \times \mathbf{y} = (x_2 y_3 - x_3 y_2, x_3 y_1 - x_1 y_3, x_1 y_2 - x_2 y_1) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix}.$$

Nótese que

$$(\mathbf{x} \times \mathbf{y})_i = x_j y_k - x_k y_j,$$

donde (i, j, k) es una permutación *cíclica* de $(1, 2, 3)$. Geométricamente, el vector $\mathbf{x} \times \mathbf{y}$ es perpendicular a los vectores \mathbf{x}, \mathbf{y} (¡demostrarlo!), y su longitud es igual a

$$\|\mathbf{x} \times \mathbf{y}\| = \|\mathbf{x}\| \|\mathbf{y}\| \sin \theta(\mathbf{x}, \mathbf{y}).$$

Por último, el *sentido* del vector $\mathbf{x} \times \mathbf{y}$ está determinado por la *regla del sacacorchos* (equivalentemente, los vectores $\mathbf{x}, \mathbf{y}, \mathbf{x} \times \mathbf{y}$ tienen la misma orientación que los vectores unitarios coordenados $\mathbf{i}, \mathbf{j}, \mathbf{k}$). En particular, los vectores \mathbf{x}, \mathbf{y} son paralelos si y solo si $\mathbf{x} \times \mathbf{y} = 0$.

Ejercicio 1. Probar que $\|\mathbf{x} \times \mathbf{y}\|^2 = \|\mathbf{x}\|^2 \|\mathbf{y}\|^2 - (\mathbf{x} \cdot \mathbf{y})^2$.

Dado un punto $\mathbf{a} = (a_1, a_2, a_3) \in \mathbb{R}^3$, la *ecuación paramétrica* de la recta que pasa por dicho punto en la dirección del vector $\mathbf{v} = (v_1, v_2, v_3) \neq 0$ es

$$\boxed{\mathbf{x} = \mathbf{a} + t\mathbf{v}, \quad t \in \mathbb{R}}. \quad (1.2)$$

En efecto, si \mathbf{x} es un vector de dicha recta entonces el vector $\mathbf{x} - \mathbf{a}$ ha de ser paralelo al vector \mathbf{v} , y por tanto $\mathbf{x} - \mathbf{a} = t\mathbf{v}$ para algún escalar t (cf. la Fig. 1.2). La *ecuación cartesiana* de dicha recta se obtiene eliminando el parámetro t de una de las ecuaciones y sustituyendo en las otras dos, lo cual proporciona dos relaciones entre las tres coordenadas (x_1, x_2, x_3) de un punto arbitrario de la recta. Con un ligero abuso de notación, podemos escribir estas dos ecuaciones como

$$\boxed{\frac{x_1 - a_1}{v_1} = \frac{x_2 - a_2}{v_2} = \frac{x_3 - a_3}{v_3}}, \quad (1.3)$$

sobreentendiéndose que si $v_i = 0$ entonces simplemente $x_i = a_i$.

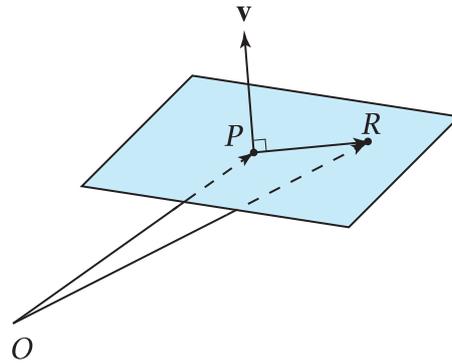


Figura 1.3: plano perpendicular al vector \mathbf{v} que pasa por el punto $\mathbf{a} = \overrightarrow{OP}$ ($\overrightarrow{OR} = \mathbf{x}$, $\overrightarrow{PR} = \mathbf{x} - \mathbf{a}$).

Ejemplo 1.1. La ecuación paramétrica de la recta que pasa por el punto $(1, -2, 0)$ en la dirección del vector $(3, 0, -1)$ es

$$\mathbf{x} = (1, -2, 0) + t(3, 0, -1) = (1 + 3t, -2, -t), \quad t \in \mathbb{R},$$

o equivalentemente,

$$x_1 = 1 + 3t, \quad x_2 = -2, \quad x_3 = -t, \quad t \in \mathbb{R}.$$

La ecuación cartesiana se obtiene fácilmente eliminando el parámetro t de la última ecuación:

$$x_1 + 3x_3 = 1, \quad x_2 = -2. \quad \square$$

Análogamente, la ecuación del plano que pasa por el punto \mathbf{a} y es perpendicular al vector $\mathbf{v} \neq 0$ es

$$(\mathbf{x} - \mathbf{a}) \cdot \mathbf{v} = 0. \quad (1.4)$$

En efecto, si el punto \mathbf{x} pertenece a dicho plano entonces $\mathbf{x} - \mathbf{a}$ es perpendicular al vector \mathbf{v} , y por tanto $(\mathbf{x} - \mathbf{a}) \cdot \mathbf{v} = 0$ (cf. la Fig. 1.3). Nótese que la ecuación (1.4) se puede escribir también en la forma

$$v_1(x_1 - a_1) + v_2(x_2 - a_2) + v_3(x_3 - a_3) = 0.$$

En general, la ecuación

$$v_1x_1 + v_2x_2 + v_3x_3 = c$$

es la ecuación de un plano perpendicular al vector (v_1, v_2, v_3) . Dicho plano pasa por el origen de coordenadas si y solo si $c = 0$.

Ejemplo 1.2. La ecuación del plano perpendicular al vector $(-1, 0, 4)$ que pasa por el punto $(1, -2, 2)$ es

$$(x_1 - 1, x_2 + 2, x_3 - 2) \cdot (-1, 0, 4) = -x_1 + 1 + 4(x_3 - 2) = -x_1 + 4x_3 - 7 = 0,$$

o equivalentemente

$$x_1 - 4x_3 = -7.$$

Ejercicio 2. ¿Cómo han de ser los vectores \mathbf{a} y \mathbf{v} para que la recta (1.2) pase por el origen?

Ejercicio 3. Hallar la ecuación del plano que contiene a las rectas $\mathbf{x} = \mathbf{a} + t\mathbf{v}_1$ y $\mathbf{x} = \mathbf{a} + t\mathbf{v}_2$ que pasan por el punto $\mathbf{a} \in \mathbb{R}^3$.

1.2 El espacio \mathbb{R}^n . Funciones de \mathbb{R}^n en \mathbb{R}^m

El conjunto \mathbb{R}^n es el *producto cartesiano*¹ $\underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{n \text{ veces}}$. En otras palabras,

$$\mathbb{R}^n = \{(x_1, \dots, x_n) \mid x_i \in \mathbb{R}, 1 \leq i \leq n\}$$

es el conjunto de las n -tuplas *ordenadas* de números reales. Es importante notar que el adjetivo “ordenadas” en la frase anterior significa que la *igualdad* de dos elementos $x \equiv (x_1, \dots, x_n), y \equiv (y_1, \dots, y_n) \in \mathbb{R}^n$ se define por

$$x = y \iff x_i = y_i, \quad 1 \leq i \leq n.$$

Los n números reales x_i ($i = 1, \dots, n$) se denominan las **componentes** de x . La suma de dos elementos $x \equiv (x_1, \dots, x_n), y \equiv (y_1, \dots, y_n) \in \mathbb{R}^n$ y el producto de un escalar $\lambda \in \mathbb{R}$ por un elemento $x \in \mathbb{R}^n$ se definen respectivamente por

$$\begin{cases} x + y = (x_1 + y_1, \dots, x_n + y_n) \\ \lambda x = (\lambda x_1, \dots, \lambda x_n). \end{cases}$$

De las propiedades de la suma de números reales se sigue fácilmente que la suma en \mathbb{R}^n es conmutativa y asociativa, su elemento neutro es $0 \equiv (0, \dots, 0)$ y el inverso de un elemento cualquiera $x = (x_1, \dots, x_n)$ es $(-x_1, \dots, -x_n) \equiv -x$. Se verifican además las siguientes identidades, que relacionan la suma y el producto por escalares:

$$\text{i) } \lambda(\mu x) = (\lambda\mu)x, \quad \forall \lambda, \mu \in \mathbb{R}, \forall x \in \mathbb{R}^n$$

$$\text{ii) } 1 \cdot x = x, \quad \forall x \in \mathbb{R}^n$$

$$\text{iii) } \lambda(x + y) = \lambda x + \lambda y, \quad \forall \lambda \in \mathbb{R}, \forall x, y \in \mathbb{R}^n$$

$$\text{iv) } (\lambda + \mu)x = \lambda x + \mu x \quad \forall \lambda, \mu \in \mathbb{R}, \forall x \in \mathbb{R}^n$$

En otras palabras, con las operaciones que acabamos de definir \mathbb{R}^n es un *espacio vectorial*.

Definición 1.3. Una **función** de \mathbb{R}^n en \mathbb{R}^m es una aplicación $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, es decir una regla que a cada elemento x de un cierto subconjunto no vacío de \mathbb{R}^n le asigna de manera *unívoca* un elemento $f(x) \in \mathbb{R}^m$. Si $m = 1$, se dice que f es una función **escalar** de n variables reales.

Definición 1.4. El conjunto $\mathcal{D}(f)$ de todos los elementos $x \in \mathbb{R}^n$ para los cuales existe (i.e., está definido) $f(x)$ se denomina **dominio** de la función f . El conjunto $\mathcal{R}(f)$ de todos los valores que toma la función f , es decir de los elementos $y \in \mathbb{R}^m$ tales que $y = f(x)$ para algún $x \in \mathbb{R}^n$, se denomina **imagen** de f .

En otras palabras,

$$\mathcal{D}(f) = \{x \in \mathbb{R}^n \mid \exists f(x) \in \mathbb{R}^m\} \subset \mathbb{R}^n, \quad \mathcal{R}(f) = \{f(x) \mid x \in \mathcal{D}(f)\} \subset \mathbb{R}^m,$$

o equivalentemente²

$$\mathcal{D}(f) = f^{-1}(\mathbb{R}^m), \quad \mathcal{R}(f) = f(\mathbb{R}^n).$$

¹Si A y B son dos conjuntos cualesquiera, su *producto cartesiano* es el conjunto de pares ordenados (a, b) con $a \in A$ y $b \in B$. En otras palabras, $A \times B = \{(a, b) \mid a \in A, b \in B\}$, siendo por definición $(a_1, b_1) = (a_2, b_2)$ si y solo si $a_1 = a_2$ y $b_1 = b_2$. El producto cartesiano de n conjuntos se define de forma análoga.

²En general, si A y B son conjuntos arbitrarios y $f : A \rightarrow B$ se define la **imagen inversa** $f^{-1}(C)$ de un subconjunto $C \subset B$ bajo f como el conjunto de todos los puntos $x \in A$ tales que $f(x) \in C$: $f^{-1}(C) \equiv \{x \in A \mid f(x) \in C\}$. Nótese que esta definición tiene sentido aunque la función f no sea invertible.

Si f es una función de \mathbb{R}^n en \mathbb{R}^m y $x \equiv (x_1, \dots, x_n) \in \mathcal{D}(f)$, como $f(x)$ es un vector de \mathbb{R}^m tendrá m componentes, que denominaremos $f_i(x)$ ($i = 1, \dots, m$). En otras palabras,

$$\boxed{f(x) = (f_1(x), \dots, f_m(x)), \quad \forall x \in \mathcal{D}(f)}, \quad (1.5)$$

donde cada f_i es una función de \mathbb{R}^n en \mathbb{R} , es decir una función *escalar* de n variables. Las m funciones f_i se denominan **funciones componentes** de f . Nótese que el vector $f(x) \in \mathbb{R}^m$ está definido si y solo si lo están cada una de sus componentes $f_i(x)$, es decir

$$\mathcal{D}(f) = \bigcap_{i=1}^m \mathcal{D}(f_i). \quad (1.6)$$

Recíprocamente, m funciones escalares $f_1, \dots, f_m : \mathbb{R}^n \rightarrow \mathbb{R}$ cuyos dominios tienen intersección no vacía definen a través de la ec. (1.5) una función $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ con dominio dado por (1.6). En otras palabras, *dar una función de \mathbb{R}^n en \mathbb{R}^m es equivalente a dar m funciones escalares de n variables*. Por este motivo, el estudio de las funciones de \mathbb{R}^n en \mathbb{R}^m podría en principio reducirse al de las funciones escalares. En general, sin embargo, una función escalar de $n > 1$ variables no puede expresarse en términos de funciones de una variable. Esta es esencialmente la causa de que el estudio de las funciones de \mathbb{R}^n en \mathbb{R}^m suponga un salto conceptual importante respecto del estudio análogo de las funciones reales de una variable real.

Ejemplo 1.5. Las funciones de \mathbb{R}^n en \mathbb{R}^m son una herramienta esencial en la descripción de los fenómenos físicos. Por ejemplo:

- i) La *temperatura* en una región del espacio $A \subset \mathbb{R}^3$ es una función escalar de 3 variables con dominio A . Otros ejemplos de funciones de \mathbb{R}^3 en \mathbb{R} son los potenciales eléctrico y gravitatorio (estáticos).
- ii) La *posición*, la *velocidad* y la *aceleración* de una partícula son ejemplos de funciones de \mathbb{R} en \mathbb{R}^3 , es decir funciones vectoriales de una variable real (el tiempo). En este caso, la imagen de la función posición es una curva en \mathbb{R}^3 llamada *trayectoria* de la partícula. Análogamente, la *línea de universo* de una partícula relativista es una función de \mathbb{R} (por ejemplo, el *tiempo propio* de la partícula) en \mathbb{R}^4 (el tiempo y las tres coordenadas espaciales medidas en un sistema de referencia inercial)
- iii) Un *campo magnético*, *eléctrico* o *gravitatorio estático* (independiente del tiempo) en una región del espacio $A \subset \mathbb{R}^3$ es un ejemplo de función de \mathbb{R}^3 (las tres coordenadas espaciales) en \mathbb{R}^3 (las tres componentes del campo) con dominio A . Otro ejemplo de función de \mathbb{R}^3 en \mathbb{R}^3 es la *velocidad* de un fluido estacionario que se mueve en una región del espacio. Un campo eléctrico, magnético o gravitatorio dependiente del tiempo viene descrito por una función de \mathbb{R}^4 (tres coordenadas espaciales más una temporal) en \mathbb{R}^3 (las tres componentes del campo).
- iv) La *métrica* de una región $A \subset \mathbb{R}^4$ del espacio-tiempo puede representarse por una función de \mathbb{R}^4 (las cuatro coordenadas espacio-temporales) en \mathbb{R}^{10} (las 10 componentes independientes de un tensor simétrico de orden 2 en 4 dimensiones³).

Una función de \mathbb{R} en \mathbb{R} se puede visualizar de manera relativamente sencilla dibujando su gráfica. En general, es mucho más complicado (si no directamente imposible) visualizar una función de \mathbb{R}^n en \mathbb{R}^m . Para ayudar a visualizar tales funciones, suelen ser de utilidad los conceptos de gráfica (análogo al caso de funciones de \mathbb{R} en \mathbb{R}) y conjunto de nivel definidos a continuación:

³Un tensor de este tipo puede en efecto representarse por una matriz simétrica 4×4 , que tiene 10 elementos de matriz independientes.

Definición 1.6. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ una función. La **gráfica** de f es el conjunto

$$\mathcal{G}(f) = \{(x, f(x)) \mid x \in \mathcal{D}(f)\} \subset \mathbb{R}^{m+n}.$$

Dado un vector $c \in \mathbb{R}^m$, el conjunto

$$L_c(f) = \{x \in \mathbb{R}^n \mid f(x) = c\} \subset \mathcal{D}(f) \subset \mathbb{R}^n$$

se denomina **conjunto de nivel** de f de valor c .

Es importante notar que la gráfica de f es un subconjunto de \mathbb{R}^{m+n} , mientras que los conjuntos de nivel de f están contenidos en \mathbb{R}^n . Es también evidente que $L_c(f) \neq \emptyset$ si y solo si $c \in \mathcal{R}(f)$, y que

$$\mathcal{R}(f) = \{c \in \mathbb{R}^m \mid L_c(f) \neq \emptyset\}, \quad \mathcal{D}(f) = \bigcup_{c \in \mathbb{R}^m} L_c(f) = \bigcup_{c \in \mathcal{R}(f)} L_c(f).$$

Obsérvese, por último, que la gráfica $\mathcal{G}(f)$ tiene la siguiente propiedad: si $x \in \mathbb{R}^n$ e $y_1, y_2 \in \mathbb{R}^m$, entonces

$$(x, y_1) \in \mathcal{G}(f), \quad (x, y_2) \in \mathcal{G}(f) \implies y_1 = y_2,$$

ya que $y_1 = y_2 = f(x)$. Recíprocamente, un subconjunto $A \subset \mathbb{R}^{m+n}$ es la gráfica de una función $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ si y solo si cumple la propiedad anterior, es decir si

$$(x, y_1) \in A, \quad (x, y_2) \in A \implies y_1 = y_2.$$

En tal caso $\mathcal{D}(f)$ es obviamente la proyección de A sobre el espacio \mathbb{R}^n de las n primeras componentes, mientras que $\mathcal{R}(f)$ es la proyección sobre el espacio \mathbb{R}^m de las últimas m componentes:

$$\mathcal{D}(f) = \{x \in \mathbb{R}^n \mid \exists y \in \mathbb{R}^m \text{ t.q. } (x, y) \in A\}, \quad \mathcal{R}(f) = \{y \in \mathbb{R}^m \mid \exists x \in \mathbb{R}^n \text{ t.q. } (x, y) \in A\}.$$

Ejemplo 1.7. Sea $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ una función escalar de dos variables. Entonces la gráfica de f es la *superficie* de \mathbb{R}^3 de ecuación

$$\boxed{z = f(x, y), \quad (x, y) \in \mathcal{D}(f)},$$

donde hemos denotado por (x, y, z) las coordenadas cartesianas en \mathbb{R}^3 . Análogamente, si $c \in \mathcal{R}(f)$ el conjunto de nivel $L_c(f)$ es el subconjunto de \mathbb{R}^2 de ecuación

$$\boxed{f(x, y) = c},$$

que en general es una *curva* en el plano (que puede degenerar, en algunos casos, a un punto).

Consideremos, por ejemplo, la función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por

$$\boxed{f(x, y) = \sqrt{\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1}},$$

con a y b constantes positivas. En primer lugar, el dominio de f es el conjunto

$$\boxed{\mathcal{D}(f) = \left\{ (x, y) \in \mathbb{R}^2 \mid \frac{x^2}{a^2} + \frac{y^2}{b^2} \geq 1 \right\}},$$

es decir la unión de la elipse $(x^2/a^2) + (y^2/b^2) = 1$ de semiejes a y b con su exterior. Es también claro que $\mathcal{R}(f) = [0, \infty)$. En efecto, por definición $f(x, y) \geq 0$, y si $c \geq 0$ la ecuación

$$f(x, y) = c \iff \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = c^2$$

tiene infinitas soluciones (por ejemplo, $y = 0$ y $x = \pm a\sqrt{1+c^2}$). El cálculo anterior también demuestra que $L_c(f) = \emptyset$ si $c < 0$, y $L_c(f) \subset \mathbb{R}^2$ es la elipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 + c^2,$$

de semiejes $a\sqrt{1+c^2}$ y $b\sqrt{1+c^2}$. Por último, la gráfica de f es el conjunto

$$\mathcal{G}(f) = \left\{ \left(x, y, \sqrt{\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1} \right) \mid (x, y) \in \mathcal{D}(f) \right\} \subset \mathbb{R}^3,$$

o equivalentemente

$$\boxed{\mathcal{G}(f) = \left\{ (x, y, z) \in \mathbb{R}^3 \mid \frac{x^2}{a^2} + \frac{y^2}{b^2} - z^2 = 1, z \geq 0 \right\}},$$

es decir la intersección del *hiperboloide elíptico de una hoja*

$$H = \left\{ (x, y, z) \in \mathbb{R}^3 \mid \frac{x^2}{a^2} + \frac{y^2}{b^2} - z^2 = 1 \right\}$$

con el semiespacio $z \geq 0$. Si $a = b$ entonces H es un hiperboloide de revolución alrededor del eje z , mientras que si $a \neq b$ las secciones de H (y por tanto $\mathcal{G}(f)$) por planos horizontales $z = c \geq 0$ son elipses de semiejes $a\sqrt{1+c^2}$ y $b\sqrt{1+c^2}$ (cf. la Fig. 1.4).

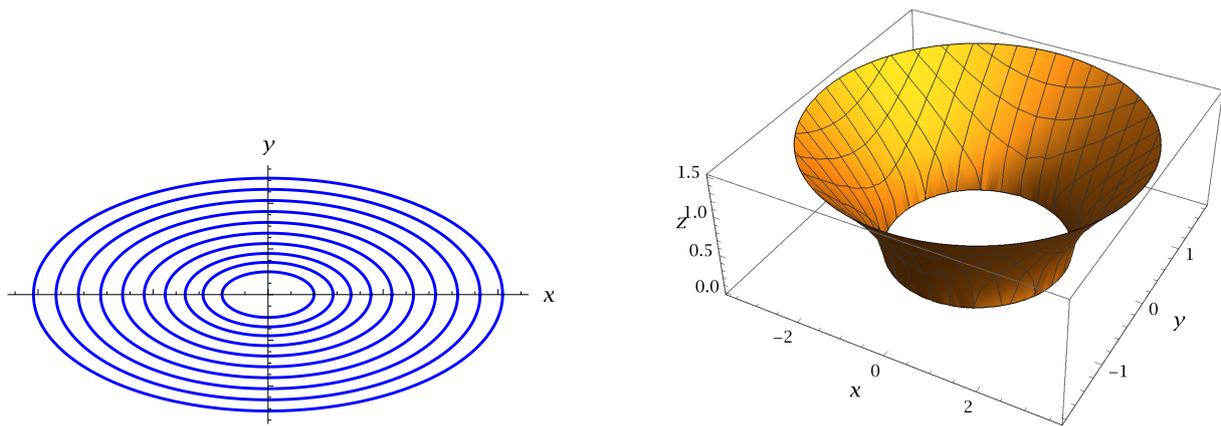


Figura 1.4: Izda.: curvas de nivel de la función del Ejemplo (1.7) (con $a = 2$, $b = 1$). Drcha.: gráfica de dicha función.

Ejercicio 4. Repetir el problema anterior para la función

$$f(x, y) = \sqrt{\frac{x^2}{a^2} + \frac{y^2}{b^2}} + 1$$

Solución. En este caso

$$\boxed{\mathcal{D}(f) = \mathbb{R}^2, \quad \mathcal{R}(f) = [1, \infty), \quad \mathcal{G}(f) = H' \cap \{(x, y, z) \in \mathbb{R}^3 \mid z \geq 0\}}$$

siendo

$$H' = \left\{ (x, y, z) \in \mathbb{R}^3 \mid \frac{x^2}{a^2} + \frac{y^2}{b^2} - z^2 = -1 \right\}.$$

un *hiperboloide elíptico de dos hojas* (cf. la Fig. 1.5). Si $c \geq 1$, los conjuntos (curvas) de nivel son las elipses

$$\boxed{\frac{x^2}{a^2} + \frac{y^2}{b^2} = c^2 - 1},$$

de semiejes $a\sqrt{c^2-1}$ y $b\sqrt{c^2-1}$ (en particular, $L_1(f) = \{(0, 0)\}$), mientras que $L_c(f) = \emptyset$ si $c < 1$.

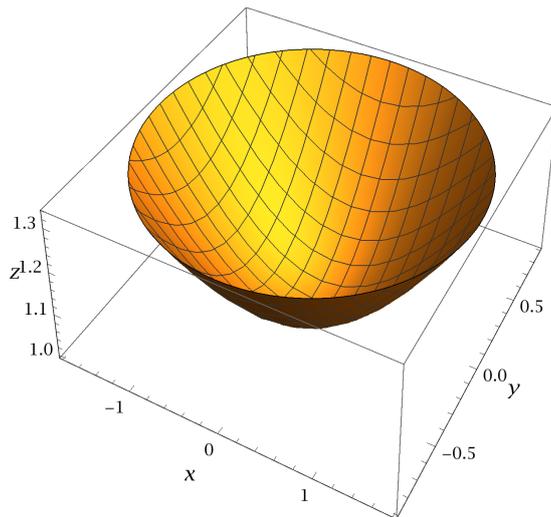


Figura 1.5: gráfica de la función del Ejercicio 4 (con $a = 1, b = 2$).

Ejemplo 1.8. Consideremos la función $f : \mathbb{R} \rightarrow \mathbb{R}^2$ dada por

$$f(t) = (\cos t, \sin t).$$

En este caso $\mathcal{D}(f) = \mathbb{R}$ y $\mathcal{R}(f)$ es la circunferencia de centro 0 y radio 1 en \mathbb{R}^2 . La gráfica de f es la curva en \mathbb{R}^3

$$\mathcal{G}(f) = \{(t, \cos t, \sin t) \mid t \in \mathbb{R}\}.$$

Se trata de una *hélice* contenida en el cilindro $y^2 + z^2 = 1$ y de paso 2π (el período de las funciones \cos y \sin).

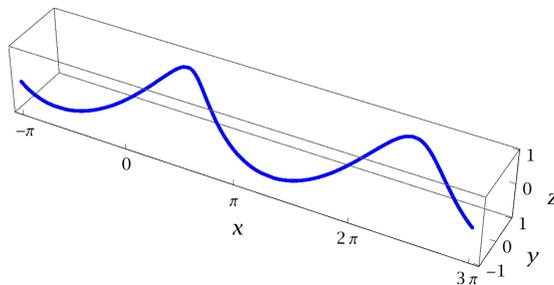


Figura 1.6: gráfica de la función del Ejemplo 1.8.

Comentario. Nótese que no es de interés considerar los conjuntos de nivel de una función vectorial de una variable, ya que son en general conjuntos de puntos aislados de la recta real. Así, en el ejemplo anterior

$$L_{(1,0)}(f) = \{2k\pi \mid k \in \mathbb{Z}\}.$$

1.3 Espacios métricos, normados y euclidianos

Para el estudio de las funciones de varias variables son fundamentales las propiedades métricas del espacio \mathbb{R}^n , que resumiremos en este apartado.

Definición 1.9. Un **espacio métrico** es un conjunto M provisto de una aplicación $d : M \times M \rightarrow \mathbb{R}$, llamada **distancia**, que verifica las siguientes propiedades:

- i) $d(x, y) \geq 0$ para todo $x, y \in M$, y $d(x, y) = 0$ si y solo si $x = y$.
- ii) $d(x, y) = d(y, x)$, $\forall x, y \in M$.
- iii) $d(x, y) \leq d(x, z) + d(z, y)$, $\forall x, y, z \in M$ (*desigualdad triangular*).

Nótese que en la definición *no* se pide que el conjunto M sea un espacio vectorial. Por ejemplo, *cualquier* subconjunto de la recta real con la distancia $d(x, y) = |x - y|$ es un espacio métrico, debido esencialmente a la propiedad

$$|x + y| \leq |x| + |y|$$

del valor absoluto (demostrarlo). Los ejemplos más importantes de espacios métricos están dados por los espacios vectoriales provistos de una norma, que definimos a continuación:

Definición 1.10. Un **espacio vectorial normado** es un espacio vectorial real V provisto de una aplicación $\| \cdot \| : V \rightarrow \mathbb{R}$, llamada **norma**, que satisface las propiedades siguientes:

- i) $\|x\| \geq 0 \forall x \in V$, y $\|x\| = 0$ si y solo si $x = 0$.
- ii) $\|\lambda x\| = |\lambda| \|x\|$, $\forall \lambda \in \mathbb{R}$, $\forall x \in V$.
- iii) $\|x + y\| \leq \|x\| + \|y\|$, $\forall x, y \in V$ (*desigualdad triangular*).

Es claro que \mathbb{R} , con la norma definida por $\|x\| = |x|$, es un ejemplo de espacio normado.

Es inmediato comprobar que *todo espacio normado es automáticamente un espacio métrico*, si definimos la distancia entre dos elementos x, y de dicho espacio mediante

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

En efecto, las propiedades i) y ii) de los espacios métricos se cumplen obviamente en virtud de las dos primeras propiedades de la norma, ya que

$$d(y, x) = \|y - x\| = \|(x - y)\| = |-1| \|x - y\| = \|x - y\|.$$

La desigualdad triangular también se verifica en virtud de la tercera propiedad de la norma, al ser

$$d(x, y) = \|x - z + z - y\| \leq \|x - z\| + \|z - y\| = d(x, z) + d(z, y).$$

Ejercicio 5. Si V es un espacio normado y $x, y \in V$, probar que

$$|\|x\| - \|y\|| \leq \|x - y\|.$$

Solución. La desigualdad propuesta es equivalente a las dos desigualdades

$$-\|x - y\| \leq \|x\| - \|y\| \leq \|x - y\|,$$

cualquiera de las cuales es consecuencia inmediata de la desigualdad triangular de la norma. Por ejemplo,

$$\|x\| \equiv \|x - y + y\| \leq \|x - y\| + \|y\| \implies \|x\| - \|y\| \leq \|x - y\|. \quad \square$$

(De hecho, la desigualdad $-\|x - y\| \leq \|x\| - \|y\|$ se obtiene de la anterior intercambiando x con y .)

Uno de los ejemplos más importantes de espacio normado es el de un espacio euclidiano, que por definición es un espacio vectorial real dotado de un producto escalar. Más precisamente:

Definición 1.11. Un **producto escalar** en un espacio vectorial real V es una aplicación $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$ que verifica:

- i) $\langle \lambda x + \mu y, z \rangle = \lambda \langle x, z \rangle + \mu \langle y, z \rangle, \quad \forall x, y, z \in V, \forall \lambda, \mu \in \mathbb{R}.$
- ii) $\langle x, y \rangle = \langle y, x \rangle, \quad \forall x, y \in V.$
- iii) $\langle x, x \rangle \geq 0 \quad \forall x \in V, \text{ y } \langle x, x \rangle = 0 \text{ si y solo si } x = 0.$

Un espacio vectorial **euclidiano** es un espacio vectorial real dotado de un producto escalar.

Nótese que de las propiedades i) y ii) se sigue que

$$\langle x, \lambda y + \mu z \rangle = \lambda \langle x, y \rangle + \mu \langle x, z \rangle, \quad \forall x, y, z \in V, \forall \lambda, \mu \in \mathbb{R};$$

en otras palabras, *el producto escalar es una aplicación bilineal* (es decir, lineal en cada uno de sus argumentos).

El ejemplo más importante de espacio euclidiano es el espacio \mathbb{R}^n , con el producto escalar definido por

$$\langle x, y \rangle = x \cdot y \equiv \sum_{i=1}^n x_i y_i, \quad (1.7)$$

donde $x \equiv (x_1, \dots, x_n), y \equiv (y_1, \dots, y_n)$. En general, *un espacio euclidiano V es automáticamente un espacio normado* (y, por tanto, *métrico*), definiendo la norma de un vector x mediante

$$\|x\| = \sqrt{\langle x, x \rangle}, \quad \forall x \in V. \quad (1.8)$$

En efecto, es inmediato comprobar que con la definición anterior de norma se verifican las propiedades i) y ii) de la norma. La propiedad iii) es consecuencia de la **desigualdad de Cauchy-Schwarz**

$$|\langle x, y \rangle| \leq \|x\| \|y\|, \quad \forall x, y \in V, \quad (1.9)$$

válida en cualquier espacio euclidiano, que demostraremos más abajo. En efecto, de (1.9) se sigue inmediatamente que

$$\|x + y\|^2 \equiv \langle x + y, x + y \rangle = \|x\|^2 + \|y\|^2 + 2\langle x, y \rangle \leq \|x\|^2 + \|y\|^2 + 2\|x\| \|y\| = (\|x\| + \|y\|)^2,$$

que es equivalente a la propiedad iii) de los espacios normados.

Dada su importancia, demostraremos a continuación la desigualdad de Cauchy-Schwarz. En primer lugar, si $y = 0$ ambos miembros de (1.9) son iguales (a cero), y por tanto la desigualdad se satisface. Supongamos a continuación que $y \neq 0$. Si $x \in \mathbb{R}^n$ es un vector cualquiera, consideremos la función $f : \mathbb{R} \rightarrow \mathbb{R}$ definida por

$$f(t) = \langle x + ty, x + ty \rangle \equiv \|x + ty\|^2, \quad \forall t \in \mathbb{R}.$$

Aplicando las propiedades del producto escalar y la definición (1.8) obtenemos

$$f(t) = \|x\|^2 + 2t\langle x, y \rangle + t^2\|y\|^2 = \left(t\|y\| + \frac{\langle x, y \rangle}{\|y\|} \right)^2 + \|x\|^2 - \frac{\langle x, y \rangle^2}{\|y\|^2}. \quad (1.10)$$

Si $t_0 = -\langle x, y \rangle / \|y\|^2$ (que está bien definido, al ser $\|y\| > 0$ en virtud de la propiedad iii) del producto escalar) entonces

$$f(t_0) = \|x\|^2 - \frac{\langle x, y \rangle^2}{\|y\|^2} \geq 0,$$

ya que $f(t) \geq 0$ para todo t en virtud de la tercera propiedad del producto escalar. Multiplicando la desigualdad anterior por $\|y\|^2 > 0$ se obtiene inmediatamente

$$\langle x, y \rangle^2 \leq (\|x\| \|y\|)^2,$$

que es equivalente a (1.9).

Ejercicio 6. Probar que la desigualdad de Cauchy-Schwarz es una igualdad si y solo si los vectores x e y son *linealmente dependientes* (es decir, uno de ellos es proporcional al otro).

Solución. En efecto, si los vectores x, y son l. d. entonces (por ejemplo) $y = \lambda x$ para algún $\lambda \in \mathbb{R}$, y se tiene

$$|\langle x, y \rangle| = |\lambda| \|x\|^2 = \|x\| \|y\|.$$

Recíprocamente, supongamos que $|\langle x, y \rangle| = \|x\| \|y\|$. Si $y = 0$, los vectores x, y son l. d. Si, por el contrario, $y \neq 0$, de la ec. (1.10) se sigue que $\|x + t_0 y\|^2 = 0$, siendo $t_0 \equiv -\langle x, y \rangle^2 / \|y\|^2$. Por tanto en este caso $x + t_0 y = 0$, y los vectores x, y son linealmente dependientes. \square

• La desigualdad de Cauchy-Schwarz permite definir el **ángulo** $\theta(x, y) \in [0, \pi]$ entre dos vectores no nulos x, y pertenecientes a un espacio euclidiano V (en particular, a \mathbb{R}^n) mediante

$$\cos \theta(x, y) = \frac{\langle x, y \rangle}{\|x\| \|y\|}$$

(cf. la ec. (1.1)), ya que el miembro derecho es siempre menor o igual que 1 en valor absoluto. Por el ejercicio anterior, $|\cos \theta(x, y)| = 1$ (es decir, $\theta(x, y) = 0$ o $\theta(x, y) = \pi$) si y solo si $y = \lambda x$ con $\lambda \neq 0$, es decir si y solo si x e y son paralelos. Más precisamente, al ser en tal caso $\cos \theta(x, y) = \operatorname{sgn} \lambda$, $\theta(x, y) = 0$ si y solo si x e y tienen el mismo sentido, y $\theta(x, y) = \pi$ si y solo si x e y tienen sentidos opuestos.

• Al ser \mathbb{R}^n un espacio euclidiano, por lo que acabamos de ver es automáticamente un espacio normado y métrico. Nótese que, en virtud de la definición (1.7), la norma de un vector $x \equiv (x_1, \dots, x_n) \in \mathbb{R}^n$ se define por

$$\|x\| = \sqrt{\sum_{i=1}^n x_i^2}, \quad (1.11)$$

y la distancia entre dos puntos x e $y \equiv (y_1, \dots, y_n) \in \mathbb{R}^n$ está dada por

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}. \quad (1.12)$$

• Nótese también que de la definición de norma euclidiana (1.11) se deduce inmediatamente que

$$|x_i| \leq \|x\|,$$

desigualdad que utilizaremos a menudo en lo que sigue.

Ejercicio 7. Probar que si $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ entonces $\|x\| \leq \sqrt{n} \max_{1 \leq i \leq n} |x_i|$.

Ejercicio 8. Sea $C[0, 1]$ el espacio vectorial de las funciones continuas en el intervalo $[0, 1]$. i) Probar que $C[0, 1]$, con el producto escalar

$$\langle f, g \rangle = \int_0^1 f(x)g(x) dx, \quad \forall f, g \in C[0, 1],$$

es un espacio euclidiano. ii) Si $a, b > 0$, calcular la norma de la función $f(x) = x^a \in C[0, 1]$ y el ángulo formado por las funciones x^a y x^b . iii) Si p y q son dos números naturales distintos, probar que las funciones $\sin(2\pi p x)$ y $\sin(2\pi q x)$ son ortogonales.

1.4 Interior, exterior, frontera

Con ayuda de la métrica euclidiana (1.12) del espacio \mathbb{R}^n , definiremos en este apartado varios conceptos topológicos esenciales para entender las nociones de límite y continuidad en funciones de varias variables. El objeto fundamental en que nos basaremos es el de bola abierta, definido a continuación:

Definición 1.12. Si $a \in \mathbb{R}^n$ y $r \in \mathbb{R}_+$ es un número real positivo, la **bola abierta** de centro a y radio r es el conjunto

$$B_r(a) = \{x \in \mathbb{R}^n \mid d(x, a) < r\} \equiv \{x \in \mathbb{R}^n \mid \|x - a\| < r\}. \quad (1.13)$$

La **bola abierta perforada** de centro $a \in \mathbb{R}^n$ y radio $r \in \mathbb{R}_+$ es el conjunto

$$B_r^*(a) = B_r(a) - \{a\} = \{x \in \mathbb{R}^n \mid 0 < d(x, a) < r\} = \{x \in \mathbb{R}^n \mid 0 < \|x - a\| < r\}. \quad (1.14)$$

En otras palabras, $B_r(a)$ es el conjunto de puntos de \mathbb{R}^n cuya distancia al punto a es *estrictamente* menor que r . Es evidente que la definición anterior es válida en un espacio métrico cualquiera, ya que se basa exclusivamente en la noción de distancia.

• En \mathbb{R} la bola abierta $B_r(a)$ es el intervalo abierto $(a - r, a + r)$ centrado en a , en \mathbb{R}^2 es el disco abierto de centro a y radio r , y en \mathbb{R}^3 es el interior de la esfera de centro a y radio r .

Ejercicio 9. Si $x, y \in B_r(a)$, probar que el **segmento** con extremos x e y , definido por

$$[x, y] \equiv \{x + s(y - x) \mid 0 \leq s \leq 1\}, \quad (1.15)$$

está contenido en $B_r(a)$.

Solución. En efecto, si $s = 0$ (resp. $s = 1$) entonces $x + s(y - x) = x \in B_r(a)$ (resp. $x + s(y - x) = y \in B_r(a)$). Por otra parte, si $0 < s < 1$ se tiene

$$\|x + s(y - x) - a\| = \|(1 - s)(x - a) + s(y - a)\| \leq (1 - s)\|x - a\| + s\|y - a\| < (1 - s)r + sr = r.$$

Con la ayuda de las bolas abiertas, dado un subconjunto $A \subset \mathbb{R}^n$ podemos clasificar los puntos de \mathbb{R}^n en tres tipos mutuamente disjuntos. Más concretamente:

Definición 1.13. Sea A un subconjunto de \mathbb{R}^n , y $x \in \mathbb{R}^n$.

i) Diremos que x es un punto **interior** de A si existe $r > 0$ tal que

$$B_r(x) \subset A.$$

ii) Diremos que x es un punto **exterior** de A si x es un punto interior de $\mathbb{R}^n - A$, es decir si existe $r > 0$ tal que

$$B_r(x) \cap A = \emptyset.$$

iii) Diremos que x es un punto **frontera** de A si x no es ni interior ni exterior a A , es decir si para todo $r > 0$

$$B_r(x) \cap A \neq \emptyset, \quad B_r(x) \cap (\mathbb{R}^n - A) \neq \emptyset.$$

Dado que la condición ii) de la definición anterior es equivalente a la relación

$$B_r(x) \subset \mathbb{R}^n - A,$$

es evidente que un punto cualquiera $x \in \mathbb{R}^n$ solo puede satisfacer *exactamente una* de las tres condiciones de la Definición 1.13. Esto conduce a la siguiente definición.

Definición 1.14. Sea, de nuevo, $A \subset \mathbb{R}^n$.

i) El **interior** de A es el conjunto $\overset{\circ}{A}$ de los puntos interiores de A , es decir

$$\overset{\circ}{A} = \{x \in \mathbb{R}^n \mid x \text{ punto interior de } A\}$$

ii) El **exterior** de A es el conjunto $\text{ext } A$ de los puntos exteriores de A , es decir

$$\text{ext } A = \{x \in \mathbb{R}^n \mid x \text{ punto exterior de } A\}.$$

iii) La **frontera** de A es el conjunto ∂A de los puntos frontera de A , es decir

$$\partial A = \{x \in \mathbb{R}^n \mid x \text{ punto frontera de } A\}.$$

Por lo visto anteriormente, \mathbb{R}^n es la unión disjunta de los conjuntos $\overset{\circ}{A}$, $\text{ext } A$ y ∂A , es decir

$$\mathbb{R}^n = \overset{\circ}{A} \cup \partial A \cup \text{ext } A, \quad \overset{\circ}{A} \cap \partial A = \overset{\circ}{A} \cap \text{ext } A = \partial A \cap \text{ext } A = \emptyset. \quad (1.16)$$

Nótese que

$$\overset{\circ}{A} \subset A, \quad \text{ext } A \subset \mathbb{R}^n - A,$$

mientras que los puntos de ∂A pueden o no pertenecer a A según los casos. Es también evidente (demostrarlo) que

$$(\mathbb{R}^n - A)^\circ = \text{ext } A, \quad \text{ext}(\mathbb{R}^n - A) = \overset{\circ}{A}, \quad \partial(\mathbb{R}^n - A) = \partial A. \quad (1.17)$$

Ejemplo 1.15. Sea $A = B_r(a)$ la bola de centro $a \in \mathbb{R}^n$ y radio $r > 0$. Es intuitivamente claro (al menos si $n \leq 3$) que

$$(B_r(a))^\circ = B_r(a), \quad \partial B_r(a) = \{x \in \mathbb{R}^n \mid \|x - a\| = r\}, \quad \text{ext } B_r(a) = \{x \in \mathbb{R}^n \mid \|x - a\| > r\}.$$

Como ejercicio sobre las propiedades de la norma euclidiana de \mathbb{R}^n , probaremos a continuación estas igualdades. Como \mathbb{R}^n es la unión disjunta de los conjuntos $B_r(a)$, $\{x \in \mathbb{R}^n \mid \|x - a\| = r\}$ y $\{x \in \mathbb{R}^n \mid \|x - a\| > r\}$, para probar las igualdades anteriores basta demostrar que

$$B_r(a) \subset (B_r(a))^\circ, \quad \{x \in \mathbb{R}^n \mid \|x - a\| = r\} \subset \partial B_r(a), \quad \{x \in \mathbb{R}^n \mid \|x - a\| > r\} \subset \text{ext } B_r(a)$$

(¿por qué?). En primer lugar, si $x \in B_r(a)$ y $s = \|x - a\|$ es claro que $r - s > 0$ (por definición de $B_r(a)$) y $B_{r-s}(x) \subset B_r(a)$. En efecto, si $y \in B_{r-s}(x)$ entonces

$$\|y - a\| \leq \|y - x\| + \|x - a\| < r - s + s = r$$

en virtud de la desigualdad triangular de la norma (propiedad iii)). Esto demuestra que $B_r(a) \subset (B_r(a))^\circ$, es decir la primera igualdad. Para probar la segunda, consideremos un vector cualquiera x tal que $\|x - a\| = r$. Si $\varepsilon > 0$, los puntos de la forma $x(t) \equiv x + t(x - a)$ con $\max(-\varepsilon/r, -1) < t < 0$ pertenecen a $B_\varepsilon(x) \cap B_r(a)$, ya que

$$\|x(t) - x\| = |t|r = -tr < \varepsilon, \quad \|x(t) - a\| = |1 + t|r = (1 + t)r < r.$$

Esto demuestra que $B_\varepsilon(x) \cap B_r(a) \neq \emptyset$. Análogamente, si $0 < t < \varepsilon/r$ entonces $x(t) \in B_\varepsilon(x) \cap (\mathbb{R}^n - B_r(a))$, ya que

$$\|x(t) - x\| = tr < \varepsilon, \quad \|x(t) - a\| = (1 + t)r > r.$$

Por tanto

$$B_\varepsilon(x) \cap B_r(a) \neq \emptyset, \quad B_\varepsilon(x) \cap (\mathbb{R}^n - B_r(a)) \neq \emptyset,$$

lo que implica (al ser $\varepsilon > 0$ arbitrario) que $x \in \partial B_r(a)$. Por último, para establecer la tercera igualdad nótese que si $s \equiv \|x - a\| > r$ entonces $B_{s-r}(x) \cap B_r(a) = \emptyset$, ya que si $y \in B_{s-r}(x)$ se tiene

$$s = \|x - a\| \leq \|x - y\| + \|y - a\| < s - r + \|y - a\| \implies \|y - a\| > r. \quad \square$$

Ejercicio 10. Si $l > 0$, el **hipercubo abierto** $C_l(a)$ de centro $a = (a_1, \dots, a_n) \in \mathbb{R}^n$ y semilado l se define por

$$C_l(a) = \{x = (x_1, \dots, x_n) \in \mathbb{R}^n \mid |x_i - a_i| < l, 1 \leq i \leq n\} \equiv [x_1 - l, x_1 + l] \times \dots \times [x_n - l, x_n + l].$$

Probar que $C_{r/\sqrt{n}}(a) \subset B_r(a) \subset C_r(a)$.

Ejercicio 11. Sea \mathbb{Q}^n el subconjunto de \mathbb{R}^n de los puntos con coordenadas racionales, es decir

$$\mathbb{Q}^n = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid x_i \in \mathbb{Q}, 1 \leq i \leq n\}.$$

Hallar el interior, el exterior y la frontera de \mathbb{Q}^n . [Ayuda: Todo intervalo de la recta real contiene infinitos puntos racionales e irracionales.]

Dado un conjunto $A \subset \mathbb{R}^n$, definimos el **cierre** de A como el conjunto

$$\overline{A} = A \cup \partial A, \quad (1.18)$$

unión de A con sus puntos frontera. Nótese que, al ser

$$\overset{\circ}{A} \subset A \subset \overset{\circ}{A} \cup \partial A \implies \overset{\circ}{A} \cup \partial A = A \cup \partial A,$$

la definición anterior es equivalente a

$$\overline{A} = \overset{\circ}{A} \cup \partial A = \mathbb{R}^n - \text{ext } A, \quad (1.19)$$

y se verifica

$$\overset{\circ}{\overline{A}} \subset A \subset \overline{A}. \quad (1.20)$$

Por ejemplo, de (1.13) se sigue inmediatamente que el cierre de la bola abierta $B_a(r)$ es la **bola cerrada**

$$\overline{B_r(a)} = \{x \in \mathbb{R}^n \mid \|x - a\| \leq r\}. \quad (1.21)$$

Ejercicio 12. i) Dar un ejemplo de conjunto A para el cual $\partial A = \overline{B_r(a)}$; ii) estudiar si $\partial(\partial A) = \partial A$.

Otros tipos de puntos interesantes en relación con un conjunto $A \subset \mathbb{R}^n$ son los aislados y los de acumulación:

Definición 1.16. Sea A un subconjunto de \mathbb{R}^n . Se dice que $x \in \mathbb{R}^n$ es un **punto aislado** de A si existe $r > 0$ tal que $B_r(x) \cap A = \{x\}$. Un punto $x \in \mathbb{R}^n$ es **punto de acumulación** de A si $B_r^*(x) \cap A \neq \emptyset$ para todo $r > 0$.

Nótese que claramente

$$x \text{ punto aislado de } A \implies x \in A.$$

Sin embargo, *un punto de acumulación de A puede o no pertenecer a dicho conjunto*. Por ejemplo, los puntos de acumulación de $B_r(a)$ son los puntos de $B_r(a)$ junto con los de su frontera, y estos últimos no pertenecen a $B_r(a)$. Si denotamos respectivamente por A' e $i(A)$ a los conjuntos de puntos de acumulación y de puntos aislados de $A \subset \mathbb{R}^n$, es decir

$$A' = \{x \in \mathbb{R}^n \mid x \text{ pto. de acumulación de } A\}, \quad i(A) = \{x \in \mathbb{R}^n \mid x \text{ pto. aislado de } A\} \subset A,$$

es inmediato comprobar las siguientes inclusiones (ejercicio):

- i) $i(A) \subset \partial A$ (los puntos aislados son puntos frontera)
- ii) $\partial A \subset i(A) \cup A'$ (los puntos frontera son puntos aislados o de acumulación)
- iii) $\overset{\circ}{A} \subset A'$ (los puntos interiores son puntos de acumulación)
- iv) $A' \subset \overset{\circ}{A} \cup \partial A$ (los puntos de acumulación son puntos interiores o frontera)

Utilizando estas inclusiones, es fácil probar que

$$\boxed{\bar{A} = A \cup A'}. \tag{1.22}$$

En efecto, de la cuarta inclusión anterior se sigue que $A' \subset \bar{A}$, y por tanto

$$A \cup A' \subset \bar{A}.$$

Recíprocamente, de la segunda inclusión y de $i(A) \subset A$ se deduce que

$$\bar{A} = A \cup \partial A \subset A \cup (i(A) \cup A') = A \cup A'.$$

Ejercicio 13. Demostrar que el cierre de A es la unión de los puntos aislados de A con sus puntos de acumulación, es decir

$$\boxed{\bar{A} = i(A) \cup A'}.$$

Del ejercicio anterior se sigue fácilmente que

$$\boxed{i(A) \cup A' \cup \text{ext } A = \mathbb{R}^n},$$

siendo los tres conjuntos del miembro izquierdo disjuntos dos a dos.

Ejercicio 14. Si $x \in \mathbb{R}^n$ es un punto de acumulación de $A \subset \mathbb{R}^n$, probar que toda bola $B_r(x)$ centrada en x contiene infinitos puntos de A .

Solución. Supongamos que la afirmación anterior fuera falsa, es decir que existiera un cierto número real positivo r_0 tal que $B_{r_0}(x) \cap A$ fuera un subconjunto finito de A . Entonces $B_{r_0}^*(x) \cap A = \{a_1, \dots, a_k\}$ sería también finito, y no vacío por definición de punto de acumulación. Si tomamos cualquier número r tal que

$$0 < r < \min(\|x - a_1\|, \dots, \|x - a_k\|),$$

(lo cual es posible, ya que $a_i \neq x$ para todo i) entonces $B_r^*(x) \cap A = \emptyset$, en contradicción con la definición de punto de acumulación. \square

Una propiedad elemental pero importante del interior y el cierre de un conjunto es la siguiente:

$$\boxed{A \subset B \subset \mathbb{R}^n \implies \overset{\circ}{A} \subset \overset{\circ}{B}, \quad A' \subset B', \quad \bar{A} \subset \bar{B}}. \tag{1.23}$$

En efecto, la primera de estas inclusiones es obvia, ya que si $x \in \overset{\circ}{A}$ por definición existe $r > 0$ tal que

$$B_r(x) \subset A \subset B,$$

y por tanto $x \in \overset{\circ}{B}$. Por otra parte, si $x \in A'$ para todo $r > 0$ se tiene

$$B_r^*(x) \cap B \supset B_r^*(x) \cap A \neq \emptyset,$$

y por tanto $x \in B'$. Esto demuestra que $A' \subset B'$, de donde se deduce que

$$\bar{A} = A \cup A' \subset B \cup B' = \bar{B}.$$

1.5 Conjuntos abiertos, cerrados y compactos

Definición 1.17. Un subconjunto $A \subset \mathbb{R}^n$ es **abierto** si $A = \overset{\circ}{A}$. Diremos que $A \subset \mathbb{R}^n$ es **cerrado** si $\mathbb{R}^n - A$ es abierto.

Como $\overset{\circ}{\overset{\circ}{A}} \subset A$, la igualdad $\overset{\circ}{A} = A$ es equivalente a $A \subset \overset{\circ}{A}$. En otras palabras,

$A \subset \mathbb{R}^n$ es abierto si todos sus puntos son puntos interiores.

Equivalentemente,

$A \subset \mathbb{R}^n$ es abierto si para todo $a \in A$ existe $r > 0$ tal que $B_r(a) \subset A$.

Definición 1.18. Un **entorno** de un punto $x \in \mathbb{R}^n$ es cualquier conjunto abierto que contiene a x . Si A es un entorno de x , diremos que $A - \{x\}$ es un **entorno perforado** (o **reducido**) de dicho punto.

Con esta definición, es inmediato comprobar (ejercicio) la siguiente caracterización de los conjuntos abiertos:

$A \subset \mathbb{R}^n$ es abierto si y solo si contiene un entorno de cualquiera de sus puntos.

Por último, nótese que al ser $A \subset \overset{\circ}{A} \cup \partial A$ y $\overset{\circ}{A} \cap \partial A = \emptyset$ es claro que $A \subset \overset{\circ}{A}$ si y solo si $A \cap \partial A = \emptyset$. En otras palabras,

$A \subset \mathbb{R}^n$ es abierto si y solo no contiene a ninguno de sus puntos frontera.

Por definición, $A \subset \mathbb{R}^n$ es **cerrado** si $(\mathbb{R}^n - A)^\circ = \mathbb{R}^n - A$. En virtud de (1.17), esto es equivalente a $\text{ext } A = \mathbb{R}^n - A$, o bien, teniendo en cuenta (1.16) y (1.19),

$$A = \mathbb{R}^n - \text{ext } A = \overset{\circ}{A} \cup \partial A \equiv \bar{A}.$$

Por tanto

$A \subset \mathbb{R}^n$ es cerrado si y solo si $A = \bar{A}$.

De (1.22) se sigue que

$A \subset \mathbb{R}^n$ es cerrado si y solo si $A' \subset A$

o, en otras palabras,

$A \subset \mathbb{R}^n$ es cerrado si y solo si contiene a sus puntos de acumulación.

Análogamente, de (1.18) se sigue que

$A \subset \mathbb{R}^n$ es cerrado si y solo si contiene a sus puntos frontera.

Comentarios.

- Es importante darse cuenta de que en \mathbb{R}^n un conjunto puede no ser ni abierto ni cerrado, como ya ocurría en \mathbb{R} . Por ejemplo, $A = B_r(a) \cup \{x\}$, donde x es cualquier punto a distancia r del centro a , no es ni abierto ni cerrado. En efecto,

$$\overset{\circ}{A} = B_r(a) \subsetneq A \subsetneq \bar{A} = \bar{B}_r(a).$$

- El conjunto vacío y el espacio \mathbb{R}^n son claramente abiertos, ya que en ambos casos es obvio que el interior del conjunto coincide con el propio conjunto. Son también cerrados, ya que son uno el complementario del otro y son ambos abiertos. Por tanto \emptyset y \mathbb{R}^n son a la vez abiertos y cerrados. Puede demostrarse (aunque no es sencillo) que estos son los únicos subconjuntos de \mathbb{R}^n que tienen esta propiedad.

- Si $A \subset \mathbb{R}^n$ es un conjunto arbitrario, entonces $\overset{\circ}{A}$ es abierto y \bar{A} es cerrado. En efecto, supongamos en primer lugar que $a \in \overset{\circ}{A}$. Por hipótesis, existe $r > 0$ tal que $B_r(a) \subset A$, y de (1.23) se sigue entonces que

$$(B_r(a))^\circ = B_r(a) \subset \overset{\circ}{A}.$$

Esto demuestra que $\overset{\circ}{A}$ es abierto, como habíamos afirmado. En cuanto a la segunda afirmación, basta notar que

$$\mathbb{R}^n - \bar{A} = \text{ext } A = (\mathbb{R}^n - A)^\circ$$

es abierto, en virtud de la primera afirmación (que acabamos de demostrar).

Ejemplo 1.19. La bola abierta $B_r(a)$ es un conjunto abierto, mientras que la bola cerrada $\bar{B}_r(a)$ es un conjunto cerrado. En efecto, del Ejemplo 1.15 se sigue que $(B_r(a))^\circ = B_r(a)$, y por tanto $B_r(a)$ es abierto, mientras que $\bar{B}_r(a) = \bar{B}_r(A)$ es cerrada en virtud del comentario anterior.

Utilizando la definición 1.17 es fácil probar las siguientes propiedades de los conjuntos abiertos:

- i) El conjunto vacío y el espacio \mathbb{R}^n son abiertos.
- ii) La unión de un número *arbitrario* (finito o infinito) de conjuntos abiertos es un conjunto abierto.
- iii) La intersección de un número *finito* de conjuntos abiertos es un conjunto abierto.

Comentarios.

- Es importante observar, en relación con la propiedad iii), que la intersección de un número *infinito* de conjuntos abiertos no necesariamente es abierto. Por ejemplo, si $x \in \mathbb{R}^n$ se tiene

$$\bigcap_{n=1}^{\infty} B_{1/n}(x) = \{x\},$$

que obviamente *no* es un conjunto abierto.

- En general, una **topología** en un conjunto arbitrario X es una familia \mathcal{T} de subconjuntos de X que cumplen las propiedades i)-iii) anteriores. Un **espacio topológico** es un conjunto arbitrario X en el que se ha seleccionado una topología \mathcal{T} . Los conjuntos abiertos del espacio topológico son por definición los elementos de la topología, y sus complementarios son los conjuntos cerrados. De las propiedades i)-iii) de los conjuntos abiertos de \mathbb{R}^n que acabamos de enunciar se sigue por tanto que \mathbb{R}^n es un espacio topológico.

De las propiedades i)-iii) de los conjuntos abiertos y las *identidades de De Morgan* (ejercicio)

$$\mathbb{R}^n - \bigcup_{i \in I} A_i = \bigcap_{i \in I} (\mathbb{R}^n - A_i), \quad \mathbb{R}^n - \bigcap_{i \in I} A_i = \bigcup_{i \in I} (\mathbb{R}^n - A_i),$$

donde I es un conjunto de índices (finito o infinito) y $A_i \subset \mathbb{R}^n$ para todo $i \in I$, se siguen inmediatamente las siguientes propiedades elementales de los conjuntos cerrados:

- i) El conjunto vacío y el espacio \mathbb{R}^n son cerrados.
- ii) La intersección de un número *arbitrario* (finito o infinito) de conjuntos cerrados es un conjunto cerrado.
- iii) La unión de un número *finito* de conjuntos cerrados es un conjunto cerrado.

Como en el caso de los conjuntos abiertos, la última propiedad no se cumple necesariamente para uniones arbitrarias. Por ejemplo,

$$\bigcup_{0 < r < 1} \bar{B}_r(x) = B_1(x),$$

que *no* es cerrado.

Ejercicio 15. Probar que la frontera de cualquier conjunto $A \subset \mathbb{R}^n$ es cerrada.

Definición 1.20. Un conjunto $A \subset \mathbb{R}^n$ es **acotado** si existe $R > 0$ tal que $A \subset B_R(0)$. Diremos que A es **compacto** si es a la vez *cerrado* y *acotado*.

Ejemplo 1.21. La bola cerrada $\bar{B}_r(a)$ es un conjunto compacto. En efecto, ya vimos en el Ejemplo (1.19) que es cerrada. Para probar que es acotada, basta notar que

$$x \in \bar{B}_r(a) \implies \|x\| = \|x - a + a\| \leq \|x - a\| + \|a\| \leq r + \|a\| \implies \bar{B}_r(a) \subset B_R(0)$$

si $R > r + \|a\|$. La bola abierta $B_r(a)$ *no* es compacta, ya que no es cerrada.

Ejercicio 16. Probar que si $A \subset \mathbb{R}^n$ es acotado, el cierre de A es compacto.

Solución. En efecto, al ser A acotado existe $R > 0$ tal que $A \subset B_R(0)$. Utilizando (1.23) se obtiene

$$\bar{A} \subset \overline{B_R(0)} = \bar{B}_R(0) \subset B_{R+1}(0),$$

y por consiguiente \bar{A} es acotado. Es también cerrado (al ser el cierre de un conjunto), y por lo tanto es compacto.

1.6 Límites

Como en el caso de funciones reales de una variable real, el cálculo diferencial de funciones de varias variables está basado en el concepto de límite, que definimos a continuación:

Definición 1.22. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, y supongamos que $a \in \mathbb{R}^n$ es un punto de acumulación del dominio de f . Diremos que un punto $b \in \mathbb{R}^m$ es el **límite** de f en el punto a , y escribiremos

$$\lim_{x \rightarrow a} f(x) = b,$$

si para todo $\varepsilon > 0$, existe $\delta(\varepsilon) > 0$ tal que

$$x \in \mathcal{D}(f), \quad 0 < \|x - a\| < \delta(\varepsilon) \implies \|f(x) - b\| < \varepsilon. \quad (1.24)$$

Comentario. Nótese que la variable x en la fórmula anterior es una variable *muda*. En otras palabras, podríamos haber utilizado en su lugar cualquier otra variable y, z , etc. que no tenga un significado especial, y escribir equivalentemente $\lim_{y \rightarrow a} f(y) = b$, $\lim_{z \rightarrow a} f(z) = b$, etc.

La definición anterior se puede formular de varias formas equivalentes. Una de las más concisas es la siguiente:

Proposición 1.23. *Sea $a \in \mathbb{R}^n$ un punto de acumulación del dominio de $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, y sea $b \in \mathbb{R}^m$. Entonces $\lim_{x \rightarrow a} f(x) = b$ si y solo si para todo entorno V del punto b , existe un entorno perforado U del punto a (que en general dependerá de U) tal que $f(U) \subset V$.*

Demostración. En primer lugar, nótese que la ecuación (1.24) es equivalente a

$$f(B_{\delta(\varepsilon)}^*(a)) \subset B_\varepsilon(b). \tag{1.25}$$

\Leftarrow) Al ser la bola $B_\varepsilon(b)$ un entorno de b , por hipótesis existe un entorno perforado U de a tal que $f(U) \subset B_\varepsilon(b)$. Por definición de entorno perforado, U contiene una bola perforada $B_{\delta(\varepsilon)}^*(a)$ de radio $\delta(\varepsilon) > 0$ suficientemente pequeño. Por tanto

$$f(B_{\delta(\varepsilon)}^*(a)) \subset f(U) \subset B_\varepsilon(b), \tag{1.26}$$

y se cumple (1.25) o, equivalentemente, (1.24).

\Rightarrow) Si V es un entorno de $b \in \mathbb{R}^m$, por definición existe $\varepsilon > 0$ tal que $B_\varepsilon(b) \subset V$. En virtud de (1.25), que es equivalente a (1.24), existe $\delta(\varepsilon) > 0$ tal que $f(B_{\delta(\varepsilon)}^*(a)) \subset B_\varepsilon(b)$. Por tanto $U = B_{\delta(\varepsilon)}^*(a)$ es un entorno reducido de a que cumple $f(U) \subset V$. \square

Comentarios.

- Intuitivamente,

$\lim_{x \rightarrow a} f(x) = b$ significa que $f(x)$ está *arbitrariamente* próximo al punto b cuando $x \neq a$ está *suficientemente* próximo al punto a .

- Es evidente de la definición que *la existencia y el valor del límite de f en a depende exclusivamente del comportamiento de f en un entorno reducido de a* (es decir, en puntos próximos al punto a pero distintos de dicho punto). En otras palabras, si $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es igual a f en una bola perforada $B_r^*(a)$, es decir si

$$f(x) = g(x), \quad \forall x \text{ t.q. } 0 < \|x - a\| < r,$$

entonces f y g tienen el mismo límite en a . En particular, *el valor que tome f en el punto a es totalmente irrelevante para la existencia o el valor del límite de f en dicho punto*. De hecho, *ni siquiera es necesario que f esté definida en a para que exista el límite de f en a* .

- Evidentemente, diremos que *no existe $\lim_{x \rightarrow a} f(x)$ si $\lim_{x \rightarrow a} f(x) \neq b$ para todo $b \in \mathbb{R}^m$* . Más precisamente, esto quiere decir lo siguiente: *dado cualquier punto $b \in \mathbb{R}^m$, existe un número $\varepsilon > 0$ tal que para todo $\delta > 0$ hay algún punto $x \in \mathcal{D}(f) - \{a\}$ tal que $\|x - a\| < \delta$ pero $\|f(x) - b\| \geq \varepsilon$* . En otras palabras, *hay puntos $x \in \mathcal{D}(f) - \{a\}$ arbitrariamente próximos al punto a cuya imagen está a una distancia mayor o igual que ε del punto b* .
- En la definición de $\lim_{x \rightarrow a} f(x)$ se pide que el punto a sea de *acumulación* del dominio de f para que la condición $f(U) \subset V$ no sea *vacua*, es decir exija *algo* a la función f . Esto es así porque si $a \in \mathcal{D}(f)'$ todo entorno perforado de a contiene necesariamente algún punto de $\mathcal{D}(f)$. Dicho de otra forma, si $a \notin \mathcal{D}(f)'$ entonces existe $r > 0$ tal que $B_r^*(a) \cap \mathcal{D}(f) = \emptyset$. Entonces $U = B_r^*(a)$ es un entorno reducido que cumple $f(U) = \emptyset \subset V$ para *cualquier* entorno V del punto b .

Ejercicio 17. Si $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, indicar cuáles de las siguientes afirmaciones son equivalentes a la Definición 1.22:

- i) $\forall \varepsilon > 0, \exists \delta \geq 0$ t.q. $(x \in \mathcal{D}(f), 0 < \|x - a\| < \delta) \Rightarrow \|f(x) - b\| < \varepsilon$.
- ii) $\forall \varepsilon > 0, \exists \delta > 0$ t.q. $(x \in \mathcal{D}(f), 0 < \|x - a\| < \delta) \Rightarrow \|f(x) - b\| \leq \varepsilon$.
- iii) $\forall \varepsilon > 0, \exists \delta > 0$ t.q. $(x \in \mathcal{D}(f), 0 < \|x - a\| \leq \delta) \Rightarrow \|f(x) - b\| < \varepsilon$.

Ejercicio 18. Probar que

$$\lim_{x \rightarrow a} f(x) = b \iff \lim_{x \rightarrow a} \|f(x) - b\| = 0. \quad \square$$

En la definición de $\lim_{x \rightarrow a} f(x)$ se habla de “el” límite de f . ¿Es posible que una función pueda tener varios límites distintos en un mismo punto a ? Como da a entender la definición de límite, esto *no* puede ocurrir:

Proposición 1.24 (unicidad del límite). *El límite de una función $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ en un punto $a \in \mathcal{D}(f)'$, si existe, es único. En otras palabras, si $b_1, b_2 \in \mathbb{R}^m$ verifican*

$$\lim_{x \rightarrow a} f(x) = b_1, \quad \lim_{x \rightarrow a} f(x) = b_2$$

entonces $b_1 = b_2$.

Demostración. En efecto, si $b_1 \neq b_2$ sea $\varepsilon = \|b_1 - b_2\|/2 > 0$. Entonces $B_\varepsilon(b_1) \cap B_\varepsilon(b_2) = \emptyset$, ya que

$$y \in B_\varepsilon(b_1) \cap B_\varepsilon(b_2) \implies \|b_1 - b_2\| \leq \|b_1 - y\| + \|y - b_2\| < 2\varepsilon \equiv \|b_1 - b_2\|.$$

Por definición de límite, existen sendos números $\delta_i(\varepsilon) > 0$ ($i = 1, 2$) tales que

$$f(B_{\delta_i(\varepsilon)}^*(a)) \subset B_\varepsilon(b_i), \quad i = 1, 2.$$

En particular, si $\delta = \min(\delta_1(\varepsilon), \delta_2(\varepsilon)) > 0$ y $x \in B_\delta^*(a) \cap \mathcal{D}(f)$ (¡nótese que tal x existe, al ser a punto de acumulación de $\mathcal{D}(f)$!) entonces $f(x)$ pertenece simultáneamente a $B_\varepsilon(b_1)$ y $B_\varepsilon(b_2)$. Pero esto es absurdo, ya que hemos visto que $B_\varepsilon(b_1) \cap B_\varepsilon(b_2) = \emptyset$. \square

El estudio de la existencia y el valor del límite de una función de \mathbb{R}^n en \mathbb{R}^m se puede reducir al de m funciones *escalares* de n variables (sus funciones componentes), como muestra la siguiente proposición:

Proposición 1.25. *Sea $f = (f_1, \dots, f_m) : \mathbb{R}^n \rightarrow \mathbb{R}^m$, y sean $a \in \mathbb{R}^n$ y $b = (b_1, \dots, b_m) \in \mathbb{R}^m$. Entonces*

$$\lim_{x \rightarrow a} f(x) = b \iff \lim_{x \rightarrow a} f_i(x) = b_i, \quad 1 \leq i \leq m.$$

Demostración. Nótese, en primer lugar, que si $a \in \mathcal{D}(f)'$ entonces $a \in \mathcal{D}(f_i)'$ para todo $i = 1, \dots, m$, en virtud de (1.6).

\implies) Si $\varepsilon > 0$, al ser $\lim_{x \rightarrow a} f(x) = b$ existe $\delta(\varepsilon) > 0$ tal que se verifica (1.24). Si $x \in \mathcal{D}(f)$ y $0 < \|x - a\| < \delta(\varepsilon)$ entonces $x \in \mathcal{D}(f_i)$ y

$$|f_i(x) - b_i| \leq \|f(x) - b\| < \varepsilon,$$

por lo que $\lim_{x \rightarrow a} f_i(x) = b_i$.

\impliedby) Si $\varepsilon > 0$, al ser $\lim_{x \rightarrow a} f_i(x) = b_i$ existe $\delta_i \equiv \delta_i(\varepsilon/\sqrt{m}) > 0$ tal que

$$x \in \mathcal{D}(f_i), \quad 0 < \|x - a\| < \delta_i \implies |f_i(x) - b_i| < \frac{\varepsilon}{\sqrt{m}}$$

para $1 \leq i \leq m$. Si tomamos $\delta(\varepsilon) = \min(\delta_1, \dots, \delta_m) > 0$ entonces

$$\begin{aligned} x \in \mathcal{D}(f), 0 < \|x - a\| < \delta(\varepsilon) &\implies x \in \mathcal{D}(f_i), 0 < \|x - a\| < \delta_i \\ &\implies |f_i(x) - b_i| < \frac{\varepsilon}{\sqrt{m}}, \quad \forall i = 1, \dots, m, \end{aligned}$$

y por tanto

$$\|f(x) - b\| = \sqrt{(f_1(x) - b_1)^2 + \dots + (f_m(x) - b_m)^2} < \sqrt{m \frac{\varepsilon^2}{m}} = \varepsilon. \quad \square$$

Proposición 1.26 (propiedades del límite). Sean $f, g : \mathbb{R}^n \rightarrow \mathbb{R}^m$, sea $a \in (\mathcal{D}(f) \cap \mathcal{D}(g))'$, y supongamos que existen $\lim_{x \rightarrow a} f(x) = l$ y $\lim_{x \rightarrow a} g(x) = l'$. Entonces se tiene:

- i) Para todo $\lambda, \mu \in \mathbb{R}$, $\lim_{x \rightarrow a} (\lambda f(x) + \mu g(x)) = \lambda l + \mu l'$.
- ii) $\lim_{x \rightarrow a} (f(x) \cdot g(x)) = l \cdot l'$.
- iii) $\lim_{x \rightarrow a} \|f(x)\| = \|l\|$.
- iv) Si $l \neq 0$, hay un entorno perforado de a en que f no se anula. Si, además, f es una función escalar (es decir, $m = 1$), hay un entorno perforado de a en que f tiene signo constante igual al del límite l .
- v) Si f es una función escalar y $l \neq 0$, $\lim_{x \rightarrow a} \frac{1}{f(x)} = \frac{1}{l}$.
- vi) f está **acotada** en un entorno de a . En otras palabras, existen $r > 0$ y $M > 0$ tales que $\|f(x)\| < M$ para todo $x \in \mathcal{D}(f) \cap B_r(a)$.

Demostración. Daremos solamente una idea esquemática de la demostración de estos resultados, que puede encontrarse en la mayoría de los libros de texto.

- i) $\|\lambda f(x) + \mu g(x) - \lambda l - \mu l'\| = \|\lambda(f(x) - l) + \mu(g(x) - l')\| \leq |\lambda| \|f(x) - l\| + |\mu| \|g(x) - l'\|$.
- vi) $\|f(x)\| \leq \|f(x) - l\| + \|l\|$.
- iii) $|\|f(x)\| - \|l\|| \leq \|f(x) - l\|$, por la desigualdad triangular de la norma (cf. el Ejercicio 5).
- iv) $\|f(x)\| \geq \|l\| - \|f(x) - l\|$, en virtud de iii). Por otra parte, si f es una función escalar de la definición de límite se sigue que

$$x \in \mathcal{D}(f), 0 < \|x - a\| < \delta(|l|/2) \implies l - \frac{|l|}{2} < f(x) < l + \frac{|l|}{2},$$

y por tanto $f(x)$ tiene el mismo signo que l para tales valores de x (¿por qué?).

- ii) En virtud de iv), existe $M > 0$ tal que $\|g(x)\| < M$ para todo $x \in \mathcal{D}(f)$ con $\|x - a\| < \delta$. Utilizando la desigualdad de Cauchy-Schwarz se obtiene entonces

$$\begin{aligned} |f(x) \cdot g(x) - l \cdot l'| &= |(f(x) - l) \cdot g(x) + l \cdot (g(x) - l')| \\ &\leq \|g(x)\| \|f(x) - l\| + \|l\| \|g(x) - l'\| < M \|f(x) - l\| + \|l\| \|g(x) - l'\|. \end{aligned}$$

- v) Como $\lim_{x \rightarrow a} |f(x)| = |l|$ en virtud de iii), si $x \in \mathcal{D}(f)$ y $0 < \|x - a\| < \delta(|l|/2)$ se verifica $|\|f(x)\| - |l|| < |l|/2$, y por tanto $|f(x)| > |l|/2 > 0$. Para tales x se tiene entonces

$$\left| \frac{1}{f(x)} - \frac{1}{l} \right| = \frac{|f(x) - l|}{|l| |f(x)|} \leq \frac{2}{|l|^2} \|f(x) - l\|. \quad \square$$

Comentarios.

- Es importante observar que puede existir alguno de los límites del miembro izquierdo de i)-iii) aun cuando no existan los límites de f o g en a . Así, por ejemplo, si no existe $\lim_{x \rightarrow a} f(x)$ y tomamos $g = -f$ y $\mu = \lambda = 1$, evidentemente sí existe el límite de $\lambda f + \mu g = f - f = 0$ en cualquier punto, y vale 0.
- Si $m = 1$, la propiedad ii) afirma simplemente que

$$\boxed{\lim_{x \rightarrow a} (f(x)g(x)) = ll'}, \quad (1.27)$$

donde en ambos miembros el producto es la multiplicación ordinaria en \mathbb{R} .

- Si $f : \mathbb{R}^n \rightarrow \mathbb{R}$ y $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$, y existen $\lim_{x \rightarrow a} f(x) = l$, $\lim_{x \rightarrow a} g(x) = l'$, aplicando la propiedad anterior a cada componente de la función $f g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ se obtiene de nuevo (1.27), donde ahora el producto es el de escalares por vectores de \mathbb{R}^m . \square

Los siguientes tres resultados son muy utilizados en la práctica a la hora de evaluar límites de funciones de varias variables:

Proposición 1.27. Sea $f = gh$, con $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ y $h : \mathbb{R}^n \rightarrow \mathbb{R}$, sea $a \in \mathbb{R}^n$ un punto de acumulación de $\mathcal{D}(f) = \mathcal{D}(g) \cap \mathcal{D}(h)$, y supongamos que $\lim_{x \rightarrow a} h(x) = 1$. Entonces el límite de f en a existe si y solo si lo mismo ocurre con el de g , y en tal caso ambos límites son iguales.

Demostración. En efecto, si $\lim_{x \rightarrow a} g(x) = l$ entonces $\lim_{x \rightarrow a} f(x) = l$ en virtud de la segunda propiedad del límite (cf. el último comentario tras la Proposición 1.26). Si, por el contrario, no existe el límite de g en a , entonces tampoco puede existir el de f . En efecto, por las propiedades iv) y v) del límite h no se anula en un entorno perforado de a y $\lim_{x \rightarrow a} (1/h(x)) = 1$. Al ser $g = f \cdot (1/h)$, si existiera el límite de f en a también existiría el de g (de nuevo por el último comentario tras la Proposición 1.26). \square

Proposición 1.28 (criterio de comparación). Sean $f, g, h : \mathbb{R}^n \rightarrow \mathbb{R}$, sea a un punto de acumulación de $\mathcal{D}(f) \cap \mathcal{D}(g) \cap \mathcal{D}(h)$, y supongamos que hay un entorno reducido U de a tal que

$$g(x) \leq f(x) \leq h(x), \quad \forall x \in \mathcal{D}(f) \cap \mathcal{D}(g) \cap \mathcal{D}(h) \cap U.$$

Entonces se tiene

$$\lim_{x \rightarrow a} g(x) = \lim_{x \rightarrow a} h(x) = l \quad \Rightarrow \quad \lim_{x \rightarrow a} f(x) = l.$$

Demostración. Por definición de límite, dado $\varepsilon > 0$ existen $\delta_1, \delta_2 > 0$ tales que

$$x \in \mathcal{D}(g), \quad 0 < \|x - a\| < \delta_1 \quad \Rightarrow \quad |g(x) - l| < \varepsilon$$

y

$$x \in \mathcal{D}(h), \quad 0 < \|x - a\| < \delta_2 \quad \Rightarrow \quad |h(x) - l| < \varepsilon.$$

Tomando entonces $\delta = \min(\delta_1, \delta_2) > 0$ y teniendo en cuenta que $|y| < \varepsilon \iff -\varepsilon < y < \varepsilon$ se obtiene

$$\begin{aligned} x \in \mathcal{D}(f) \cap \mathcal{D}(g) \cap \mathcal{D}(h) \cap U, \quad 0 < \|x - a\| < \delta \\ \Rightarrow \quad -\varepsilon < g(x) - l < f(x) - l < h(x) - l < \varepsilon \quad \Rightarrow \quad |f(x) - l| < \varepsilon. \quad \square \end{aligned}$$

Proposición 1.29. Sean $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$, y sea a punto de acumulación de $\mathcal{D}(f) \cap \mathcal{D}(g)$. Si g está acotada en un entorno de a y $\lim_{x \rightarrow a} f(x) = 0$ entonces

$$\lim_{x \rightarrow a} (f(x)g(x)) = 0.$$

Demostración. En efecto, por hipótesis existen $r > 0$ y $M > 0$ tales que

$$|g(x)| < M, \quad \forall x \in \mathcal{D}(g) \cap B_r(a),$$

y por tanto

$$x \in \mathcal{D}(f) \cap \mathcal{D}(g) \cap B_r(a) \implies |f(x)g(x)| = |f(x)| |g(x)| \leq M |f(x)|.$$

De la proposición anterior con $g = 0$ y $h = M|f|$ se sigue que

$$\lim_{x \rightarrow a} |f(x)g(x)| = 0,$$

que es equivalente al resultado anunciado (cf. el Ejercicio 18). □

Ejemplo 1.30. Consideremos la función $h : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por

$$h(x, y) = \frac{\sqrt{|x|} y}{|x|^{2/3} + 2|y|}, \quad (x, y) \neq (0, 0).$$

Entonces $h(x, y) = f(x, y)g(x, y)$, donde $g(x, y) = y/(|x|^{2/3} + 2|y|)$ está acotada (por 1/2) y $f(x, y) = \sqrt{|x|}$ tiende a cero en el origen. Por la proposición anterior, $\lim_{(x,y) \rightarrow (0,0)} h(x, y) = 0$.

1.7 Continuidad

Una vez introducido el concepto de límite, estamos en disposición de dar la definición de función continua en un punto:

Definición 1.31. Una función $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es **continua** en un punto $a \in \mathbb{R}^n$ si el límite de f en a existe y coincide con el valor de la función en dicho punto, es decir si

$$\lim_{x \rightarrow a} f(x) = f(a).$$

Diremos que f es continua en un conjunto $A \subset \mathbb{R}^n$ si f es continua en todos los puntos de A . □

*Comentario**⁴. Nótese, en particular, que si f es continua en a entonces $a \in \mathcal{D}(f) \cap \mathcal{D}(f)'$, es decir a pertenece al dominio de f y es un *punto de acumulación* de dicho dominio. En particular, f es automáticamente discontinua en cualquier punto *aislado* de su dominio. (De hecho, algunos autores exigen aún más restrictivamente que a sea un punto *interior* de $\mathcal{D}(f)$ para que f sea continua en dicho punto.) □

La definición anterior se puede escribir de forma algo más detallada como sigue:

⁴Denotaremos por un asterisco aquellas partes de estas notas de tipo más técnico o nivel avanzado, que pueden omitirse en una primera lectura.

f es continua en $a \in \mathcal{D}(f) \cap \mathcal{D}(f)'$ si para todo $\varepsilon > 0$, existe $\delta(\varepsilon) > 0$ tal que

$$x \in \mathcal{D}(f), \|x - a\| < \delta(\varepsilon) \implies \|f(x) - f(a)\| < \varepsilon.$$

Esta es la tradicional “definición ε - δ ” de la continuidad. Equivalentemente,

f es continua en $a \in \mathcal{D}(f) \cap \mathcal{D}(f)'$ si para todo entorno V de $f(a)$ existe un entorno U de a tal que $f(U) \subset V$.

Más informalmente,

f es continua en a si $f(x)$ está *arbitrariamente* próximo a $f(a)$ cuando x está *suficientemente* próximo al punto a .

La Proposición 1.25 se traduce en la siguiente propiedad fundamental de las funciones continuas:

Proposición 1.32. $f = (f_1, \dots, f_m) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es continua en un punto $a \in \mathbb{R}^n$ si y solo si lo son sus m funciones componentes $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$.

Ejemplo 1.33. Las funciones constantes son continuas en todo punto (en este caso se puede tomar $\delta(\varepsilon)$ arbitrariamente en la definición). La función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ definida por $f(x_1, \dots, x_n) = x_i$ (**proyección** sobre la i -ésima componente) es continua en cualquier punto $a \in \mathbb{R}^n$, ya que al ser

$$|f(x) - f(a)| = |x_i - a_i| \leq \|x - a\|$$

se satisface la definición con $\delta(\varepsilon) = \varepsilon$. La función $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por

$$g(x, y) = \frac{2xy^2}{x^2 + y^2}, \quad (x, y) \neq (0, 0),$$

no es continua en el origen, ya que $g(0, 0)$ no está definido. Sin embargo, para todo $(x, y) \neq (0, 0)$ se tiene

$$|g(x, y)| = 2|x| \frac{y^2}{x^2 + y^2} \leq 2|x|,$$

y por tanto $\lim_{(x,y) \rightarrow (0,0)} g(x, y) = 0$ (en efecto, podemos tomar $\delta(\varepsilon) = \varepsilon/2$ en la definición de límite, o alternativamente aplicar la Proposición 1.29). En consecuencia, si definimos $g(0, 0) = 0$ la función g es continua en el origen. \square

De la Proposición 1.26 se deducen inmediatamente las siguientes propiedades elementales de las funciones continuas:

Proposición 1.34 (propiedades de las funciones continuas). Sean $f, g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ y $h : \mathbb{R}^n \rightarrow \mathbb{R}$ funciones continuas en un punto $a \in \mathbb{R}^n$. Entonces se verifica:

- i) Las funciones $\lambda f + \mu g$ (con $\lambda, \mu \in \mathbb{R}$ constantes), $f \cdot g$, hf y $\|f\|$ son continuas en a .
- ii) f está acotada en un entorno de a .
- iii) Si $f(a) \neq 0$, hay un entorno de a en que f no se anula. Si, además, f es una función escalar, hay un entorno de a en que f tiene signo constante igual al de $f(a)$.
- iv) Si f es una función escalar y $f(a) \neq 0$, la función $1/f$ es continua en a .

*Comentario**. En el apartado i), hay que imponer que a pertenezca a $\mathcal{D}(f) \cap \mathcal{D}(g)$ y además sea punto de acumulación de dicho conjunto para que $\lambda f + \mu g$ (con $\mu \neq 0$) y $f \cdot g$ sean continuas en a . Análogamente, para poder asegurar que hf es continua en a dicho punto ha de pertenecer a $\mathcal{D}(h) \cap \mathcal{D}(f)$ y ser punto de acumulación de este conjunto.

Ejemplo 1.35. La función escalar de n variables $f(x) = x_i^k$, siendo k un número natural, es continua en todo punto en virtud de la primera propiedad de las funciones continuas (ya que, como se vio en el Ejemplo 1.33, $g(x) = x_i$ lo es). De esta propiedad se sigue que cualquier *polinomio* en n variables (x_1, \dots, x_n) , que es una suma de constantes por *monomios* de la forma $x_1^{k_1} \dots x_n^{k_n}$ (con k_1, \dots, k_n enteros no negativos), es continuo en todo punto. De la cuarta propiedad de las funciones continuas se sigue inmediatamente que una *función racional* (es decir, el cociente de dos polinomios) es continua en todo \mathbb{R}^n menos en los puntos en que se anula el denominador. De la Proposición 1.32 se deduce entonces que *cualquier función $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ cuyas componentes f_i sean funciones racionales es continua en todo \mathbb{R}^n menos en los puntos en que se anula el denominador de alguna de sus componentes f_i* . En particular, una *aplicación lineal* $L = (L_1, \dots, L_m) : \mathbb{R}^n \rightarrow \mathbb{R}^m$, cuyas funciones componentes son de la forma

$$L_i(x) = \sum_{j=1}^n l_{ij}x_j, \quad 1 \leq i \leq m, \tag{1.28}$$

con l_{ij} constantes reales, es continua en todo \mathbb{R}^n . □

Otro resultado fundamental para determinar la continuidad de una función de varias variables es la siguiente proposición, que afirma esencialmente que *la composición de funciones continuas es una función continua*:

Proposición 1.36. Sean $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ y $g : \mathbb{R}^m \rightarrow \mathbb{R}^p$ dos funciones, y supongamos que f es continua en $a \in \mathbb{R}^n$ y g es continua en $f(a) \in \mathbb{R}^m$. Si a es punto de acumulación de $\mathcal{D}(g \circ f)$, entonces la función compuesta $g \circ f$ es continua en a , es decir

$$\lim_{x \rightarrow a} g(f(x)) = g(f(a)).$$

Demostración. En efecto, sea V un entorno cualquiera de $g(f(a)) \in \mathbb{R}^p$. Al ser g continua en $f(a)$, hay un entorno W de dicho punto tal que $g(W) \subset V$. Por otro lado, al ser f continua en a , existe un entorno U de a tal que $f(U) \subset W$. Por tanto

$$(g \circ f)(U) = g(f(U)) \subset g(W) \subset V,$$

y se satisface la definición de continuidad de $(g \circ f)$ en a (cf. segundo comentario tras la Definición 1.31). □

*Comentario**. La condición $a \in \mathcal{D}(g \circ f)'$ es necesaria, es decir, *no* es consecuencia de las condiciones $a \in \mathcal{D}(f)'$ y $f(a) \in \mathcal{D}(g)'$ que se han de cumplir en virtud de la continuidad de f en a y g en $f(a)$. En efecto, consideremos, por ejemplo, las funciones $f : \mathbb{R} \rightarrow \mathbb{R}$ y $g : \mathbb{R} \rightarrow \mathbb{R}$ definidas por

$$f(x) = -|x|, \quad g(x) = \sqrt{x},$$

con dominios $\mathcal{D}(f) = \mathbb{R}$ y $\mathcal{D}(g) = [0, \infty)$. Entonces f es continua en $a = 0$ y g es continua en $f(a) = 0$, pero $\mathcal{D}(g \circ f) = \{0\}$, y por tanto 0 no es punto de acumulación de $\mathcal{D}(g \circ f)$. En particular, en este caso la función compuesta $g \circ f$ *no* es continua en 0.

*Ejercicio 19**. Probar que si g está definida en un entorno de $f(a)$ entonces a es automáticamente un punto de acumulación de $\mathcal{D}(g \circ f)$.

Solución. En efecto, supongamos que g está definida en un entorno V de $f(a)$. Por continuidad de g en $f(a)$, existe un entorno U de a tal que $f(U) \subset V$. Dado $r > 0$, al ser $B_r(a) \cap U$ un abierto (intersección finita de abiertos) que contiene al punto a , existe un número real $0 < \delta < r$ tal que $B_\delta(a) \subset B_r(a) \cap U$. Como a es por hipótesis punto de acumulación de $\mathcal{D}(f)$ (en virtud de la existencia de $\lim_{x \rightarrow a} f(x)$), la bola perforada $B_\delta^*(a) \subset B_r^*(a)$ contiene algún punto $x \in \mathcal{D}(f)$. Tal punto también pertenece al dominio de $g \circ f$, ya que al ser $x \in B_\delta(a) \subset U$ se tiene que $x \in \mathcal{D}(f)$ y $f(x) \in f(U) \subset V \subset \mathcal{D}(g)$. Por tanto $x \in B_r^*(a) \cap \mathcal{D}(g \circ f)$, lo que por definición implica (al ser $r > 0$ arbitrario) que a es punto de acumulación de $\mathcal{D}(g \circ f)$.

Ejemplo 1.37. Consideremos la función $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por

$$f(x, y) = e^{x^2+y^2} \sin(xy).$$

Entonces $f = g \circ h$, con $g: \mathbb{R} \rightarrow \mathbb{R}$ y $h: \mathbb{R}^2 \rightarrow \mathbb{R}$ definidas respectivamente por $g(t) = e^t$ y $h(x, y) = x^2 + y^2 \sin(xy)$. La función $(x, y) \mapsto \sin(xy)$ es continua en \mathbb{R}^2 , ya que es la composición de \sin (continua en \mathbb{R}) con $(x, y) \mapsto xy$ (continua en \mathbb{R}^2 por ser un polinomio). A su vez, la función $(x, y) \mapsto y^2 \sin(xy)$ es continua en \mathbb{R}^2 (producto de un polinomio por una función continua), y por tanto también lo es h (suma de un polinomio con una función continua). Por la Proposición 1.36, $f = g \circ h$ es continua en \mathbb{R}^2 .

1.8 Límite a lo largo de una curva

Consideremos una función vectorial $\gamma: \mathbb{R} \rightarrow \mathbb{R}^n$ continua en $t_0 \in \mathbb{R}$, y sea $\gamma(t_0) = a \in \mathbb{R}^n$. Intuitivamente, la imagen de γ es una *curva* en \mathbb{R}^n que pasa por el punto a . Sea $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$, y supongamos que $\mathcal{R}(\gamma) \subset \mathcal{D}(f)$. Entonces t_0 es punto de acumulación del dominio de $f \circ \gamma$, ya que $\mathcal{D}(f \circ \gamma) = \mathcal{D}(\gamma)$ y existe por hipótesis el límite de γ en t_0 . Llamaremos **límite de f en el punto a a lo largo de la curva γ** al límite

$$\lim_{t \rightarrow t_0} f(\gamma(t)). \quad (1.29)$$

Por la Proposición 1.36, si f es continua en el punto a la composición $f \circ \gamma$ es continua en t_0 , es decir

$$\lim_{t \rightarrow t_0} f(\gamma(t)) = f(a). \quad (1.30)$$

Por tanto si f es continua en a existe el límite de f en a a lo largo de cualquier curva continua γ que pase por a , y es independiente de la curva considerada. Esta afirmación implica el importante resultado siguiente:

Para probar que una función f no es continua en un punto $a \in \mathbb{R}^n$ basta encontrar una curva continua γ que pase por a para la que no exista el límite (1.29), o hallar sendas curvas de este tipo $\gamma_1, \gamma_2: \mathbb{R} \rightarrow \mathbb{R}^n$ para las cuales

$$\lim_{t \rightarrow t_0} f(\gamma_1(t)) \neq \lim_{t \rightarrow t_0} f(\gamma_2(t)).$$

El ejemplo anterior se puede extender al *límite* de una función (no necesariamente continua) en un punto. En efecto, supongamos que

$$\lim_{x \rightarrow a} f(x) = l,$$

y sea $\gamma : \mathbb{R} \rightarrow \mathbb{R}^n$ cualquier curva continua en $t_0 \in \mathbb{R}$ tal que $\gamma(t_0) = a$ y $\mathcal{R}(\gamma) - \{a\} \subset \mathcal{D}(f)$. Si $\gamma(t) \neq a$ en un entorno reducido de t_0 , entonces existe el límite de f en a a lo largo de γ y vale l , es decir

$$\lim_{t \rightarrow t_0} f(\gamma(t)) = l.$$

Para demostrar este resultado, basta notar que la función $\tilde{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ definida por $\tilde{f}(a) = l$ y $\tilde{f}(x) = f(x)$ para todo $x \in \mathcal{D}(f) - \{a\}$ es continua en a , y por tanto verifica

$$\lim_{t \rightarrow t_0} \tilde{f}(\gamma(t)) = \tilde{f}(a) = l.$$

Pero el límite anterior coincide con el límite de f en a a lo largo de γ , al ser $f \circ \gamma = \tilde{f} \circ \gamma$ en el entorno reducido de t_0 en que $\gamma(t) \neq a$. Al igual que antes, de esta observación se deduce el siguiente resultado:

Para probar que una función f no tiene límite en el punto $a \in \mathbb{R}^n$ basta encontrar una curva continua γ que pase por $a = \gamma(t_0)$ (con $\gamma(t) \neq a$ en un entorno reducido de t_0) a lo largo de la cual no exista el límite de f en a , o hallar sendas curvas de este tipo a lo largo de las cuales f tenga límites *distintos* en a .

*Ejercicio 20**. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ y $a \in \mathcal{D}(f) \cap \mathcal{D}(f)'$, y supongamos que para toda curva $\gamma : \mathbb{R} \rightarrow \mathbb{R}^n$ que pasa por a para un cierto $t = t_0$ y es continua en t_0 se cumple (1.30). Probar que f es continua en a .

Solución. En efecto, supongamos que $\lim_{x \rightarrow a} f(x) \neq f(a)$. En tal caso, existe $\varepsilon > 0$ tal que para todo $n \in \mathbb{N}$ existe algún punto $x_n \in B_{1/n}^*(a)$ tal que $\|f(x_n) - f(a)\| \geq \varepsilon$. Definimos entonces una curva $\gamma : \mathbb{R} \rightarrow \mathbb{R}^n$ mediante $\gamma(0) = a$ y

$$\gamma(t) = x_n + (n+1)(x_{n+1} - x_n)(1-nt), \quad \text{si } \frac{1}{n+1} < t \leq \frac{1}{n}.$$

En otras palabras, la imagen de γ es la *poligonal* (quebrada) que une los puntos $x_1, x_2, \dots, x_n, \dots$ y el punto a . Entonces γ es continua en $t_0 = 0$, ya que es fácil ver que⁵

$$0 < t \leq \frac{1}{n} \implies \gamma(t) \in B_{1/n}^*(a).$$

Sin embargo, $\lim_{t \rightarrow 0} f(\gamma(t)) \neq f(a)$, ya que para todo $n \in \mathbb{N}$ se cumple

$$\gamma(1/n) = x_n \implies \|f(\gamma(1/n)) - f(a)\| = \|f(x_n) - f(a)\| \geq \varepsilon.$$

Como esto contradice la hipótesis, f ha de ser continua en el punto a .

Nótese que la curva γ que hemos construido es, de hecho, continua en los demás puntos de su dominio $\mathcal{D}(\gamma) = [0, 1]$, ya que en un entorno de todo $t \in (0, 1]$ que no sea de la forma $1/(n+1)$ (con $n \in \mathbb{N}$) es una función *lineal* de t , y en los puntos de la forma $1/(n+1)$ estas funciones lineales empalman de forma continua. Aunque γ solo está definida a la derecha de $t_0 = 0$, es inmediato definirla (sin afectar al argumento anterior) en un intervalo simétrico alrededor de 0, sin más que definir $\gamma(t) = \gamma(-t)$ para $t \in [-1, 0)$.

Ejemplo 1.38. Sea $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ la función dada por

$$f(x, y) = \frac{|x|^a |y|^b}{(x^2 + y^2)^c}, \quad (x, y) \neq (0, 0),$$

⁵En efecto, si $0 < t \leq 1/n$, existe $k \geq n$ tal que $1/(k+1) < t \leq 1/k$. Por otra parte, si x, y son dos puntos de una bola abierta el segmento $\{x + s(y-x) \mid 0 \leq s \leq 1\}$ está contenido en dicha bola (cf. el Ejercicio 9). Aplicando el resultado anterior con $x = x_k, y = x_{k+1}$ y $s = (k+1)(1-kt)$ se deduce que $\gamma(t) \in B_{1/k}^*(a) \subset B_{1/n}^*(a)$.

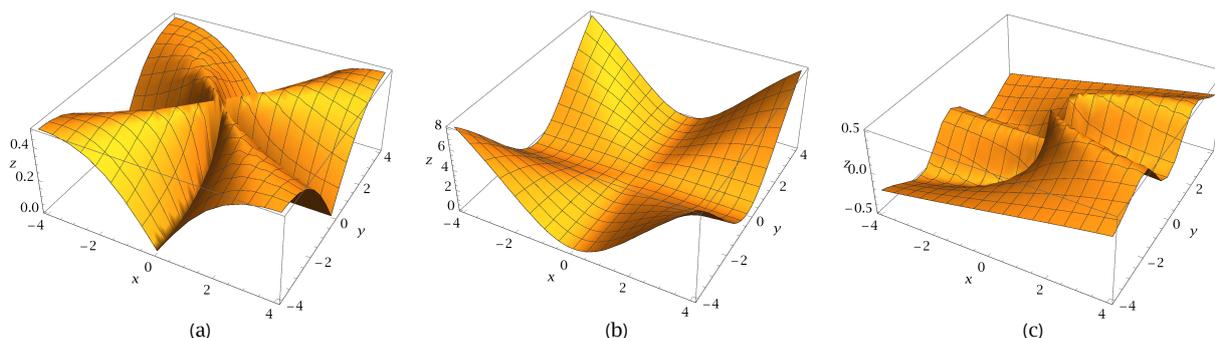


Figura 1.7: gráficas de las funciones $f(x, y) = |xy|/(x^2 + y^2)$ (a), $f(x, y) = x^2y^2/(x^2 + y^2)$ (b), $f(x, y) = xy^2/(x^2 + y^4)$ (c).

donde a , b y c son parámetros reales positivos. La función $(x, y) \mapsto |x|^a$ es continua en todo punto, ya que es la composición de la función de una variable $t \mapsto |t|^a$, continua en \mathbb{R} al ser $a > 0$, con la proyección sobre la primera componente. Lo mismo ocurre con la función $(x, y) \mapsto |y|^b$. Por otra parte, $(x, y) \mapsto (x^2 + y^2)^{-c}$ es continua en todo \mathbb{R}^2 menos el origen, al ser composición de la función continua $t \mapsto |t|^c$ con la función racional $(x, y) \mapsto (x^2 + y^2)^{-1}$, cuyo denominador solo se anula si $(x, y) = (0, 0)$. Por tanto f es continua en $\mathbb{R}^2 - \{(0, 0)\}$, al ser el producto de tres funciones continuas.

Veamos a continuación qué ocurre en el origen. Para ello consideramos la curva (recta por el origen)

$$y(t) = tv \equiv (tv_1, tv_2),$$

(siendo $v \in \mathbb{R}^2$ un vector constante no nulo), que es continua en \mathbb{R} y cumple $y(0) = (0, 0)$. Como

$$\lim_{t \rightarrow 0} f(y(t)) = \frac{|v_1|^a |v_2|^b}{(v_1^2 + v_2^2)^c} \lim_{t \rightarrow 0} |t|^{a+b-2c}, \quad (1.31)$$

para que el límite del miembro izquierdo exista para todo $v \neq 0$ (y, por tanto, para que f pueda tener límite en el origen) es *necesario* que $a + b \geq 2c$. Si $a + b = 2c$ el límite anterior existe, pero depende de la recta escogida para aproximarnos al origen. Por tanto tampoco en este caso existe el límite de f en el origen. ¿Qué ocurre si $a + b > 2c$? Nótese que si existe el límite de f en $(0, 0)$ ha de valer necesariamente 0, en virtud de (1.31). Para probar que cuando $a + b > 2c$ el límite de f en el origen es efectivamente igual a cero, basta notar que si llamamos $r = \sqrt{x^2 + y^2}$ entonces

$$a + b > 2c \quad \Rightarrow \quad |f(x, y)| = \left(\frac{|x|}{r}\right)^a \left(\frac{|y|}{r}\right)^b r^{a+b-2c} \leq r^{a+b-2c} \xrightarrow{(x,y) \rightarrow (0,0)} 0.$$

Por tanto en este caso f tiende a 0 en el origen, en virtud de la Proposición 1.29. En definitiva, f es continua en el origen si y solo si $a + b > 2c$ y definimos $f(0, 0) = 0$.

Ejemplo 1.39. Sea ahora $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ la función definida por

$$f(x, y) = \frac{xy^2}{x^2 + y^4}, \quad (x, y) \neq (0, 0).$$

Como f es una función racional cuyo denominador solo se anula en el origen, f es continua en $\mathbb{R}^2 - \{(0, 0)\}$. ¿Qué ocurre en el origen? Como en el ejemplo anterior, calculemos primero el límite de f a lo largo de una recta cualquiera que pase por el origen

$$y(t) = t(v_1, v_2).$$

Al ser

$$\lim_{t \rightarrow 0} f(y(t)) = \lim_{t \rightarrow 0} \frac{t v_1 v_2^2}{v_1^2 + t^2 v_2^4} = 0,$$

el límite de f en el origen a lo largo de *cualquier* recta que pase por el origen existe y vale 0, independientemente de la recta considerada. Sin embargo, esto *no* significa que f tiene límite 0 en el origen. De hecho, en este caso es fácil ver que este límite *no* existe. En efecto, al ser

$$f(x, y) = \frac{y^2/x}{1 + (y^2/x)^2}, \quad x \neq 0,$$

las curvas de nivel de f son los ejes coordenados y las parábolas $y^2 = cx$ (con $c \neq 0$ constante), excluido el origen (donde la función no está definida). A lo largo de estas parábolas, es decir de las curvas $y(t) = (t^2/c, t)$ (todas las cuales pasan por el origen para $t = 0$), el valor de la función está dado por

$$f(t^2/c, t) = \frac{c}{1 + c^2}, \quad \forall t \neq 0.$$

En particular, como

$$\lim_{t \rightarrow 0} f(t^2/c, t) = \frac{c}{1 + c^2}$$

depende de la curva considerada (ya que es función del parámetro c), f *no* es continua en el origen. En otras palabras, f toma *cualquier* valor de la forma $c/(1 + c^2)$ arbitrariamente cerca del origen⁶.

1.9 Notación o y O de Landau

Definición 1.40. Sean $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ sendas funciones definidas en un entorno de $a \in \mathbb{R}^n$. Diremos que $f : \mathbb{R}^n \rightarrow \mathbb{R}$ es **de menor o igual orden que g en a** , y escribiremos⁷ $f = O(g)$ en a , si existen $M > 0$ y $r > 0$ tales que

$$\|x - a\| < r \implies |f(x)| \leq M |g(x)|.$$

Análogamente, diremos que f es **de menor orden que g en a** , y escribiremos $f = o(g)$ en a , si para todo $\varepsilon > 0$ existe $\delta(\varepsilon) > 0$ tal que

$$\|x - a\| < \delta(\varepsilon) \implies |f(x)| \leq \varepsilon |g(x)|.$$

Intuitivamente, $f = O(g)$ en a si $|f|$ es a lo sumo del mismo orden de magnitud que $|g|$ en un entorno de a , y $f = o(g)$ en a si $|f|$ es mucho menor que $|g|$ en un entorno de dicho punto. Por ejemplo, $5x^2 - 6x = O(x)$ y $5x^2 = o(x)$ en $x = 0$.

Si (como suele ser el caso) g no se anula en un entorno reducido de a entonces

$$\begin{aligned} f = O(g) \text{ en } a &\iff f/g \text{ acotada en un entorno de } a, \\ f = o(g) \text{ en } a &\iff \lim_{x \rightarrow a} \frac{f(x)}{g(x)} = 0. \end{aligned} \tag{1.32}$$

En particular, $f = O(1)$ en a si y solo si f está *acotada* en un entorno de a , y $f = o(1)$ en a si y solo si $\lim_{x \rightarrow a} f(x) = 0$.

Es fácil comprobar que se verifican las siguientes relaciones (ejercicio):

⁶La función de una variable $g(c) = c/(1 + c^2)$ tiene imagen $\mathcal{R}(g) = [-1/2, 1/2]$ (ejercicio), y por tanto $\mathcal{R}(f) = [-1/2, 1/2]$.

⁷La notación o y O fue introducida por el matemático alemán Edmund Landau (1877-1938).

- i) $f = o(g) \Rightarrow f = O(g)$
 ii) $f_1 = o(g), f_2 = o(g) \Rightarrow f_1 + f_2 = o(g)$
 iii) $f_1 = o(g_1), f_2 = O(g_2) \Rightarrow f_1 f_2 = o(g_1 g_2)$
 iv) $f_1 = O(g_1), f_2 = O(g_2) \Rightarrow f_1 f_2 = O(g_1 g_2)$

Con un ligero abuso de notación, las propiedades ii)-iv) se suelen escribir como sigue:

ii) $o(g) + o(g) = o(g)$

iii) $o(g_1)O(g_2) = o(g_1 g_2)$

iv) $O(g_1)O(g_2) = O(g_1 g_2)$.

En particular, de la propiedades iii) y iv) se sigue que $\lambda o(g) = o(g)$ y $\lambda O(g) = O(g)$ para toda constante $\lambda \in \mathbb{R}$. Nótese también que la Proposición 1.29 se obtiene haciendo $g_1 = g_2 = 1$ en iii).

Ejercicio 21. Probar que si g no se anula en un entorno reducido de $a \in \mathbb{R}^n$ entonces $o(g) = go(1)$ en a .

Solución. Si $f = o(g)$ y g no se anula en un entorno reducido de a entonces $f/g = o(1)$ en a (cf. (1.32)), y por tanto $f = go(1)$ en a .

Si $f : \mathbb{R} \rightarrow \mathbb{R}$ es una función r veces derivable en $a \in \mathbb{R}$, el teorema de Taylor afirma que

$$f(x) = \sum_{k=0}^r \frac{f^{(k)}(a)}{k!} (x-a)^k + o((x-a)^r) \quad \text{en } x = a. \quad (1.33)$$

Así, por ejemplo, en $x = 0$ se cumple

$$\begin{aligned} e^x &= \sum_{k=0}^r \frac{x^k}{k!} + o(x^r), \\ \cos x &= \sum_{k=0}^r (-1)^k \frac{x^{2k}}{(2k)!} + o(x^{2r}), \\ \sin x &= \sum_{k=0}^r (-1)^k \frac{x^{2k+1}}{(2k+1)!} + o(x^{2r+1}), \\ \log(1+x) &= \sum_{k=1}^r (-1)^{k+1} \frac{x^k}{k} + o(x^r), \\ \frac{1}{1-x} &= \sum_{k=0}^r x^k + o(x^r), \\ (1+x)^\alpha &= \sum_{k=0}^r \binom{\alpha}{k} x^k + o(x^r), \quad \alpha \in \mathbb{R}, \end{aligned}$$

donde el coeficiente binomial $\binom{\alpha}{k}$ se define para cualquier $\alpha \in \mathbb{R}$ mediante $\binom{\alpha}{0} = 1$ y

$$\binom{\alpha}{k} = \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!}.$$

Ejemplo 1.41. Consideremos el límite cuando $(x, y) \rightarrow (0, 0)$ de la función

$$f(x, y) = \frac{\log(1 + \sin^2 x) \tan y}{(1 - \cos \sqrt{y}) \arcsin(x^2)}, \quad (1.34)$$

definida en el conjunto $\{(x, y) \in \mathbb{R}^2 \mid |x| \leq 1, y > 0\}$ menos las rectas $x = 0$, $y = (2k - 1)\pi/2$ e $y = 4k^2\pi^2$ (con $k \in \mathbb{N}$). En particular, el origen es un punto de acumulación de $\mathcal{D}(f)$. El límite de f en el origen es indeterminado, ya que tanto el numerador como el denominador de (1.34) son claramente funciones continuas en el origen que se anulan en dicho punto. Para resolver la indeterminación, basta utilizar los desarrollos anteriores junto con los de $\tan t = t(1 + o(1))$ y $\arcsin t = t(1 + o(1))$ (que se siguen fácilmente de (1.33) con $r = 1$):

$$f(x, y) = \frac{\sin^2 x(1 + o(1)) \cdot y(1 + o(1))}{\frac{y}{2}(1 + o(1)) \cdot x^2(1 + o(1))} \equiv \frac{2 \sin^2 x}{x^2} g(x, y),$$

donde $g(x, y) \rightarrow 1$ si $(x, y) \rightarrow (0, 0)$ y hemos aplicado que $2o(1) = o(1)$. Por tanto

$$\boxed{\lim_{(x,y) \rightarrow (0,0)} f(x, y) = 2 \lim_{(x,y) \rightarrow (0,0)} \left(\frac{\sin x}{x} \right)^2 = 2.}$$

Ejemplo 1.42. Consideremos a continuación el límite en el punto $(1, \pi/2)$ de la función

$$\boxed{f(x, y) = \frac{x^{\sin y} - 1}{a(x - 1) + b\left(y - \frac{\pi}{2}\right)},}$$

donde a y b son parámetros reales que no se anulan simultáneamente. Nótese que, al ser $x^{\sin y} = e^{\log x \sin y}$, f está definida en el semiplano $x > 0$ menos la recta $a(x - 1) + b\left(y - \frac{\pi}{2}\right) = 0$. De nuevo, tanto el numerador como el denominador son funciones continuas en $(1, \pi/2)$ que se anulan simultáneamente en dicho punto, por lo que el límite pedido es indeterminado. En este caso se tiene

$$x^{\sin y} - 1 = e^{\sin y \log x} - 1 = \log x \cdot \sin y(1 + o(1)) \quad \text{en } (1, \pi/2).$$

Como el término que multiplica a $\log x$ tiende a 1 cuando $(x, y) \rightarrow (1, \pi/2)$, por la Proposición 1.27 el comportamiento de f en este último punto es el mismo que el de

$$g(x, y) = \frac{\log x}{a(x - 1) + b\left(y - \frac{\pi}{2}\right)}.$$

A su vez, al ser

$$\log x = x - 1 + o(x - 1) = (x - 1)(1 + o(1))$$

en el punto $x = 1$, g (y por tanto f) se comporta en $(1, \pi/2)$ como

$$h(x, y) = \frac{x - 1}{a(x - 1) + b\left(y - \frac{\pi}{2}\right)}.$$

Para estudiar el límite de esta función en el punto $(1, \pi/2)$ la evaluamos primero a lo largo de las rectas $y(t) = \left(1 + tv_1, \frac{\pi}{2} + tv_2\right)$ con⁸ $av_1 + bv_2 \neq 0$, que pasan por dicho punto para $t = 0$, obteniendo

$$h(y(t)) = \frac{v_1}{av_1 + bv_2}.$$

Si $b \neq 0$ el miembro derecho depende del vector v considerado, y por tanto en este caso el límite no existe. Si $b = 0$ entonces $a \neq 0$ y $h(x, y) = 1/a$ para $x \neq 1$, y por tanto

$$\boxed{b = 0 \implies \lim_{(x,y) \rightarrow (1, \pi/2)} f(x, y) = \frac{1}{a}.}$$

En definitiva, el límite de f en $(1, \pi/2)$ solo existe si $b = 0$ y $a \neq 0$, y en tal caso es igual a $1/a$.

⁸Es necesario tomar v tal que $av_1 + bv_2 \neq 0$, ya que si $av_1 + bv_2 = 0$ se anula la recta en cuestión es la de ecuación $a(x - 1) + b\left(y - \frac{\pi}{2}\right) = 0$ en que h es singular.

1.10 Propiedades topológicas de las funciones continuas

Estudiaremos en esta sección varias propiedades importantes de las funciones continuas estrechamente relacionadas con la topología del espacio \mathbb{R}^n . Empezaremos por la siguiente propiedad fundamental:

Proposición 1.43. Si $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es continua en \mathbb{R}^n , la imagen inversa $f^{-1}(V)$ de cualquier abierto $V \subset \mathbb{R}^m$ es un abierto de \mathbb{R}^n .

Demostración. En efecto, si $V = f^{-1}(V)$ es vacío no hay nada que demostrar (ya que el conjunto vacío es abierto). En caso contrario, sea x un elemento cualquiera de $f^{-1}(V)$, es decir un punto de \mathbb{R}^n en que $f(x) \in V$. Como f es continua en x (al ser continua en todo \mathbb{R}^n por hipótesis) y V es un entorno de $f(x)$, hay un entorno U de x tal que $f(U) \subset V$, o equivalentemente (al ser $\mathcal{D}(f) = \mathbb{R}^n$, ya que f es continua en \mathbb{R}^n por hipótesis) $U \subset f^{-1}(V)$. Esto demuestra que $f^{-1}(V)$ contiene un entorno de cualquiera de sus puntos, y por lo tanto es abierto. \square

Corolario 1.44. Si $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ continua en \mathbb{R}^n , la imagen inversa $f^{-1}(V)$ de cualquier cerrado $V \subset \mathbb{R}^m$ es un cerrado de \mathbb{R}^n .

Demostración. En efecto, si $V \subset \mathbb{R}^m$ es cerrado entonces

$$\mathbb{R}^n - f^{-1}(V) = \{x \in \mathbb{R}^n \mid f(x) \notin V\} = \{x \in \mathbb{R}^n \mid f(x) \in \mathbb{R}^m - V\} = f^{-1}(\mathbb{R}^m - V),$$

donde la segunda igualdad se debe a que $\mathcal{D}(f) = \mathbb{R}^n$. Por la proposición anterior, $f^{-1}(\mathbb{R}^m - V)$ es abierto (al serlo $\mathbb{R}^m - V$ por hipótesis), y en consecuencia $f^{-1}(V)$ es cerrado. \square

Comentarios.*

- De hecho, es fácil probar que los recíprocos de la proposición y el corolario anteriores son ciertos. En otras palabras, una función $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ con $\mathcal{D}(f) = \mathbb{R}^n$ es continua en \mathbb{R}^n si y solo si la imagen inversa bajo f de cualquier abierto (resp. cerrado) de \mathbb{R}^m es abierta (resp. cerrada) en \mathbb{R}^n .
- Los resultados anteriores pueden ser generalizados al caso en que f es continua en un conjunto $A \neq \mathbb{R}^n$. Más precisamente, en tal caso la imagen inversa $f^{-1}(V)$ de un conjunto abierto (resp. cerrado) $V \subset \mathbb{R}^m$ es la intersección de un conjunto abierto (resp. cerrado) con A . En particular, si A es abierto (resp. cerrado) entonces $f^{-1}(V)$ es abierto (resp. cerrado).
- La imagen de un conjunto abierto (resp. cerrado) bajo una función continua *no* tiene por que ser un conjunto abierto (resp. cerrado). Por ejemplo, si $A = (0, 2\pi) \subset \mathbb{R}$ y $f(x) = \sin x$, entonces f es continua en A pero $f(A) = [-1, 1]$ no es abierto.

Sean $f_1, \dots, f_m : \mathbb{R}^n \rightarrow \mathbb{R}$ continuas en \mathbb{R}^n , y consideremos los subconjuntos de \mathbb{R}^n

$$\begin{aligned} A &= \{x \in \mathbb{R}^n \mid f_i(x) > 0, \quad 1 \leq i \leq m\} \\ C_1 &= \{x \in \mathbb{R}^n \mid f_i(x) \geq 0, \quad 1 \leq i \leq m\} \\ C_2 &= \{x \in \mathbb{R}^n \mid f_i(x) = 0, \quad 1 \leq i \leq m\}. \end{aligned}$$

Entonces A es abierto, y $C_{1,2}$ son cerrados, en virtud del comentario anterior y las propiedades de los conjuntos abiertos y cerrados. En efecto,

$$A = \bigcap_{i=1}^m f_i^{-1}((0, \infty)), \quad C_1 = \bigcap_{i=1}^m f_i^{-1}([0, \infty)), \quad C_2 = \bigcap_{i=1}^m f_i^{-1}(\{0\}),$$

siendo $(0, \infty)$ abierto y $[0, \infty)$, $\{0\}$ cerrados en \mathbb{R} . Así, por ejemplo, el (hiper)paralelepípedo abierto $P = (a_1, b_1) \times \cdots \times (a_n, b_n) \subset \mathbb{R}^n$ (donde se admite que $a_i = -\infty$ o $b_i = +\infty$) es abierto. En efecto, $P = A_1 \cap A_2$, con

$$A_1 = \{x \in \mathbb{R}^n \mid x_i > a_i, 1 \leq i \leq n\}, \quad A_2 = \{x \in \mathbb{R}^n \mid x_i < b_i, 1 \leq i \leq n\}$$

abiertos en virtud de las consideraciones anteriores aplicadas a las funciones $f_i(x) = x_i - a_i$ ($1 \leq i \leq n$) en el primer caso y $f_i(x) = b_i - x_i$ ($1 \leq i \leq n$) en el segundo. Del mismo modo se demuestra que el *paralelepípedo cerrado* $\bar{P} = [a_1, b_1] \times \cdots \times [a_n, b_n]$ es cerrado.

Otra propiedad fundamental de las funciones continuas está relacionada con los conjuntos compactos:

Proposición 1.45. Si $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es continua en un conjunto compacto K , entonces $f(K)$ es compacto. En particular, si $m = 1$ (es decir, si f es una función escalar) entonces existen $a, b \in K$ tales que

$$f(a) \leq f(x) \leq f(b), \quad \forall x \in K. \tag{1.35}$$

En otras palabras, f alcanza sus valores máximo y mínimo en K .

Demostración. La demostración de que $f(K)$ es compacto, que omitiremos por su complejidad, es consecuencia del *teorema de Heine-Borel-Lebesgue*. La demostración del caso particular $m = 1$ a partir de este resultado es relativamente sencilla. En efecto, al ser $f(K) \subset \mathbb{R}$ acotado existe $M > 0$ tal que $|f(x)| < M$ para todo $x \in K$, es decir

$$-M < f(x) < M, \quad \forall x \in K.$$

Como $f(K)$ está acotado tanto inferior como superiormente, existen el ínfimo y el supremo de este conjunto, que denotaremos respectivamente por A y B :

$$A = \inf f(K) \equiv \inf\{f(x) \mid x \in K\}, \quad B = \sup f(K).$$

Por definición de ínfimo y supremo, A y B son puntos de frontera de $f(K)$ (ejercicio). Pero, al ser $f(K)$ cerrado (por ser compacto), A y B han de pertenecer a dicho conjunto. Por tanto $A = f(a)$ y $B = f(b)$, con $a, b \in K$. Como el ínfimo y el supremo de un conjunto son cotas inferior y superior, respectivamente, de dicho conjunto, para todo $x \in K$ se verifica (1.35). \square

Nótese que para la validez de la proposición anterior son fundamentales tanto la continuidad de f como la compacidad del conjunto K . De hecho, una función continua en un conjunto no compacto no tiene ni siquiera por qué estar acotada en dicho conjunto. Por ejemplo, la función $\tan : (-\pi/2, \pi/2) \rightarrow \mathbb{R}$ es continua en el intervalo abierto (acotado pero no compacto) $(-\pi/2, \pi/2)$, pero no está acotada en dicho conjunto ni superior ni inferiormente.

La siguiente propiedad topológica fundamental de las funciones continuas está relacionada con los conjuntos conexos, que definiremos a continuación.

Definición 1.46. Un conjunto $A \subset \mathbb{R}^n$ es **disconexo** si existen dos abiertos $U, V \subset \mathbb{R}^n$ tales que $A = (A \cap U) \cup (A \cap V)$, con $A \cap U$ y $A \cap V$ disjuntos y no vacíos. En caso contrario, se dice que A es **conexo**.

Ejemplo 1.47. El conjunto $A = [-1, 1) \cup (1, 2] = [-1, 2] - \{1\}$ es desconexo. En efecto,

$$A = (A \cap (-2, 1)) \cup (A \cap (1, 3)),$$

con $A \cap (-2, 1) = [-1, 1)$ y $A \cap (1, 3) = (1, 2]$ disjuntos y no vacíos. \square

Ejercicio 22. Probar que el conjunto \mathbb{Q} de los números racionales es desconexo.

Solución. En efecto, (por ejemplo) $\mathbb{Q} = (\mathbb{Q} \cap (-\infty, \sqrt{2})) \cup (\mathbb{Q} \cap (\sqrt{2}, \infty))$. De forma semejante se demuestra que \mathbb{Q}^n tampoco es conexo para ningún n , y lo mismo ocurre con $(\mathbb{R} - \mathbb{Q})^n$ (ejercicio).

En general, no es fácil demostrar que un conjunto es conexo aplicando la definición. Una forma sencilla de probar la conexión de un conjunto es demostrando que es arco-conexo, ya que la conexión por arcos implica la conexión. Más precisamente:

Definición 1.48. Un conjunto $A \subset \mathbb{R}^n$ es **arco-conexo** si dos puntos *cualesquiera* de A se pueden unir por una curva *continua* cuya imagen esté enteramente *contenida en* A . En otras palabras, A es arco-conexo si para todo $x, y \in A$ hay una curva continua $\gamma : [0, 1] \rightarrow A$ tal que $\gamma(0) = x, \gamma(1) = y$.

Se demuestra entonces el siguiente resultado:

Proposición 1.49. *Un conjunto arco-conexo $A \subset \mathbb{R}^n$ es automáticamente conexo. Un conjunto abierto $A \subset \mathbb{R}^n$ es conexo si y solo si es arco-conexo.*

Ejemplo 1.50.

- Se demuestra que *los únicos conjuntos conexos en \mathbb{R} son los intervalos* (abiertos o cerrados, finitos o infinitos).
- El espacio \mathbb{R}^n es claramente conexo, ya que es arco-conexo. En efecto, si $x, y \in \mathbb{R}^n$ la curva $\gamma : [0, 1] \rightarrow \mathbb{R}^n$ definida por $\gamma(t) = (1-t)x + t y$ (es decir, el *segmento* de extremos x e y) es claramente continua (por ser una función lineal de t) y tiene por extremos los puntos x e y .
- Por el mismo motivo, cualquier conjunto *convexo* (es decir, que contiene el segmento que tiene por extremos dos cualesquiera de sus puntos) es conexo. Ejemplos de conjuntos convexos son las bolas (abiertas o cerradas), los paralelepípedos $(a_1, b_1) \times \cdots \times (a_n, b_n)$ y $[a_1, b_1] \times \cdots \times [a_n, b_n]$ (ejercicio), etc.
- También son claramente (arco-)conexos muchos conjuntos no convexos como \mathbb{R}^n menos un punto (si $n > 1$), \mathbb{R}^n menos un segmento (si $n > 1$), \mathbb{R}^n menos una recta (si $n > 2$), etc.
- No es conexo, sin embargo, el abierto $A = \{(x, y) \in \mathbb{R}^2 \mid y \neq 0\}$. En efecto, $A = U \cup V$, siendo U y V respectivamente los semiplanos abiertos disjuntos $y > 0$ e $y < 0$.

Ejercicio 23. Probar que los únicos subconjuntos de \mathbb{R}^n que son a la vez abiertos y cerrados son \mathbb{R}^n y \emptyset .

Solución. En efecto, supongamos que $A \subset \mathbb{R}^n$ fuera abierto y cerrado simultáneamente. Como $\mathbb{R}^n = A \cup (\mathbb{R}^n - A)$, con A y $\mathbb{R}^n - A$ ambos abiertos y obviamente disjuntos, si tanto A como $\mathbb{R}^n - A$ fueran ambos no vacíos \mathbb{R}^n sería desconexo. Por tanto o bien $A = \emptyset$, o bien $\mathbb{R}^n - A = \emptyset$ (es decir, $A = \mathbb{R}^n$).

Ejercicio 24. Un abierto $A \subset \mathbb{R}^n$ se denomina *estrellado* si existe un punto $a \in A$ tal que para todo $x \in A$, el segmento de extremos a y x está contenido en A . Probar que cualquier abierto estrellado es conexo.

Otra propiedad fundamental de las funciones continuas es que transforman conjuntos conexos en conjuntos conexos:

Proposición 1.51. *Si $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es continua en un conjunto conexo $A \subset \mathbb{R}^n$ entonces $f(A)$ es conexo.*

Demostración. Supondremos, por sencillez, que f es continua en todo \mathbb{R}^n (el caso general se demuestra de forma semejante utilizando el segundo comentario tras el Corolario 1.44). Si $f(A)$ fuera disconexo, existirían sendos abiertos $U, V \subset \mathbb{R}^m$ tales que

$$f(A) = (f(A) \cap U) \cup (f(A) \cap V),$$

con $f(A) \cap U$ y $f(A) \cap V$ disjuntos y no vacíos. Pero entonces

$$A = (A \cap f^{-1}(U)) \cup (A \cap f^{-1}(V))$$

sería disconexo, al ser $f^{-1}(U)$ y $f^{-1}(V)$ ambos abiertos (por la Proposición 1.43) y $A \cap f^{-1}(U)$, $A \cap f^{-1}(V)$ disjuntos y no vacíos (por serlo $f(A) \cap U$ y $f(A) \cap V$). \square

Corolario 1.52. *La imagen de un conjunto conexo $A \subset \mathbb{R}^n$ bajo una función escalar continua en A es un intervalo (finito o infinito).*

Demostración. Por la proposición anterior $f(A)$ ha de ser conexo, y los únicos subconjuntos conexos de la recta real son precisamente los intervalos (cf. el Ejemplo 1.50). \square

Comentario. Nótese, sin embargo, que si f es continua en el abierto conexo $A \subset \mathbb{R}^n$ el conjunto $f(A)$ no tiene por qué ser un intervalo abierto (cf. el último comentario tras el Corolario 1.44).

El siguiente resultado es una generalización del teorema de los valores intermedios para funciones reales de una variable a funciones escalares de varias variables:

Proposición 1.53 (teorema de los valores intermedios). *Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ continua en el conjunto conexo $A \subset \mathbb{R}^n$, sean $x, y \in A$, y supongamos que $f(x) \leq f(y)$. Entonces f toma todos los valores entre $f(x)$ y $f(y)$ (es decir, $[f(x), f(y)] \subset f(A)$).*

Demostración. Por hipótesis $f(x) \leq f(y)$ y $f(x), f(y) \in f(A)$. Al ser $f(A)$ un intervalo (por la proposición anterior), el intervalo $[f(x), f(y)]$ ha de estar contenido en $f(A)$, como habíamos afirmado. (En efecto, un intervalo es por definición un conjunto $I \subset \mathbb{R}$ con la siguiente propiedad: si $a, b \in I$ y $a \leq b$ entonces $[a, b] \subset I$). \square

Corolario 1.54. *Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ continua en un conjunto conexo y compacto $A \subset \mathbb{R}^n$. Entonces $f(A) = [m, M]$, donde m y M son respectivamente los valores mínimo y máximo de f en A .*

Demostración. Por las Proposiciones 1.52 y 1.53, $f(A)$ ha de ser un intervalo compacto $[m, M]$. En particular, esto implica que m y M son respectivamente los valores mínimo y máximo de f en A . \square

Definición 1.55. Un **homeomorfismo** entre dos conjuntos $A \subset \mathbb{R}^n, B \subset \mathbb{R}^m$ es una aplicación biyectiva $f : A \rightarrow B$ tal que f es continua en A y f^{-1} es continua en $B = f(A)$. Dos subconjuntos $A \subset \mathbb{R}^n, B \subset \mathbb{R}^m$ se dicen **homeomorfos** si existe algún homeomorfismo $f : A \rightarrow B$.

Puede demostrarse que dos conjuntos homeomorfos $A \subset \mathbb{R}^n, B \subset \mathbb{R}^m$ son *equivalentes* desde el punto de vista topológico, en el sentido de que tienen exactamente las *mismas* propiedades topológicas; por ejemplo, A es conexo o compacto si y solo si lo es B (ejercicio).

*Comentario**. Es importante observar que en la definición de homeomorfismo es imprescindible que f^{-1} , y no solo f , sea continua. En otras palabras, f puede ser continua y biyectiva sin que necesariamente f^{-1} sea continua, como muestra el siguiente ejemplo:

Ejemplo 1.56*. Sea $f : [0, 2\pi) \rightarrow S^1 \equiv \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\} \subset \mathbb{R}^2$ definida por

$$f(t) = (\cos t, \sin t).$$

Entonces f es claramente biyectiva, y también es continua (al serlo sus componentes). Sin embargo, f^{-1} no es continua en el punto $(1, 0)$. En efecto, nótese en primer lugar que $f^{-1}(1, 0) = 0$. Sin embargo, si $0 < \varepsilon < 2\pi$ no es posible encontrar un disco $B_\delta(1, 0)$ cuya imagen bajo f^{-1} esté contenida en el intervalo $(-\varepsilon, \varepsilon)$ centrado en 0. En efecto, las imágenes bajo f^{-1} de los puntos de $B_\delta(1, 0)$ con segunda componente negativa están arbitrariamente próximas al punto 2π , que por construcción no pertenece al intervalo $(-\varepsilon, \varepsilon)$. (Más informalmente, un punto $(\cos \theta, \sin \theta) \in S^1$ con $\theta < 2\pi$ próximo a 2π está próximo al punto $(1, 0)$ en S^1 , pero las imágenes bajo f^{-1} de estos puntos (los “ángulos” θ y 0) no están próximas en \mathbb{R}).

Capítulo 2

Diferenciabilidad

2.1 Derivada de una función vectorial

Una función $f : \mathbb{R} \rightarrow \mathbb{R}^m$ es derivable en un punto $a \in \mathbb{R}$ si está definida en un entorno de a , y existe el límite

$$\lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h} \equiv f'(a). \quad (2.1)$$

El valor de este límite es por definición la **derivada** de f en el punto a . Esta definición se generaliza inmediatamente a una *función vectorial* $f = (f_1, \dots, f_m) : \mathbb{R} \rightarrow \mathbb{R}^m$. En tal caso, la existencia del límite anterior es equivalente a la existencia de los m límites

$$\lim_{h \rightarrow 0} \frac{f_i(a+h) - f_i(a)}{h} = f'_i(a), \quad 1 \leq i \leq m,$$

siendo $f'(a) = (f'_1(a), \dots, f'_m(a))$. Geométricamente, la función vectorial $f : \mathbb{R} \rightarrow \mathbb{R}^m$ parametriza una curva en \mathbb{R}^m (la imagen de f). Si $f'(a) \neq 0$, el vector $f'(a) \in \mathbb{R}^m$ es un **vector tangente** a la curva en el punto $f(a)$. En efecto, $f'(a)$ es la posición límite del vector $(f(a+h) - f(a))/h$ cuando $h \rightarrow 0$, y este último vector es siempre paralelo a la secante a la curva de extremos $f(a)$ (fijo) y $f(a+h)$ (que tiende a $f(a)$ cuando $h \rightarrow 0$). En problemas físicos, la función vectorial $f(t)$ se interpreta a menudo como la *trayectoria* recorrida por una partícula, y el parámetro t suele representar el *tiempo*. En tal caso el vector $f'(t)$ es el *vector velocidad* de la partícula que recorre la curva.

2.2 Derivada de una función de \mathbb{R}^n en \mathbb{R}^m

Si f es una función de \mathbb{R}^n en \mathbb{R}^m con $n > 1$, ya no podemos utilizar directamente la igualdad (2.1) para definir la derivada de f en a , ya que en este caso h es un *vector*. Sin embargo, dicha igualdad se puede reescribir en la forma equivalente

$$\lim_{h \rightarrow 0} \frac{f(a+h) - f(a) - hf'(a)}{h} = 0,$$

es decir

$$f(a+h) - f(a) - hf'(a) = o(h),$$

o equivalentemente,

$$\Delta f(a; h) \equiv f(a+h) - f(a) = f'(a)h + o(h).$$

El miembro izquierdo de esta última igualdad es el **incremento** de la función f en el punto a cuando la variable dependiente se incrementa en h . Lo que afirma la igualdad anterior es que $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es derivable en el punto a si su incremento en a es una *función lineal* del incremento

h (el término $f'(a)h$) más un término (“infinitésimo”) *de orden menor que h* . Esta definición equivalente de derivabilidad sí es susceptible de ser generalizada al caso de funciones de \mathbb{R}^n en \mathbb{R}^m . Para ello es necesario revisar algunas nociones sobre funciones lineales de \mathbb{R}^n en \mathbb{R}^m , que serán ampliadas en el curso de Álgebra Lineal.

Definición 2.1. Un función $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es **lineal** si

$$A(\lambda x + \mu y) = \lambda A(x) + \mu A(y), \quad \forall \lambda, \mu \in \mathbb{R}, \forall x, y \in \mathbb{R}^n.$$

Nótese que en el caso de funciones lineales es muy frecuente omitir los paréntesis, y escribir $A \cdot x$ (o incluso Ax) en lugar de $A(x)$.

Las aplicaciones lineales admiten una representación particularmente sencilla en términos de *matrices*. Para ver esto, sea

$$\mathcal{B} = \{e_1, \dots, e_n\}$$

la *base*¹ *canónica* de \mathbb{R}^n , donde

$$e_i = (0, \dots, 0, \underset{\downarrow}{1}, 0, \dots, 0) \in \mathbb{R}^n.$$

Cualquier vector $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ se puede expresar en términos de los vectores de la base \mathcal{B} como

$$x = \sum_{j=1}^n x_j e_j;$$

en otras palabras, las *coordenadas* de x en la base canónica son simplemente sus *componentes* x_i . Si $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es una función lineal, el vector

$$y \equiv Ax = (y_1, \dots, y_m)$$

se expresa en términos de los vectores

$$e'_i = (0, \dots, 0, \underset{\downarrow}{1}, 0, \dots, 0) \in \mathbb{R}^m$$

de la base canónica $\mathcal{B}' = \{e'_1, \dots, e'_m\}$ de \mathbb{R}^m como

$$y = \sum_{i=1}^m y_i e'_i.$$

Las componentes $y_i = A_i(x) \in \mathbb{R}$ del vector $y = Ax$ se calculan fácilmente aplicando la linealidad de A :

$$y = Ax = A\left(\sum_{j=1}^n x_j e_j\right) = \sum_{j=1}^n x_j A e_j.$$

Cada uno de los n vectores $A e_j \in \mathbb{R}^m$, a su vez, ha de ser combinación lineal de los vectores de la base \mathcal{B}' , es decir

$$A e_j = \sum_{i=1}^m a_{ij} e'_i, \quad (2.2)$$

¹Una *base* de \mathbb{R}^n es un conjunto de n vectores linealmente independientes. Un subconjunto de \mathbb{R}^n es *linealmente independiente* si ninguno de sus elementos es combinación lineal de los restantes. Si $\mathcal{B} = \{v_1, \dots, v_n\}$ es una base de \mathbb{R}^n , todo vector $x \in \mathbb{R}^n$ se expresa de manera *única* como combinación lineal $x = \sum_{i=1}^n \lambda_i v_i$ de los vectores de \mathcal{B} . Por definición, los coeficientes λ_i de esta combinación lineal son las *coordenadas* del vector x respecto de la base \mathcal{B} .

con $a_{ij} \in \mathbb{R}$. Sustituyendo estas igualdades en la expresión anterior para Ax se obtiene inmediatamente

$$y = \sum_{i=1}^m \sum_{j=1}^n a_{ij} x_j e'_i \implies \boxed{y_i = \sum_{j=1}^n a_{ij} x_j, \quad 1 \leq i \leq m.} \quad (2.3)$$

En otras palabras, las coordenadas y_i del vector $y = Ax$ son *funciones lineales* de las x_i , como era de esperar. Los mn números reales a_{ij} ($1 \leq i \leq m$, $1 \leq j \leq n$) son por definición los *elementos de matriz* de A respecto de las bases canónicas \mathcal{B} y \mathcal{B}' . Más precisamente, si

$$A_{\mathcal{B}\mathcal{B}'} = \begin{pmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \dots & a_{mj} & \dots & a_{mn} \end{pmatrix} \quad (2.4)$$

es la matriz real $m \times n$ cuyos elementos de matriz son los números a_{ij} , las m igualdades (2.3) son equivalentes a la igualdad matricial

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = A_{\mathcal{B}\mathcal{B}'} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}. \quad (2.5)$$

La matriz $A_{\mathcal{B}\mathcal{B}'}$ se denomina **matriz de la aplicación lineal** A respecto de las bases canónicas \mathcal{B} y \mathcal{B}' . Nótese que la j -ésima *columna* de dicha matriz $(a_{ij})_{1 \leq i \leq m}$ no es otra cosa que la matriz columna formada por las componentes del vector Ae_j (cf. (2.2)). De la ec. (2.3) se deduce que

$$\boxed{Ax = (A_1(x), \dots, A_m(x)), \quad \text{con } A_i(x) = \sum_{j=1}^n a_{ij} x_j.} \quad (2.6)$$

En otras palabras, las componentes de una función lineal son funciones lineales de las componentes del vector $x \in \mathbb{R}^n$. Recíprocamente, es fácil ver que cualquier función de este tipo es lineal. Por tanto:

Una función $A \equiv (A_1, \dots, A_m) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es *lineal* si y solo si sus componentes son de la forma (2.6).

En otras palabras, la definición algebraica de función lineal 2.1 coincide con la dada anteriormente en el Ejemplo 1.35.

Notación.

- Es frecuente omitir los subíndices $\mathcal{B}, \mathcal{B}'$ y denotar (con un ligero abuso de notación) por A a la matriz $A_{\mathcal{B}\mathcal{B}'}$. En otras palabras, en este caso *se identifica la aplicación lineal (2.6) con su matriz (2.4) respecto de las bases canónicas de \mathbb{R}^n y \mathbb{R}^m .*
- En lo que sigue, denotaremos por $M_{\mathbb{R}}(m, n)$ el conjunto (espacio vectorial de dimensión mn) de las matrices reales $m \times n$.

Ejemplo 2.2. Sea $A : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ la función definida por

$$A(x) = (2x_1 - 3x_2, x_1 - x_2, x_1 + x_2). \quad (2.7)$$

Por el comentario anterior, A es una aplicación lineal y las componentes del vector $y = A(x)$ están dadas por

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 2 & -3 \\ 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

La matriz 3×2 que aparece en la fórmula anterior es la matriz de A en las bases canónicas de \mathbb{R}^2 y \mathbb{R}^3 , que como ya hemos dicho identificaremos con la aplicación lineal A . En otras palabras, escribiremos normalmente

$$A = \begin{pmatrix} 2 & -3 \\ 1 & -1 \\ 1 & 1 \end{pmatrix}$$

para denotar la aplicación (2.7).

Comentarios.

- Las aplicaciones lineales son funciones *continuas en cualquier punto*, ya que sus componentes son funciones lineales de las componentes x_i del vector x (cf. la ec. (2.6) y el Ejemplo 1.35).
- Si $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es una aplicación lineal, la función $f(x) = \|Ax\|$ es continua por serlo A , y por tanto alcanza un máximo y un mínimo en la esfera unidad $\partial B_1(0)$ (ya que dicho conjunto es compacto²). En particular, existe $v \in \partial B_1(0)$ tal que

$$\|Ax\| \leq \|Av\| \equiv \|A\|, \quad \forall x \in \mathbb{R}^n \text{ t.q. } \|x\| = 1.$$

Al número $\|A\|$ se le denomina **norma**³ de la matriz A . Si $h \in \mathbb{R}^n$ es un vector no nulo se tiene entonces

$$\|Ah\| = \|A(\|h\| \cdot h/\|h\|)\| = \|A(h/\|h\|)\| \cdot \|h\| \leq \|A\| \|h\|.$$

La desigualdad anterior es obviamente cierta también para $h = 0$, y por tanto es válida para todo vector $h \in \mathbb{R}^n$. \square

Estamos ya en condiciones de definir el concepto de diferenciabilidad para funciones de \mathbb{R}^n en \mathbb{R}^m . Para formular dicha definición de forma más sencilla, utilizaremos la notación $o(\|h\|)$ para denotar *cualquier* función $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ de la variable $h \in \mathbb{R}^n$ tal que

$$\lim_{h \rightarrow 0} \frac{g(h)}{\|h\|} = 0.$$

Equivalentemente, $g = o(\|h\|)$ si cada una de las m funciones componentes $g_i(h)$ es $o(\|h\|)$ en el sentido de la Definición 1.40. Nótese que de la definición anterior se deduce inmediatamente que

$$\lim_{h \rightarrow 0} o(\|h\|) = \lim_{h \rightarrow 0} \left(\frac{o(\|h\|)}{\|h\|} \|h\| \right) = 0 \cdot 0 = 0. \quad (2.8)$$

Definición 2.3. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, y sea $a \in \mathbb{R}^n$ un punto *interior* de $\mathcal{D}(f)$. Diremos que f es **diferenciable** en el punto a si existe una aplicación *lineal* $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ (que en general dependerá del punto a) tal que

$$f(a+h) = f(a) + Ah + o(\|h\|), \quad \forall h \in \mathbb{R}^n. \quad (2.9)$$

La función f es diferenciable en un *abierto* $U \subset \mathbb{R}^n$ si es diferenciable en todos los puntos de U .

²En efecto, es obviamente acotado, y también es cerrado al ser la imagen inversa de $\{1\}$ bajo la función continua $f(x) = \|x\|$.

³Con más propiedad, *norma espectral*, para distinguirla de otras normas que también se suelen utilizar en el espacio $M_{\mathbb{R}}(m, n)$.

En otras palabras,

f es diferenciable en $a \in \mathcal{D}(f)$ si el incremento $\Delta f(a; h)$ se puede aproximar por una función lineal de h cometiendo un error de orden menor que $\|h\|$.

Nótese también que la ec. (2.9) se puede expresar más explícitamente como

$$\lim_{h \rightarrow 0} \frac{f(a+h) - f(a) - Ah}{\|h\|} = 0, \quad (2.10)$$

o también

$$\lim_{h \rightarrow 0} \frac{\|f(a+h) - f(a) - Ah\|}{\|h\|} = 0. \quad (2.11)$$

Proposición 2.4. Si $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es diferenciable en $a \in \mathbb{R}^n$, la aplicación lineal A que aparece en (2.9) es única. En otras palabras, si $A_1, A_2 : \mathbb{R}^n \rightarrow \mathbb{R}^m$ son dos aplicaciones lineales que verifican

$$f(a+h) = f(a) + A_1 h + o(\|h\|) = f(a) + A_2 h + o(\|h\|), \quad \forall h \in \mathbb{R}^n,$$

entonces $A_1 = A_2$.

Demostración. De las igualdades anteriores se sigue que

$$A_1 h - A_2 h = o(\|h\|).$$

Si $v \neq 0$ es un vector fijo, sustituyendo $h = tv$ (con $t \neq 0$) en la igualdad anterior se obtiene

$$t(A_1 v - A_2 v) = o(|t|\|v\|) = o(|t|) \implies A_1 v - A_2 v = \frac{o(|t|)}{t}.$$

Haciendo t tender a 0 en esta igualdad se obtiene $A_1 v - A_2 v = 0$. Esto demuestra que $A_1 v = A_2 v$ para todo vector $v \neq 0$, y por tanto $A_1 = A_2$ (ya que $A(0) = 0$ para cualquier aplicación lineal A). \square

La proposición anterior motiva la siguiente definición:

Definición 2.5. Si $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es diferenciable en $a \in \mathbb{R}^n$, llamaremos **derivada de f en a** a la *única* aplicación lineal que satisface (2.9), y la denotaremos por $Df(a)$.

En otras palabras, si f es diferenciable en a su derivada es la *única* aplicación lineal $Df(a) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ que satisface

$$f(a+h) = f(a) + Df(a) \cdot h + o(\|h\|), \quad (2.12)$$

es decir

$$\lim_{h \rightarrow 0} \frac{f(a+h) - f(a) - Df(a) \cdot h}{\|h\|} = 0, \quad (2.13)$$

o bien

$$\lim_{h \rightarrow 0} \frac{\|f(a+h) - f(a) - Df(a) \cdot h\|}{\|h\|} = 0. \quad (2.14)$$

Nótese también que, al ser la derivada una aplicación lineal de \mathbb{R}^n en \mathbb{R}^m , se puede identificar con una *matriz* $m \times n$ (su matriz en las bases canónicas de \mathbb{R}^n y \mathbb{R}^m).

Ejemplo 2.6. Sea la función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por $f(x, y) = xy$. En este caso se tiene

$$f(x+h, y+k) - f(x, y) = (x+h)(y+k) - xy = xk + yh + hk.$$

Si fijamos el punto $(x, y) \in \mathbb{R}^2$, los dos primeros sumandos $xk + yh$ son una función lineal del vector (h, k) . En cuanto al término restante,

$$\frac{hk}{\|(h, k)\|} = \frac{hk}{\sqrt{h^2 + k^2}} = \underbrace{\frac{h}{\sqrt{h^2 + k^2}}}_{\text{acotado por 1}} k$$

tiende a cero cuando $(h, k) \rightarrow (0, 0)$ en virtud de la Proposición 1.29. Por tanto $hk = o(\|(h, k)\|)$, y se satisface la definición de diferenciabilidad en (x, y) con

$$Df(x, y) \cdot (h, k) = yh + xk = (y \ x) \begin{pmatrix} h \\ k \end{pmatrix}. \quad (2.15)$$

Esto demuestra que f es diferenciable en todo \mathbb{R}^2 , siendo

$$\boxed{Df(x, y) = (y \ x)}.$$

Nótese que en la igualdad anterior hemos identificado la derivada $Df(x, y)$ con su matriz en las bases canónicas de \mathbb{R}^2 y \mathbb{R} , como haremos frecuentemente. \square

Ejemplo 2.7. Sea $f = (f_1, \dots, f_m) : \mathbb{R} \rightarrow \mathbb{R}^m$ una *función vectorial*. Es inmediato probar que f es diferenciable en $t \in \mathbb{R}$ si y solo si existe $f'(t) = (f'_1(t), \dots, f'_m(t))$ (cf. la ec. (2.1)), y en tal caso

$$\boxed{Df(t) = \begin{pmatrix} f'_1(t) \\ \vdots \\ f'_m(t) \end{pmatrix}}. \quad (2.16)$$

En efecto, en este caso $h \in \mathbb{R}$, $\|h\| = |h|$, $Df(t) \cdot h = h Df(t) \cdot 1$ y

$$\frac{\|f(t+h) - f(t) - Df(t) \cdot h\|}{\|h\|} = \left\| \frac{f(t+h) - f(t) - Df(t) \cdot h}{h} \right\| = \left\| \frac{f(t+h) - f(t)}{h} - Df(t) \cdot 1 \right\|$$

tiende a 0 cuando $h \rightarrow 0$ si y solo si

$$\exists \lim_{h \rightarrow 0} \frac{f(t+h) - f(t)}{h} = Df(t) \cdot 1.$$

Como

$$\lim_{h \rightarrow 0} \frac{f(t+h) - f(t)}{h} = f'(t),$$

esto demuestra que f es diferenciable en t si y solo existe $f'(t)$, y en tal caso

$$Df(t) \cdot 1 = f'(t) = (f'_1(t), \dots, f'_m(t)).$$

Al ser la derivada $Df(t)$ una aplicación lineal de \mathbb{R} en \mathbb{R}^m , la podemos identificar con la matriz columna $m \times 1$ formada por las m componentes del vector $Df(t) \cdot 1$ (ya que 1 es el único elemento de la base canónica de \mathbb{R} !). En virtud de la igualdad anterior, dichas componentes son las derivadas $f'_i(t)$, $1 \leq i \leq m$, lo que demuestra (2.16). Nótese por último que, con un ligero abuso de notación, podemos escribir esta ecuación en la forma más sintética

$$Df(t) = f'(t),$$

si identificamos el vector $f'(t)$ con la matriz *columna* formada por sus componentes. \square

Al igual que en el caso de funciones de \mathbb{R} en \mathbb{R} , la diferenciabilidad de una función en un punto implica su continuidad:

Proposición 2.8. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ diferenciable en $a \in \mathbb{R}^n$. Entonces f es continua en dicho punto.

Demostración. En efecto,

$$\lim_{x \rightarrow a} f(x) = \lim_{h \rightarrow 0} f(a + h) = \lim_{h \rightarrow 0} [f(a) + Df(a) \cdot h + o(\|h\|)] = f(a),$$

ya que $\lim_{h \rightarrow 0} Df(a) \cdot h = Df(a) \cdot 0 = 0$ al ser continua la aplicación lineal $Df(a)$, y $\lim_{h \rightarrow 0} o(\|h\|) = 0$ en virtud de (2.8). □

La derivada de una función tiene las propiedades elementales siguientes, que se prueban fácilmente a partir de su definición:

i) Si $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ y $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ son ambas diferenciables en a , entonces $\lambda f + \mu g$ es diferenciable en $a \forall \lambda, \mu \in \mathbb{R}$, y se tiene

$$D(\lambda f + \mu g)(a) = \lambda Df(a) + \mu Dg(a).$$

ii) Si f es constante es diferenciable en cualquier punto, siendo $Df(a) = 0$ para todo $a \in \mathbb{R}^n$.

iii) Si $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es una aplicación lineal, entonces L es diferenciable en cualquier punto, con derivada

$$DL(a) = L, \quad \forall a \in \mathbb{R}^n.$$

iv) $f = (f_1, \dots, f_m) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es diferenciable en a si y solo si lo son sus componentes $f_i : \mathbb{R}^n \rightarrow \mathbb{R} (1 \leq i \leq m)$. Además, $Df_i(a)$ coincide con la i -ésima fila de $Df(a)$.

Dada su importancia, demostraremos en detalle la última propiedad. Para ello, nótese que si

$$g(h) \equiv f(a + h) - f(a) - Df(a) \cdot h$$

entonces $g(h) = o(\|h\|)$ si y solo si $g_i(h) = o(\|h\|)$ para todo $i = 1, \dots, m$, en virtud de la Proposición 1.25. Si denotamos por A_{ij} (con $1 \leq i \leq m, 1 \leq j \leq n$) a los elementos de matriz de $Df(a)$ en las bases canónicas de \mathbb{R}^n y \mathbb{R}^m entonces

$$g_i(h) = f_i(a + h) - f_i(a) - \sum_{j=1}^n A_{ij} h_j$$

es $o(\|h\|)$ si y solo si f_i es diferenciable en a , con derivada

$$Df_i(a) \cdot h = \sum_{j=1}^n A_{ij} h_j = (A_{i1} \ \cdots \ A_{in}) \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix},$$

o equivalentemente,

$$Df_i(a) = (A_{i1} \ \cdots \ A_{in}) = i\text{-ésima fila de } Df(a). \quad \square$$

2.3 Derivadas direccionales y parciales

Un inconveniente de la definición de derivada es que, a diferencia de lo que ocurre en el caso $n = 1$ (cf. el Ejemplo 2.7), no se reduce a calcular un límite concreto, ya que en ella aparece la

aplicación lineal $A \equiv Df(a)$ que no se conoce a priori. Para resolver este problema, estudiemos el límite (2.14) a lo largo de la recta por el origen $h = tv$ (con $v \neq 0$ fijo y $t \in \mathbb{R}$). Si f es diferenciable en a , este límite debe existir y valer cero cualquiera que sea la recta considerada. Como

$$\frac{\|f(a + tv) - f(a) - tDf(a) \cdot v\|}{|t|\|v\|} = \frac{1}{\|v\|} \left\| \frac{f(a + tv) - f(a)}{t} - Df(a) \cdot v \right\|,$$

si f es diferenciable en a el límite del miembro derecho cuando $t \rightarrow 0$ existe y vale cero. Por tanto

$$f \text{ diferenciable en } a \implies Df(a) \cdot v = \lim_{t \rightarrow 0} \frac{f(a + tv) - f(a)}{t}, \quad \forall v \in \mathbb{R}^n \quad (2.17)$$

(ya que la igualdad anterior es obviamente válida también para $v = 0$). Esto motiva la siguiente definición:

Definición 2.9. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ una función definida en un entorno de $a \in \mathbb{R}^n$. La **derivada direccional** de f en el punto a según la dirección del vector $v \in \mathbb{R}^n$, que denotaremos por $D_v f(a)$, es el vector

$$D_v f(a) = \lim_{t \rightarrow 0} \frac{f(a + tv) - f(a)}{t} \in \mathbb{R}^m. \quad (2.18)$$

Nótese que

$$D_v f(a) = \left. \frac{d}{dt} f(a + tv) \right|_{t=0} = (D_v f_1(a), \dots, D_v f_m(a)). \quad (2.19)$$

El resultado que acabamos de demostrar (ec. (2.17)) se puede entonces enunciar como sigue:

Proposición 2.10. Si $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es diferenciable en a , entonces existen las derivadas direccionales de f en a según todas las direcciones, y se tiene

$$D_v f(a) = Df(a) \cdot v, \quad \forall v \in \mathbb{R}^n. \quad (2.20)$$

Comentarios.

- De (2.20) se deduce que si f es diferenciable en a entonces $D_v f(a)$ es una función *lineal* del vector $v \in \mathbb{R}^n$.
- Es importante observar que el recíproco de la Proposición 2.10 *no* tiene por qué ser cierto. En otras palabras:

La existencia de las derivadas direccionales de f en a según todas las direcciones *no* implica la diferenciable de f en el punto a .

- De hecho, la existencia de las derivadas direccionales en un punto según todas las direcciones ni siquiera implica la *continuidad* de la función en dicho punto, como demuestra el siguiente ejemplo.

Ejemplo 2.11. Sea $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ la función definida por $f(0, 0) = 0$ y

$$f(x, y) = \frac{xy^2}{x^2 + y^4}, \quad \forall (x, y) \neq (0, 0).$$

Vimos en el Ejemplo 1.39 que f es discontinua en $(0, 0)$, ya que el límite de f a lo largo de parábolas $x = cy^2$ depende del parámetro c . Por tanto f no es diferenciable en el origen, ya que si lo fuera sería continua en dicho punto en virtud de la Proposición 2.8. Sin embargo, es fácil comprobar que f tiene derivada direccional en el origen según cualquier dirección $v = (v_1, v_2)$. En efecto, si $v \neq 0$ se tiene⁴

$$D_v f(0, 0) = \lim_{t \rightarrow 0} \frac{f(tv_1, tv_2)}{t} = \lim_{t \rightarrow 0} \frac{v_1 v_2^2}{v_1^2 + t^2 v_2^2} = \begin{cases} v_2^2/v_1, & v_1 \neq 0 \\ 0, & v_1 = 0. \end{cases}$$

Nótese también que $D_v f(0, 0)$, aunque está definido para todo $v \in \mathbb{R}^2$, no es lineal en v . De esto también se sigue que f no es diferenciable en $(0, 0)$, ya que si lo fuera $D_v f(0, 0)$ sería lineal en v . \square

Si $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es diferenciable en a , la ecuación (2.20) permite calcular la derivada $Df(a)$ en función de las derivadas direccionales de f en dicho punto. En particular, para calcular la matriz de $Df(a)$ en las bases canónicas de \mathbb{R}^n y \mathbb{R}^m solo necesitamos conocer las derivadas direccionales de f en las direcciones de los vectores e_i ($1 \leq i \leq n$) de la base canónica de \mathbb{R}^n . Estas derivadas direccionales, que son especialmente importantes, reciben el nombre de derivadas parciales:

Definición 2.12. La i -ésima **derivada parcial** de f en el punto $a \in \mathbb{R}^n$, que denotaremos indistintamente por $D_i f(a)$ o $\frac{\partial f}{\partial x_i}(a)$, es la derivada direccional en a según la dirección del vector e_i , es decir

$$D_i f(a) \equiv \frac{\partial f}{\partial x_i}(a) = D_{e_i} f(a), \quad 1 \leq i \leq n.$$

Nótese que, en virtud de la segunda igualdad en la ec. (2.19),

$$\frac{\partial f}{\partial x_i}(a) = \left(\frac{\partial f_1}{\partial x_i}(a), \dots, \frac{\partial f_m}{\partial x_i}(a) \right).$$

Notación. Si $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ es una función de dos variables reales (x, y, z) , denotaremos frecuentemente sus derivadas parciales respecto de dichas variables por f_x, f_y y f_z . En otras palabras, en este caso

$$f_x \equiv \frac{\partial f}{\partial x} = D_1 f, \quad f_y \equiv \frac{\partial f}{\partial y} = D_2 f, \quad f_z \equiv \frac{\partial f}{\partial z} = D_3 f.$$

Esta notación se suele extender a funciones escalares de n variables reales (x_1, \dots, x_n) , denotando

$$f_{,i} \equiv \frac{\partial f}{\partial x_i} = D_i f, \quad 1 \leq i \leq n.$$

En lo que sigue no utilizaremos esta última notación, ya que la derivada parcial $f_{,i}$ podría confundirse con la i -ésima componente f_i de una función $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$. \square

Utilizando la definición de derivada direccional (2.18) se obtiene la siguiente expresión explícita para la derivada parcial $D_i f(a)$:

$$\begin{aligned} \frac{\partial f}{\partial x_i}(a) &= \lim_{h \rightarrow 0} \frac{1}{h} [f(a_1, \dots, a_{i-1}, a_i + h, a_{i+1}, \dots, a_n) - f(a)] \\ &\equiv \frac{d}{dx_i} f(a_1, \dots, a_{i-1}, x_i, a_{i+1}, \dots, a_n) \Big|_{x_i=a_i}. \end{aligned}$$

⁴Es inmediato comprobar a partir de la Definición 2.9 que $D_0 f(a) = 0$ para toda función f y para todo $a \in \mathbb{R}^n$.

En otras palabras, $\frac{\partial f}{\partial x_i}(a)$ se obtiene fijando las variables x_j con $j \neq i$ y calculando la derivada de f respecto de la variable x_i en el punto a_i .

- En particular, nótese que para calcular $\frac{\partial f}{\partial x_i}(a)$ podemos sustituir las variables x_j con $j \neq i$ por sus correspondientes valores a_j antes de derivar respecto de x_i , lo cual en ocasiones simplifica considerablemente el cálculo. Por ejemplo, si

$$f(x, y) = x^{x^y} + \arctan(\arctan(\cos(xy) - \log(x + y))) \log x,$$

entonces

$$\frac{\partial f}{\partial y}(1, y) = \frac{d}{dy} f(1, y) = \frac{d}{dy} 1 = 0. \quad \square$$

Si f es diferenciable en $a \in \mathbb{R}^n$, la matriz de $Df(a)$ en las bases canónicas de \mathbb{R}^n y \mathbb{R}^m tiene por j -ésima columna las m componentes del vector $Df(a) \cdot e_j$. Pero, al ser f diferenciable en a , en virtud de (2.20) se tiene

$$Df(a) \cdot e_j = D_{e_j} f(a) \equiv \frac{\partial f}{\partial x_j}(a).$$

Por tanto la j -ésima columna de la matriz de $Df(a)$ en las bases canónicas de \mathbb{R}^n y \mathbb{R}^m es

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_j}(a) \\ \vdots \\ \frac{\partial f_i}{\partial x_j}(a) \\ \vdots \\ \frac{\partial f_m}{\partial x_j}(a) \end{pmatrix}.$$

Identificando, como solemos hacer, $Df(a)$ con su matriz en las bases canónicas de \mathbb{R}^n y \mathbb{R}^m , podemos escribir

$$Df(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \cdots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \cdots & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix}. \quad (2.21)$$

La matriz $m \times n$ que aparece en la ecuación anterior recibe el nombre de **matriz jacobiana** de f en a . Nótese que la i -ésima fila de la matriz jacobiana está formada por las n derivadas parciales en a de la i -ésima componente f_i de f (cf. también la cuarta propiedad elemental de la derivada). En otras palabras, el elemento de matriz en la fila i y columna j de la matriz jacobiana es la derivada parcial $\frac{\partial f_i}{\partial x_j}(a)$.

Al ser $Df(a)$ lineal, su actuación sobre cualquier vector $h = (h_1, \dots, h_n) \in \mathbb{R}^n$ se puede expresar en términos de las derivadas parciales de f en a como sigue:

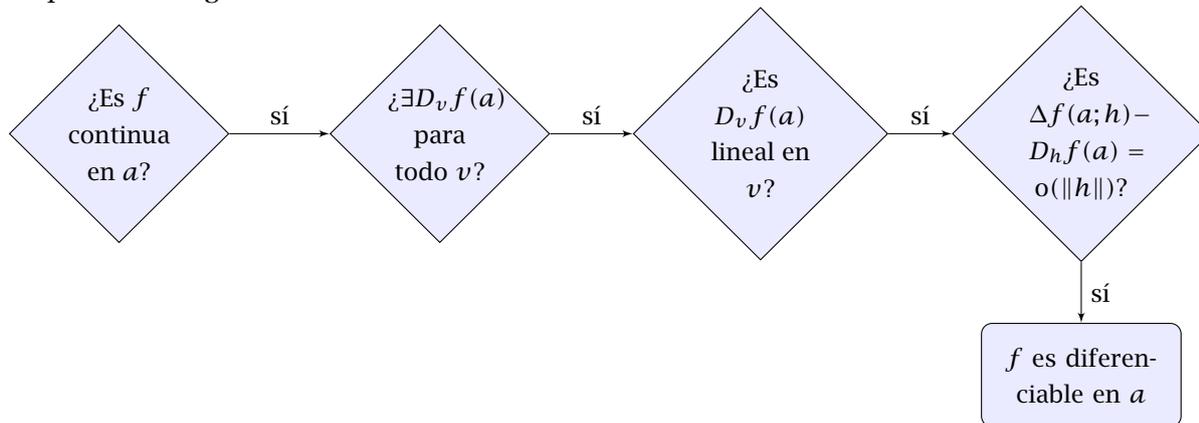
$$\begin{aligned} Df(a) \cdot h &= Df(a) \cdot (h_1 e_1 + \cdots + h_n e_n) = h_1 Df(a) \cdot e_1 + \cdots + h_n Df(a) \cdot e_n \\ &= h_1 \frac{\partial f}{\partial x_1}(a) + \cdots + h_n \frac{\partial f}{\partial x_n}(a). \end{aligned} \quad (2.22)$$

Por tanto si $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ es diferenciable en $a \in \mathbb{R}^n$ se tiene

$$f(a + h) = f(a) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) h_i + o(\|h\|). \quad (2.23)$$

Comentario. Es importante observar que pueden existir todas las derivadas parciales, y por tanto la matriz jacobiana, de f en un punto a sin que f sea diferenciable (¡ni siquiera continua!) en dicho punto (véase 2.11 o el siguiente ejemplo).

Los resultados anteriores proporcionan un método sencillo para determinar si una función es diferenciable en un punto y , en tal caso, calcular su derivada. Básicamente, hay que ir respondiendo a las preguntas que aparecen en el siguiente diagrama de flujo, y si en algún momento la respuesta es negativa la función *no* es diferenciable en a .



Ejemplo 2.13. Sea $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ la función definida por $f(0, 0) = 0$ y

$$f(x, y) = \frac{|x|^a |y|^b}{x^2 + y^2}, \quad \forall (x, y) \neq (0, 0),$$

con $a, b > 0$. Estudiemos para qué valores de los parámetros a y b es f diferenciable en el origen. En primer lugar, vimos en el Ejemplo 1.38 que f es continua en $(0, 0)$ si y solo si $a + b > 2$. Suponiendo que se cumple esta condición, estudiemos a continuación si existen las derivadas direccionales en $(0, 0)$ según cualquier dirección $v = (v_1, v_2) \neq 0$:

$$D_v f(0, 0) = \lim_{t \rightarrow 0} \frac{f(tv)}{t} = \frac{|v_1|^a |v_2|^b}{v_1^2 + v_2^2} \lim_{t \rightarrow 0} (\text{sgn } t \cdot |t|^{a+b-3}). \quad (2.24)$$

El límite anterior se anula si $v_1 v_2 = 0$, y por tanto

$$\frac{\partial f}{\partial x}(0, 0) = \frac{\partial f}{\partial y}(0, 0) = 0$$

para todo $a, b > 0$. En particular, si $a + b \leq 2$ existe la matriz jacobiana de f en el origen, pero f no es ni siquiera continua en dicho punto.

Estudiemos a continuación la existencia de las derivadas direccionales de f con $v_1 v_2 \neq 0$. En tal caso el límite (2.24) no existe (es $\pm\infty$) si $a + b < 3$, tampoco existe si $a + b = 3$ (vale $|v_1|^a |v_2|^b / (v_1^2 + v_2^2)$ por la derecha y $-|v_1|^a |v_2|^b / (v_1^2 + v_2^2)$ por la izquierda), y existe y vale 0 si $a + b > 3$. Por tanto todas las derivadas direccionales existen (y valen cero) si y solo si $a + b > 3$. En este último caso $D_v f(0, 0) = 0$ para todo $v \in \mathbb{R}^2$, y por tanto de existir $Df(0, 0)$ tendría que anularse. Por consiguiente, para determinar si f es diferenciable en el origen cuando $a + b > 3$ hay que estudiar si $f(h) / \|h\|$ tiende a 0 para $h \rightarrow 0$. Pero

$$\frac{f(h)}{\|h\|} = \frac{|h_1|^a |h_2|^b}{\|h\|^3} = \left(\frac{|h_1|}{\|h\|}\right)^a \left(\frac{|h_2|}{\|h\|}\right)^b \|h\|^{a+b-3} \xrightarrow{h \rightarrow 0} 0,$$

ya que los dos primeros factores están acotados (por 1) y el último tiende a 0 en virtud de la condición $a + b - 3 > 0$. Por tanto f es diferenciable en el origen si y solo si $a + b > 3$, y en tal caso $Df(0, 0) = 0$. □

El caso de *funciones escalares* es particularmente importante, ya que $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es diferenciable si y solo si lo son sus componentes $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ ($1 \leq i \leq m$). La matriz jacobiana de una función escalar $f : \mathbb{R}^n \rightarrow \mathbb{R}$ en un punto $a \in \mathbb{R}^n$ es de orden $1 \times n$, es decir es una *matriz fila* cuyas componentes son las derivadas parciales de f en el punto a :

$$Df(a) = \left(\frac{\partial f}{\partial x_1}(a) \quad \cdots \quad \frac{\partial f}{\partial x_n}(a) \right).$$

Por tanto en este caso

$$\boxed{Df(a) \cdot h = \nabla f(a) \cdot h}, \quad (2.25)$$

donde el punto del miembro derecho significa *producto escalar*, y

$$\boxed{\nabla f(a) \equiv \left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a) \right) \in \mathbb{R}^n}$$

es un vector de \mathbb{R}^n que recibe el nombre de **gradiente** de la función f en el punto a . Por ejemplo, si $f(x) = \|x\|^2 \equiv x_1^2 + \cdots + x_n^2$ entonces

$$\nabla f(x) = (2x_1, \dots, 2x_n) = 2x.$$

Nótese, por último, que en el caso general de una función $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ la i -ésima fila de la matriz⁵ $Df(a)$ es la matriz $Df_i(a)$, que a su vez se puede identificar con $\nabla f_i(a)$ considerado como un vector fila.

Notación. Es preciso distinguir entre $x \equiv (x_1, \dots, x_n) \in \mathbb{R}^n$, que es un *vector* de \mathbb{R}^n , y la *matriz fila* $(x_1 \ \cdots \ x_n) \in M_{1 \times n}(\mathbb{R})$ (nótese la ausencia de comas en este último caso). En general, si f es una función escalar las componentes del vector $\nabla f(a)$ y los elementos de matriz de $Df(a)$ solo coinciden en la base canónica de \mathbb{R}^n , ya que vectores y matrices columna se transforman de forma distinta bajo cambios de base. De hecho, la matriz asociada de forma natural al vector x

es la matriz *columna* $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ (cf., por ejemplo, la ec. (2.5)).

2.4 Funciones de clase C^1

Es evidente que determinar la diferenciabilidad de una función en un punto aplicando únicamente la definición es casi siempre un procedimiento muy laborioso. Afortunadamente, hay varios resultados generales que facilitan considerablemente esta tarea. El primero de ellos, que veremos a continuación, está basado en la continuidad de las derivadas parciales. La demostración de este resultado es una simple aplicación del teorema del valor medio para funciones de \mathbb{R} en \mathbb{R} , cuyo enunciado recordamos a continuación:

Teorema del valor medio. Si $f : \mathbb{R} \rightarrow \mathbb{R}$ es continua en $[a, b]$ y derivable en (a, b) , entonces existe $c \in (a, b)$ tal que $f(b) = f(a) + (b - a)f'(c)$.

Proposición 2.14. Supongamos que una de las derivadas parciales de $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ existe en $a \in \mathbb{R}^n$, y las restantes $n - 1$ derivadas parciales existen en un entorno de a y son continuas en dicho punto. Entonces f es diferenciable en a .

⁵Como de costumbre, identificamos las aplicaciones lineales con sus matrices en las bases canónicas correspondientes.

Demostración. En primer lugar, el caso general se sigue del caso $m = 1$ aplicado a cada una de las componentes de f (recuérdese que f es diferenciable en a si y solo si lo son cada una de sus componentes). Por tanto, podemos restringirnos sin pérdida de generalidad al caso en que f es una función escalar. Supongamos que (por ejemplo) $D_1f(a)$ existe, y que las derivadas parciales $D_i f$ con $2 \leq i \leq n$ existen en un entorno de a y son continuas en dicho punto. Nótese que si f es diferenciable en a , necesariamente su derivada tiene que venir dada en términos de las n derivadas parciales $D_i f(a)$ por (2.22). Por tanto, para demostrar que f es diferenciable en a hay que probar que

$$\lim_{h \rightarrow 0} \frac{f(a+h) - f(a) - \sum_{i=1}^n D_i f(a) h_i}{\|h\|} = 0.$$

Para ello, escribimos primero $f(a+h) - f(a)$ como la *suma telescópica*

$$\begin{aligned} f(a+h) - f(a) &= f(a_1 + h_1, a_2, \dots, a_n) - f(a) \\ &+ [f(a_1 + h_1, a_2 + h_2, a_3, \dots, a_n) - f(a_1 + h_1, a_2, \dots, a_n)] \\ &+ \dots + [f(a+h) - f(a_1 + h_1, \dots, a_{n-1} + h_{n-1}, a_n)] \\ &\equiv f(a_1 + h_1, a_2, \dots, a_n) - f(a) + \sum_{i=2}^n [g_i(a_i + h_i) - g_i(a_i)], \end{aligned} \quad (2.26)$$

donde las funciones $g_i : \mathbb{R} \rightarrow \mathbb{R}$ están definidas por

$$g_i(t) = f(a_1 + h_1, \dots, a_{i-1} + h_{i-1}, t, a_{i+1}, \dots, a_n), \quad 2 \leq i \leq n.$$

Las hipótesis hechas implican que para $\|h\|$ suficientemente pequeño g_i es continua en el intervalo $[a_i, a_i + h_i]$ (o $[a_i + h_i, a_i]$, si $h_i < 0$) y diferenciable en el interior de dicho intervalo, siendo

$$g'_i(t) = D_i f(a_1 + h_1, \dots, a_{i-1} + h_{i-1}, t, a_{i+1}, \dots, a_n).$$

Por el teorema del valor medio, existe $s_i(h_i) \equiv s_i \in (0, 1)$ tal que

$$\begin{aligned} g_i(a_i + h_i) - g_i(a_i) &= h_i g'_i(a_i + s_i h_i) \\ &= h_i D_i f(a_1 + h_1, \dots, a_{i-1} + h_{i-1}, a_i + s_i h_i, a_{i+1}, \dots, a_n). \end{aligned}$$

Utilizando estas expresiones en (2.26) se obtiene

$$\begin{aligned} \frac{f(a+h) - f(a) - \sum_{i=1}^n D_i f(a) h_i}{\|h\|} &= \frac{f(a_1 + h_1, a_2, \dots, a_n) - f(a) - h_1 D_1 f(a)}{\|h\|} \\ &+ \sum_{i=2}^n [D_i f(a_1 + h_1, \dots, a_{i-1} + h_{i-1}, a_i + s_i h_i, a_{i+1}, \dots, a_n) - D_i f(a)] \frac{h_i}{\|h\|}. \end{aligned}$$

Cada término en el sumatorio tiende claramente a 0 cuando $h \rightarrow 0$, ya que $h_i/\|h\|$ está acotado y

$$D_i f(a_1 + h_1, \dots, a_{i-1} + h_{i-1}, a_i + s_i h_i, a_{i+1}, \dots, a_n) - D_i f(a) \xrightarrow{h \rightarrow 0} 0$$

por la continuidad de $D_i f$ en a para $2 \leq i \leq n$ (nótese que $0 < s_i(h_i) < 1$). En cuanto al término restante, la existencia de

$$D_1 f(a) = \lim_{h_1 \rightarrow 0} \frac{f(a_1 + h_1, a_2, \dots, a_n) - f(a)}{h_1}$$

implica que

$$f(a_1 + h_1, a_2, \dots, a_n) - f(a) - h_1 D_1 f(a) = h_1 o(1),$$

y por tanto

$$\frac{f(a_1 + h_1, a_2, \dots, a_n) - f(a) - h_1 D_1 f(a)}{\|h\|} = o(1) \frac{h_1}{\|h\|} \xrightarrow{h \rightarrow 0} 0$$

en virtud de la Proposición 1.29

□

Comentario. La proposición anterior proporciona una condición *suficiente*, pero no necesaria, de diferenciabilidad. En otras palabras, *hay funciones diferenciables en un punto cuyas derivadas parciales son discontinuas en dicho punto*. Por ejemplo, la función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $f(0,0) = 0$ y $f(x,y) = (x^2 + y^2) \sin((x^2 + y^2)^{-1})$ es diferenciable en el origen, pero sus derivadas parciales no son continuas en dicho punto (ejercicio).

Definición 2.15. Sea U un abierto de \mathbb{R}^n . Se dice que $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es **de clase C^0** en U , y se escribe $f \in C^0(U)$, si f es continua en U . Si todas las derivadas parciales de f existen y son continuas en U , se dice que f es **de clase C^1** en U , lo que se denota por $f \in C^1(U)$.

Evidentemente, $f \in C^1(U)$ si y solo si $f_i \in C^1(U)$ para todo $i = 1, \dots, m$.

Corolario 2.16. Si $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es de clase C^1 en un abierto U , entonces f es diferenciable en U .

- Nótese que del corolario anterior y la Proposición 2.8 se sigue que

$$C^1(U) \subset C^0(U).$$

Ejemplo 2.17. Un *polinomio* es una función diferenciable en cualquier punto. En efecto, las derivadas parciales de un polinomio respecto de cualquier coordenada son polinomios, y por tanto son continuas en todo punto. De la misma forma se demuestra que una *función racional* es diferenciable en todos los puntos en que no se anula su denominador (que es un conjunto *abierto*, al ser la imagen inversa de $\mathbb{R} - \{0\}$ bajo un polinomio).

2.5 Regla de la cadena

Otro de los resultados generales que más se utilizan para establecer la diferenciabilidad de una función en un punto y hallar su derivada es la regla de la cadena, que establece la diferenciabilidad de la composición de funciones diferenciables:

Regla de la cadena. Supongamos que $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es diferenciable en $a \in \mathbb{R}^n$, y $f : \mathbb{R}^m \rightarrow \mathbb{R}^p$ es diferenciable en $g(a) \in \mathbb{R}^m$. Entonces la composición $f \circ g : \mathbb{R}^n \rightarrow \mathbb{R}^p$ es diferenciable en a , y su derivada en dicho punto está dada por

$$D(f \circ g)(a) = Df(g(a)) \cdot Dg(a). \quad (2.27)$$

Comentario. Nótese que el punto en el miembro derecho denota la *composición* de las aplicaciones lineales $Df(g(a))$ y $Dg(a)$, o equivalentemente (identificando, como de costumbre, una aplicación lineal con su matriz en las bases canónicas) el *producto* de las correspondientes *matrices jacobianas*.

Demostración. En primer lugar, nótese que la composición $f \circ g$ está definida en un entorno de a . En efecto, al ser f diferenciable en $g(a)$ ha de estar definida en una cierta bola abierta $B_r(g(a))$. Como, a su vez, g es continua en a , hay un cierto $\delta > 0$ tal que $g(B_\delta(a)) \subset B_r(g(a)) \subset \mathcal{D}(f)$. En particular, esto implica que $B_\delta(a) \subset \mathcal{D}(f \circ g)$. En segundo lugar, al ser g y f por hipótesis diferenciables en a y en $g(a)$, respectivamente, se tiene

$$g(a+h) = g(a) + Dg(a) \cdot h + \|h\|\lambda(h), \quad f(g(a)+k) = f(g(a)) + Df(g(a)) \cdot k + \|k\|\mu(k),$$

con

$$\lim_{h \rightarrow 0} \lambda(h) = \lim_{k \rightarrow 0} \mu(k) = 0.$$

De estas ecuaciones se deduce que

$$\begin{aligned} f(g(a+h)) &= f(g(a) + Dg(a) \cdot h + \|h\|\lambda(h)) \\ &= f(g(a)) + Df(g(a))Dg(a) \cdot h + \|h\| Df(g(a)) \cdot \lambda(h) + \|k(h)\| \mu(k(h)), \end{aligned}$$

con $k(h) \equiv Dg(a) \cdot h + \|h\|\lambda(h)$. Para concluir la demostración, basta probar que los dos últimos sumandos son $o(\|h\|)$. En el caso del primero esto es inmediato, ya que

$$\lim_{h \rightarrow 0} [Df(g(a)) \cdot \lambda(h)] = Df(g(a)) \cdot 0 = 0$$

por la continuidad de las aplicaciones lineales. En cuanto al segundo, nótese en primer lugar que $k(h) \rightarrow 0$ cuando $h \rightarrow 0$ (de nuevo por la continuidad de las aplicaciones lineales). Además,

$$\|k(h)\| \leq (\|Dg(a)\| + \|\lambda(h)\|)\|h\|,$$

y por tanto si $h \neq 0$ se tiene

$$\frac{\|k(h)\| \| \mu(k(h)) \|}{\|h\|} \leq (\|Dg(a)\| + \|\lambda(h)\|) \| \mu(k(h)) \| \xrightarrow{h \rightarrow 0} \|Dg(a)\| \cdot 0 = 0. \quad \square$$

La regla de la cadena es particularmente importante en el caso en que f , y por tanto $f \circ g$, es una función escalar (es decir, si $p = 1$). Si $F = f \circ g$ entonces

$$F(x_1, \dots, x_n) = f(g_1(x_1, \dots, x_n), \dots, g_m(x_1, \dots, x_n)),$$

y (2.27) se escribe

$$\left(\frac{\partial F}{\partial x_1}(a) \quad \dots \quad \frac{\partial F}{\partial x_n}(a) \right) = \left(\frac{\partial f}{\partial y_1}(g(a)) \quad \dots \quad \frac{\partial f}{\partial y_m}(g(a)) \right) \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(a) & \dots & \frac{\partial g_1}{\partial x_n}(a) \\ \vdots & \ddots & \vdots \\ \frac{\partial g_m}{\partial x_1}(a) & \dots & \frac{\partial g_m}{\partial x_n}(a) \end{pmatrix},$$

donde hemos denotado por (y_1, \dots, y_m) las m variables de las cuales depende f . Efectuando la multiplicación de matrices se obtiene la siguiente expresión para las derivadas parciales de $F(x) \equiv f(g(x))$:

$$\boxed{\frac{\partial F}{\partial x_i}(a) = \sum_{j=1}^m \frac{\partial f}{\partial y_j}(g(a)) \frac{\partial g_j}{\partial x_i}(a)}, \quad 1 \leq i \leq n. \quad (2.28)$$

Con un ligero abuso de notación (muy frecuente en los libros clásicos) podemos utilizar la notación $y(x)$, $y_j(x)$ para la función g y sus componentes. En tal caso la función compuesta es $f(y(x))$, y la igualdad anterior (en un punto genérico x , en lugar de a) se puede escribir como sigue:

$$\boxed{\frac{\partial}{\partial x_i} f(y(x)) = \sum_{j=1}^m \frac{\partial f}{\partial y_j}(y(x)) \frac{\partial y_j}{\partial x_i}(x)}, \quad 1 \leq i \leq n. \quad (2.29)$$

En los textos clásicos esta igualdad se abrevia aún más suprimiendo los argumentos de las funciones, y se escribe en la forma más sencilla

$$\frac{\partial f}{\partial x_i} = \sum_{j=1}^m \frac{\partial f}{\partial y_j} \frac{\partial y_j}{\partial x_i}, \quad 1 \leq i \leq n.$$

El inconveniente de esta notación (que procuraremos evitar) es que la f que aparece en el miembro izquierdo es en realidad $f(y(x))$ considerada como función de $x \in \mathbb{R}^n$ a través de y ,

mientras que la f del miembro derecho es $f(y)$ considerada como función de la variable *independiente* $y \in \mathbb{R}^m$.

Por supuesto, lo anterior se generaliza fácilmente al caso $p > 1$. En este caso $F(x) = f(g(x))$ tiene p componentes $F_k(x) = f_k(g(x))$ ($1 \leq k \leq p$), y las ecs. (2.28)-(2.29) aplicadas a cada una de ellas proporcionan

$$\frac{\partial F_k}{\partial x_i}(a) = \sum_{j=1}^m \frac{\partial f_k}{\partial y_j}(g(a)) \frac{\partial g_j}{\partial x_i}(a), \quad 1 \leq i \leq n, \quad 1 \leq k \leq p \quad (2.30)$$

y

$$\frac{\partial}{\partial x_i} f_k(y(x)) = \sum_{j=1}^m \frac{\partial f_k}{\partial y_j}(y(x)) \frac{\partial y_j}{\partial x_i}(x), \quad 1 \leq i \leq n, \quad 1 \leq k \leq p. \quad (2.31)$$

Corolario 2.18. Sean $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ y $f: \mathbb{R}^m \rightarrow \mathbb{R}^p$ de clase C^1 en sendos entornos U de a y V de $g(a)$, y supongamos que $g(U) \subset V$. Entonces $f \circ g$ es de clase C^1 en U .

Demostración. En efecto, g y f son diferenciables respectivamente en U y V por ser de clase C^1 en dichos conjuntos. Por la regla de la cadena, $f \circ g$ es diferenciable en todo punto $a \in U$ y sus derivadas parciales están dadas por la ec. (2.31). De esta ecuación se sigue que las derivadas parciales de $f \circ g$ son continuas en U , y por tanto $f \circ g$ es de clase C^1 en U . \square

Ejemplo 2.19. Sean $f, g: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ las funciones definidas por

$$f(u, v) = (e^u \cos v, e^u \sin v), \quad g(x, y) = (x^2 - y^2, 2xy).$$

La función compuesta $F = f \circ g$ está dada por

$$F(x, y) = f(g(x, y)) = f(x^2 - y^2, 2xy) = (e^{x^2 - y^2} \cos(2xy), e^{x^2 - y^2} \sin(2xy)). \quad (2.32)$$

Más informalmente (y con un ligero abuso de notación), podemos escribir

$$F(x, y) = f(u, v),$$

donde se sobreentiende que en el miembro derecho sustituimos u y v por las expresiones

$$u = x^2 - y^2, \quad v = 2xy$$

(es decir, por las dos componentes de la función $g(x, y)$). Las funciones f y g son diferenciables en todo \mathbb{R}^2 , al ser obviamente de clase C^1 , y sus derivadas están dadas por

$$Df(u, v) = \begin{pmatrix} e^u \cos v & -e^u \sin v \\ e^u \sin v & e^u \cos v \end{pmatrix}, \quad Dg(x, y) = \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix}.$$

Aplicando la regla de la cadena se obtiene:

$$\begin{aligned} DF(x, y) &= Df(x^2 - y^2, 2xy) \cdot Dg(x, y) = 2e^{x^2 - y^2} \begin{pmatrix} \cos(2xy) & -\sin(2xy) \\ \sin(2xy) & \cos(2xy) \end{pmatrix} \begin{pmatrix} x & -y \\ y & x \end{pmatrix} \\ &= \boxed{2e^{x^2 - y^2} \begin{pmatrix} x \cos(2xy) - y \sin(2xy) & -y \cos(2xy) - x \sin(2xy) \\ x \sin(2xy) + y \cos(2xy) & -y \sin(2xy) + x \cos(2xy) \end{pmatrix}}. \end{aligned}$$

Este resultado se puede comprobar fácilmente derivando directamente la fórmula explícita para $F(x, y)$ obtenida más arriba (ec. (2.32)). Así, por ejemplo,

$$\frac{\partial F_1}{\partial y}(x, y) = \frac{\partial}{\partial y} \left(e^{x^2-y^2} \cos(2xy) \right) = -2ye^{x^2-y^2} \cos(2xy) - 2xe^{x^2-y^2} \sin(2xy),$$

que efectivamente coincide con el elemento de matriz (1, 2) de la derivada $DF(x, y)$ calculada más arriba. Nótese que $\frac{\partial F_1}{\partial y}(x, y)$ también se podría haber calculado teniendo en cuenta que $F_1(x, y) = f_1(u, v) = e^u \cos v$ y utilizando la regla de la cadena:

$$\begin{aligned} \frac{\partial F_1}{\partial y}(x, y) &= \frac{\partial}{\partial y} f_1(u, v) = \frac{\partial f_1}{\partial u} u_y + \frac{\partial f_1}{\partial v} v_y = e^u \cos v (-2y) - e^u \sin v \cdot 2x \\ &= -2e^u (y \cos v + x \sin v), \end{aligned}$$

que coincide con el resultado anterior teniendo en cuenta la definición de u y v . \square

Ejemplo 2.20. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función diferenciable, y sea $\gamma : \mathbb{R} \rightarrow \mathbb{R}^n$ una curva diferenciable. Evaluando f a lo largo de la curva γ se obtiene la función $F = f \circ \gamma : \mathbb{R} \rightarrow \mathbb{R}$, que es diferenciable en \mathbb{R} por la regla de la cadena. Aplicando dicha regla se obtiene

$$F'(t) \equiv \frac{d}{dt} f(\gamma_1(t), \dots, \gamma_n(t)) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\gamma(t)) \gamma'_i(t),$$

es decir

$$F'(t) = \nabla f(\gamma(t)) \cdot \gamma'(t) \equiv Df(\gamma(t)) \cdot \gamma'(t). \quad (2.33)$$

Nótese que la fórmula anterior es una generalización de (2.20) (tómese $\gamma(t) = a + tv$ y evalúese F' en $t = 0$).

La regla de la cadena proporciona una demostración muy sencilla del siguiente resultado importante:

Proposición 2.21. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ una función diferenciable en un abierto conexo A tal que $Df(x) = 0$ para todo $x \in A$. Entonces f es constante en A .

Demostración. Sea x_0 un punto arbitrario pero fijo de A , y consideremos cualquier otro punto $x \in A$. Al ser A arco-conexo (por ser abierto y conexo), se puede probar que existe una curva continua $\gamma : [0, 1] \rightarrow A$ diferenciable en $(0, 1)$ tal que $\gamma(0) = x_0$ y $\gamma(1) = x$. Por las hipótesis hechas sobre la función f , la función $f \circ \gamma$ es continua en $[0, 1]$ (composición de funciones continuas) y diferenciable en $(0, 1)$ (regla de la cadena). Además, la derivada de $f \circ \gamma$ se anula idénticamente en el intervalo $(0, 1)$, ya que

$$(f \circ \gamma)'(t) = Df(\gamma(t)) \cdot \gamma'(t) = 0, \quad \forall t \in (0, 1),$$

al ser $\gamma(t) \in A$ para todo $t \in (0, 1)$. Por el teorema del valor medio para funciones de \mathbb{R} en \mathbb{R} , cada componente $f_k \circ \gamma$ de $f \circ \gamma$ es constante en $[0, 1]$. En particular,

$$f_k(x_0) = \gamma(0) = \gamma(1) = f_k(x), \quad 1 \leq k \leq m,$$

o lo que es lo mismo $f(x_0) = f(x)$. \square

Ejemplo 2.22. Sea $g(x) = \varphi(f(x))$, con $f : \mathbb{R}^n \rightarrow \mathbb{R}$ y $\varphi : \mathbb{R} \rightarrow \mathbb{R}$. Si f es diferenciable en x y φ es diferenciable en $f(x)$ entonces $g = \varphi \circ f$ es diferenciable en x (al ser composición de funciones diferenciables), y la regla de la cadena proporciona

$$Dg(x) = \varphi'(f(x))Df(x),$$

o equivalentemente

$$\boxed{\nabla g(x) = \varphi'(f(x)) \nabla f(x)}.$$

Aplicando lo anterior a $\varphi(t) = 1/t$, diferenciable para todo $t \neq 0$ con derivada $\varphi'(t) = -1/t^2$, se obtiene el siguiente resultado: si $f : \mathbb{R}^n \rightarrow \mathbb{R}$ es diferenciable en $a \in \mathbb{R}^n$ y $f(a) \neq 0$, entonces $1/f$ es diferenciable en a , con derivada

$$D(1/f)(a) = -\frac{1}{f(a)^2} Df(a),$$

o, en términos del gradiente,

$$\boxed{\nabla(1/f)(a) = -\frac{1}{f(a)^2} \nabla f(a)}.$$

Ejemplo 2.23 (regla de Leibniz). Si $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ son diferenciables en $a \in \mathbb{R}^n$, el producto fg es diferenciable en dicho punto, y se tiene

$$D(fg)(a) = f(a)Dg(a) + g(a)Df(a) \quad (\text{regla de Leibniz}). \quad (2.34)$$

Este importante resultado es una sencilla consecuencia de la regla de la cadena. En efecto, $fg = p \circ (f, g)$, siendo $p : \mathbb{R}^2 \rightarrow \mathbb{R}$ el producto ($p(u, v) = uv$). Como (f, g) es diferenciable en a por serlo sus componentes f y g , y su matriz jacobiana es la matriz $2 \times n$ cuyas filas son las derivadas (gradientes) de f y g , de la regla de la cadena y la ec. (2.15) se sigue que

$$D(fg)(a) = Dp(f(a), g(a)) \cdot D(f, g)(a) = (g(a) \ f(a)) \begin{pmatrix} Df(a) \\ Dg(a) \end{pmatrix} = f(a)Dg(a) + g(a)Df(a).$$

Nótese que (2.34) se puede expresar equivalentemente en términos del gradiente como

$$\boxed{\nabla(fg)(a) = f(a)\nabla g(a) + g(a)\nabla f(a)}. \quad (2.35)$$

Por ejemplo, si $g = P/Q$ es una función racional, de la regla de Leibniz y el ejemplo anterior se sigue de nuevo que g es diferenciable en el conjunto de puntos en que no se anula el denominador Q (que, como ya hemos visto, es un abierto de \mathbb{R}^n), con gradiente

$$\nabla\left(\frac{P}{Q}\right) = \frac{\nabla P}{Q} - \frac{P}{Q^2} \nabla Q = \frac{Q\nabla P - P\nabla Q}{Q^2}.$$

Ejemplo 2.24 (teorema de Euler). Una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ es **homogénea de grado p** si existe $p \in \mathbb{R}$ tal que

$$f(\lambda x) = \lambda^p f(x), \quad \forall \lambda > 0, \quad \forall x \in \mathbb{R}^n.$$

(Si $\mathcal{D}(f) \neq \mathbb{R}^n$, basta que se cumpla lo anterior para todo $x \in \mathcal{D}(f)$ y $\lambda > 0$ tales que $\lambda x \in \mathcal{D}(f)$.) Por ejemplo, las funciones constantes son homogéneas de grado 0, las funciones lineales son homogéneas de grado 1, y en general un *polinomio homogéneo* de la forma

$$P(x) = \sum_{\substack{i_1, \dots, i_n \geq 0 \\ i_1 + \dots + i_n = k}} a_{i_1 \dots i_n} x_1^{i_1} \cdots x_n^{i_n}$$

(con $a_{i_1 \dots i_n} \in \mathbb{R}$) es homogéneo de grado k . También es obvio que el cociente P_1/P_2 de dos polinomios homogéneos de grados respectivos k_1 y k_2 es una función homogénea de grado $k_1 - k_2$. Un ejemplo de función homogénea de grado p , con $p \in \mathbb{R}$ arbitrario, es $f(x) = \|x\|^p$. En efecto, si $\lambda > 0$ se tiene

$$f(\lambda x) = \|\lambda x\|^p = (\lambda \|x\|)^p = \lambda^p f(x), \quad \forall x \in \mathbb{R}^n.$$

Supongamos que f es una función homogénea de grado p diferenciable en un entorno de un punto fijo $x \neq 0$. Entonces f es diferenciable en todos los puntos de la forma λx con $|\lambda - 1|$ suficientemente pequeño⁶, y podemos aplicar la regla de la cadena (2.33) a la función $g(\lambda) = f(\lambda x)$:

$$g'(\lambda) = \nabla f(\lambda x) \cdot \frac{d}{d\lambda}(\lambda x) = \nabla f(\lambda x) \cdot x.$$

Por otra parte, como $g(\lambda) = \lambda^p f(x)$ por ser f homogénea de grado p se tiene

$$g'(\lambda) = \frac{d}{d\lambda}(\lambda^p f(x)) = p\lambda^{p-1} f(x).$$

Igualando ambas expresiones para $g'(\lambda)$ se deduce que

$$\nabla f(\lambda x) \cdot x = p\lambda^{p-1} f(x),$$

y haciendo $\lambda = 1$ obtenemos el **teorema de Euler** para funciones homogéneas:

$$\nabla f(x) \cdot x \equiv x_1 \frac{\partial f}{\partial x_1}(x) + \dots + x_n \frac{\partial f}{\partial x_n}(x) = pf(x).$$

2.6 Derivadas parciales de orden superior

Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, y supongamos que existe la derivada parcial $D_j f$ en un entorno de $a \in \mathbb{R}^n$. Si existe la derivada parcial $D_i(D_j f)$ en el punto a , es decir

$$\left. \frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right) \right|_{x=a},$$

se denomina **derivada parcial de segundo orden** de f en el punto a respecto de las variables x_i y x_j , y se denota por

$$D_{ij} f(a) \equiv \frac{\partial^2 f}{\partial x_i \partial x_j}(a).$$

En otras palabras,

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \lim_{h \rightarrow 0} \frac{1}{h} \left[\frac{\partial f}{\partial x_j}(a + he_i) - \frac{\partial f}{\partial x_j}(a) \right] = \left. \frac{d}{dt} \frac{\partial f}{\partial x_j}(a + te_i) \right|_{t=0}.$$

Comentarios.

- Por las propiedades elementales de las derivadas parciales, $D_{ij} f$ existe si y sólo si existen las derivadas parciales segundas $D_{ij} f_k$ de todas las componentes f_k de f , y se tiene

$$D_{ij} f = (D_{ij} f_1, \dots, D_{ij} f_m).$$

- En principio, $D_{ij} f \neq D_{ji} f$ si $i \neq j$, y puede incluso existir una de estas derivadas parciales en algún punto sin que exista la otra. Sin embargo, veremos más adelante una condición suficiente que garantiza la igualdad de estas derivadas parciales cruzadas bajo hipótesis muy generales.

Notación. Si $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ es una función de las variables $(x, y) \in \mathbb{R}^2$, denotaremos habitualmente sus derivadas parciales mediante subíndices:

$$f_{xx} \equiv \frac{\partial^2 f}{\partial x^2} \equiv D_{11} f, \quad f_{xy} \equiv \frac{\partial^2 f}{\partial y \partial x} \equiv D_{21} f, \quad f_{yx} \equiv \frac{\partial^2 f}{\partial x \partial y} \equiv D_{12} f, \quad f_{yy} \equiv \frac{\partial^2 f}{\partial y^2} \equiv D_{22} f.$$

Análogamente para funciones $f : \mathbb{R}^3 \rightarrow \mathbb{R}$.

⁶En efecto, si f es diferenciable en $B_r(x)$ entonces f es diferenciable en todos los puntos de la forma λx con $|\lambda - 1| < r/\|x\|$ (ejercicio).

Definición 2.25. Si $U \subset \mathbb{R}^n$ es abierto y $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, diremos que f es **de clase C^2** en U , y escribiremos $f \in C^2(U)$, si existen y son continuas en U las n^2 derivadas parciales de segundo orden $D_{ij}f$, $1 \leq i, j \leq n$.

Evidentemente, $f = (f_1, \dots, f_m) \in C^2(U)$ si y sólo si $f_i \in C^2(U)$, para todo $i = 1, 2, \dots, m$.

Si $f \in C^2(U)$, las n funciones $D_i f$ son por definición de clase C^1 en U , y en particular son continuas (pues $C^1(U) \subset C^0(U)$). Por tanto, si f es de clase C^2 en U automáticamente f es de clase C^1 en U (y, en consecuencia, diferenciable), es decir

$$\boxed{C^2(U) \subset C^1(U) \subset C^0(U)}.$$

En general, una **derivada parcial de orden p** de f es una derivada parcial del tipo

$$\boxed{D_{i_1 i_2 \dots i_p} f = D_{i_1} (D_{i_2} (\dots (D_{i_p} f) \dots)) \equiv \frac{\partial^p f}{\partial x_{i_1} \dots \partial x_{i_p}}}, \quad 1 \leq i_1, \dots, i_p \leq n.$$

Como en el caso $p = 2$, es evidente que $D_{i_1 \dots i_p} f$ existe si y sólo si existen las derivadas parciales $D_{i_1 \dots i_p} f_k$, $1 \leq k \leq m$, de las componentes de f , y se tiene

$$\boxed{D_{i_1 \dots i_p} f = (D_{i_1 \dots i_p} f_1, \dots, D_{i_1 \dots i_p} f_m)}.$$

Definición 2.26. Diremos que $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es **de clase C^p** en un abierto U de \mathbb{R}^n , y escribiremos $f \in C^p(U)$, si existen y son continuas en U las n^p derivadas parciales de orden p de f . La función f es **de clase C^∞** en U (lo que denotaremos por $f \in C^\infty(U)$) si es de clase C^p en U para todo $p \in \mathbb{N}$.

Comentarios.

- Al igual que antes, es evidente que $f = (f_1, \dots, f_m) \in C^p(U)$ (resp. $f \in C^\infty(U)$) si y solo si $f_i \in C^p(U)$ (resp. $f_i \in C^\infty(U)$) para todo $i = 1, \dots, m$.
- Por definición,

$$C^\infty(U) = \bigcap_{p=0}^{\infty} C^p(U).$$

En otras palabra, f es de clase $C^\infty(U)$ si existen y son continuas las derivadas parciales de *cualquier* orden de f en U . Por ejemplo, *las funciones racionales son de clase C^∞ en el abierto en que no se anula su denominador*, ya que cualquier derivada parcial de una función racional P/Q es a su vez una función racional cuyo denominador es igual a una potencia de Q . En particular, *los polinomios son funciones C^∞ en todo \mathbb{R}^n .*

- Es inmediato comprobar que $C^{p+1}(U) \subset C^p(U)$ para todo $p \in \mathbb{N}$. En efecto, si f es de clase $p + 1$ en U , como las derivadas parciales de orden $p + 1$ de f son las derivadas parciales de primer orden de las derivadas parciales de orden p , todas las derivadas parciales de orden p de f son de clase C^1 , y por tanto continuas, en U . En consecuencia $f \in C^p(U)$, como queríamos demostrar. En otras palabras,

$$\boxed{C^\infty(U) \subset C^{p+1}(U) \subset C^p(U)}, \quad \forall p \in \mathbb{N}.$$

2.7 Lema de Schwarz

Probaremos a continuación que si f es de clase C^2 en un entorno de $a \in \mathbb{R}^n$ entonces las derivadas parciales cruzadas de segundo orden de f en a son iguales: $D_{ij}f(a) = D_{ji}f(a)$, $\forall i \neq j$. Esto será consecuencia del resultado ligeramente más general siguiente:

Lema de Schwarz. Sea $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ una función definida en un entorno U de $(x_0, y_0) \in \mathbb{R}^2$. Si las derivadas parciales cruzadas $D_{12}f$ y $D_{21}f$ existen en U y son continuas en (x_0, y_0) , entonces

$$D_{12}f(x_0, y_0) = D_{21}f(x_0, y_0).$$

Demostración. Antes de empezar, obsérvese que la existencia de $D_{12}f = D_1(D_2f)$ y $D_{21}f = D_2(D_1f)$ en U garantiza la existencia de las derivadas parciales de primer orden de f en U (¡aunque no su continuidad!). La demostración se basa en probar que el cociente incremental de segundo orden

$$\Delta(u) = \frac{1}{u^2} [f(x_0 + u, y_0 + u) - f(x_0 + u, y_0) - f(x_0, y_0 + u) + f(x_0, y_0)]$$

tiende tanto a $D_{12}f(x_0, y_0)$ como a $D_{21}f(x_0, y_0)$ cuando $u \rightarrow 0+$. Si $B_\delta(x_0, y_0) \subset U$, entonces $\Delta(u)$ está definido si $2u^2 < \delta^2$; como en lo que sigue sólo estamos interesados en valores positivos de u , tomaremos u en el intervalo $(0, \delta/\sqrt{2})$. Fijado un u en dicho intervalo, sea

$$g(t) = f(t, y_0 + u) - f(t, y_0), \quad x_0 - \varepsilon < t < x_0 + u + \varepsilon,$$

con $\varepsilon > 0$ suficientemente pequeño para que $(t, y_0), (t, y_0 + u) \in U$ para $t \in (x_0 - \varepsilon, x_0 + u + \varepsilon)$ (basta tomar $\varepsilon < \sqrt{\delta^2 - u^2} - u$). Entonces g es derivable en su dominio, pues

$$g'(t) = D_1f(t, y_0 + u) - D_1f(t, y_0)$$

existe si $t \in (x_0 - \varepsilon, x_0 + u + \varepsilon)$, al existir D_1f en U por hipótesis. En particular, g es continua en el intervalo $[x_0, x_0 + u]$, y derivable en el interior de dicho intervalo. Aplicando el teorema del valor medio a g en $[x_0, x_0 + u]$ se obtiene

$$\Delta(u) = \frac{g(x_0 + u) - g(x_0)}{u^2} = \frac{g'(\xi_u)}{u} = \frac{D_1f(\xi_u, y_0 + u) - D_1f(\xi_u, y_0)}{u},$$

para algún $\xi_u \in (x_0, x_0 + u)$. A su vez, la función $h : \mathbb{R} \rightarrow \mathbb{R}$ definida mediante $h(t) = D_1f(\xi_u, t)$ tiene derivada

$$h'(t) = D_{21}f(\xi_u, t), \quad t \in (y_0 - \varepsilon, y_0 + u + \varepsilon),$$

con $\varepsilon > 0$ suficientemente pequeño (basta tomar de nuevo $\varepsilon^2 < \sqrt{\delta^2 - u^2} - u$), debido a que $D_{21}f$ por hipótesis existe en U . Aplicando de nuevo el teorema del valor medio a h en el intervalo $[y_0, y_0 + u]$ queda

$$\Delta(u) = \frac{h(y_0 + u) - h(y_0)}{u} = h'(\eta_u) = D_{21}f(\xi_u, \eta_u)$$

para algún $\eta_u \in (y_0, y_0 + u)$. Por otra parte, también podemos escribir

$$\Delta(u) = \frac{G(y_0 + u) - G(y_0)}{u^2},$$

con $G(t) = f(x_0 + u, t) - f(x_0, t)$. Aplicando el teorema del valor medio a G en el intervalo $[y_0, y_0 + u]$ obtenemos

$$\Delta(u) = \frac{G'(\tilde{\eta}_u)}{u} = \frac{D_2f(x_0 + u, \tilde{\eta}_u) - D_2f(x_0, \tilde{\eta}_u)}{u},$$

para algún $\tilde{\eta}_u \in (y_0, y_0 + u)$. Finalmente, aplicando otra vez el teorema del valor medio a $D_2f(t, \tilde{\eta}_u)$ en el intervalo $[x_0, x_0 + u]$, lo cual puede hacerse por la existencia de $D_{12}f$ en U , queda

$$\Delta(u) = D_{12}f(\tilde{\xi}_u, \tilde{\eta}_u)$$

para algún $\tilde{\xi}_u \in (x_0, x_0 + u)$. En definitiva, hemos probado que para todo $u \in (0, \delta/\sqrt{2})$ existen dos puntos $(\xi_u, \eta_u), (\tilde{\xi}_u, \tilde{\eta}_u) \in (x_0, x_0 + u) \times (y_0, y_0 + u) \subset U$ tales que

$$\Delta(u) = D_{12}f(\xi_u, \eta_u) = D_{12}f(\tilde{\xi}_u, \tilde{\eta}_u).$$

Si ahora hacemos u tender a $0+$ se tiene $(\xi_u, \eta_u) \rightarrow (x_0, y_0)$, $(\tilde{\xi}_u, \tilde{\eta}_u) \rightarrow (x_0, y_0)$, y por la continuidad de $D_{12}f$ y $D_{21}f$ en $(x_0, y_0) \in U$ queda

$$\lim_{u \rightarrow 0+} \Delta(u) = D_{21}f(x_0, y_0) = D_{12}f(x_0, y_0). \quad \square$$

Corolario 2.27. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ una función definida en un entorno U de $a \in \mathbb{R}^n$. Si las derivadas parciales $D_{ij}f$ y $D_{ji}f$ (con $i \neq j$) existen en U y son continuas en dicho punto, entonces $D_{ij}f(a) = D_{ji}f(a)$.

Demostración. Sin pérdida de generalidad, podemos suponer que (por ejemplo) $i < j$. Evidentemente, basta probar que $D_{ij}f_k(a) = D_{ji}f_k(a)$ para todo $k = 1, \dots, m$. A tal efecto, definimos la función $g_k : \mathbb{R}^2 \rightarrow \mathbb{R}$ por

$$g_k(x, y) = f_k(a_1, \dots, a_{i-1}, x, a_{i+1}, \dots, a_{j-1}, y, a_{j+1}, \dots, a_n).$$

Al ser

$$\begin{cases} D_{12}g_k(x, y) = D_{ij}f_k(a_1, \dots, a_{i-1}, x, a_{i+1}, \dots, a_{j-1}, y, a_{j+1}, \dots, a_n), \\ D_{21}g_k(x, y) = D_{ji}f_k(a_1, \dots, a_{i-1}, x, a_{i+1}, \dots, a_{j-1}, y, a_{j+1}, \dots, a_n) \end{cases}$$

por las hipótesis hechas sobre f se puede aplicar el lema de Schwarz a la función $g_k(x, y)$ en el punto (a_i, a_j) , y por tanto

$$D_{ij}f_k(a) = D_{12}g_k(a_i, a_j) = D_{21}g_k(a_i, a_j) = D_{ji}f_k(a).$$

□

En particular, del corolario anterior se sigue que si f es de clase C^2 en un abierto $U \subset \mathbb{R}^n$ entonces todas las derivadas parciales cruzadas de f son iguales en U :

$$f \in C^2(U) \implies D_{ij}f(x) = D_{ji}f(x), \quad \forall x \in U, \quad 1 \leq i, j \leq n.$$

Más generalmente, si f es de clase $p > 1$ se verifica el resultado siguiente:

Proposición 2.28. Si $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es de clase C^p con $p > 1$ en un abierto $U \subset \mathbb{R}^n$, entonces todas las derivadas parciales cruzadas de orden $q \leq p$ de f son iguales en U , es decir

$$D_{i_1 \dots i_q} f(x) = D_{i_{\sigma(1)} \dots i_{\sigma(q)}} f(x), \quad \forall x \in U,$$

para toda permutación σ de $\{1, 2, \dots, q\}$.

Mencionaremos, por último, sin demostración dos condiciones *suficientes* algo más débiles que la continuidad en a de las derivadas parciales cruzadas $D_{ij}f$ y $D_{ji}f$ de una función $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ que garantizan la igualdad $D_{ij}f(a) = D_{ji}f(a)$:

- i) Si $D_i f$ y $D_j f$ son diferenciables en a , entonces $D_{ij} f(a) = D_{ji} f(a)$.
- ii) Si $D_i f$, $D_j f$ y $D_{ij} f$ son continuas en un entorno de a , entonces existe $D_{ji} f(a)$ y es igual a $D_{ij} f(a)$.

Ejercicio 25. ¿Cuántas derivadas parciales independientes de orden p posee una función escalar de clase C^p ?

Solución. Si $f : \mathbb{R}^n \rightarrow \mathbb{R}$ es de clase C^p , en virtud de la Proposición 2.28 las derivadas parciales independientes de orden p de f están dadas por

$$\frac{\partial^p f}{\partial x_{i_1} \cdots \partial x_{i_p}}, \quad 1 \leq i_1 \leq \dots \leq i_p \leq n.$$

Por tanto el número de derivadas parciales independientes de orden p de f es el cardinal del conjunto

$$\{(i_1, \dots, i_p) \mid 1 \leq i_1 \leq \dots \leq i_p \leq n\}.$$

Para evaluar este cardinal, efectuamos el cambio de índices

$$j_k = i_k + k - 1, \quad 1 \leq k \leq p,$$

de forma que

$$\{(i_1, \dots, i_p) \mid 1 \leq i_1 \leq \dots \leq i_p \leq n\} = \{(j_1, \dots, j_p) \mid 1 < j_1 < \dots < j_p \leq n + p - 1\}.$$

El cardinal de este último conjunto, y por tanto el número de derivadas parciales independientes de orden p de la función f , es el número de formas de escoger p elementos distintos j_1, \dots, j_p del conjunto de índices $\{1, \dots, n + p - 1\}$, es decir $\binom{n+p-1}{p}$. \square

Ejemplo 2.29. Consideremos la función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $f(0, 0) = 0$ y

$$f(x, y) = \frac{xy(x^2 - y^2)}{x^2 + y^2}, \quad (x, y) \neq (0, 0).$$

Esta función es claramente de clase C^∞ en $\mathbb{R}^2 - \{(0, 0)\}$, ya que es una función racional cuyo denominador solo se anula en el origen. ¿Qué ocurre en $(0, 0)$? Para responder a esta pregunta, nótese en primer lugar que

$$|f(x, y)| \leq |xy|,$$

y por tanto f es continua en el origen. Las derivadas parciales de f en el origen existen y son nulas:

$$f_x(0, 0) = \lim_{h \rightarrow 0} \frac{f(h, 0)}{h} = 0, \quad f_y(0, 0) = \lim_{h \rightarrow 0} \frac{f(0, h)}{h} = 0.$$

Por otra parte, si $(x, y) \neq (0, 0)$ se tiene

$$f_x(x, y) = \frac{(x^2 + y^2)(3x^2y - y^3) - 2x^2y(x^2 - y^2)}{(x^2 + y^2)^2} = \frac{y(x^4 + 4x^2y^2 - y^4)}{(x^2 + y^2)^2}.$$

Al ser $f(x, y) = -f(y, x)$, la regla de la cadena proporciona

$$f_y(x, y) = -D_1 f(y, x) = -f_x(y, x) = \frac{x(x^4 - 4x^2y^2 - y^4)}{(x^2 + y^2)^2}.$$

Como

$$x^4 \leq (x^2 + y^2)^2, \quad y^4 \leq (x^2 + y^2)^2, \quad x^2y^2 \leq (x^2 + y^2)^2,$$

de las fórmulas anteriores para f_x y f_y se sigue que

$$|f_x(x, y)| \leq 6|y|, \quad |f_y(x, y)| \leq 6|x|,$$

de modo que

$$\lim_{(x,y) \rightarrow (0,0)} f_x(x,y) = \lim_{(x,y) \rightarrow (0,0)} f_y(x,y) = 0.$$

Por tanto f tiene derivadas parciales continuas en $(0,0)$, lo que implica (ya que f es de clase C^∞ en $\mathbb{R}^2 - \{(0,0)\}$) que $f \in C^1(\mathbb{R}^2)$. En particular, f es diferenciable en \mathbb{R}^2 .

Al ser f de clase C^∞ en $\mathbb{R}^2 - \{(0,0)\}$, las derivadas parciales cruzadas de f son iguales para $(x,y) \neq (0,0)$. ¿Existen y son iguales las derivadas parciales f_{xy} y f_{yx} en el origen? Para dilucidar esta cuestión, nótese que para todo $(x,y) \in \mathbb{R}^2$ se tiene

$$f_x(0,y) = -y, \quad f_y(x,0) = -f_x(0,x) = x,$$

y por tanto

$$f_{xy}(0,0) = \left. \frac{d}{dy} f_x(0,y) \right|_{y=0} = -1, \quad f_{yx}(0,0) = \left. \frac{d}{dx} f_y(x,0) \right|_{x=0} = 1.$$

En particular, en este ejemplo

$$\boxed{f_{xy}(0,0) \neq f_{yx}(0,0)}.$$

¿Por qué no se verifica el lema de Schwarz en este caso? Nótese que, al existir f_{xy} y f_{yx} en todo \mathbb{R}^2 , dicho lema implica que al menos una de estas dos derivadas parciales ha de ser *discontinua* en el origen. En efecto, el cálculo directo demuestra que

$$f_{xy}(x,y) = f_{yx}(x,y) = \frac{(x^2 - y^2)(x^4 + 10x^2y^2 + y^4)}{(x^2 + y^2)^3}, \quad (x,y) \neq (0,0).$$

Si hacemos (x,y) tender al origen a lo largo de la recta $y(t) = tv$ (con $v \neq 0$), al ser f_{xy} homogénea de grado 0 (cociente de dos polinomios homogéneos de grado 6) se tiene

$$\lim_{t \rightarrow 0} f_{xy}(tv) = f_{xy}(v),$$

y análogamente para f_{yx} (ya que $f_{yx} = f_{xy}$ en $\mathbb{R}^2 - \{(0,0)\}$!). Como el límite depende de la recta escogida (es función del vector v), *tanto f_{xy} como f_{yx} son discontinuas en el origen* (ya que no tienen límite en dicho punto).

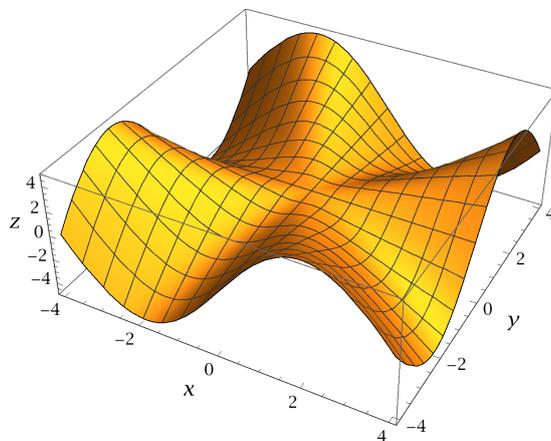


Figura 2.1: Gráfica de la función del Ejemplo 2.29.

2.8 Interpretación geométrica del gradiente y la derivada

Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$, sea $a \in \mathbb{R}^n$, y sea $v \in \mathbb{R}^n$ un vector **unitario** (es decir, tal que $\|v\| = 1$). Como ya sabemos (cf. la ec. (2.19)), la derivada direccional de f en a según el vector v (si existe) está dada por

$$D_v f(a) = \left. \frac{d}{dt} f(a + tv) \right|_{t=0},$$

y por tanto representa el *ritmo de variación* de f en a a lo largo de la recta que pasa por dicho punto en la dirección del vector v . ¿Cómo varía $D_v f(a)$ en función de la dirección v ? Si f no es diferenciable en a , no es fácil en general responder a esta pregunta. Sin embargo, cuando f es *diferenciable* en el punto a hemos visto que

$$D_v f(a) = Df(a) \cdot v \equiv \nabla f(a) \cdot v$$

(cf. las ecs. (2.20) y (2.25)). Supongamos que $\nabla f(a) \neq 0$ (ya que si $\nabla f(a) = 0$ el ritmo de variación de f en a es nulo en cualquier dirección⁷). Como $\|v\| = 1$, si llamamos $\theta(v) \in [0, \pi]$ al ángulo formado por los vectores $\nabla f(a)$ (que está fijo) y v , la fórmula anterior se escribe

$$\boxed{D_v f(a) = \|\nabla f(a)\| \cos \theta(v)}. \quad (2.36)$$

De esta ecuación se sigue inmediatamente que la derivada direccional $D_v f(a)$ es máxima si $\theta(v) = 0$, es decir si v tiene la misma dirección y sentido que $\nabla f(a)$ (i.e., si $v = \nabla f(a) / \|\nabla f(a)\|$), y el valor máximo de $D_v f(a)$ es igual a $\|\nabla f(a)\|$. En otras palabras:

Si f es diferenciable en a y $\nabla f(a) \neq 0$, la dirección del vector $\nabla f(a)$ es la de *máximo crecimiento* de f en a , y su norma es igual al *valor máximo del ritmo de crecimiento* de f en dicho punto.

Evidentemente, el ritmo de crecimiento de f en a es *mínimo* en la dirección *opuesta* a $\nabla f(a)$, es decir la del vector $v = -\nabla f(a) / \|\nabla f(a)\|$, y el valor mínimo de dicho ritmo de crecimiento es igual a $-\|\nabla f(a)\|$.

Comentario. El resultado anterior se generaliza fácilmente a cualquier *curva* diferenciable que pase por el punto a . En efecto, sea $\gamma : \mathbb{R} \rightarrow \mathbb{R}^n$ una curva diferenciable en t_0 tal que $\gamma(t_0) = a$. El valor de la función f a lo largo de la curva γ está dado por la función de una variable $f(\gamma(t))$. Por ejemplo, si f es una magnitud física escalar (temperatura, potencial eléctrico, gravitatorio, etc.) y $\gamma(t)$ es la *trayectoria* de un observador, entonces $f(\gamma(t))$ representa el valor de dicha magnitud medido por el observador en el instante t . Al ser γ diferenciable en t_0 , de la regla de la cadena se deduce inmediatamente que

$$\left. \frac{d}{dt} f(\gamma(t)) \right|_{t=t_0} = \nabla f(a) \cdot \gamma'(t_0) = \|\nabla f(a)\| \|\gamma'(t_0)\| \cos \theta$$

(cf. la ec. (2.33)), donde θ es el ángulo formado por los vectores $\gamma'(t_0)$ (tangente a la curva en a) y $\nabla f(a)$. Para $\|\gamma'(t_0)\|$ fijo⁸, el ritmo de crecimiento de f a lo largo de γ en a es máximo si $\theta = 0$, es decir si el vector tangente a la curva tiene la misma dirección y sentido que $\nabla f(a)$, y el valor máximo de dicho ritmo de crecimiento es igual a $\|\nabla f(a)\| \|\gamma'(t_0)\|$. \square

Sea, de nuevo; $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función escalar, y consideremos uno de sus conjuntos de nivel $L_c(f)$. Veremos más adelante (cuando estudiemos el *teorema de la función implícita*) que si f es de clase C^1 en un abierto que contiene a $L_c(f)$ y

$$\nabla f(x) \neq 0, \quad \forall x \in L_c(f),$$

⁷Veremos en el capítulo próximo que en tal caso a es un punto crítico de f , en el cual la función podría tener un extremo (máximo o mínimo).

⁸Si $\gamma(t)$ es la trayectoria de una partícula, $\|\gamma'(t_0)\|$ es el módulo del vector velocidad en el instante t_0 .

entonces $L_c(f)$ es una *hipersuperficie* de \mathbb{R}^n de dimensión $n - 1$ (superficie si $n = 3$, curva si $n = 2$). Esto significa esencialmente que $L_c(f)$ se puede parametrizar en un entorno de cada punto como la gráfica de una función suficientemente regular de $n - 1$ variables. Dado un punto $x_0 \in L_c(f)$, es de esperar entonces que $L_c(f)$ se pueda aproximar en las proximidades de x_0 por un cierto subespacio afín de dimensión $n - 1$ (es decir, por un *hiperplano* afín) que pase por dicho punto. ¿Cómo se determina dicho hiperplano?

Para responder a esta pregunta, nótese que si f es diferenciable en $x_0 \in L_c(f)$ por definición de derivada se tiene

$$f(x) - c = f(x) - f(x_0) = Df(x_0) \cdot (x - x_0) + o(\|x - x_0\|) = \nabla f(x_0) \cdot (x - x_0) + o(\|x - x_0\|),$$

donde hemos tenido en cuenta que

$$x_0 \in L_c(f) \iff f(x_0) = c.$$

Por tanto la ecuación de $L_c(f)$ (es decir, $f(x) - c = 0$) en las proximidades de x_0 se puede aproximar por

$$\nabla f(x_0) \cdot (x - x_0) = 0 \quad (2.37)$$

cometiendo un error que tiende a 0 más rápido que $\|x - x_0\|$ cuando x tiende a x_0 . Más aún (por la unicidad de la derivada), si

$$f(x) - c = v \cdot (x - x_0) + o(\|x - x_0\|)$$

entonces $v = \nabla f(x_0)$. Por tanto la mejor aproximación lineal a la ecuación de $L_c(f)$ en las proximidades de x_0 es la ec. (2.37). Si $\nabla f(x_0) \neq 0$, esta es la ecuación de un hiperplano afín que pasa por el punto x_0 y es *perpendicular al vector* $\nabla f(x_0)$, al que llamaremos **hiperplano tangente** a $L_c(f)$ en el punto x_0 y denotaremos por $H_{x_0}(L_c(f))$. En otras palabras,

El hiperplano tangente a $L_c(f)$ en el punto $x_0 \in L_c(f)$ es el hiperplano que más se aproxima al conjunto de nivel $L_c(f)$ en las proximidades de x_0 .

Supongamos, de nuevo, que f es diferenciable en x_0 y $\nabla f(x_0) \neq 0$. Si $\gamma: \mathbb{R} \rightarrow \mathbb{R}^n$ es cualquier curva diferenciable tal que $\gamma(t_0) = x_0$ cuya imagen esté contenida en el conjunto de nivel $L_c(f)$, es decir

$$f(\gamma(t)) = c, \quad \forall t \in \mathcal{D}(\gamma),$$

derivando la igualdad anterior respecto de t en $t = t_0$ se obtiene

$$\nabla f(x_0) \cdot \gamma'(t_0) = 0.$$

Por tanto el vector tangente en x_0 a *cualquier* curva diferenciable contenida en $L_c(f)$ que pase por dicho punto pertenece al subespacio vectorial (hiperplano) $T_{x_0}(L_c(f))$ de ecuación

$$\nabla f(x_0) \cdot x = 0.$$

Puede demostrarse que el recíproco también es cierto, es decir que *cualquier* vector de $T_{x_0}(L_c(f))$ es el vector tangente en x_0 a alguna curva diferenciable contenida en $L_c(f)$ que pasa por x_0 . Por tanto $T_{x_0}(L_c(f))$, que se denomina **espacio tangente** a $L_c(f)$ en x_0 , es el subespacio vectorial formado por los vectores tangentes en x_0 a las curvas diferenciables contenidas en $L_c(f)$ que pasan por dicho punto. Nótese que $T_{x_0}(L_c(f))$ es *paralelo* a $H_{x_0}(L_c(f))$, ya que ambos hiperplanos son perpendiculares a $\nabla f(x_0)$, y por tanto

$$H_{x_0}(L_c(f)) = x_0 + T_{x_0}(L_c(f)).$$

Diremos que un vector es **normal** a $L_c(f)$ en el punto x_0 si es perpendicular a los vectores tangentes en x_0 a *todas* las curvas diferenciables contenidas en $L_c(f)$ que pasan por dicho punto. De lo anterior se deduce que un vector es normal a $L_c(f)$ en x_0 si y solo si es perpendicular al espacio tangente $T_{x_0}(L_c(f))$, es decir si y solo si es proporcional a $\nabla f(x_0)$. Llamaremos **recta normal** a $L_c(f)$ en x_0 a la recta que pasa por x_0 y es normal a $L_c(f)$ en dicho punto. Por lo que acabamos de ver, esta recta ha de ser paralela al vector $\nabla f(x_0)$, y por tanto su ecuación paramétrica es

$$x = x_0 + t\nabla f(x_0), \quad t \in \mathbb{R}.$$

Ejemplo 2.30. Hallemos la ecuación del plano tangente y la recta normal al elipsoide

$$(x - 1)^2 + (y - 2)^2 + 2z^2 = 1$$

en el punto $(x_0, y_0, z_0) = (3/2, 2, \sqrt{6}/4)$.

En primer lugar, nótese que el punto dado pertenece al elipsoide, ya que

$$\frac{1}{4} + 0 + 2 \cdot \frac{6}{16} = \frac{1}{4} + \frac{3}{4} = 1.$$

En este caso $f(x, y, z) = (x - 1)^2 + (y - 2)^2 + 2z^2$, y por consiguiente

$$\nabla f(x, y, z) = (2(x - 1), 2(y - 2), 4z) \implies \nabla f(x_0, y_0, z_0) = (1, 0, \sqrt{6}).$$

Por tanto las ecuaciones paramétricas de la recta normal al elipsoide en el punto (x_0, y_0, z_0) dado son

$$(x, y, z) = (3/2, 2, \sqrt{6}/4) + t(1, 0, \sqrt{6}), \quad t \in \mathbb{R},$$

o equivalentemente

$$x = \frac{3}{2} + t, \quad y = 2, \quad z = \frac{\sqrt{6}}{4}(1 + 4t), \quad t \in \mathbb{R}.$$

Las ecuaciones implícitas de dicha recta se obtienen despejando t de una de las ecuaciones anteriores y sustituyendo en las otras dos:

$$y = 2, \quad z = \frac{\sqrt{6}}{4}(4x - 5).$$

Por último, la ecuación del plano tangente pedido es

$$(1, 0, \sqrt{6}) \cdot \left(x - \frac{3}{2}, y - 2, z - \frac{\sqrt{6}}{4}\right) = x - \frac{3}{2} + \sqrt{6}\left(z - \frac{\sqrt{6}}{4}\right) = 0,$$

es decir

$$x + \sqrt{6}z = 3$$

(cf. la Fig. 2.2, izda.).

Un caso particular importante de lo anterior es el de la *gráfica* $\mathcal{G}(f)$ de una función f de \mathbb{R}^n en \mathbb{R} , que como sabemos es un subconjunto de \mathbb{R}^{n+1} . Más precisamente, si denotamos las coordenadas de los puntos de \mathbb{R}^{n+1} por (x, x_{n+1}) , con $x \equiv (x_1, \dots, x_n) \in \mathbb{R}^n$, entonces la ecuación cartesiana de $\mathcal{G}(f)$ es simplemente

$$x_{n+1} = f(x).$$

En otras palabras,

$$\mathcal{G}(f) = L_0(F),$$

con $F : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ definida por $F(x, x_{n+1}) \equiv f(x) - x_{n+1}$. Claramente, F es diferenciable en un punto $(x_0, f(x_0)) \in \mathcal{G}(f)$ si y solo si f es diferenciable en x_0 (ya que x_{n+1} es diferenciable en cualquier punto, al ser lineal), y se tiene

$$\nabla F(x_0, f(x_0)) = (\nabla f(x_0), -1) \neq 0.$$

Por tanto la ecuación del hiperplano tangente a $\mathcal{G}(f)$ en el punto $(x_0, f(x_0)) \in \mathcal{G}(f)$ es

$$\begin{aligned} (\nabla f(x_0), -1) \cdot [(x, x_{n+1}) - (x_0, f(x_0))] &= (\nabla f(x_0), -1) \cdot (x - x_0, x_{n+1} - f(x_0)) \\ &= \nabla f(x_0) \cdot (x - x_0) - x_{n+1} + f(x_0) = 0, \end{aligned}$$

es decir

$$x_{n+1} = f(x_0) + \nabla f(x_0) \cdot (x - x_0). \quad (2.38)$$

Por otra parte, al ser el vector $(\nabla f(x_0), -1)$ normal a la gráfica de f en $(x_0, f(x_0))$, la ecuación de la recta normal a $\mathcal{G}(f)$ en dicho punto es

$$(x, x_{n+1}) = (x_0, f(x_0)) + t(\nabla f(x_0), -1), \quad t \in \mathbb{R},$$

o equivalentemente

$$x = x_0 + t\nabla f(x_0), \quad x_{n+1} = f(x_0) - t, \quad t \in \mathbb{R}.$$

Despejando t en esta ecuación se obtiene la ecuación cartesiana de la recta normal:

$$x = x_0 + (f(x_0) - x_{n+1})\nabla f(x_0).$$

La ecuación (2.38) proporciona una interpretación geométrica de la diferenciable de la función f en el punto x_0 . En efecto, consideremos un hiperplano afín (no vertical) H_v cualquiera que pase por el punto $(x_0, f(x_0))$, de ecuación

$$x_{n+1} = f(x_0) + v \cdot (x - x_0), \quad (2.39)$$

siendo $v \in \mathbb{R}^n$ un vector arbitrario. La *distancia* entre un punto cualquiera $(x, f(x)) \in \mathcal{G}(f)$ y el punto correspondiente (en la misma vertical) $(x, f(x_0) + v \cdot (x - x_0)) \in H_v$ está dada por

$$|f(x) - f(x_0) - v \cdot (x - x_0)|.$$

El hiperplano (2.39) es tangente a la gráfica de la función f si esta distancia tiende a cero *más rápido que* $\|x - x_0\|$ cuando x tiende a x_0 . Como vimos en la Sección (2.1), esto ocurre si y solo si la función f es *diferenciable* en x_0 , y tomamos $v = \nabla f(x_0)$ (ya que para una función escalar $Df(x_0) \cdot h = \nabla f(x_0) \cdot h$). En otras palabras:

La diferenciable de una función escalar $f : \mathbb{R}^n \rightarrow \mathbb{R}$ en un punto x_0 es *equivalente* a la existencia del hiperplano tangente a la gráfica de f en dicho punto.

Ejemplo 2.31. Calculemos la ecuación del plano tangente y la recta normal a la gráfica de la función $f(x, y) = x^3 + y^2$ en el punto $(1, 3, 10)$.

En primer lugar, el punto dado pertenece a la gráfica de la función, ya que $1 + 3^2 = 10$. Como

$$\nabla f(x, y) = (3x^2, 2y) \quad \Rightarrow \quad \nabla f(1, 3) = (3, 6),$$

la ecuación del plano tangente es

$$z = 10 + (3, 6) \cdot (x - 1, y - 3) = 10 + 3(x - 1) + 6(y - 3),$$

o equivalentemente

$$3x + 6y - z = 11$$

(cf. la Fig. 2.2, drcha.). Las ecuaciones paramétricas de la recta normal a la gráfica por el punto (x_0, y_0, z_0) son

$$(x, y, z) = (x_0, y_0, z_0) + t(\nabla f(x_0, y_0), -1), \quad t \in \mathbb{R},$$

es decir

$$x = 1 + 3t, \quad y = 3 + 6t, \quad z = 10 - t, \quad t \in \mathbb{R},$$

de donde se obtienen fácilmente las correspondientes ecuaciones implícitas:

$$x = 1 + 3(10 - z) = 31 - 3z, \quad y = 3 + 6(10 - z) = 63 - 6z.$$

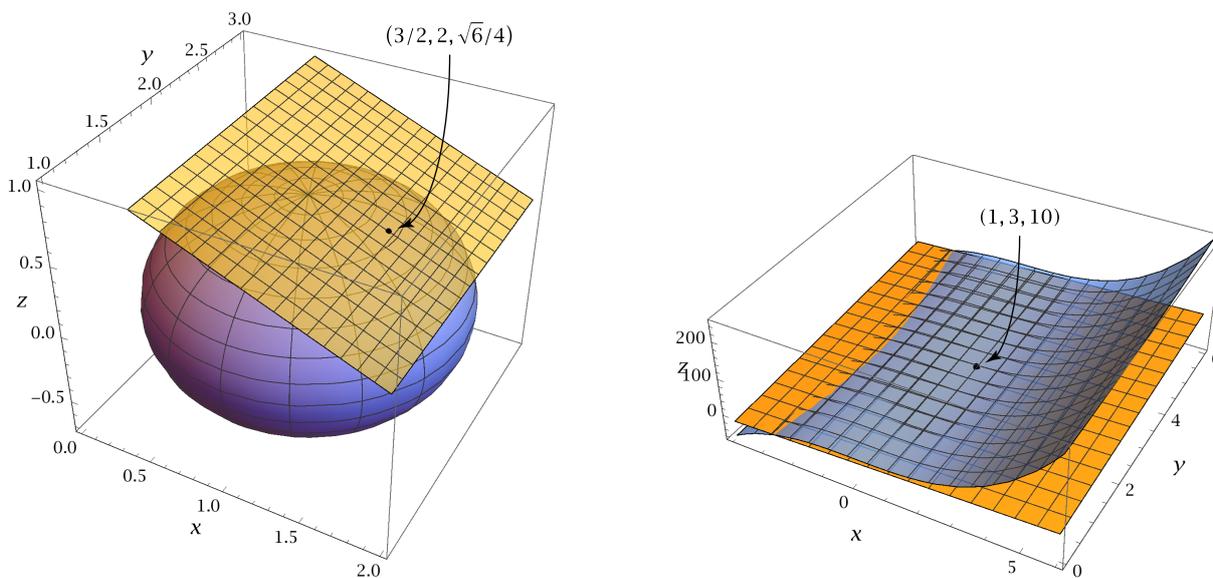


Figura 2.2: Plano tangente al elipsoide $(x - 1)^2 + (y - 2)^2 + 2z^2 = 1$ en el punto $(3/2, 2, \sqrt{6}/4)$ (izda.) y a la gráfica de la función $x^3 + y^2$ en el punto $(1, 3, 10)$ (drcha.).

2.9 Cálculo diferencial vectorial

En general, un **campo vectorial** en \mathbb{R}^n es una función $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$, es decir una función que a cada vector $x \in \mathcal{D}(F) \subset \mathbb{R}^n$ le asigna otro vector del mismo tipo $F(x) \in \mathbb{R}^n$. Si $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ es un campo vectorial diferenciable, su derivada (o, más propiamente, su matriz jacobiana) en un punto cualquiera $x \in \mathbb{R}^n$ es la matriz *cuadrada* $n \times n$

$$DF(x) = \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(x) & \cdots & \frac{\partial F_1}{\partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial F_n}{\partial x_1}(x) & \cdots & \frac{\partial F_n}{\partial x_n}(x) \end{pmatrix},$$

cuyo elemento de matriz (i, j) es la derivada parcial $\frac{\partial F_i}{\partial x_j}(x)$. Por definición, la **divergencia** del campo vectorial F , que denotaremos por $\nabla \cdot F$, es la *traza* de la matriz jacobiana, es decir

$$\nabla \cdot F = \sum_{i=1}^n \frac{\partial F_i}{\partial x_i}. \tag{2.40}$$

Evidentemente, $\nabla \cdot F : \mathbb{R}^n \rightarrow \mathbb{R}$ es una función escalar de n variables.

Notación. La notación $\nabla \cdot F$ está motivada por el hecho de que *formalmente* el miembro derecho de (2.40) es el producto escalar del *operador diferencial*

$$\nabla \equiv \left(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n} \right), \quad (2.41)$$

que se suele llamar **nabla**, con el vector $F(x)$. Es también muy habitual la notación $\operatorname{div} F$ para designar la divergencia del campo vectorial F .

Ejemplo 2.32. Sea $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ un campo vectorial de la forma $F(x) = f(r)x$, donde $r \equiv \|x\|$ y $f : \mathbb{R} \rightarrow \mathbb{R}$ es una función diferenciable (este tipo de campos se denominan **centrales**). La función $x \mapsto r = \sqrt{x_1^2 + \dots + x_n^2}$ es diferenciable para todo $x \neq 0$ por la regla de la cadena⁹, ya que es la composición de las funciones $t \mapsto \sqrt{t}$ (diferenciable para todo $t > 0$) con el polinomio $x_1^2 + \dots + x_n^2$. Si $x \neq 0$, derivando la igualdad

$$r^2 = \sum_{i=1}^n x_i^2$$

respecto de una variable cualquiera x_i y aplicando la regla de la cadena se obtiene

$$\frac{\partial}{\partial x_i} r^2 = 2r \frac{\partial r}{\partial x_i} = 2x_i \implies \frac{\partial r}{\partial x_i} = \frac{x_i}{r}, \quad x \neq 0, \quad (2.42)$$

o equivalentemente

$$\nabla r = \frac{x}{r}.$$

Por las reglas de Leibniz y de la cadena, $F_i(x) = f(r)x_i$ es diferenciable para todo $x \neq 0$, y se tiene

$$\frac{\partial F_i}{\partial x_i}(x) = f(r) + f'(r) \frac{x_i^2}{r}, \quad 1 \leq i \leq n.$$

Sumando estas igualdades respecto de i se obtiene

$$\nabla \cdot F(x) = rf'(r) + nf(r). \quad (2.43)$$

□

Un caso particular importante de campo vectorial es el *gradiente*¹⁰ $F(x) = \nabla f(x)$ de una función escalar (diferenciable) $f : \mathbb{R}^n \rightarrow \mathbb{R}$, cuyas componentes son las n derivadas parciales de f . Si f es de clase C^2 , ∇f es de clase C^1 , y por tanto diferenciable. En tal caso la divergencia del campo vectorial ∇f está dada por

$$\nabla \cdot (\nabla f) = \sum_{i=1}^n \frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_i} \right) = \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2}.$$

El miembro derecho de esta ecuación recibe el nombre de **laplaciano** de la función f , y se denota por $\nabla^2 f$:

$$\nabla^2 f \equiv \nabla \cdot (\nabla f) = \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2}.$$

Otra notación muy habitual para el laplaciano de f es Δf . Evidentemente, el laplaciano de una función escalar $f : \mathbb{R}^n \rightarrow \mathbb{R}$ es otra función escalar de \mathbb{R}^n en \mathbb{R} .

⁹Es fácil ver que $\|x\|$ no es diferenciable en $x = 0$ (ejercicio).

¹⁰El gradiente de una función f se denota a veces por $\operatorname{grad} f$, aunque nosotros no utilizaremos esta notación.

Definición 2.33. Una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ es **armónica** en un abierto $U \subset \mathbb{R}^n$ si su laplaciano se anula idénticamente en U , es decir si $\nabla^2 f(x) = 0$ para todo $x \in U$. La ecuación $\nabla^2 f = 0$, cuyas soluciones son las funciones armónicas, recibe el nombre de **ecuación de Laplace**.

Las funciones armónicas (en especial en \mathbb{R}^2 y \mathbb{R}^3) tienen una gran importancia en Física, ya que por ejemplo el potencial electrostático y el gravitatorio son funciones armónicas en las regiones del espacio donde no hay cargas o masas, respectivamente. Por este motivo, la ecuación de Laplace es una de las ecuaciones más importantes de la Física Matemática.

Ejemplo 2.34. Calculemos el laplaciano de una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ que solo depende de r , es decir tal que $f(x) = g(r)$, con $g : \mathbb{R} \rightarrow \mathbb{R}$ de clase C^2 . En primer lugar, aplicando la regla de la cadena y la ec. (2.42) se obtiene

$$\frac{\partial f}{\partial x_i}(x) = g'(r) \frac{x_i}{r}, \quad x \neq 0,$$

o equivalentemente

$$\nabla f(x) = \frac{g'(r)}{r} x, \quad x \neq 0.$$

De (2.43) con $f(r)$ reemplazada por $g'(r)/r$ se obtiene:

$$\nabla^2 f(x) = n \frac{g'(r)}{r} + r \frac{d}{dr} \left(\frac{g'(r)}{r} \right) = \boxed{g''(r) + (n-1) \frac{g'(r)}{r}}, \quad x \neq 0. \quad (2.44)$$

¿Cuándo es f armónica? Por la igualdad anterior, la condición necesaria y suficiente para que esto ocurra es que la función g satisfaga la ecuación diferencial de segundo orden

$$r g''(r) + (n-1)g'(r) = 0, \quad r > 0. \quad (2.45)$$

La ecuación anterior tiene la solución obvia $g(r) = c$, con $c \in \mathbb{R}$ constante. Por otra parte, si suponemos que $g'(r)$ no se anula en un entorno de un punto r entonces dicha ecuación se escribe:

$$\frac{g''(r)}{g'(r)} \equiv \frac{d}{dr} (\log |g'(r)|) = \frac{1-n}{r}, \quad r > 0.$$

Integrando respecto de r obtenemos

$$\log |g'(r)| = (1-n) \log r + k \iff |g'(r)| = e^k r^{1-n} \iff g'(r) = a r^{1-n}, \quad (2.46)$$

con $a = \pm e^k$ constante real no nula (el caso $a = 0$ corresponde a la solución $g(r) = \text{const.}$ hallada anteriormente). Si $n = 2$, integrando (2.46) respecto de r se obtiene

$$g(r) = a \log r + b \quad (n = 2),$$

con b constante real arbitraria. Si $n > 2$, integrando (2.46) se obtiene

$$g(r) = \frac{a}{2-n} r^{2-n} + b \equiv \tilde{a} r^{2-n} + b \quad (n > 2),$$

con \tilde{a}, b constantes reales arbitrarias. En definitiva (renombrando a la constante arbitraria \tilde{a} en este último caso), las funciones armónicas de n variables que dependen solo de r son

$$f(x) = \begin{cases} a \log r + b, & n = 2, \\ \frac{a}{r^{n-2}} + b, & n > 2, \end{cases}$$

con a, b constantes arbitrarias. En particular, para $n = 3$ se obtiene el *potencial coulombiano* (o *gravitatorio*) $f(x) = (a/r) + b$. \square

El caso $n = 3$ es particularmente importante en las aplicaciones físicas, y merece por ello un tratamiento especial. En este caso utilizaremos la notación más habitual

$$\mathbf{r} = (x, y, z) \in \mathbb{R}^3, \quad r = \|\mathbf{r}\|$$

y denotaremos los campos de vectores mediante letras en negrita (como \mathbf{F} , \mathbf{G} , etc.), reservando las letras en cursiva (como f , g , etc.) para las funciones escalares. Así, por ejemplo, en esta notación las ecuaciones (2.43) y (2.44) se escriben

$$\nabla \cdot (f(r)\mathbf{r}) = 3f(r) + rf'(r), \quad \nabla^2 f(r) = f''(r) + \frac{2f'(r)}{r}.$$

Si $\mathbf{F} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ es un campo vectorial diferenciable, su **rotacional** es el campo vectorial definido por

$$\nabla \times \mathbf{F} = \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}, \frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}, \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right). \quad (2.47)$$

Otra notaciones utilizadas para el rotacional del campo vectorial \mathbf{F} son $\text{rot } \mathbf{F}$ y (en el mundo anglosajón) $\text{curl } \mathbf{F}$. Nótese que, como la propia notación indica, $\nabla \times \mathbf{F}$ es formalmente el producto vectorial del operador nabla (2.41) por el campo \mathbf{F} , es decir

$$\nabla \times \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_1 & F_2 & F_3 \end{vmatrix}, \quad (2.48)$$

donde

$$\mathbf{i} = (1, 0, 0), \quad \mathbf{j} = (0, 1, 0), \quad \mathbf{k} = (0, 0, 1)$$

son los vectores de la base canónica de \mathbb{R}^3 . En otras palabras, las componentes de $\nabla \times \mathbf{F}$ están dadas por

$$(\nabla \times \mathbf{F})_i = \frac{\partial F_k}{\partial x_j} - \frac{\partial F_j}{\partial x_k}, \quad (2.49)$$

donde (i, j, k) es una permutación par (i.e., *cíclica*) de $(1, 2, 3)$ (es decir, $(1, 2, 3)$, $(2, 3, 1)$ o $(3, 1, 2)$).

Ejercicio 26. Se define el **símbolo (tensor) completamente antisimétrico de Levi-Civita** ε_{ijk} mediante

$$\varepsilon_{ijk} = \begin{cases} 1, & (i, j, k) \text{ permutación par de } (1, 2, 3) \\ -1, & (i, j, k) \text{ permutación impar de } (1, 2, 3) \\ 0, & \text{en cualquier otro caso,} \end{cases}$$

donde los índices i, j, k varían en el conjunto $\{1, 2, 3\}$. En otras palabras, si i, j, k son distintos entre sí ε_{ijk} es el *signo* de la permutación (i, j, k) , y si algunos de los índices i, j, k son iguales $\varepsilon_{ijk} = 0$. Probar que

$$(\nabla \times \mathbf{F})_i = \sum_{j,k=1}^3 \varepsilon_{ijk} \frac{\partial F_k}{\partial x_j}$$

y, en general,

$$(\mathbf{a} \times \mathbf{b})_i = \sum_{j,k=1}^3 \varepsilon_{ijk} a_j b_k, \quad \mathbf{a}, \mathbf{b} \in \mathbb{R}^3.$$

i) Utilizando la relación

$$\sum_{i=1}^3 \varepsilon_{ijk} \varepsilon_{ilm} = \delta_{jl} \delta_{km} - \delta_{jm} \delta_{kl}$$

(donde el símbolo δ_{ij} es la llamada **delta de Kronecker**, igual a 1 si $i = j$ y a 0 si $i \neq j$), demostrar la identidad vectorial

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}.$$

ii) Si $\mathbf{F} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ es un campo vectorial de clase C^2 , probar que

$$\nabla \times (\nabla \times \mathbf{F}) = \nabla(\nabla \cdot \mathbf{F}) - \nabla^2 \mathbf{F}, \tag{2.50}$$

donde $\nabla^2 \mathbf{F} = (\nabla^2 F_1, \nabla^2 F_2, \nabla^2 F_3)$. □

Ejemplo 2.35. Las ecuaciones de Maxwell en el vacío (es decir, en cualquier región del espacio que no contenga cargas eléctricas ni corrientes) se pueden escribir (en unidades SI) en la forma

$$\boxed{\nabla \cdot \mathbf{E} = 0, \quad \nabla \cdot \mathbf{B} = 0, \quad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad \nabla \times \mathbf{B} = \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t}},$$

siendo $c \equiv 299792.458 \text{ Km/s}$ la velocidad de la luz en el vacío¹¹. En dichas ecuaciones $\mathbf{E}(t, \mathbf{r})$ y $\mathbf{B}(t, \mathbf{r})$ denotan respectivamente el valor de los campos eléctrico y magnético en el punto $\mathbf{r} \in \mathbb{R}^3$ y en el instante $t \in \mathbb{R}$, y

$$\nabla \equiv \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right).$$

Tomando el rotacional de la tercera ecuación y aplicando la primera ecuación y la identidad (2.50) se obtiene:

$$\nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = -\nabla^2 \mathbf{E} = -\nabla \times \frac{\partial \mathbf{B}}{\partial t}.$$

Si los campos son de clase C^2 podemos permutar los operadores ∇ y $\partial/\partial t$ en el miembro derecho de esta ecuación, lo que proporciona

$$\nabla^2 \mathbf{E} = \frac{\partial}{\partial t} (\nabla \times \mathbf{B}) = \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2}.$$

Análogamente,

$$\nabla(\nabla \cdot \mathbf{B}) - \nabla^2 \mathbf{B} = -\nabla^2 \mathbf{B} = \frac{1}{c^2} \frac{\partial}{\partial t} (\nabla \times \mathbf{E}) = -\frac{1}{c^2} \frac{\partial^2 \mathbf{B}}{\partial t^2}.$$

Por tanto los campos \mathbf{E} y \mathbf{B} verifican en el vacío las ecuaciones

$$\boxed{\nabla^2 \mathbf{E} - \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0, \quad \nabla^2 \mathbf{B} - \frac{1}{c^2} \frac{\partial^2 \mathbf{B}}{\partial t^2} = 0}.$$

La ecuación

$$\boxed{\nabla^2 u - \frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} = 0} \tag{2.51}$$

para una función escalar $u(t, \mathbf{r})$ es la llamada **ecuación de ondas**, ya que describe la propagación de una onda en el espacio con velocidad c . Por ejemplo, dicha ecuación admite soluciones de la forma

$$u(t, \mathbf{r}) = f(x - ct),$$

con $f \in C^2(\mathbb{R})$, que describen la propagación de una onda de amplitud $f(s)$ que se propaga con velocidad c en la dirección del eje x . (En efecto, si $c = f(x_0)$ la superficie de nivel $L_c(f)$ en

¹¹Recuérdese que en unidades SI la velocidad de la luz en el vacío tiene *por definición* el valor *exacto* anterior, y la unidad de longitud (el kilómetro) se define como la distancia recorrida por la luz en $1/c$ segundos.

un cierto instante t es el plano $x = x_0 + ct$.) Por lo que acabamos de ver, las componentes de los campos \mathbf{E} y \mathbf{B} verifican la ecuación de ondas en el vacío, con velocidad de propagación c . En otras palabras, las ecuaciones de Maxwell predicen la existencia de *ondas electromagnéticas* que se propagan (en el vacío) con velocidad c en cualquier sistema de referencia inercial.

Ejercicio 27. Hallar bajo qué condiciones la función $u(t, \mathbf{r}) = f(\mathbf{k} \cdot \mathbf{r} - \omega t)$, donde $f: \mathbb{R} \rightarrow \mathbb{R}$ es una función de clase C^2 , $\omega \in (0, \infty)$ y $\mathbf{k} \in \mathbb{R}^3 - \{0\}$, es solución de la ecuación de ondas (2.51). Describir cualitativamente dicha solución (se trata de una *onda plana* de frecuencia ω y longitud de onda $2\pi/\|\mathbf{k}\|$ que se propaga en la dirección del vector \mathbf{k} con velocidad $\omega/\|\mathbf{k}\|$). \square

La siguiente proposición establece una importante relación entre el gradiente, la divergencia y el rotacional:

Proposición 2.36. Sean $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ y $\mathbf{F}: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ respectivamente una función escalar y un campo vectorial de clase C^2 . Entonces se verifican las identidades

$$\nabla \times (\nabla f) = 0, \quad \nabla \cdot (\nabla \times \mathbf{F}) = 0. \quad (2.52)$$

Demostración. En primer lugar, utilizando la ec. (2.49) y el lema de Schwarz se obtiene

$$[\nabla \times (\nabla f)]_i = \frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_k} \right) - \frac{\partial}{\partial x_k} \left(\frac{\partial f}{\partial x_j} \right) = \frac{\partial^2 f}{\partial x_j \partial x_k} - \frac{\partial^2 f}{\partial x_k \partial x_j} = 0.$$

La segunda identidad (2.52) es también consecuencia directa del lema de Schwarz:

$$\begin{aligned} \nabla \cdot (\nabla \times \mathbf{F}) &= \frac{\partial}{\partial x} \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z} \right) + \frac{\partial}{\partial y} \left(\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x} \right) + \frac{\partial}{\partial z} \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) \\ &= \frac{\partial^2 F_3}{\partial x \partial y} - \frac{\partial^2 F_2}{\partial x \partial z} + \frac{\partial^2 F_1}{\partial y \partial z} - \frac{\partial^2 F_3}{\partial y \partial x} + \frac{\partial^2 F_2}{\partial z \partial x} - \frac{\partial^2 F_1}{\partial z \partial y} = 0. \end{aligned} \quad \square$$

La Proposición anterior afirma que si $\mathbf{F}: U \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3$ es un campo vectorial de clase C^1 en U entonces

$$\mathbf{F} = \nabla f \implies \nabla \times \mathbf{F} = 0, \quad \mathbf{F} = \nabla \times \mathbf{A} \implies \nabla \cdot \mathbf{F} = 0.$$

donde $f: U \rightarrow \mathbb{R}$ es una función escalar y $\mathbf{A}: U \rightarrow \mathbb{R}^3$ es un campo vectorial. ¿Son ciertos los recíprocos de estos resultados? En otras palabras,

$$\nabla \times \mathbf{F} = 0 \stackrel{?}{\implies} \exists f: U \rightarrow \mathbb{R} \text{ t.q. } \mathbf{F} = \nabla f, \quad \nabla \cdot \mathbf{F} = 0 \stackrel{?}{\implies} \exists \mathbf{A}: U \rightarrow \mathbb{R}^3 \text{ t.q. } \mathbf{F} = \nabla \times \mathbf{A}.$$

La respuesta a ambas preguntas depende esencialmente de las propiedades *topológicas* del abierto U donde está definido (y es de clase C^1) el campo vectorial \mathbf{F} . Más precisamente:

- i) Si $\nabla \times \mathbf{F} = 0$ en un abierto U el recíproco de la Proposición 2.36 es cierto (es decir, \mathbf{F} es el gradiente de una función escalar en U) si U es **simplemente conexo**, es decir es conexo y tiene la propiedad de que cualquier curva *cerrada* continua contenida en U puede deformarse de forma continua a un punto sin salirse de U . Ejemplos de abiertos simplemente conexos son \mathbb{R}^3 , una bola abierta, un paralelepípedo abierto, \mathbb{R}^3 menos un punto, \mathbb{R}^3 menos una semirecta, etc., mientras que \mathbb{R}^3 menos una recta es conexo pero no simplemente conexo (en efecto, cualquier curva cerrada que rodee a la recta *no* puede deformarse a un punto de forma continua sin cortar a dicha recta en algún momento).
- ii) Análogamente, si $\mathbf{F} \in C^1(U)$ la condición $\nabla \cdot \mathbf{F} = 0$ en un abierto U garantiza la existencia de un campo vectorial $\mathbf{A}: U \rightarrow \mathbb{R}^3$ tal que $\mathbf{F} = \nabla \times \mathbf{A}$ si U tiene la propiedad de que cualquier *superficie cerrada*¹² (suave) contenida en U se puede deformar de forma continua a un punto sin salirse de U . Por ejemplo, \mathbb{R}^3 , una bola abierta o un paralelepípedo abierto tienen esta propiedad, mientras que \mathbb{R}^3 menos un punto no la tiene (¡aunque es simplemente conexo!).

¹²En este contexto, una superficie cerrada es una superficie sin borde, como por ejemplo una esfera, un elipsoide, etc.

En resumen:

i) Si $\mathbf{F} : U \rightarrow \mathbb{R}^3$ es un campo vectorial de clase C^1 en un abierto *simplemente conexo* U , entonces

$$\mathbf{F} = \nabla f \iff \nabla \times \mathbf{F} = \mathbf{0}.$$

ii) Si $\mathbf{F} : U \rightarrow \mathbb{R}^3$ es de clase C^1 en un abierto U en que *cualquier superficie cerrada continua es deformable a un punto*, entonces

$$\mathbf{F} = \nabla \times \mathbf{A} \iff \nabla \cdot \mathbf{F} = 0.$$

Comentario. El resultado del ejercicio anterior es de gran importancia en Física. En efecto, de las ecuaciones de Maxwell (cf. el Ejemplo 2.35) se deduce que en un campo electromagnético estático el rotacional de $\mathbf{E}(\mathbf{r})$ es nulo. Por lo que acabamos de ver (al ser \mathbb{R}^3 simplemente conexo) $\mathbf{E} = -\nabla\phi(\mathbf{r})$, donde $\phi(\mathbf{r})$ es el llamado *potencial electrostático*. Análogamente, en Mecánica newtoniana el rotacional del campo gravitatorio \mathbf{g} es nulo, y por tanto $\mathbf{g} = -\nabla V$, siendo V el *potencial gravitatorio*. Por otra parte, otra de las ecuaciones de Maxwell afirma que la divergencia del campo magnético \mathbf{B} es nula. Por tanto ha de existir un campo vectorial \mathbf{A} , llamado *potencial vector*, tal que $\mathbf{B} = \nabla \times \mathbf{A}$.

Ejercicio 28. Sea $\mathbf{F} : U \rightarrow \mathbb{R}^3$ un campo vectorial de clase C^1 en el paralelepípedo $U = (a_1, b_1) \times (a_2, b_2) \times (a_3, b_3)$, donde se admite que $a_i = -\infty$ o $b_i = +\infty$. (En particular, los resultados de este ejercicio son válidos si $U = \mathbb{R}^3$.)

i) Si $\nabla \times \mathbf{F} = \mathbf{0}$ en U , probar que $\mathbf{F} = \nabla f$, siendo $f : U \rightarrow \mathbb{R}$ la función definida por

$$f(x, y, z) = \int_{x_0}^x F_1(s, y, z) ds + \int_{y_0}^y F_2(x_0, s, z) ds + \int_{z_0}^z F_3(x_0, y_0, s) ds, \quad (2.53)$$

donde (x_0, y_0, z_0) es un punto arbitrario (pero fijo) de U . Demostrar que cualquier otra función g tal que $\mathbf{F} = \nabla g$ difiere de f en una constante.

ii) Si $\nabla \cdot \mathbf{F} = 0$ en U , probar que $\mathbf{F} = \nabla \times \mathbf{A}$, siendo $\mathbf{A} : U \rightarrow \mathbb{R}^3$ el campo vectorial definido por

$$\begin{cases} A_1(x, y, z) = \int_{z_0}^z F_2(x, y, s) ds, & A_2(x, y, z) = \int_{x_0}^x F_3(s, y, z_0) ds - \int_{z_0}^z F_1(x, y, s) ds, \\ A_3(x, y, z) = 0. \end{cases}$$

Demostrar que cualquier otro campo vectorial \mathbf{B} tal que $\mathbf{F} = \nabla \times \mathbf{B}$ difiere de \mathbf{A} en el gradiente de una función escalar.

Ejercicio 29. Probar que los campos $\mathbf{E}(t, \mathbf{r})$ y $\mathbf{B}(t, \mathbf{r})$ se pueden expresar en la forma

$$\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} - \nabla \phi, \quad \mathbf{B} = \nabla \times \mathbf{A},$$

donde $\phi(t, \mathbf{r})$ (función escalar) y $\mathbf{A}(t, \mathbf{r})$ (campo vectorial) pueden depender de t . ¿Son ϕ y \mathbf{A} únicos? Deducir a partir de las ecuaciones de Maxwell las ecuaciones satisfechas por los campos ϕ y \mathbf{A} .

Ejercicio 30. Hallar una condición necesaria para que el campo vectorial $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ de clase C^1 sea el gradiente de una función escalar $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Probar que dicha condición es también suficiente (por ejemplo) en el paralelepípedo $(a_1, b_1) \times \dots \times (a_n, b_n)$.

Solución. La condición pedida es

$$\boxed{\frac{\partial F_i}{\partial x_j}(x) = \frac{\partial F_j}{\partial x_i}(x)}, \quad 1 \leq i \neq j \leq n,$$

equivalente a $\nabla \times \mathbf{F} = 0$ si $n = 3$. Si dicha condición se cumple en el paralelepípedo $(a_1, b_1) \times \cdots \times (a_n, b_n)$, entonces $F = \nabla f$ si y solo si

$$\boxed{f(x) = \sum_{i=1}^n \int_{c_i}^{x_i} F_i(c_1, \dots, c_{i-1}, s, x_{i+1}, \dots, x_n) ds + k}, \quad (2.54)$$

donde k es constante y $c \equiv (c_1, \dots, c_n)$ es un punto cualquiera del paralelepípedo (cf. la ec. (2.53)).

Ejemplo 2.37 (*gradiente, divergencia y laplaciano en coordenadas polares*). Como es sabido, cualquier punto del plano \mathbb{R}^2 se puede representar en la forma

$$\boxed{(x, y) = (r \cos \theta, r \sin \theta)},$$

donde $r = \sqrt{x^2 + y^2} \geq 0$ y $0 \leq \theta < 2\pi$. Si $r > 0$ (es decir, si $(x, y) \neq (0, 0)$), también el ángulo θ está unívocamente determinado por el punto (x, y) ; por ejemplo, si $x > 0$ e $y \geq 0$ entonces $\theta = \arctan(y/x)$. Más precisamente,

$$\theta = \begin{cases} \arctan(y/x), & x > 0, y \geq 0 \\ \arctan(y/x) + \pi, & x < 0, \\ \arctan(y/x) + 2\pi, & x > 0, y < 0 \\ \pi/2, & x = 0, y > 0 \\ 3\pi/2, & x = 0, y < 0 \end{cases} \quad (2.55)$$

(recuérdese que, por definición, $-\pi/2 < \arctan t < \pi/2$ para todo $t \in \mathbb{R}$). En otras palabras, la función $g : (0, \infty) \times [0, 2\pi) \rightarrow \mathbb{R}^2$ definida por

$$g(r, \theta) = (r \cos \theta, r \sin \theta)$$

es una *biyección* de la franja $(0, \infty) \times [0, 2\pi)$ en $\mathbb{R}^2 - \{(0, 0)\}$. Se dice que $(r, \theta) \in (0, \infty) \times [0, 2\pi)$ son las **coordenadas polares** del punto $(x, y) = (r \cos \theta, r \sin \theta) \neq (0, 0)$. A partir de ahora, escribiremos por sencillez (con un ligero abuso de notación) $\mathbf{r}(r, \theta)$ en lugar de $g(r, \theta)$. El **vector unitario coordenado** \mathbf{e}_r en la dirección radial es

$$\boxed{\mathbf{e}_r = \frac{\partial \mathbf{r}}{\partial r} = \frac{\mathbf{r}}{r} = (\cos \theta, \sin \theta) \equiv \cos \theta \mathbf{i} + \sin \theta \mathbf{j}}. \quad (2.56)$$

El vector unitario coordenado \mathbf{e}_θ es tangente a las circunferencias $r = \text{const.}$, es decir

$$\boxed{\mathbf{e}_\theta = \frac{1}{r} \frac{\partial \mathbf{r}}{\partial \theta} = (-\sin \theta, \cos \theta) \equiv -\sin \theta \mathbf{i} + \cos \theta \mathbf{j}}, \quad (2.57)$$

donde hemos dividido por r para que $\|\mathbf{e}_\theta\| = 1$. Es inmediato comprobar que \mathbf{e}_r y \mathbf{e}_θ son *ortogonales* en cualquier punto, es decir

$$\boxed{\mathbf{e}_r \cdot \mathbf{e}_\theta = 0},$$

y por tanto los vectores $\{\mathbf{e}_r, \mathbf{e}_\theta\}$ forman una *base ortonormal* de \mathbb{R}^2 en cada punto. Nótese sin embargo que, a diferencia de los vectores \mathbf{i}, \mathbf{j} de la base cartesiana, los vectores \mathbf{e}_r y \mathbf{e}_θ no son constantes sino que *dependen del punto considerado*. (En otras palabras, \mathbf{e}_r y \mathbf{e}_θ son campos de

vectores no constantes.) Al ser $\{\mathbf{e}_r, \mathbf{e}_\theta\}$ una base ortonormal, todo vector $\mathbf{v} = (v_1, v_2) \in \mathbb{R}^2$ se puede expresar en la forma

$$\mathbf{v} = v_r \mathbf{e}_r + v_\theta \mathbf{e}_\theta, \quad (2.58)$$

con

$$v_r = \mathbf{v} \cdot \mathbf{e}_r = v_1 \cos \theta + v_2 \sin \theta, \quad v_\theta = \mathbf{v} \cdot \mathbf{e}_\theta = -v_1 \sin \theta + v_2 \cos \theta. \quad (2.59)$$

En particular,

$$\mathbf{i} = \cos \theta \mathbf{e}_r - \sin \theta \mathbf{e}_\theta, \quad \mathbf{j} = \sin \theta \mathbf{e}_r + \cos \theta \mathbf{e}_\theta. \quad (2.60)$$

Recíprocamente, las componentes cartesianas (v_1, v_2) de $\mathbf{v} \in \mathbb{R}^2$ se expresan en términos de (v_r, v_θ) mediante las fórmulas

$$v_1 = \mathbf{v} \cdot \mathbf{i} = v_r \cos \theta - v_\theta \sin \theta, \quad v_2 = \mathbf{v} \cdot \mathbf{j} = v_r \sin \theta + v_\theta \cos \theta, \quad (2.61)$$

donde se han utilizado las ecs. (2.56)-(2.57) para calcular los productos escalares de \mathbf{v} con los vectores \mathbf{i} y \mathbf{j} .

Si $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ es una función escalar de clase C^1 , el gradiente de f está dado por

$$\nabla f = f_x \mathbf{i} + f_y \mathbf{j},$$

donde los subíndices denotan derivadas parciales. Las coordenadas de ∇f en la base $\{\mathbf{e}_r, \mathbf{e}_\theta\}$ se calculan fácilmente teniendo en cuenta de nuevo que dicha base es *ortonormal*, y por tanto

$$\nabla f = (\nabla f \cdot \mathbf{e}_r) \mathbf{e}_r + (\nabla f \cdot \mathbf{e}_\theta) \mathbf{e}_\theta = \left(\nabla f \cdot \frac{\partial \mathbf{r}}{\partial r} \right) \mathbf{e}_r + \frac{1}{r} \left(\nabla f \cdot \frac{\partial \mathbf{r}}{\partial \theta} \right) \mathbf{e}_\theta.$$

Por la regla de la cadena,

$$\begin{cases} \nabla f \cdot \frac{\partial \mathbf{r}}{\partial r} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial r} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial r} = \frac{\partial}{\partial r} f(r \cos \theta, r \sin \theta) \equiv f_r, \\ \nabla f \cdot \frac{\partial \mathbf{r}}{\partial \theta} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial \theta} = \frac{\partial}{\partial \theta} f(r \cos \theta, r \sin \theta) \equiv f_\theta, \end{cases}$$

donde se sobreentiende que las derivadas parciales de f están evaluadas en $\mathbf{r}(r, \theta)$ y $x = r \cos \theta$, $y = r \sin \theta$. Por tanto la expresión de ∇f en coordenadas polares es

$$\nabla f = f_r \mathbf{e}_r + \frac{f_\theta}{r} \mathbf{e}_\theta. \quad (2.62)$$

Multiplicando escalarmente la expresión anterior por \mathbf{i} y \mathbf{j} —o, equivalentemente, utilizando la ec. (2.61)— se obtienen fácilmente las derivadas parciales f_x y f_y en términos de f_r y f_θ :

$$f_x = \nabla f \cdot \mathbf{i} = \cos \theta f_r - \frac{\sin \theta}{r} f_\theta, \quad f_y = \nabla f \cdot \mathbf{j} = \sin \theta f_r + \frac{\cos \theta}{r} f_\theta. \quad (2.63)$$

Simbólicamente, podemos escribir las ecuaciones anteriores en la forma

$$\frac{\partial}{\partial x} = \cos \theta \frac{\partial}{\partial r} - \frac{\sin \theta}{r} \frac{\partial}{\partial \theta}, \quad \frac{\partial}{\partial y} = \sin \theta \frac{\partial}{\partial r} + \frac{\cos \theta}{r} \frac{\partial}{\partial \theta}. \quad (2.64)$$

La divergencia del campo vectorial (de clase C^1) $\mathbf{A} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ se expresa fácilmente en coordenadas polares utilizando las ecuaciones (2.59) y (2.64). En efecto, escribiendo de nuevo (con un ligero abuso de notación)

$$A_{i,r} \equiv \frac{\partial}{\partial r} A_i(r \cos \theta, r \sin \theta), \quad A_{i,\theta} \equiv \frac{\partial}{\partial \theta} A_i(r \cos \theta, r \sin \theta)$$

se tiene

$$\nabla \cdot \mathbf{A} = A_{1,x} + A_{2,y} = \cos \theta A_{1,r} - \frac{\sin \theta}{r} A_{1,\theta} + \sin \theta A_{2,r} + \frac{\cos \theta}{r} A_{2,\theta} \quad (2.65)$$

$$\begin{aligned} &= \frac{\partial}{\partial r} (\cos \theta A_1 + \sin \theta A_2) + \frac{1}{r} \frac{\partial}{\partial \theta} (-\sin \theta A_1 + \cos \theta A_2) + \frac{1}{r} (\cos \theta A_1 + \sin \theta A_2) \\ &\equiv \boxed{A_{r,r} + \frac{A_{\theta,\theta}}{r} + \frac{A_r}{r}}. \end{aligned} \quad (2.66)$$

Si $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ es de clase C^2 , para expresar el laplaciano de f en coordenadas polares basta utilizar la fórmula anterior con $\mathbf{A} = \nabla f$, es decir $A_r = f_r$ y $A_\theta = f_\theta/r$ (cf. la ec. (2.62)). De esta forma se obtiene:

$$\boxed{\nabla^2 f = f_{rr} + \frac{f_{\theta\theta}}{r^2} + \frac{f_r}{r}}. \quad (2.67)$$

En definitiva, el gradiente, la divergencia y el rotacional en coordenadas polares se pueden calcular mediante las fórmulas

$$\boxed{\nabla f = f_r \mathbf{e}_r + \frac{f_\theta}{r} \mathbf{e}_\theta, \quad \nabla \cdot \mathbf{A} = A_{r,r} + \frac{A_{\theta,\theta}}{r} + \frac{A_r}{r}, \quad \nabla^2 f = f_{rr} + \frac{f_{\theta\theta}}{r^2} + \frac{f_r}{r}}. \quad (2.68)$$

Comentario. De los cálculos anteriores se deducen fácilmente la expresión del gradiente, la divergencia y el laplaciano en *coordenadas cilíndricas* (ρ, θ, z) , en términos de las cuales

$$\boxed{\mathbf{r} = (\rho \cos \theta, \rho \sin \theta, z); \quad \rho \equiv \sqrt{x^2 + y^2} > 0, \quad 0 \leq \theta < 2\pi, \quad z \in \mathbb{R}},$$

sin más que añadir el correspondiente término para la coordenada vertical z :

$$\begin{aligned} \nabla f &= f_\rho \mathbf{e}_\rho + \frac{f_\theta}{\rho} \mathbf{e}_\theta + f_z \mathbf{e}_z, & \nabla \cdot \mathbf{A} &= A_{\rho,\rho} + \frac{A_{\theta,\theta}}{\rho} + A_{z,z} + \frac{A_\rho}{\rho}, \\ \nabla^2 f &= f_{\rho\rho} + \frac{f_{\theta\theta}}{\rho^2} + f_{zz} + \frac{f_\rho}{\rho}. \end{aligned}$$

Ejercicio 31. Repetir el cálculo anterior en *coordenadas esféricas* en \mathbb{R}^3

$$\boxed{\mathbf{r} = (r \sin \theta \cos \varphi, r \sin \theta \sin \varphi, r \cos \theta); \quad r > 0, \quad 0 < \theta < \pi, \quad 0 \leq \varphi < 2\pi}.$$

Solución. En este caso los vectores unitarios coordenados son

$$\begin{cases} \mathbf{e}_r = \frac{\partial \mathbf{r}}{\partial r} = \frac{\mathbf{r}}{r}, & \mathbf{e}_\theta = \frac{1}{r} \frac{\partial \mathbf{r}}{\partial \theta} = (\cos \theta \cos \varphi, \cos \theta \sin \varphi, -\sin \theta), \\ \mathbf{e}_\varphi = \frac{1}{r \sin \theta} \frac{\partial \mathbf{r}}{\partial \varphi} = (-\sin \varphi, \cos \varphi, 0), \end{cases}$$

y de nuevo forman una base ortonormal en cada punto (ejercicio). Por tanto las coordenadas de un vector $\mathbf{v} \in \mathbb{R}^3$ respecto de la base $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\varphi\}$ están dadas por

$$v_r = \mathbf{v} \cdot \mathbf{e}_r, \quad v_\theta = \mathbf{v} \cdot \mathbf{e}_\theta, \quad v_\varphi = \mathbf{v} \cdot \mathbf{e}_\varphi.$$

En particular,

$$\begin{aligned} \nabla f &= (\nabla f \cdot \mathbf{e}_r) \mathbf{e}_r + (\nabla f \cdot \mathbf{e}_\theta) \mathbf{e}_\theta + (\nabla f \cdot \mathbf{e}_\varphi) \mathbf{e}_\varphi \\ &= \left(\nabla f \cdot \frac{\partial \mathbf{r}}{\partial r} \right) \mathbf{e}_r + \frac{1}{r} \left(\nabla f \cdot \frac{\partial \mathbf{r}}{\partial \theta} \right) \mathbf{e}_\theta + \frac{1}{r \sin \theta} \left(\nabla f \cdot \frac{\partial \mathbf{r}}{\partial \varphi} \right) \mathbf{e}_\varphi = f_r \mathbf{e}_r + \frac{f_\theta}{r} \mathbf{e}_\theta + \frac{f_\varphi}{r \sin \theta} \mathbf{e}_\varphi. \end{aligned}$$

De esta ecuación se deduce que

$$\begin{cases} \frac{\partial}{\partial x} = \sin \theta \cos \varphi \frac{\partial}{\partial r} + \frac{1}{r} \cos \theta \cos \varphi \frac{\partial}{\partial \theta} - \frac{\sin \varphi}{r \sin \theta} \frac{\partial}{\partial \varphi}, \\ \frac{\partial}{\partial y} = \sin \theta \sin \varphi \frac{\partial}{\partial r} + \frac{1}{r} \cos \theta \sin \varphi \frac{\partial}{\partial \theta} + \frac{\cos \varphi}{r \sin \theta} \frac{\partial}{\partial \varphi}, \\ \frac{\partial}{\partial z} = \cos \theta \frac{\partial}{\partial r} - \frac{\sin \theta}{r} \frac{\partial}{\partial \theta}. \end{cases}$$

Si $\mathbf{A} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ es un campo vectorial, de las ecuaciones anteriores se sigue que

$$\begin{aligned} \nabla \cdot \mathbf{A} &= \sin \theta \cos \varphi A_{1,r} + \cos \theta \cos \varphi \frac{A_{1,\theta}}{r} - \frac{\sin \varphi}{r \sin \theta} A_{1,\varphi} \\ &\quad + \sin \theta \sin \varphi A_{2,r} + \cos \theta \sin \varphi \frac{A_{2,\theta}}{r} + \frac{\cos \varphi}{r \sin \theta} A_{2,\varphi} + \cos \theta A_{3,r} - \frac{\sin \theta}{r} A_{3,\theta} \\ &= A_{r,r} + \frac{A_{\theta,\theta}}{r} + \frac{A_r}{r} + \frac{A_{\varphi,\varphi}}{r \sin \theta} + \frac{1}{r \sin \theta} (A_1 \cos \varphi + A_2 \sin \varphi) \\ &= A_{r,r} + \frac{A_{\theta,\theta}}{r} + \frac{A_{\varphi,\varphi}}{r \sin \theta} + \frac{2A_r}{r} + \frac{\cot \theta}{r} A_\theta. \end{aligned}$$

Para obtener la ecuación anterior hemos tenido en cuenta que

$$A_1 \cos \varphi + A_2 \sin \varphi = A_r \sin \theta + A_\theta \cos \theta,$$

al ser

$$\begin{cases} A_1 = \mathbf{A} \cdot \mathbf{i} = A_r \mathbf{e}_r \cdot \mathbf{i} + A_\theta \mathbf{e}_\theta \cdot \mathbf{i} + A_\varphi \mathbf{e}_\varphi \cdot \mathbf{i} = \sin \theta \cos \varphi A_r + \cos \theta \cos \varphi A_\theta - \sin \varphi A_\varphi, \\ A_2 = \mathbf{A} \cdot \mathbf{j} = A_r \mathbf{e}_r \cdot \mathbf{j} + A_\theta \mathbf{e}_\theta \cdot \mathbf{j} + A_\varphi \mathbf{e}_\varphi \cdot \mathbf{j} = \sin \theta \sin \varphi A_r + \cos \theta \sin \varphi A_\theta + \cos \varphi A_\varphi, \\ A_3 = \mathbf{A} \cdot \mathbf{k} = A_r \mathbf{e}_r \cdot \mathbf{k} + A_\theta \mathbf{e}_\theta \cdot \mathbf{k} + A_\varphi \mathbf{e}_\varphi \cdot \mathbf{k} = \cos \theta A_r - \sin \theta A_\theta. \end{cases}$$

A partir de la ecuación para $\nabla \cdot \mathbf{A}$ se calcula fácilmente $\nabla^2 f$ sustituyendo A_r , A_θ y A_φ respectivamente por f_r , f_θ/r y $f_\varphi/(r \sin \theta)$. En definitiva, el gradiente, la divergencia y el laplaciano se expresan en coordenadas polares por las fórmulas siguientes:

$$\begin{aligned} \nabla f &= f_r \mathbf{e}_r + \frac{f_\theta}{r} \mathbf{e}_\theta + \frac{f_\varphi}{r \sin \theta} \mathbf{e}_\varphi, \\ \nabla \cdot \mathbf{A} &= A_{r,r} + \frac{A_{\theta,\theta}}{r} + \frac{A_{\varphi,\varphi}}{r \sin \theta} + \frac{2A_r}{r} + \frac{\cot \theta}{r} A_\theta, \\ \nabla^2 f &= f_{rr} + \frac{f_{\theta\theta}}{r^2} + \frac{f_{\varphi\varphi}}{r^2 \sin^2 \theta} + \frac{2f_r}{r} + \frac{\cot \theta}{r^2} f_\theta. \end{aligned}$$

Para finalizar este capítulo, enunciaremos a continuación una serie de identidades básicas satisfechas por el gradiente, la divergencia y el rotacional:

$$\begin{aligned} \text{i)} \quad &\nabla \cdot (f\mathbf{F}) = \nabla f \cdot \mathbf{F} + f \nabla \cdot \mathbf{F} \\ \text{ii)} \quad &\nabla \times (f\mathbf{F}) = \nabla f \times \mathbf{F} + f \nabla \times \mathbf{F} \\ \text{iii)} \quad &\nabla \cdot (\mathbf{F} \times \mathbf{G}) = \mathbf{G} \cdot \nabla \times \mathbf{F} - \mathbf{F} \cdot \nabla \times \mathbf{G} \\ \text{iv)} \quad &\nabla (\mathbf{F} \cdot \mathbf{G}) = (\mathbf{F} \cdot \nabla) \mathbf{G} + (\mathbf{G} \cdot \nabla) \mathbf{F} + \mathbf{F} \times (\nabla \times \mathbf{G}) + \mathbf{G} \times (\nabla \times \mathbf{F}) \\ \text{v)} \quad &\nabla \times (\mathbf{F} \times \mathbf{G}) = (\nabla \cdot \mathbf{G}) \mathbf{F} - (\nabla \cdot \mathbf{F}) \mathbf{G} + (\mathbf{G} \cdot \nabla) \mathbf{F} - (\mathbf{F} \cdot \nabla) \mathbf{G} \\ \text{vi)} \quad &\nabla \times (\nabla \times \mathbf{F}) = \nabla (\nabla \cdot \mathbf{F}) - \nabla^2 \mathbf{F}. \end{aligned}$$

En las expresiones anteriores, f es una función escalar de \mathbb{R}^3 en \mathbb{R} , \mathbf{F} y \mathbf{G} son campos vectoriales en \mathbb{R}^3 , $\mathbf{F} \cdot \nabla$ es el operador diferencial

$$\mathbf{F} \cdot \nabla \equiv F_1 \frac{\partial}{\partial x} + F_2 \frac{\partial}{\partial y} + F_3 \frac{\partial}{\partial z},$$

y análogamente $\mathbf{G} \cdot \nabla$.

Capítulo 3

Fórmula de Taylor. Extremos. Funciones inversas e implícitas

3.1 Fórmula de Taylor

Comenzaremos recordando el teorema de Taylor para funciones reales de una variable real. Para ello, necesitamos primero la siguiente definición:

Definición 3.1. Una función $f : \mathbb{R} \rightarrow \mathbb{R}$ es de clase C^k en el intervalo cerrado $[a, b]$ si f es de clase C^k en (a, b) y todas las derivadas de orden $\leq k$ de f son continuas en los extremos a y b , es decir si existen los límites

$$\lim_{h \rightarrow 0^+} f^{(i)}(a+h) \equiv f^{(i)}(a), \quad \lim_{h \rightarrow 0^+} f^{(i)}(b-h) \equiv f^{(i)}(b), \quad 0 \leq i \leq k.$$

Teorema de Taylor. Sea $f : \mathbb{R} \rightarrow \mathbb{R}$ una función de clase C^{p-1} (con $p \geq 1$) en $[a, b]$, y supongamos que $f^{(p)}$ existe en (a, b) . Entonces se verifica:

$$f(b) = \sum_{i=0}^{p-1} \frac{f^{(i)}(a)}{i!} (b-a)^i + R_p, \quad (3.1a)$$

con

$$R_p = \frac{f^{(p)}(c)}{p!} (b-a)^p \quad (3.1b)$$

para algún $c \in (a, b)$.

Nótese, en particular, que cualquier función de clase C^p en $[a, b]$ cumple las hipótesis del teorema, y por tanto verifica la ec. (3.6), que recibe el nombre de *fórmula de Taylor de orden $p-1$ con resto de Lagrange*.

El teorema de Taylor admite una sencilla generalización a funciones *escalares* de varias variables que demostraremos a continuación. Antes de ello, necesitaremos algunas definiciones previas. Recordemos, en primer lugar, que dados dos puntos $x, y \in \mathbb{R}^n$, el **segmento cerrado** que tiene por extremos x e y es el conjunto

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

Análogamente, si $x \neq y$ el **segmento abierto** de extremos x e y es por definición el segmento cerrado excluyendo sus extremos, es decir

$$(x, y) = \{x + s(y-x) \mid 0 < s < 1\}.$$

Comentario. Nótese que el segmento cerrado $[x, y]$ es *compacto* (y por tanto cerrado), al ser la imagen del compacto $[0, 1]$ bajo la función continua (lineal) $\gamma(t) = x + t(y - x)$. Por el contrario, si $n > 1$ el segmento abierto (x, y) *no* es un conjunto abierto (ni, por supuesto, cerrado, ya que su frontera es el segmento cerrado).

Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función escalar de clase C^p (con $p \geq 1$) en un abierto U , y supongamos que el segmento cerrado $[x, x + h]$ está contenido en U . La idea fundamental que conduce a la generalización del teorema de Taylor es estudiar los valores de la función f a lo largo del segmento $[x, x + h]$, es decir considerar la función

$$\phi(t) = f(x + th), \quad 0 \leq t \leq 1.$$

La función ϕ es la composición de f con la función $\gamma : \mathbb{R} \rightarrow \mathbb{R}^n$ definida por $\gamma(t) = x + th$. Por la regla de la cadena, $\phi = f \circ \gamma$ es diferenciable en el intervalo $[0, 1]$. En efecto, γ es de clase C^∞ en todo \mathbb{R} (lineal), y si $t \in [0, 1]$ entonces $\gamma(t) \in [x, x + h] \subset U$, siendo f diferenciable en U por hipótesis (al ser de clase C^p con $p \geq 1$). La derivada de ϕ se calcula directamente aplicando la regla de la cadena:

$$\phi'(t) = Df(x + th) \cdot h = \nabla f(x + th) \cdot h = \sum_{i,j=1}^n \frac{\partial f}{\partial x_j}(x + th) h_j.$$

Si $p \geq 2$, las derivadas parciales $D_i f$ son de clase C^{p-1} , y por tanto diferenciables (al ser $p-1 \geq 1$) en el abierto U . Aplicando de nuevo la regla de la cadena a cada sumando de la fórmula anterior se obtiene

$$\phi''(t) = \sum_{j=1}^n \left(\sum_{i=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(x + th) h_i \right) h_j \equiv \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(x + th) h_i h_j.$$

Procediendo de esta forma (o, más rigurosamente, por inducción) se demuestra que ϕ es p veces derivable en el intervalo $[0, 1]$, y que sus derivadas de orden $\leq p$ están dadas por

$$\phi^{(k)}(t) = \sum_{i_1, \dots, i_k=1}^n \frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}(x + th) h_{i_1} \dots h_{i_k}, \quad 1 \leq k \leq p. \quad (3.2)$$

Por tanto ϕ es de clase C^p en el intervalo $[0, 1]$, ya que $x + th \in [x, x + h] \subset U$ y las derivadas parciales de orden $\leq p$ de f son continuas en U por hipótesis. Es conveniente en este punto introducir la notación abreviada

$$D^k f(y) \cdot h \equiv \sum_{i_1, \dots, i_k=1}^n \frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}(y) h_{i_1} \dots h_{i_k}, \quad (3.3)$$

en términos de la cual las derivadas de orden $\leq p$ de ϕ están dadas por

$$\phi^{(k)}(t) = D^k f(x + th) \cdot h, \quad 1 \leq k \leq p, \quad (3.4a)$$

y análogamente

$$\phi^{(k)}(0) = D^k f(x) \cdot h, \quad 1 \leq k \leq p. \quad (3.4b)$$

Nótese que, al estar el vector x fijo, podemos considerar a $D^k f(x) \cdot h$ como una función (polinomio homogéneo de grado k) del vector h . En particular, $D^1 f(x) \equiv Df(x)$, y $D^2 f(x)$ es la forma cuadrática en h dada por

$$D^2 f(x) \cdot h = \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(x) h_i h_j. \quad (3.5)$$

Estamos ya en condiciones de demostrar el teorema de Taylor para funciones escalares de n variables:

Proposición 3.2 (fórmula de Taylor con resto). Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función de clase C^p (con $p \geq 1$) en un abierto U , y supongamos que el segmento cerrado $[x, x + h]$ está contenido en U . Entonces se tiene

$$f(x + h) = f(x) + \sum_{k=1}^{p-1} \frac{1}{k!} D^k f(x) \cdot h + R_p(x; h), \quad (3.6)$$

siendo

$$R_p(x; h) = \frac{1}{p!} D^p f(x + sh) \cdot h \quad (3.7)$$

para un cierto $s \in (0, 1)$.

Demostración. Basta aplicar el teorema de Taylor a la función ϕ en el intervalo $[0, 1]$, teniendo en cuenta las ecs. (3.1) (con $a = 0$, $b = 1$) y (3.4). \square

Comentario. En particular, si $p = 1$ la fórmula de Taylor con resto (3.6)-(3.7)

$$\boxed{f(x + h) = f(x) + Df(x + sh) \cdot h \equiv f(x) + \nabla f(x + sh) \cdot h}, \quad \text{para algún } s \in (0, 1),$$

se conoce como el *teorema del valor medio* o *de los incrementos finitos*. De hecho, repasando la demostración de las ecuaciones (3.6)-(3.7) para $p = 1$ se comprueba fácilmente que el teorema de valor medio es válido si f es simplemente diferenciable (y no necesariamente de clase C^1) en el segmento abierto $(x, x + h)$ y continua en sus extremos. \square

Estudiemos a continuación con más detalle el resto (3.7) de la fórmula de Taylor. Por la continuidad de las derivadas parciales de orden p de f en U , $R_p(x; h)$ debería ser aproximadamente igual a $\frac{1}{p!} D^p f(x) \cdot h$ para h suficientemente pequeño. Más precisamente, probaremos a continuación que

$$\boxed{R_p(x; h) = \frac{1}{p!} D^p f(x) \cdot h + o(\|h\|^p)}. \quad (3.8)$$

En efecto

$$\begin{aligned} \left| R_p(x; h) - \frac{1}{p!} D^p f(x) \cdot h \right| &= \frac{1}{p!} \left| \sum_{i_1, \dots, i_p=1}^n \left(\frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_p}}(x + sh) - \frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_p}}(x) \right) h_{i_1} \dots h_{i_p} \right| \\ &\leq \frac{1}{p!} \sum_{i_1, \dots, i_p=1}^n \left| \frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_p}}(x + sh) - \frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_p}}(x) \right| |h_{i_1}| \dots |h_{i_p}| \\ &\leq \frac{\|h\|^p}{p!} \sum_{i_1, \dots, i_p=1}^n \left| \frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_p}}(x + sh) - \frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_p}}(x) \right|, \end{aligned}$$

al ser $|h_i| \leq \|h\|$ para todo i . Por tanto

$$\frac{1}{\|h\|^p} \left| R_p(x; h) - \frac{1}{p!} D^p f(x) \cdot h \right| \leq \frac{1}{p!} \sum_{i_1, \dots, i_p=1}^n \left| \frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_p}}(x + sh) - \frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_p}}(x) \right|,$$

donde cada sumando de la suma del miembro derecho tiende a cero cuando $h \rightarrow 0$ por la continuidad de las derivadas parciales de orden p de f en U . Esto concluye la demostración de (3.8). Combinando (3.6) con (3.8) se obtiene la siguiente versión del teorema de Taylor para funciones escalares de varias variables:

Proposición 3.3 (fórmula de Taylor de orden p). Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función de clase C^p (con $p \geq 1$) en un abierto U , y supongamos que el segmento cerrado $[x, x + h]$ está contenido en U .

Entonces se tiene:

$$f(x+h) = f(x) + \sum_{k=1}^p \frac{1}{k!} D^k f(x) \cdot h + o(\|h\|^p). \quad (3.9)$$

Recuérdese que un subconjunto $A \subset \mathbb{R}^n$ es **convexo** si para todo par de puntos $x, y \in A$ el segmento cerrado $[x, y]$ está contenido en A . En particular:

Si $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ es de clase C^p en un abierto convexo U la ecuación (3.9) es válida para cualquier par de puntos $x, x+h \in U$.

La ecuación (3.3) puede simplificarse considerablemente teniendo en cuenta que, al ser f de clase C^p por hipótesis, sus derivadas parciales cruzadas de orden $\leq p$ son iguales en virtud del lema de Schwarz. Por ejemplo, para $p = 2$ se tiene:

$$D^2 f(x) \cdot h = \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2}(x) h_i^2 + \sum_{1 \leq i \neq j \leq n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x) h_i h_j = \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2}(x) h_i^2 + 2 \sum_{1 \leq i < j \leq n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x) h_i h_j.$$

En general, si $k \leq p$

$$\frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}}(x) h_{i_1} \cdots h_{i_k} = \frac{\partial^k f}{\partial x_1^{j_1} \cdots \partial x_n^{j_n}}(x) h_1^{j_1} \cdots h_n^{j_n}$$

si el multiíndice (i_1, \dots, i_k) es una permutación de

$$\underbrace{(1, \dots, 1)}_{j_1}, \underbrace{2, \dots, 2}_{j_2}, \dots, \underbrace{n, \dots, n}_{j_n},$$

con $0 \leq j_i \leq k$ y $j_1 + \dots + j_n = k$. Como hay $k!/(j_1! \cdots j_n!)$ permutaciones distintas de este último multiíndice, la ec. (3.3) puede escribirse en la forma simplificada

$$D^k f(x) \cdot h = \sum_{\substack{0 \leq j_1, \dots, j_n \leq k \\ j_1 + \dots + j_n = k}} \frac{k!}{j_1! \cdots j_n!} \frac{\partial^k f}{\partial x_1^{j_1} \cdots \partial x_n^{j_n}}(x) h_1^{j_1} \cdots h_n^{j_n}. \quad (3.10)$$

En particular, para $n = 2$ la fórmula anterior se reduce a

$$D^k f(x, y) \cdot h = \sum_{i=0}^k \binom{k}{i} \frac{\partial^k f}{\partial x^i \partial y^{k-i}}(x, y) h_1^i h_2^{k-i}.$$

En la práctica, en lugar de memorizar la ec. (3.10) se puede proceder como sigue. En primer lugar, nótese que la fórmula para la derivada de la función $\phi(t) \equiv f(x+th)$ se puede escribir en la forma

$$\phi'(t) = ((h \cdot \nabla) f)(x+th),$$

donde $h \cdot \nabla$ es el operador diferencial

$$h \cdot \nabla = h_1 \frac{\partial}{\partial x_1} + \cdots + h_n \frac{\partial}{\partial x_n}.$$

De esta ecuación se deduce fácilmente que

$$\phi^{(k)}(t) = \left(\frac{d}{dt} \right)^k \phi(t) = ((h \cdot \nabla)^k f)(x+th), \quad 1 \leq k \leq p,$$

y por tanto

$$\phi^{(k)}(0) = \boxed{D^k f(x) \cdot h = (h \cdot \nabla)^k f(x)}, \quad 1 \leq k \leq p. \quad (3.11)$$

Utilizando esta igualdad en la fórmula de Taylor (3.9) se obtiene la fórmula equivalente

$$\boxed{f(x+h) = f(x) + \sum_{k=1}^p \frac{1}{k!} (h \cdot \nabla)^k f(x) + o(\|h\|^p)}. \quad (3.12)$$

Nótese por último que, al ser f de clase C^p , si $k \leq p$ en virtud del lema de Schwarz podemos desarrollar $(h \cdot \nabla)^k$ tratando a las derivadas parciales que aparecen en esta expresión como variables que *conmutan*, exactamente igual que si fueran números reales. Así, por ejemplo, si $n = 2$ y $k = 3$ se tiene

$$(h \cdot \nabla)^3 \equiv \left(h_1 \frac{\partial}{\partial x} + h_2 \frac{\partial}{\partial y} \right)^3 = h_1^3 \frac{\partial^3}{\partial x^3} + 3h_1^2 h_2 \frac{\partial^3}{\partial x^2 \partial y} + 3h_1 h_2^2 \frac{\partial^3}{\partial x \partial y^2} + h_2^3 \frac{\partial^3}{\partial y^3}.$$

Ejemplo 3.4. Hallems el desarrollo de Taylor de orden 2 de la función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por

$$f(x, y) = e^{x+y} \cos(x - y^2)$$

en el punto $(0, 0)$. Es claro, en primer lugar, que f es de clase C^∞ en todo \mathbb{R}^2 , y por tanto existe el desarrollo de Taylor de cualquier orden en el origen. Calculando las derivadas parciales de primer orden de f obtenemos

$$\begin{cases} f_x = e^{x+y} [\cos(x - y^2) - \sin(x - y^2)] & \Rightarrow f_x(0, 0) = 1 \\ f_y = e^{x+y} [\cos(x - y^2) + 2y \sin(x - y^2)] & \Rightarrow f_y(0, 0) = 1. \end{cases}$$

Las derivadas segundas también se calculan fácilmente:

$$\begin{cases} f_{xx} = -2e^{x+y} \sin(x - y^2) \\ f_{xy} = e^{x+y} [\cos(x - y^2) - \sin(x - y^2) + 2y (\sin(x - y^2) + \cos(x - y^2))] \\ f_{yy} = e^{x+y} [\cos(x - y^2) + 2(1 + 2y) \sin(x - y^2) - 4y^2 \cos(x - y^2)] \end{cases}$$

de donde

$$f_{xx}(0, 0) = 0, \quad f_{xy}(0, 0) = f_{yy}(0, 0) = 1.$$

Por tanto, el desarrollo de Taylor de orden 2 de f en el origen es

$$\begin{aligned} f(x, y) &= f(0, 0) + f_x(0, 0)x + f_y(0, 0)y \\ &\quad + \frac{1}{2} (f_{xx}(0, 0)x^2 + 2f_{xy}(0, 0)xy + f_{yy}(0, 0)y^2) + o(x^2 + y^2) \\ &= \boxed{1 + x + y + \frac{1}{2}(2xy + y^2) + o(x^2 + y^2)}. \end{aligned}$$

Nótese que este resultado también se podía haber calculado (algo más fácilmente) multiplicando los desarrollos de e^{x+y} y $\cos(x - y^2)$. En efecto, utilizando los desarrollos de Taylor a segundo orden de e^t y $\cos t$ en $t = 0$ se obtiene

$$\begin{aligned} e^{x+y} &= 1 + x + y + \frac{1}{2}(x+y)^2 + o((x+y)^2) = 1 + x + y + \frac{1}{2}(x+y)^2 + o(x^2 + y^2), \\ \cos(x - y^2) &= 1 - \frac{1}{2}(x - y^2)^2 + o((x - y^2)^2) = 1 - \frac{1}{2}x^2 + o(x^2 + y^2), \end{aligned}$$

donde hemos utilizado el hecho de que xy^2 , y^4 , $o((x+y)^2)$ y $o((x - y^2)^2)$ son $o(x^2 + y^2)$. (Ejercicio: probar que, en general, si $f = O(g)$ entonces $o(f) = o(g)$).

Multiplicando ambos desarrollos (y descartando términos $o(x^2 + y^2)$) se sigue que

$$e^{x+y} \cos(x-y^2) = 1+x+y+\frac{1}{2}(x+y)^2-\frac{1}{2}x^2+o(x^2+y^2) = 1+x+y+\frac{1}{2}(2xy+y^2)+o(x^2+y^2),$$

como antes.

3.2 Extremos

La fórmula de Taylor es particularmente útil para hallar los extremos de una función escalar de varias variables, como veremos en esta sección. Comenzaremos con las siguientes definiciones:

Definición 3.5. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función definida en un conjunto $A \subset \mathbb{R}^n$, y sea $x_0 \in A$.

i) x_0 es un **máximo absoluto** (ó **global**) de f en A si

$$f(x) \leq f(x_0), \quad \forall x \in A.$$

ii) x_0 es un **máximo absoluto estricto** de f en A si

$$f(x) < f(x_0), \quad \forall x \neq x_0, \quad x \in A.$$

iii) x_0 es un **máximo relativo** (ó **local**) de f en A si existe un entorno U de x_0 tal que x_0 es un máximo absoluto de f en $U \cap A$, es decir

$$f(x) \leq f(x_0), \quad \forall x \in U \cap A.$$

iv) x_0 es un **máximo relativo estricto** de f en A si existe un entorno U de x_0 tal que x_0 es un máximo absoluto estricto de f en $U \cap A$, es decir

$$f(x) < f(x_0), \quad \forall x \neq x_0, \quad x \in U \cap A.$$

v) x_0 es un **mínimo** de cualquiera de los tipos anteriores de f en A si es un máximo del mismo tipo de $-f$ en A .

vi) x_0 es un **extremo** de cualquiera de los tipos anteriores de f en A si es un máximo o un mínimo de ese tipo de f en A .

vii) x_0 es un **punto de silla** de f en A si $\nabla f(x_0) = 0$ y x_0 *no* es un extremo relativo de f en A , es decir si para todo $r > 0$ hay sendos puntos $x_1, x_2 \in B_r(x_0) \cap A$ tales que

$$f(x_1) < f(x_0) < f(x_2).$$

Comentario. Cuando no se menciona el conjunto A en cualquiera de las definiciones anteriores se sobreentiende que $A = \mathbb{R}^n$. Así, por ejemplo, se dice que f tiene un máximo relativo en $x_0 \in \mathbb{R}^n$ si hay un entorno U de x_0 tal que

$$f(x) \leq f(x_0), \quad \forall x \in U.$$

Ejemplo 3.6. Veamos algunos ejemplos de funciones de \mathbb{R}^2 en \mathbb{R} que ilustran las definiciones anteriores (cf. la Fig. 3.1):

i) $f(x, y) = -\sqrt{x^2 + y^2}$ tiene un máximo absoluto estricto en $(0, 0)$ que *no* se puede hallar con el Cálculo diferencial, ya que f no tiene derivadas parciales en el origen. En efecto,

$$f_x(0, 0) = \lim_{h \rightarrow 0} \frac{f(h, 0) - f(0, 0)}{h} = -\lim_{h \rightarrow 0} \frac{\sqrt{h^2}}{h} = -\lim_{h \rightarrow 0} \operatorname{sgn} h$$

no existe (vale -1 por la derecha y 1 por la izquierda), y análogamente $f_y(0, 0)$.

ii) $f(x, y) = -(x + y)^2$ tiene un máximo absoluto, pero no estricto, en $(0, 0)$. En efecto, es claro que $f(x, y) \leq 0 = f(0, 0)$, pero $f(x, y) = 0$ si y solo si $x + y = 0$. Nótese, en particular, que todo punto de la forma $(x_0, -x_0)$ con $x_0 \in \mathbb{R}$ es un máximo absoluto (no estricto) de la función.

iii) $f(x, y) = -x^2 - y^2 + x^2 y$ tiene un máximo relativo estricto, pero no un máximo absoluto, en $(0, 0)$. En efecto, $f(x, y)/(x^2 + y^2)$ tiende a -1 para $(x, y) \rightarrow (0, 0)$. Esto implica que $f(x, y)/(x^2 + y^2)$, y por tanto $f(x, y)$, es negativa en un entorno reducido del origen. Como $f(0, 0) = 0$, f tiene un máximo relativo estricto en el origen. Este máximo no es absoluto, ya que f es positiva si $x^2(y - 1) > y^2$ (por ejemplo, en puntos de la forma $(x, 2)$ con $|x| > 2$).

iv) $f(x, y) = x^2 - y^4$ tiene un punto de silla en $(0, 0)$ (aunque el término de orden más bajo en $(0, 0)$, $g(x, y) = x^2$, tiene un mínimo absoluto no estricto en el origen). En efecto, $\nabla f(x, y) = (2x, -4y^3)$ se anula en el origen, y por otra parte $f(0, 0) = 0$, $f(x, 0) > 0$ si $x \neq 0$ y $f(0, y) < 0$ si $y \neq 0$. La razón por la cual en este caso el término de orden 2 en el origen (x^2) no determina el comportamiento de f cerca de dicho punto es que a lo largo de la recta $x = 0$ este término se anula idénticamente, y por lo tanto en la recta $x = 0$ el término dominante es $-y^4$.

Naturalmente, cambiando f por $-f$ en los tres primeros ejemplos se obtienen ejemplos análogos reemplazando la palabra “máximo” por “mínimo”.

Definición 3.7. Un **punto crítico** de una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ es un punto $x_0 \in \mathbb{R}^n$ tal que $Df(x_0) = 0$.

Equivalentemente, $x_0 \in \mathbb{R}^n$ es punto crítico de f si y solo si

$$\boxed{\nabla f(x_0) = 0},$$

o bien si se anulan *todas* las derivadas parciales de f en x_0 :

$$\boxed{\frac{\partial f}{\partial x_i}(x_0) = 0, \quad 1 \leq i \leq n}.$$

El resultado fundamental que relaciona los extremos de una función con sus puntos críticos es el siguiente:

Proposición 3.8. Sea x_0 un extremo relativo de $f : \mathbb{R}^n \rightarrow \mathbb{R}$ en el conjunto $A \subset \mathbb{R}^n$. Si $x_0 \in \overset{\circ}{A}$ y f es diferenciable en dicho punto, entonces x_0 es un punto crítico de f .

Demostración. Como f es diferenciable en x_0 está definida en un entorno U de dicho punto, que podemos suponer contenido en A por ser x_0 un punto interior a este conjunto. Además, al ser x_0 un extremo relativo de f en A hay una bola $B_r(x_0) \subset U \subset A$ tal que x_0 es un extremo de f en $B_r(x_0)$. Si $v \in \mathbb{R}^n$ es un vector no nulo, la función $\phi(t) = f(x_0 + tv)$ está bien definida

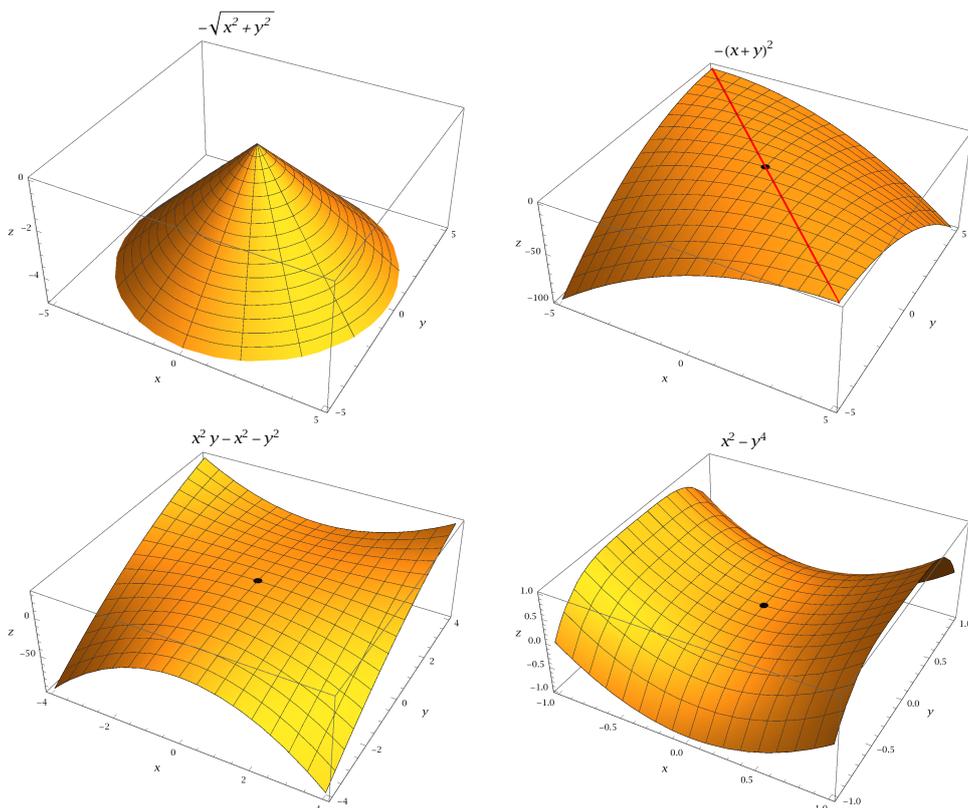


Figura 3.1: Gráfica de las funciones del Ejemplo 3.6. El punto en negro en las tres últimas gráficas es el origen, y la curva en rojo es la intersección de la gráfica de la función $-(x + y)^2$ con el plano $x + y = 0$.

para $|t| < r/\|v\|$ (ya que $x_0 + tv \in B_r(x_0)$ si $|t| < r/\|v\|$) y tiene obviamente un extremo en $t = 0$. En consecuencia, se cumple

$$\phi'(0) = Df(x_0) \cdot v = 0, \quad \forall v \neq 0, v \in \mathbb{R}^n,$$

y por tanto $Df(x_0) = 0$. □

Comentarios.

- Evidentemente, el resultado anterior *no* tiene por qué ser cierto si x_0 *no* es un punto interior de A . Por ejemplo, la función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $f(x, y) = x^2 + y^2$ tiene un mínimo y cuatro máximos (todos ellos absolutos) en el cuadrado cerrado $A = [-1, 1] \times [-1, 1]$, respectivamente en el origen (mínimo) y en los cuatro vértices $(\pm 1, \pm 1)$ del cuadrado (máximos). Sin embargo,

$$\nabla f(\pm 1, \pm 1) = 2(\pm 1, \pm 1) \neq 0.$$

Esto no contradice el resultado anterior, ya que los vértices del cuadrado A *no* son puntos interiores de A (de hecho, son puntos frontera).

- Es muy importante observar que la anulación de $\nabla f(x_0)$ en un punto interior al dominio de f es condición *necesaria*, pero no *suficiente*, para que f tenga un extremo relativo en x_0 . En otras palabras, *los puntos críticos de una función no son necesariamente extremos relativos* de dicha función. Por ejemplo, la función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por $f(x, y) = x^2 - y^4$ no tiene un extremo en el origen, aunque $\nabla f(x, y) = (2x, -4y^3)$ se anula en dicho punto.

3.3 Formas cuadráticas

En el caso de funciones de una variable, es posible utilizar las derivadas de orden superior para averiguar si un punto crítico es un máximo o un mínimo. Estos resultados se pueden generalizar a funciones escalares de varias variables, como veremos a continuación. Para ello, necesitaremos antes algunas definiciones previas.

Definición 3.9 (formas cuadráticas).

- i) Una **forma cuadrática** en n variables $(h_1, \dots, h_n) \equiv h \in \mathbb{R}^n$ es un polinomio homogéneo de grado 2 en h de la forma

$$Q(h) = \sum_{i,j=1}^n a_{ij} h_i h_j, \quad (3.13)$$

donde $a_{ij} \in \mathbb{R}$ y podemos suponer sin pérdida de generalidad que $a_{ij} = a_{ji}$ para todo $i \neq j$ (en efecto, si esta igualdad no se cumpliera bastaría reemplazar a_{ij} por $(a_{ij} + a_{ji})/2$, lo que no altera $Q(h)$). La matriz *simétrica*

$$A = (a_{ij})_{1 \leq i, j \leq n} = A^T$$

es por definición la matriz de la forma cuadrática Q . Si (como es habitual) representamos el vector h por la matriz columna $\begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}$ entonces (3.13) es equivalente a

$$Q(h) = h^T A h.$$

Nótese que, al ser Q homogénea de grado 2 en h , se tiene

$$Q(\lambda h) = \lambda^2 Q(h), \quad \forall \lambda \in \mathbb{R}; \quad (3.14)$$

en particular, $Q(0) = 0$ para toda forma cuadrática Q .

- ii) La forma cuadrática Q es **semidefinida positiva** si

$$Q(h) \geq 0, \quad \forall h \in \mathbb{R}^n,$$

y **definida positiva** si

$$Q(h) > 0, \quad \forall h \neq 0, \quad h \in \mathbb{R}^n.$$

- iii) Análogamente, Q es **semidefinida negativa** (resp. **definida negativa**) si $-Q$ es semidefinida positiva (resp. definida positiva), es decir si $Q(h) \leq 0$ para todo $h \in \mathbb{R}^n$ (resp. $Q(h) < 0$ para todo $h \neq 0$).

- iv) La forma cuadrática Q es **indefinida** si no es semidefinida positiva ni negativa, es decir si existen sendos vectores $v_1, v_2 \in \mathbb{R}^n$ tales que

$$Q(v_1) > 0, \quad Q(v_2) < 0.$$

- v) Por último, Q es **degenerada** si el determinante de su matriz es nulo, es decir si $\det A = 0$, y **no degenerada** en caso contrario (es decir, si $\det A \neq 0$).

Comentarios.

- Se demostrará en el curso de Álgebra Lineal que el determinante de una matriz A es el producto de sus *autovalores*, es decir, las raíces $\lambda_1, \dots, \lambda_n$ (en general complejas) del *polinomio característico* $\det(A - \lambda I)$ (donde I denota la matriz unidad). Por tanto:

La forma cuadrática Q es *no degenerada* si y solo si *todos* los autovalores de su matriz A son *no nulos*.

- Se demostrará también en el curso de Álgebra Lineal que los autovalores de una matriz simétrica real son todos reales, junto con la siguiente caracterización de las formas cuadráticas (semi)definidas positivas o negativas en términos de los autovalores¹ de su matriz A :

- i) Q es *semidefinida positiva* (resp. *negativa*) si todos los autovalores de A son no negativos (resp. no positivos).
- ii) Q es *definida positiva* (resp. *negativa*) si todos los autovalores de A son positivos (resp. negativos).
- iii) Q es *indefinida* si A tiene al menos un autovalor positivo y al menos otro negativo.

- De los dos comentarios anteriores se sigue el siguiente resultado:

Una forma cuadrática *no degenerada* solo puede ser *definida positiva* (si todos los autovalores de su matriz son positivos), *definida negativa* (si todos los autovalores de su matriz son negativos), o *indefinida* (si su matriz tiene tanto autovalores positivos como negativos).

- La forma cuadrática $Q(h)$ es una función continua de $h \in \mathbb{R}^n$, al ser un polinomio. En particular, ha de alcanzar un máximo y un mínimo en la esfera unidad $\partial B_1(0)$, ya que este conjunto es compacto. En otras palabras, existen sendos vectores $v_{1,2}$ tales que $\|v_{1,2}\| = 1$ y

$$Q(v_1) \leq Q(v) \leq Q(v_2), \quad \forall v \text{ t.q. } \|v\| = 1. \quad (3.15)$$

3.4 Criterio del hessiano

Si $f : \mathbb{R}^n \rightarrow \mathbb{R}$ es de clase C^2 en un abierto $U \subset \mathbb{R}^n$ y fijamos $x \in U$, la función $D^2f(x) \cdot h$ definida por la ec. (3.5) es claramente una forma cuadrática en h de matriz

$$\begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(x) \end{pmatrix} \equiv Hf(x) \quad (3.16)$$

denominada **matriz hessiana** de f en x . Nótese que $Hf(x)$ es efectivamente una matriz *simétrica*, al ser $D_{ij}f(x) = D_{ji}f(x)$ en virtud del lema de Schwarz.

Con los resultados y definiciones anteriores, estamos ya en disposición de probar el resultado fundamental para clasificar puntos críticos de una función de varias variables:

Proposición 3.10. *Sea x_0 un punto crítico de $f : \mathbb{R}^n \rightarrow \mathbb{R}$, y supongamos que f es de clase C^2 en un entorno de x_0 . Entonces se verifica:*

- i) *Si x_0 es un mínimo relativo de f , entonces $D^2f(x_0)$ es semidefinida positiva.*

¹Esta caracterización es consecuencia de la siguiente propiedad: para toda forma cuadrática Q , existe un cambio de coordenadas lineal $(h_1, \dots, h_n) \mapsto (h'_1, \dots, h'_n)$ tal que $Q(h) = \sum_{i=1}^n \lambda_i h_i'^2$, donde $\lambda_1, \dots, \lambda_n$ son los autovalores de la matriz de Q .

ii) Si $D^2f(x_0)$ es definida positiva, entonces x_0 es un mínimo relativo de f .

iii) Si $D^2f(x_0)$ es indefinida, entonces x_0 es un punto de silla de f .

Demostración. En efecto, sea $x_0 \in \mathbb{R}^n$ un punto crítico de una función escalar $f : \mathbb{R}^n \rightarrow \mathbb{R}$, y supongamos que f es de clase C^2 en la bola abierta $B_r(x_0)$. Por el teorema de Taylor, si $\|h\| < r$ se tiene

$$f(x_0 + h) = f(x_0) + \frac{1}{2} D^2f(x_0) \cdot h + o(\|h\|^2),$$

donde hemos tenido en cuenta que $Df(x_0) = 0$ al ser x_0 un punto crítico de f . Dividiendo la identidad anterior por $\|h\|^2$ y teniendo en cuenta que $D^2f(x_0)$ es homogénea de grado 2 se obtiene

$$\boxed{\frac{f(x_0 + h) - f(x_0)}{\|h\|^2} = \frac{1}{2} D^2f(x_0) \cdot \left(\frac{h}{\|h\|}\right) + o(1), \quad 0 < \|h\| < r.} \quad (3.17)$$

i) Si x_0 es un mínimo relativo de f , existe $0 < \delta < r$ tal que

$$\|h\| < \delta < r \implies f(x_0 + h) - f(x_0) \geq 0.$$

Tomando $h = tv$ con $v \neq 0$ y $0 < |t| < \delta/\|v\|$, de modo que $0 < \|h\| < \delta < r$, y sustituyendo en la ec. (3.17) se tiene

$$\frac{f(x_0 + tv) - f(x_0)}{t^2\|v\|^2} = \frac{1}{2} D^2f(x_0) \cdot \left(\frac{v}{\|v\|}\right) + o(1) \geq 0.$$

Haciendo t tender a 0 se obtiene

$$D^2f(x_0) \cdot \left(\frac{v}{\|v\|}\right) = \frac{1}{\|v\|^2} D^2f(x_0) \cdot v \geq 0 \implies D^2f(x_0) \cdot v \geq 0,$$

y por tanto $D^2f(x_0)$ es semidefinida positiva.

ii) Supongamos a continuación que la forma cuadrática $D^2f(x_0)$ es definida positiva. En virtud de la ec. (3.15), existe un vector v_1 de norma 1 tal que

$$D^2f(x_0) \cdot v \geq D^2f(x_0) \cdot v_1 \equiv m > 0, \quad \forall v \in \mathbb{R}^n \text{ t.q. } \|v\| = 1,$$

ya que $D^2f(x_0) \cdot v_1 > 0$ al ser $D^2f(x_0)$ definida positiva. Dado que el término $o(1)$ tiende a 0 cuando h tiende a cero, existe $0 < \delta < r$ tal que

$$\|h\| < \delta \implies |o(1)| < \frac{m}{4}.$$

Aplicando de nuevo la ec. (3.17) se obtiene

$$0 < \|h\| < \delta \implies \frac{f(x_0 + h) - f(x_0)}{\|h\|^2} \geq \frac{m}{2} - \frac{m}{4} = \frac{m}{4} > 0 \implies f(x_0 + h) > f(x_0),$$

y por tanto f tiene un mínimo relativo en x_0 .

iii) Supongamos, por último, que $D^2f(x_0)$ es indefinida. En virtud del apartado i), f no tiene un mínimo relativo en x_0 . Tampoco tiene un máximo relativo en dicho punto, ya que si lo tuviera $-f$ tendría un mínimo relativo en x_0 , y por tanto $D^2(-f)(x_0) \cdot h = -D^2f(x_0) \cdot h$ sería semidefinida positiva, o lo que es lo mismo $D^2f(x_0)$ sería semidefinida negativa. En consecuencia, x_0 es un punto de silla de f . \square

Comentario. Evidentemente, cambiando “positiva” por “negativa” en la proposición anterior se obtiene un resultado análogo para *máximos* relativos de f .

- De la Proposición 3.10 se sigue que la naturaleza del punto crítico x_0 (es decir, si se trata de un máximo, un mínimo o un punto de silla) está determinada por el comportamiento de la forma cuadrática $D^2f(x_0)$ (es decir, del término de segundo orden en el desarrollo de Taylor de f centrado en el punto x_0) *excepto* cuando dicha forma cuadrática es *semidefinida*. Por ejemplo, la función $f(x, y) = x^2 - y^4$ del apartado iv) del Ejemplo 3.6 tiene un punto de silla en el origen, aunque $D^2f(0,0) \cdot h = 2h_1^2$ es semidefinida positiva. Por el contrario, la función $f(x, y) = x^2 + y^4$, que también cumple $D^2f(0,0) \cdot h = 2h_1^2$, obviamente tiene un mínimo (absoluto) en $(0,0)$.
- Por otra parte, la forma cuadrática $D^2f(x_0)$ solo puede ser semidefinida si es degenerada, es decir si el determinante $\det Hf(x)$, que se denomina **hessiano** de la función f en el punto x , se anula en x_0 . Esto motiva la siguiente definición:

Definición 3.11. Un punto crítico $x_0 \in \mathbb{R}^n$ de una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ se dice **no degenerado** si f es de clase C^2 en un entorno de x_0 y $\det Hf(x_0) \neq 0$.

Recordando que una forma cuadrática no degenerada solo puede ser definida positiva, definida negativa o indefinida y aplicando la Proposición 3.10 se obtiene inmediatamente el siguiente resultado fundamental para la clasificación de puntos críticos no degenerados:

Proposición 3.12 (clasificación de puntos críticos no degenerados). *Supongamos que la función escalar $f : \mathbb{R}^n \rightarrow \mathbb{R}$ tiene un punto crítico no degenerado en $x_0 \in \mathbb{R}^n$. Entonces se verifica exactamente una de las tres opciones siguientes:*

- i) Si $D^2f(x_0)$ es definida positiva, f tiene un mínimo relativo en x_0 .
- ii) Si $D^2f(x_0)$ es definida negativa, f tiene un máximo relativo en x_0 .
- iii) Si $D^2f(x_0)$ es indefinida, f tiene un punto de silla en x_0 .

Las proposiciones anteriores hacen que sea de gran interés el encontrar criterios lo más sencillos posibles para averiguar si una forma cuadrática es definida positiva, definida negativa o indefinida. Uno de los principales resultados de este tipo es el llamado *criterio de Sylvester*. Para formular este criterio, consideremos una forma cuadrática $Q(h)$ y sea $A \in M_{n \times n}(\mathbb{R})$ su matriz. Por definición, los **menores principales** de A son los n determinantes

$$d_k = \begin{vmatrix} a_{11} & \dots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \dots & a_{kk} \end{vmatrix}, \quad k = 1, \dots, n.$$

En particular, el último menor principal es el determinante de la matriz A .

Criterio de Sylvester. *Una forma cuadrática es definida positiva si y solo si todos los menores principales de su matriz son positivos.*

Evidentemente, Q es definida negativa si y solo si $-Q$ es definida positiva. Por el criterio de Sylvester, esto ocurrirá si y solo si los menores principales de la matriz $-A$ son todos positivos. Al ser

$$\begin{vmatrix} -a_{11} & \dots & -a_{1k} \\ \vdots & \ddots & \vdots \\ -a_{k1} & \dots & -a_{kk} \end{vmatrix} = (-1)^k d_k,$$

esta condición se cumplirá si y solo si d_k tiene el signo de $(-1)^k$, es decir si los menores principales de A tienen signo alternado y $d_1 \equiv a_{11} < 0$. En otras palabras:

Corolario 3.13. La forma cuadrática $Q(h)$ es definida negativa si y solo si $a_{11} < 0$ y los menores principales de su matriz tienen signo alternado.

En particular, como una forma cuadrática no degenerada solo puede ser definida positiva, definida negativa o indefinida, del criterio de Sylvester se sigue el siguiente corolario:

Proposición 3.14. Sea Q una forma cuadrática no degenerada. Entonces se verifica:

- i) Si todos los menores principales de la matriz de Q son positivos, Q es definida positiva.
- ii) Si los menores principales de la matriz A de Q tienen signo alternado y $a_{11} < 0$, Q es definida negativa.
- iii) En cualquier otro caso, Q es indefinida.

Aplicando este criterio a la forma cuadrática $D^2f(x_0)$ y utilizando la Proposición 3.12 se obtiene el siguiente resultado, que permite clasificar de forma exhaustiva los puntos críticos *no degenerados* de una función de clase C^2 en términos de los menores principales de la matriz hessiana $Hf(x_0)$:

Proposición 3.15 (criterio del hessiano). Supongamos que $f : \mathbb{R}^n \rightarrow \mathbb{R}$ tiene un punto crítico no degenerado en $x_0 \in \mathbb{R}^n$. Entonces se cumple:

- i) Si todos los menores principales de $Hf(x_0)$ son positivos, x_0 es un mínimo relativo.
- ii) Si $\frac{\partial^2 f}{\partial x_1^2}(x_0) < 0$ y los menores principales de $Hf(x_0)$ tienen signo alternado, x_0 es un máximo relativo.
- iii) En cualquier otro caso, x_0 es un punto de silla.

El criterio anterior adopta una forma particularmente sencilla para una función de dos variables $f : \mathbb{R}^2 \rightarrow \mathbb{R}$. En efecto, en este caso los menores principales de la matriz hessiana en un punto crítico (x_0, y_0) son

$$d_1 = f_{xx}(x_0, y_0), \quad d_2 = \det Hf(x_0, y_0) \equiv f_{xx}(x_0, y_0)f_{yy}(x_0, y_0) - f_{xy}^2(x_0, y_0),$$

con $d_2 \neq 0$ si el punto crítico es no degenerado. Además, en este caso

$$d_2 > 0 \implies f_{xx}(x_0, y_0) \neq 0.$$

Por tanto, para funciones escalares de dos variables la Proposición 3.15 se reduce a la siguiente:

Proposición 3.16. Sea (x_0, y_0) un punto crítico no degenerado de una función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$. Entonces se verifica:

- i) Si $\det Hf(x_0, y_0) > 0$ y $f_{xx}(x_0, y_0) > 0$, (x_0, y_0) es un mínimo relativo.
- ii) Si $\det Hf(x_0, y_0) > 0$ y $f_{xx}(x_0, y_0) < 0$, (x_0, y_0) es un máximo relativo.
- iii) Si $\det Hf(x_0, y_0) < 0$, (x_0, y_0) es un punto de silla.

Nótese que si $\det Hf(x_0, y_0) > 0$ entonces $f_{xx}(x_0, y_0)$ y $f_{yy}(x_0, y_0)$ tienen el mismo signo. Obsérvese también que en este caso si $\det Hf(x_0, y_0) = 0$, es decir si (x_0, y_0) es un punto crítico *degenerado*, la forma cuadrática $D^2f(x_0, y_0)$ es siempre *semidefinida* (¿por qué?), y por tanto la naturaleza del punto crítico (x_0, y_0) depende de los términos de orden superior en el desarrollo de Taylor de f centrado en dicho punto.

Ejemplo 3.17. Hallemos los extremos de la función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por

$$f(x, y) = x^3 - 3xy^2 - 2y^2 + x^4, \quad \forall (x, y) \in \mathbb{R}^2.$$

Calculemos, en primer lugar, los puntos críticos de f . Para ello, hay que resolver el sistema

$$\begin{cases} f_x(x, y) = 3x^2 - 3y^2 + 4x^3 = 0 \\ f_y(x, y) = -2y(3x + 2) = 0. \end{cases}$$

De la segunda ecuación se sigue que $y = 0$ o $x = -2/3$. Si $y = 0$, sustituyendo en la primera ecuación se obtiene $x = 0$ o $x = -3/4$. Por otra parte, si $x = -2/3$ la primera ecuación proporciona $y = \pm 2/9$. Por tanto, f tiene exactamente cuatro puntos críticos, a saber

$$\boxed{(0, 0), \left(-\frac{3}{4}, 0\right), \left(-\frac{2}{3}, -\frac{2}{9}\right), \left(-\frac{2}{3}, \frac{2}{9}\right)},$$

cuya naturaleza vamos a determinar a continuación.

En general, la matriz hessiana de f es igual a

$$Hf(x, y) = \begin{pmatrix} 6x + 12x^2 & -6y \\ -6y & -6x - 4 \end{pmatrix}.$$

En el origen

$$Hf(0, 0) = \begin{pmatrix} 0 & 0 \\ 0 & -4 \end{pmatrix},$$

de donde se sigue que $(0, 0)$ es un *punto crítico degenerado*, al que no se puede aplicar la Proposición 3.16. Más precisamente, la forma cuadrática

$$D^2f(0, 0) \cdot h = -4h_2^2$$

es semidefinida negativa, por lo que el origen solo puede ser un máximo relativo o un punto de silla en virtud de la Proposición 3.10. En este caso es fácil ver que el origen es un *punto de silla* de f , ya que $f(0, 0) = 0$ y (por ejemplo) $f(x, 0) = x^3(x + 1)$ tiene el signo de x si $|x| < 1$. En el segundo punto crítico

$$Hf\left(-\frac{3}{4}, 0\right) = \begin{pmatrix} \frac{9}{4} & 0 \\ 0 & \frac{1}{2} \end{pmatrix},$$

de donde se sigue que $(-3/4, 0)$ es un *mínimo relativo* de f . Finalmente, en los dos puntos críticos restantes

$$Hf\left(-\frac{2}{3}, \pm\frac{2}{9}\right) = \frac{4}{3} \begin{pmatrix} 1 & \mp 1 \\ \mp 1 & 0 \end{pmatrix}$$

tiene determinante negativo, por lo que ambos puntos críticos son *puntos de silla*.

Aunque f tiene un mínimo relativo en $(-3/4, 0)$, es fácil ver que f no tiene extremos absolutos en \mathbb{R}^2 . En efecto, es claro que f no puede tener ningún máximo absoluto, pues no tiene máximos relativos. De hecho, $f(x, 0) = x^3 + x^4$ tiende a $+\infty$ si $x \rightarrow \pm\infty$. Por otra parte, $f(0, y) = -2y^2$ tiende a $-\infty$ si $y \rightarrow \pm\infty$, por lo que f tampoco tiene ningún mínimo absoluto. Nótese que, para funciones diferenciables de una variable, la existencia de un mínimo local y tres puntos de inflexión implica que el mínimo local es de hecho un mínimo absoluto. La diferencia con el caso multidimensional es que en este último en un punto de silla hay direcciones a lo largo de las cuales f crece y otras en las que decrece.

Ejercicio 32. Hallar los valores máximo y mínimo de la función $f(x, y) = x^2 + y^2 - x - y$ en el disco cerrado $x^2 + y^2 \leq 1$.

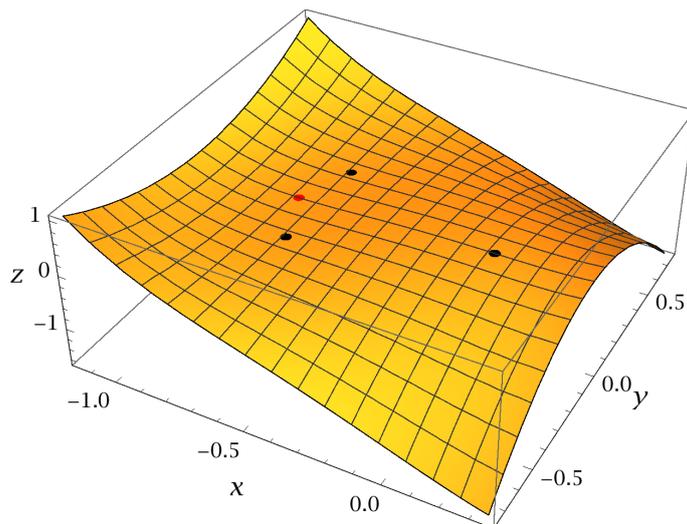


Figura 3.2: Gráfica de la función $f(x, y) = x^3 - 3xy^2 - 2y^2 + x^4$ (en rojo, el punto de la gráfica correspondiente al mínimo de la función en $(-3/4, 0)$, en negro los puntos de la gráfica correspondientes a los tres puntos de silla en el origen y en $(-2/3, \pm 2/9)$).

Solución. En primer lugar nótese que, al ser f continua en el disco cerrado (y, por tanto, compacto) $D \equiv \overline{B}_1(0)$, necesariamente ha de alcanzar un máximo y un mínimo en dicho conjunto. En segundo lugar, si (x_0, y_0) es un extremo absoluto de f en D , o bien (x_0, y_0) pertenece a la circunferencia unidad $x^2 + y^2 = 1$, o bien (x_0, y_0) pertenece a $\overset{\circ}{D} = B_1(0)$ y $\nabla f(x_0, y_0) = 0$, en virtud de la Proposición 3.8. Comencemos, por tanto, calculando los puntos en que se anula el gradiente de f :

$$\nabla f(x, y) = (2x - 1, 2y - 1) = (0, 0) \iff (x, y) = \left(\frac{1}{2}, \frac{1}{2}\right).$$

El punto $(1/2, 1/2)$ pertenece a $\overset{\circ}{D}$, y por tanto podría ser un extremo absoluto de f en D . Nótese que

$$Hf(1/2, 1/2) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

es definida positiva, y por tanto $(1/2, 1/2)$ es un mínimo relativo de f . Veamos a continuación cuáles son los extremos de f en la frontera de D , es decir en la circunferencia unidad $x^2 + y^2 = 1$. Para ello basta notar que estos puntos son de la forma $(x, y) = (\cos t, \sin t)$, con $0 \leq t < 2\pi$. La función

$$g(t) \equiv f(\cos t, \sin t) = 1 - \cos t - \sin t$$

tiene derivada

$$g'(t) = \sin t - \cos t = 0 \iff \tan t = 1 \iff t = \frac{\pi}{4} + k\pi, \quad k \in \mathbb{Z}.$$

Por tanto los únicos puntos críticos de g en el intervalo $[0, 2\pi)$ son $t = \pi/4$ y $t = 5\pi/4$. (Es fácil ver que el primero es un mínimo relativo y el segundo un máximo relativo de g , por ejemplo calculando $g''(t)$ en dichos puntos.) Además de en estos puntos, g podría tener un extremo en $[0, 2\pi)$ en $t = 0$ o 2π . Este no es el caso, ya que

$$g(\pi/4) = 1 - \sqrt{2} < g(0) = g(2\pi) = 0 < g(5\pi/4) = 1 + \sqrt{2}.$$

En definitiva, los extremos de f en D solo pueden estar entre los siguientes tres puntos:

$$\left(\frac{1}{2}, \frac{1}{2}\right), \quad (\cos(\pi/4), \sin(\pi/4)) = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right), \quad (\cos(5\pi/4), \sin(5\pi/4)) = -\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right).$$

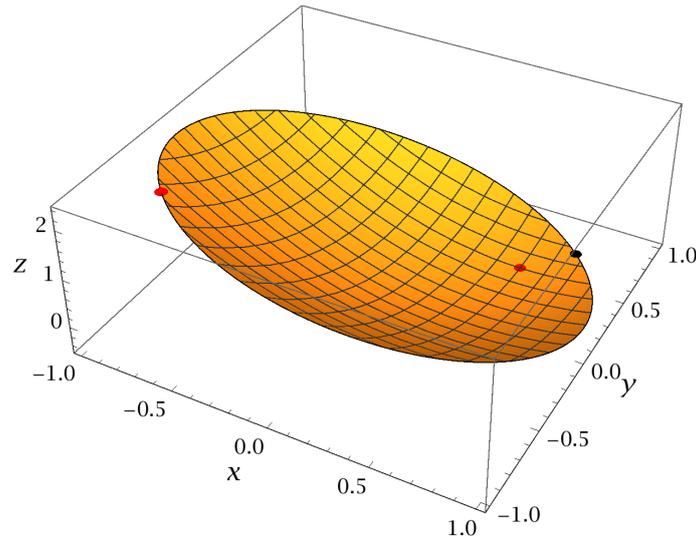


Figura 3.3: Gráfica de la función $f(x, y) = x^2 + y^2 - x - y$ en el disco cerrado $x^2 + y^2 \leq 1$. Se han indicado en rojo los puntos de la gráfica correspondientes a los extremos de la función en $(1/2, 1/2)$ y $-(\sqrt{2}/2, \sqrt{2}/2)$, y en negro el punto de la gráfica correspondiente al punto $(\sqrt{2}/2, \sqrt{2}/2)$ (mínimo de f en la circunferencia unidad).

Como

$$f\left(\frac{1}{2}, \frac{1}{2}\right) = -\frac{1}{2}, \quad f\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right) = 1 - \sqrt{2}, \quad f\left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right) = 1 + \sqrt{2}$$

y $-1/2 < 1 - \sqrt{2}$, el mínimo de f en D se alcanza en el punto $(1/2, 1/2)$, y su valor es $-1/2$, mientras que su máximo se alcanza en el punto $-(\sqrt{2}/2, \sqrt{2}/2)$ y su valor es $1 + \sqrt{2}$. Nótese que f no tiene un máximo absoluto en \mathbb{R}^2 , y por tanto no está acotada superiormente. En efecto, si f tuviera un máximo absoluto en \mathbb{R}^2 el gradiente de f se anularía en dicho punto, y acabamos de ver que ∇f solo se anula en el mínimo relativo $(1/2, 1/2)$. También es fácil comprobar que $(1/2, 1/2)$ es de hecho un mínimo absoluto de f en \mathbb{R}^2 , ya que

$$f(x, y) = \left(x - \frac{1}{2}\right)^2 + \left(y - \frac{1}{2}\right)^2 - \frac{1}{2} \geq -\frac{1}{2} = f\left(\frac{1}{2}, \frac{1}{2}\right).$$

3.5 Teorema de la función inversa

Uno de los problemas más importantes del Análisis Matemático consiste en determinar si una función $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ es invertible en un cierto abierto U , es decir si existe una función $f^{-1} : f(U) \rightarrow U$ tal que

$$\boxed{f^{-1} \circ f = I_U, \quad f \circ f^{-1} = I_{f(U)}}, \quad (3.18)$$

donde I_A denota la *función identidad* en un conjunto A ($I_A(x) = x$, para todo $x \in A$). (Recuérdese que $f : U \rightarrow f(U)$ es automáticamente suprayectiva, y por tanto es invertible si y solo si es inyectiva.) En la mayor parte de las aplicaciones la función f es (como mínimo) diferenciable, e interesa hallar bajo qué condiciones la función inversa f^{-1} es también diferenciable. Nótese que en tal caso f^{-1} es automáticamente *continua*, y por tanto $f(U) = (f^{-1})^{-1}(U)$ es *abierto*.

Definición 3.18. Un **difeomorfismo** (de clase C^p) entre dos abiertos $U, V \subset \mathbb{R}^n$ es una aplicación *biyectiva* $f : U \rightarrow V$ tal que tanto f como $f^{-1} : V \rightarrow U$ son diferenciables (de clase C^p). Dos abiertos $U, V \subset \mathbb{R}^n$ son **difeomorfos** si existe un difeomorfismo $f : U \rightarrow V$.

Comentario. Como una aplicación diferenciable es automáticamente continua, todo difeomorfismo es en particular un homeomorfismo.

Ejemplo 3.19. Una aplicación *lineal* $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ es un difeomorfismo (de clase C^∞) si y solo si $\det A \neq 0$. En efecto, A es invertible si y solo si $\det A \neq 0$, y en tal caso tanto A como A^{-1} son C^∞ (al ser lineales). Consideremos, en particular, la aplicación lineal $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definida por $A(x, y) = (ax, by)$, que es un difeomorfismo si $ab \neq 0$. Si, por ejemplo, U es el disco abierto $x^2 + y^2 < 1$, $A(U)$ es el conjunto $(x/a)^2 + (y/b)^2 < 1$, es decir el interior de la elipse de centro el origen y semiejes $|a|$ y $|b|$. En particular, estos dos conjuntos son difeomorfos.

En la práctica, suele ser suficiente para las aplicaciones que la función f sea invertible en el entorno de un cierto punto $x_0 \in \mathbb{R}^n$, lo que da lugar al concepto de invertibilidad local:

Definición 3.20. Una función $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ es **localmente invertible** en un punto $x_0 \in \mathbb{R}^n$ si es invertible en un entorno de x_0 . Análogamente, f es localmente invertible en un abierto U si es localmente invertible en todos los puntos de U . Diremos que $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ es un **difeomorfismo local** (de clase C^p) en $x_0 \in \mathbb{R}^n$ si es un difeomorfismo (de clase C^p) en un entorno U de x_0 .

Es importante observar que una función puede ser localmente invertible en un abierto U sin ser invertible en dicho conjunto (o ser un difeomorfismo local en cada punto de U sin ser un difeomorfismo), como muestra el siguiente ejemplo:

Ejemplo 3.21. Consideremos la aplicación $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definida por

$$f(x, y) = e^x (\cos y, \sin y).$$

Esta aplicación *no* es biyectiva en \mathbb{R}^2 , ya que es *periódica* (y, por tanto, *no* inyectiva):

$$f(x, y + 2k\pi) = f(x, y), \quad \forall k \in \mathbb{Z}.$$

Sin embargo, f es inyectiva en cualquier conjunto en que la coordenada y esté restringida a un intervalo semiabierto de longitud 2π , ya que es fácil comprobar que

$$e^{x_1} (\cos y_1, \sin y_1) = e^{x_2} (\cos y_2, \sin y_2) \iff x_1 = x_2, \quad y_1 - y_2 = 2k\pi \quad \text{con } k \in \mathbb{Z}.$$

Por ejemplo, f es inyectiva en la banda abierta $U = \mathbb{R} \times (0, 2\pi)$. Es fácil ver que

$$V \equiv f(U) = \mathbb{R}^2 - \{(u, 0) \mid u \geq 0\},$$

que también es un abierto (ya que la semirrecta $\{(x, 0) \mid x \geq 0\}$ es cerrada). Es también inmediato comprobar que la inversa de f en U está dada por

$$f^{-1}(u, v) = \left(\frac{1}{2} \log(u^2 + v^2), \theta(u, v) \right), \quad (3.19)$$

donde $\theta(u, v)$ es el ángulo polar θ del punto (u, v) (cf. la ec. (2.55) con $x = u$, $y = v$). En efecto, basta tener en cuenta que $f^{-1}(u, v) = (x, y)$ si y solo si

$$(u, v) = f(x, y) = e^x (\cos y, \sin y),$$

de donde se sigue que $u^2 + v^2 = e^{2x}$ e $y = \theta(u, v)$ (ya que $0 < y < 2\pi$, al ser $(x, y) \in U$). De estas fórmulas se sigue que $f : U \rightarrow V$ es un difeomorfismo de clase C^∞ . Por otra parte, es fácil ver que aunque f no sea biyectiva es *localmente invertible en todo punto* (x_0, y_0) , ya que f es claramente inyectiva en la banda abierta $\mathbb{R} \times (y_0 - \pi, y_0 + \pi)$. La inversa de f en esta banda está dada por

$$f^{-1}(u, v) = \left(\frac{1}{2} \log(u^2 + v^2), \theta_{(y_0 - \pi, y_0 + \pi)}(u, v) \right),$$

donde $\theta_{(y_0 - \pi, y_0 + \pi)}(u, v)$ es el ángulo polar de (u, v) en el intervalo $(y_0 - \pi, y_0 + \pi)$, y es por tanto de clase C^∞ en $\mathbb{R} \times (y_0 - \pi, y_0 + \pi)$. Luego en este caso f es un difeomorfismo local (de clase C^∞) en cada punto.

Comentario. Si en el ejemplo anterior utilizamos las variables complejas $z = x + iy$ y $w = u + iv$ entonces

$$w = e^x \cos y + i e^x \sin y = e^x (\cos y + i \sin y) = e^x e^{iy} = e^{x+iy} = e^z,$$

y la función inversa es

$$z = \frac{1}{2} \log(u^2 + v^2) + i \theta(u, v) = \log |w| + i \arg_{[0, 2\pi)} w = \log_{[0, 2\pi)} z,$$

donde $\arg_{[0, 2\pi)}$ y $\log_{[0, 2\pi)}$ denotan respectivamente la determinación $[0, 2\pi)$ del argumento y del logaritmo complejo. \square

Nuestro objetivo es encontrar condiciones sencillas que nos permitan determinar si una función $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ es un difeomorfismo local en un punto. Para ello, nótese en primer lugar que si f es un difeomorfismo en un abierto U , es decir si f es biyectiva y tanto f como su inversa son diferenciables (respectivamente en U y $f(U)$), derivando las relaciones (3.18) se obtiene

$$D(f^{-1})(f(x)) \cdot Df(x) = I, \quad \forall x \in U.$$

Por tanto $Df(x)$ es *invertible* en todo punto x de U , con inversa

$$\boxed{Df(x)^{-1} = D(f^{-1})(f(x))}. \quad (3.20)$$

De esta ecuación se deduce (haciendo $f(x) = y$) que

$$\boxed{D(f^{-1})(y) = Df(f^{-1}(y))^{-1}}, \quad \forall y \in f(U). \quad (3.21)$$

También se sigue de la ec. (3.20) que el determinante de $Df(x)$ no se anula en U . En otras palabras,

Una condición *necesaria* para que una aplicación diferenciable $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ sea un difeomorfismo en un abierto U es que

$$\det Df(x) \neq 0, \quad \forall x \in U. \quad (3.22)$$

Notación. El determinante $\det Df(x)$ se suele denominar **jacobiano** de la función f en el punto x .

¿Es la condición (3.22) también *suficiente*? En general *no*, como muestra el Ejemplo 3.21. En efecto, en este caso

$$Df(x, y) = \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix} \Rightarrow \det Df(x, y) = e^{2x} \neq 0, \quad \forall (x, y) \in \mathbb{R}^2,$$

y sin embargo f *no* es un difeomorfismo en \mathbb{R}^2 (no es ni siquiera inyectiva). De todas formas, hemos visto en este caso que f es un difeomorfismo local en cualquier punto de \mathbb{R}^2 . El teorema de la función inversa, cuyo enunciado formularemos a continuación, nos garantiza que esta situación es completamente general:

Teorema de la función inversa. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ una función de clase C^p (con $p \geq 1$) en un entorno de un punto $x_0 \in \mathbb{R}^n$, y supongamos que $\det Df(x_0) \neq 0$. Entonces f es un difeomorfismo local de clase C^p en x_0 , es decir existe un entorno U de x_0 tal que $f(U)$ es abierto, $f : U \rightarrow f(U)$ es biyectiva y $f^{-1} : f(U) \rightarrow U$ es de clase C^p en $f(U)$. Además, la derivada de f^{-1} está dada por

$$D(f^{-1})(y) = Df(f^{-1}(y))^{-1}, \quad \forall y \in f(U). \quad (3.23)$$

Idea de la demostración: la demostración de este resultado está basada en la siguiente observación. Al ser f de clase C^p en un entorno de x_0 en particular es diferenciable en dicho punto, y por tanto

$$f(x) = f(x_0) + Df(x_0) \cdot (x - x_0) + o(\|x - x_0\|)$$

para x suficientemente próximo a x_0 . Por consiguiente la ecuación $y = f(x)$ se puede aproximar cerca de x_0 por el sistema lineal (en x)

$$y = f(x_0) + Df(x_0) \cdot (x - x_0).$$

Si $Df(x_0)$ es *invertible*, este sistema tiene la solución única

$$x = x_0 + Df(x_0)^{-1} \cdot (y - f(x_0)).$$

Esto que *sugiere* que para x en un entorno de x_0 la ecuación $y = f(x)$ tiene solución única dada aproximadamente por el miembro derecho de la ecuación anterior. En otras palabras, f debería ser invertible en un entorno de x_0 , siendo

$$f^{-1}(y) \simeq x_0 + Df(x_0)^{-1} \cdot (y - f(x_0)).$$

Nótese, por último, que la fórmula para la derivada de la función inversa f^{-1} es la ec. (3.21) deducida anteriormente \square

Corolario 3.22. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ una función biyectiva de clase C^p (con $p \geq 1$) en un abierto $U \subset \mathbb{R}^n$, y supongamos que $\det Df(x) \neq 0$ para todo $x \in U$. Entonces $f : U \rightarrow f(U)$ es un difeomorfismo de clase C^p .

Demostración. Por el teorema de la función inversa $f^{-1} : f(U) \rightarrow U$ (que existe al ser f invertible por hipótesis) es de clase C^p en un entorno de cada punto de $f(U)$, y por tanto en $f(U)$. \square

Ejercicio 33. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ de clase C^1 en un abierto U , y supongamos que $\det Df(x) \neq 0$ para todo $x \in U$. Entonces f transforma subconjuntos abiertos de U en conjuntos abiertos.

Solución. En efecto, sea $A \subset U$ abierto y sea $x_0 \in A$. Por el teorema de la función inversa aplicado al punto x_0 y al abierto A , existe un entorno $W \subset A$ de x_0 tal que $f : W \rightarrow f(W)$ es biyectiva y $f(W) \subset f(A)$ es abierto. Por tanto $f(W)$ es un entorno de $f(x_0)$ contenido en $f(A)$. Como $f(x_0)$ es un punto arbitrario de $f(A)$ (al ser x_0 un punto arbitrario de A), esto demuestra que $f(A)$ contiene un entorno de cualquiera de sus puntos, y es por tanto abierto.

Ejemplo 3.23. Sea $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ la función definida por

$$f(x, y, z) = (e^{2y} + e^{2z}, e^{2x} - e^{2z}, x - y).$$

La matriz jacobiana de f está dada por

$$Df(x, y, z) = \begin{pmatrix} 0 & 2e^{2y} & 2e^{2z} \\ 2e^{2x} & 0 & -2e^{2z} \\ 1 & -1 & 0 \end{pmatrix},$$

y su determinante es igual a

$$\det Df(x, y, z) = -4e^{2(y+z)} - 4e^{2(x+z)}.$$

Como $\det Df(x, y, z)$ es negativo para todo $(x, y, z) \in \mathbb{R}^3$, la función f es un difeomorfismo local (de clase C^∞) en todo punto de \mathbb{R}^3 . ¿Es $f : \mathbb{R}^3 \rightarrow f(\mathbb{R}^3)$ un difeomorfismo global? Para ello basta con que f sea inyectiva en \mathbb{R}^3 , ya que entonces es biyectiva considerada como función de

\mathbb{R}^3 en $f(\mathbb{R}^3)$ y podemos aplicar el Corolario 3.22. Para ver si f es inyectiva hay que estudiar si la ecuación $f(x_1, y_1, z_1) = f(x_2, y_2, z_2)$, es decir el sistema

$$e^{2y_1} + e^{2z_1} = e^{2y_2} + e^{2z_2}, \quad e^{2x_1} - e^{2z_1} = e^{2x_2} - e^{2z_2}, \quad x_1 - y_1 = x_2 - y_2,$$

tiene como única solución $(x_1, y_1, z_1) = (x_2, y_2, z_2)$. Sumando las dos primeras ecuaciones obtenemos

$$\begin{aligned} e^{2x_1} + e^{2y_1} = e^{2x_2} + e^{2y_2} &\implies e^{2x_1}(1 + e^{2(y_1-x_1)}) = e^{2x_2}(1 + e^{2(y_2-x_2)}) \\ &\implies e^{2x_1} = e^{2x_2} \implies x_1 = x_2, \end{aligned}$$

al ser $x_1 - y_1 = x_2 - y_2$. De esta última ecuación y de la igualdad $x_1 = x_2$ se sigue que $y_1 = y_2$. El sistema se reduce entonces a la ecuación $e^{2z_1} = e^{2z_2}$, y por tanto también se cumple que $z_1 = z_2$. Esto prueba que f es un difeomorfismo (de clase C^∞) de \mathbb{R}^3 en $f(\mathbb{R}^3)$.

Ejercicio 34. Probar que $f(\mathbb{R}^3)$ es el abierto

$$\{(u, v, w) \in \mathbb{R}^3 \mid u > 0, u + v > 0, e^{2w} > v/u\},$$

y que

$$f^{-1}(u, v, w) = \frac{1}{2} \left(\log \left(\frac{u+v}{1+e^{-2w}} \right), \log \left(\frac{u+v}{1+e^{2w}} \right), \log \left(\frac{ue^{2w}-v}{1+e^{2w}} \right) \right).$$

3.5.1 Sistemas de coordenadas

Si $f: U \subset \mathbb{R}^n \rightarrow V \equiv f(U) \subset \mathbb{R}^n$ es un difeomorfismo, podemos considerar a las n componentes (u_1, \dots, u_n) de un punto de U como las **coordenadas curvilíneas** del punto imagen $x = f(u) \in V$ en el sistema de coordenadas definido por la función f . Dado un punto $x_0 = f(u_0)$ con coordenadas curvilíneas $u_0 \equiv (u_{01}, \dots, u_{0n})$, la i -ésima **línea coordenada** que pasa por dicho punto es la curva

$$y_i(t) = f(u_{01}, \dots, u_{0,i-1}, u_{0i} + t, u_{0,i+1}, \dots, u_{0n})$$

que se obtiene “congelando” todas las coordenadas curvilíneas del punto x_0 menos la i -ésima. Un vector unitario tangente a dicha curva en $x_0 = y_i(0)$ está dado por

$$\frac{y_i'(0)}{\|y_i'(0)\|} \equiv e_i(u_0) = \frac{1}{h_i(u_0)} \frac{\partial f}{\partial u_i}(u_0), \quad 1 \leq i \leq n,$$

donde hemos definido

$$h_i(u_0) \equiv \left\| \frac{\partial f}{\partial u_i}(u_0) \right\|.$$

Nótese que $h_i(u_0) \neq 0$ para todo $i = 1, \dots, n$. En efecto, los n vectores $\frac{\partial f}{\partial u_i}(u_0)$ son linealmente independientes al ser las columnas de la matriz $Df(u_0)$, que por hipótesis es no singular por ser f un difeomorfismo. El vector $e_i(u_0)$ se denomina **i -ésimo vector unitario coordenado** en el sistema de coordenadas curvilíneas (u_1, \dots, u_n) definido por la función f . Los n vectores unitarios coordenados $\{e_1(u_0), \dots, e_n(u_0)\}$ son linealmente independientes y forman por tanto una base de \mathbb{R}^n . En efecto, si $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ se tiene

$$\begin{aligned} \lambda_1 e_1(u_0) + \dots + \lambda_n e_n(u_0) = 0 &\iff \frac{\lambda_1}{h_1(u_0)} \frac{\partial f}{\partial u_1}(u_0) + \dots + \frac{\lambda_n}{h_n(u_0)} \frac{\partial f}{\partial u_n}(u_0) = 0 \\ \implies \frac{\lambda_1}{h_1(u_0)} = \dots = \frac{\lambda_n}{h_n(u_0)} = 0 &\implies \lambda_1 = \dots = \lambda_n = 0, \end{aligned}$$

al ser los vectores $\frac{\partial f}{\partial u_i}(u_0)$ ($1 \leq i \leq n$) linealmente independientes y $h_i(u_0) \neq 0$ para todo i . Nótese que la base $\{e_1(u_0), \dots, e_n(u_0)\}$ formada por los n vectores unitarios coordenados varía de punto a punto, ya que los vectores $e_i(u_0)$ dependen de u_0 .

Definición 3.24. El sistema de coordenadas curvilíneas definido por el difeomorfismo f es **ortogonal** si sus vectores unitarios coordenados son mutuamente ortogonales, es decir si

$$e_i(u) \cdot e_j(u) = 0, \quad \forall u \in U, \quad 1 \leq i \neq j \leq n,$$

o equivalentemente

$$\frac{\partial f}{\partial u_i}(u) \cdot \frac{\partial f}{\partial u_j}(u) = 0, \quad \forall u \in U, \quad 1 \leq i \neq j \leq n.$$

En otras palabras, el sistema de coordenadas definido por f es ortogonal si las columnas de Df son ortogonales entre sí.

Ejemplo 3.25 (coordenadas esféricas). En este caso $u \equiv (r, \theta, \varphi)$ y

$$f(r, \theta, \varphi) = r(\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta),$$

inyectiva si

$$r > 0, \quad 0 < \theta < \pi, \quad 0 \leq \varphi < 2\pi.$$

Nótese que en este caso las líneas coordenadas son respectivamente semirrectas por el origen, semicircunferencias centradas en el origen situadas en un plano vertical (*meridianos*) y circunferencias situadas en un plano horizontal y centradas en un punto del eje z (*paralelos*). La matriz jacobiana de f es

$$Df(r, \theta, \varphi) = \begin{pmatrix} \sin \theta \cos \varphi & r \cos \theta \cos \varphi & -r \sin \theta \sin \varphi \\ \sin \theta \sin \varphi & r \cos \theta \sin \varphi & r \sin \theta \cos \varphi \\ \cos \theta & -r \sin \theta & 0 \end{pmatrix},$$

y su determinante es igual a

$$\det Df(r, \theta, \varphi) = \cos \theta \cdot r^2 \sin \theta \cos \theta + r \sin \theta \cdot r \sin^2 \theta = r^2 \sin \theta.$$

Como $\det Df(r, \theta, \varphi) \neq 0$ en el abierto U definido por

$$r > 0, \quad 0 < \theta < \pi, \quad 0 < \varphi < 2\pi,$$

en virtud del Corolario 3.22 $f : U \rightarrow V \equiv f(U)$ es un difeomorfismo, y por tanto define un sistema de coordenadas curvilíneas en el abierto

$$V = \mathbb{R}^3 - \{(x, y, z) \in \mathbb{R}^3 \mid y = 0, x \geq 0\}.$$

En este caso las funciones $h_i(r, \theta, \varphi)$ están dadas por

$$h_r(r, \theta, \varphi) = 1, \quad h_\theta(r, \theta, \varphi) = r, \quad h_\varphi(r, \theta, \varphi) = r \sin \theta,$$

y por tanto los vectores unitarios coordenados son

$$e_r = \frac{\partial f}{\partial r} = (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta) = \frac{\mathbf{r}}{r}, \quad e_\theta = \frac{1}{r} \frac{\partial f}{\partial \theta} = (\cos \theta \cos \varphi, \cos \theta \sin \varphi, -\sin \theta),$$

$$e_\varphi = \frac{1}{r \sin \theta} \frac{\partial f}{\partial \varphi} = (-\sin \varphi, \cos \varphi, 0).$$

Es inmediato comprobar que estos vectores son mutuamente ortogonales, y por tanto las coordenadas esféricas constituyen un sistema de coordenadas ortogonales.

Ejemplo 3.26 (coordenadas cilíndricas). Ahora $u \equiv (\rho, \varphi, z)$, siendo

$$f(\rho, \varphi, z) = (\rho \cos \varphi, \rho \sin \varphi, z)$$

inyectiva para

$$\rho > 0, \quad 0 \leq \varphi < 2\pi, \quad z \in \mathbb{R}.$$

Las líneas coordenadas son respectivamente semirrectas horizontales que pasan por el eje z , circunferencias en un plano horizontal centradas en un punto del eje z y rectas verticales. La matriz jacobiana de f está dada por

$$Df(\rho, \varphi, z) = \begin{pmatrix} \cos \varphi & -\rho \sin \varphi & 0 \\ \sin \varphi & \rho \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

y por tanto

$$\det Df(\rho, \varphi, z) = \rho.$$

Al ser este determinante no nulo en el abierto U de ecuación

$$\rho > 0, \quad 0 < \varphi < 2\pi, \quad z \in \mathbb{R},$$

la función f define un sistema de coordenadas curvilíneas en el abierto

$$f(U) = \mathbb{R}^3 - \{(x, y, z) \in \mathbb{R}^3 \mid y = 0, x \geq 0\}.$$

En este caso las funciones $h_i(\rho, \varphi, z)$ están dadas por

$$h_\rho(\rho, \varphi, z) = 1, \quad h_\varphi(\rho, \varphi, z) = \rho, \quad h_z(\rho, \varphi, z) = 1,$$

y por tanto los vectores unitarios coordenados están dados por

$$e_\rho = \frac{\partial f}{\partial \rho} = (\cos \varphi, \sin \varphi, 0), \quad e_\varphi = \frac{1}{\rho} \frac{\partial f}{\partial \varphi} = (-\sin \varphi, \cos \varphi, 0), \quad e_z = \frac{\partial f}{\partial z} = (0, 0, 1).$$

De nuevo, se comprueba fácilmente que estos tres vectores son mutuamente ortogonales, y por tanto las coordenadas cilíndricas son un sistema de coordenadas ortogonales.

3.6 Teorema de la función implícita

Es muy frecuente en Física que una función $y = f(x)$ esté definida *implícitamente*, es decir mediante una relación de la forma $\Phi(x, y) = 0$. Más precisamente, la relación anterior define en un abierto $V \subset \mathbb{R}^2$ a y como una función $f : U \rightarrow \mathbb{R}$ (donde $U \subset \mathbb{R}$ es abierto) si

$$(x, y) \in V, \quad \Phi(x, y) = 0 \quad \Leftrightarrow \quad x \in U, \quad y = f(x).$$

En otras palabras,

$$\{(x, y) \in V \mid \Phi(x, y) = 0\} = \{(x, f(x)) \mid x \in U\}.$$

Como $(x, f(x))$ es un punto cualquiera de la *gráfica* de la función $f : U \rightarrow \mathbb{R}$, la ecuación anterior significa que la gráfica de dicha función coincide con la curva $\Phi(x, y) = 0$. Esto implica que cualquier recta vertical $x = x_0$, con $x_0 \in U$, solo puede cortar a dicha curva en un punto.

Ejemplo 3.27. Consideremos la relación

$$\Phi(x, y) \equiv x^2 + y^2 - 1 = 0,$$

que es la ecuación de una circunferencia centrada en el origen y de radio unidad. Esta relación define *dos* funciones *continuas* $f_{\pm} : [-1, 1] \rightarrow \mathbb{R}$, dadas por

$$f_{\pm}(x) = \pm\sqrt{1 - x^2}.$$

Nótese que, aunque la función Φ es de clase C^{∞} , las dos funciones f_{\pm} definidas por la relación $\Phi(x, y) = 0$ *no* son derivables en los puntos ± 1 , ya que en dichos sus gráficas tienen tangente vertical. Sin embargo, si $(x_0, y_0) \neq (\pm 1, 0)$ es un punto de la circunferencia $x^2 + y^2 = 1$ hay un entorno V de dicho punto en que la relación $\Phi(x, y) = 0$ define una *única* función $f(x)$ de clase C^{∞} . En efecto, si $y_0 > 0$ y tomamos $V = \mathbb{R} \times (0, \infty)$ (es decir, el semiplano superior) entonces

$$x^2 + y^2 = 1, (x, y) \in \mathbb{R} \times (0, \infty) \iff x \in (-1, 1), y = \sqrt{1 - x^2},$$

siendo $f(x) = \sqrt{1 - x^2}$ de clase C^{∞} en $(-1, 1)$. Del mismo modo, si $y_0 < 0$ entonces

$$x^2 + y^2 = 1, (x, y) \in \mathbb{R} \times (-\infty, 0) \iff x \in (-1, 1), y = -\sqrt{1 - x^2}.$$

No ocurre lo mismo en los puntos $(\pm 1, 0)$, ya que en cualquier entorno de uno de dichos puntos la circunferencia $x^2 + y^2 = 1$ *no* es una gráfica. (Obsérvese, sin embargo, que en un entorno de $(\pm 1, 0)$ la relación $\Phi(x, y) = 0$ define a x como función de y .)

Nótese, por último, que los puntos $(\pm 1, 0)$ son los puntos de la circunferencia $\Phi(x, y) = 0$ en que $\frac{\partial \Phi}{\partial y} = 0$. Este hecho también es fácil de entender geoméricamente, ya que los puntos $(\pm 1, 0)$ son aquellos en que la tangente a la circunferencia $\Phi(x, y) = 0$ es *vertical*, y por tanto $\nabla \Phi(x, y)$, que es *normal* a dicha curva, ha de ser *horizontal*. \square

El ejemplo anterior sugiere que si la función $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}$ es de clase C^1 entonces la relación $\Phi(x, y) = 0$ define una única función $y = f(x)$ de clase C^1 en un entorno de cada punto (x_0, y_0) tal que $\frac{\partial \Phi}{\partial y}(x_0, y_0) \neq 0$. Más generalmente, consideremos m relaciones implícitas en $m + n$ variables de la forma

$$\Phi_i(x_1, \dots, x_n, y_1, \dots, y_m) = 0, \quad 1 \leq i \leq m, \quad (3.24)$$

que podemos escribir abreviadamente en la forma

$$\boxed{\Phi(x, y) = 0}, \quad (3.25)$$

siendo

$$\boxed{x \equiv (x_1, \dots, x_n) \in \mathbb{R}^n, \quad y \equiv (y_1, \dots, y_m) \in \mathbb{R}^m, \quad \Phi \equiv (\Phi_1, \dots, \Phi_m) : \mathbb{R}^{m+n} \rightarrow \mathbb{R}^m}.$$

Definición 3.28. Diremos que las relaciones (3.24) —o, equivalentemente, (3.25)—, definen implícitamente en el abierto $V \subset \mathbb{R}^{m+n}$ a y como una función $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, donde U es abierto, si

$$\{(x, y) \in V \mid \Phi(x, y) = 0\} = \{(x, f(x)) \mid x \in U\}.$$

En otras palabras, el conjunto de puntos $(x, y) \in V$ que satisfacen la relación $\Phi(x, y) = 0$ coincide exactamente con la gráfica de la función $f: U \rightarrow \mathbb{R}^m$. Nótese, en particular, que si esto ocurre se tiene

$$\boxed{\Phi(x, f(x)) = 0, \quad \forall x \in U}.$$

Veamos a continuación bajo qué condiciones la relación $\Phi(x, y) = 0$ define a $y \in \mathbb{R}^m$ como función de las variables $x \in \mathbb{R}^n$ en un entorno de un cierto punto $(x_0, y_0) \in \mathbb{R}^{m+n}$ que satisface dicha relación, es decir tal que $\Phi(x_0, y_0) = 0$. Si Φ es diferenciable en un entorno de $(x_0, y_0) \in \mathbb{R}^{m+n}$, la matriz jacobiana $D\Phi(x_0, y_0)$ es una matriz $m \times (m+n)$ que puede escribirse como

$$D\Phi(x_0, y_0) = \left(\frac{\partial\Phi}{\partial x}(x_0, y_0) \quad \frac{\partial\Phi}{\partial y}(x_0, y_0) \right),$$

siendo respectivamente $\frac{\partial\Phi}{\partial x}(x_0, y_0)$ y $\frac{\partial\Phi}{\partial y}(x_0, y_0)$ las matrices $m \times n$ y $m \times m$ dadas por

$$\frac{\partial\Phi}{\partial x}(x_0, y_0) = \begin{pmatrix} \frac{\partial\Phi_1}{\partial x_1}(x_0, y_0) & \cdots & \frac{\partial\Phi_1}{\partial x_n}(x_0, y_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial\Phi_m}{\partial x_1}(x_0, y_0) & \cdots & \frac{\partial\Phi_m}{\partial x_n}(x_0, y_0) \end{pmatrix} \in M_{m \times n}(\mathbb{R}),$$

$$\frac{\partial\Phi}{\partial y}(x_0, y_0) = \begin{pmatrix} \frac{\partial\Phi_1}{\partial y_1}(x_0, y_0) & \cdots & \frac{\partial\Phi_1}{\partial y_m}(x_0, y_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial\Phi_m}{\partial y_1}(x_0, y_0) & \cdots & \frac{\partial\Phi_m}{\partial y_m}(x_0, y_0) \end{pmatrix} \in M_{m \times m}(\mathbb{R}).$$

Al ser $\Phi(x_0, y_0) = 0$, en un entorno del punto (x_0, y_0) la relación $\Phi(x, y) = 0$ puede escribirse en la forma

$$\Phi(x, y) = \frac{\partial\Phi}{\partial x}(x_0, y_0) \cdot (x - x_0) + \frac{\partial\Phi}{\partial y}(x_0, y_0) \cdot (y - y_0) + o((\|x - x_0\|^2 + \|y - y_0\|^2)^{1/2}) = 0,$$

que a su vez se puede aproximar cerca del punto (x_0, y_0) por el sistema *lineal* (en y)

$$\frac{\partial\Phi}{\partial x}(x_0, y_0) \cdot (x - x_0) + \frac{\partial\Phi}{\partial y}(x_0, y_0) \cdot (y - y_0) = 0. \quad (3.26)$$

Si la matriz cuadrada $\frac{\partial\Phi}{\partial y}(x_0, y_0) \in M_{m \times m}(\mathbb{R})$ es *invertible*, es decir si

$$\boxed{\det \frac{\partial\Phi}{\partial y}(x_0, y_0) \neq 0}, \quad (3.27)$$

el sistema (3.26) tiene la solución única

$$y = y_0 - \left[\frac{\partial\Phi}{\partial y}(x_0, y_0) \right]^{-1} \cdot \frac{\partial\Phi}{\partial x}(x_0, y_0) \cdot (x - x_0).$$

Esta discusión *sugiere* que si se verifica la condición (3.27) entonces la relación $\Phi(x, y) = 0$ define en un entorno del punto (x_0, y_0) una única función $y = f(x)$, dada aproximadamente por

$$f(x) \simeq y_0 - \left[\frac{\partial\Phi}{\partial y}(x_0, y_0) \right]^{-1} \cdot \frac{\partial\Phi}{\partial x}(x_0, y_0) \cdot (x - x_0).$$

En particular, f debería ser diferenciable en x_0 , con derivada

$$Df(x_0) = - \left[\frac{\partial\Phi}{\partial y}(x_0, y_0) \right]^{-1} \cdot \frac{\partial\Phi}{\partial x}(x_0, y_0). \quad (3.28)$$

Las afirmaciones anteriores son esencialmente correctas. Más precisamente:

Teorema de la función implícita. Sea $\Phi : \mathbb{R}^{m+n} \rightarrow \mathbb{R}^m$ una función de clase C^p (con $p \geq 1$) en un entorno de un punto $(x_0, y_0) \in \mathbb{R}^{m+n}$ tal que $\Phi(x_0, y_0) = 0$. Si

$$\det \frac{\partial \Phi}{\partial y}(x_0, y_0) \neq 0,$$

entonces la relación $\Phi(x, y) = 0$ define implícitamente en un entorno de (x_0, y_0) una función $y = f(x)$, donde $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, U es un entorno de x_0 y f es de clase C^p en U .

Comentarios.

- Aunque no lo haremos en estas notas, se puede dar una demostración elemental sencilla de este resultado utilizando el teorema de la función inversa. De hecho, ambos teoremas son *equivalentes*, ya que también es fácil probar el teorema de la función inversa a partir del teorema de la función implícita. Véase, por ejemplo, el libro de Fleming citado en la bibliografía.
- En los puntos (x_0, y_0) en que $\det \frac{\partial \Phi}{\partial y}(x_0, y_0) = 0$, no puede afirmarse nada acerca de si la relación $\Phi(x, y) = 0$ define o no a y como función de x . Por ejemplo, $\Phi(x, y) = y^2 - x = 0$ no define a y como función de x en un entorno de $(0, 0)$, mientras que $\Phi(x, y) = y^3 - x^3 = 0$ define globalmente a y como función de x ($y = -x$). Sin embargo, en ambos casos $\frac{\partial \Phi}{\partial y}(0, 0) = 0$.

3.6.1 Derivación implícita

Cuando el teorema de la función implícita es aplicable en un punto (x_0, y_0) , puede calcularse el valor de las derivadas parciales de orden menor o igual que p de la función implícita f en el punto x_0 sin necesidad de determinar f explícitamente. Para ello, basta derivar las m identidades

$$\Phi_k(x, f(x)) = 0, \quad \forall x \in U, \quad 1 \leq k \leq m, \quad (3.29)$$

respecto de una cualquiera de las coordenadas x_i de x aplicando la *regla de la cadena*, obteniéndose

$$\frac{\partial \Phi_k}{\partial x_i}(x, f(x)) + \sum_{j=1}^m \frac{\partial \Phi_k}{\partial y_j}(x, f(x)) \frac{\partial f_j}{\partial x_i}(x) = 0, \quad 1 \leq i \leq n, \quad 1 \leq k \leq m. \quad (3.30)$$

Nótese que las derivadas parciales en las ecuaciones anteriores se calculan considerando a Φ_k como una función de las variables *independientes* (x, y) , y a *continuación* sustituyendo y por $f(x)$. Las nm igualdades escalares (3.30) son equivalente a la igualdad matricial

$$\frac{\partial \Phi}{\partial x}(x, f(x)) + \frac{\partial \Phi}{\partial y}(x, f(x)) \cdot Df(x) = 0, \quad \forall x \in U.$$

Como, por hipótesis, el determinante de $\frac{\partial \Phi}{\partial y}(x_0, y_0)$ no se anula, sustituyendo $x = x_0$ en la igualdad anterior y despejando $Df(x_0)$ se obtiene finalmente

$$Df(x_0) = - \left[\frac{\partial \Phi}{\partial y}(x_0, y_0) \right]^{-1} \cdot \frac{\partial \Phi}{\partial x}(x_0, y_0).$$

Nótese que esta es la fórmula para $Df(x_0)$ que obtuvimos más arriba (cf. la ec. (3.28)) por un procedimiento heurístico. Si $p \geq 2$, para calcular las derivadas parciales de segundo orden de f en x_0 derivamos (3.30) una vez más y evaluamos el resultado para $x = x_0$, obteniendo una expresión del tipo

$$\sum_{j=1}^m \frac{\partial \Phi_k}{\partial y_j}(x_0, y_0) \frac{\partial^2 f_j}{\partial x_i \partial x_i}(x_0) + \dots = 0, \quad 1 \leq i \leq n, \quad 1 \leq k \leq m,$$

o equivalentemente

$$\frac{\partial \Phi}{\partial y}(x_0, y_0) \cdot \frac{\partial^2 f}{\partial x_l \partial x_i}(x_0) + \dots = 0,$$

donde los términos omitidos dependen de las derivadas parciales de primer orden de f en x_0 y de las derivadas parciales de orden menor o igual que 2 de Φ en (x_0, y_0) , ambas conocidas. De nuevo por la invertibilidad de $\frac{\partial \Phi}{\partial y}(x_0, y_0)$, se pueden despejar en la ecuación anterior las derivadas parciales de segundo orden de f en x_0 , y así sucesivamente. En la práctica, *no es necesario recordar las fórmulas anteriores, sino que es preferible derivar directamente las m igualdades (3.29).*

Comentario. Al derivar implícitamente se suele sustituir (con un ligero abuso de notación) $f(x)$ por y y $\frac{\partial f_j}{\partial x_i}(x)$ por $\frac{\partial y_j}{\partial x_i}$, escribiendo (3.30) en la forma

$$\frac{\partial \Phi_k}{\partial x_i}(x, y) + \sum_{j=1}^m \frac{\partial \Phi_k}{\partial y_j}(x, y) \frac{\partial y_j}{\partial x_i} = 0, \quad 1 \leq i \leq n, \quad 1 \leq k \leq m,$$

o incluso

$$\boxed{\frac{\partial \Phi_k}{\partial x_i} + \sum_{j=1}^m \frac{\partial \Phi_k}{\partial y_j} \frac{\partial y_j}{\partial x_i} = 0}, \quad 1 \leq i \leq n, \quad 1 \leq k \leq m. \quad (3.31)$$

Notación. Es muy frecuente utilizar la notación

$$\boxed{\frac{\partial(\Phi_1, \dots, \Phi_m)}{\partial(y_1, \dots, y_m)}}$$

para denotar a la matriz $\frac{\partial \Phi}{\partial y}$ cuyos elementos de matriz son las derivadas parciales $\frac{\partial \Phi_i}{\partial y_j}$ (con $1 \leq i, j \leq m$).

Ejemplo 3.29. Consideremos las relaciones

$$\begin{cases} \Phi_1(x, y, z, u, v) \equiv xy^2 + xzu + yu^2 - 3 = 0 \\ \Phi_2(x, y, z, u, v) \equiv u^3yz + 2xy - 2u^2v^2 - 1 = 0. \end{cases}$$

En primer lugar, veamos en qué puntos puede asegurarse que estas relaciones definen implícitamente a u y v como funciones de (x, y, z) . Como Φ es evidentemente de clase C^∞ en \mathbb{R}^5 , basta imponer la condición

$$\det \frac{\partial(\Phi_1, \Phi_2)}{\partial(u, v)} = \begin{vmatrix} xz + 2yu & 0 \\ 3u^2yz - 4uv^2 & -4u^2v \end{vmatrix} = -4u^2v(xz + 2yu) \neq 0.$$

Por el teorema de la función implícita, la relación $\Phi = 0$ define implícitamente a (u, v) como función de clase C^∞ de (x, y, z) en un entorno de cualquier punto (x, y, z, u, v) que satisfaga las condiciones

$$\boxed{\Phi(x, y, z, u, v) = 0, \quad uv(xz + 2yu) \neq 0};$$

por ejemplo, esto ocurre en un entorno del punto $(1, 1, 1, 1, 1)$. Calculemos a continuación la derivada parcial v_y en el punto $(1, 1, 1)$. Para ello, basta derivar implícitamente respecto de y las relaciones $\Phi_i = 0$, $1 \leq i \leq 2$, obteniéndose

$$\begin{cases} 2xy + u^2 + (xz + 2yu)u_y = 0 \\ u^3z + 2x + (3u^2yz - 4uv^2)u_y - 4u^2vv_y = 0. \end{cases}$$

Sustituyendo en estas ecuaciones $(x, y, z, u, v) = (1, 1, 1, 1, 1)$ se obtiene inmediatamente

$$\begin{cases} 3 + 3u_y(1, 1, 1) = 0 \\ 3 - u_y(1, 1, 1) - 4v_y(1, 1, 1) = 0. \end{cases}$$

Despejando $u_y(1, 1, 1)$ y $v_y(1, 1, 1)$ de estas relaciones se llega a

$$\boxed{u_y(1, 1, 1) = -1, \quad v_y(1, 1, 1) = 1}.$$

Para calcular, por ejemplo, $v_{yy}(1, 1, 1)$, basta derivar las relaciones anteriores respecto de y :

$$\begin{cases} 2x + 2uu_y + (2u + 2yu_y)u_y + (xz + 2yu)u_{yy} = 0 \\ 3u^2zu_y + (3u^2z + (6uyz - 4v^2)u_y - 8uvv_y)u_y \\ + (3u^2yz - 4uv^2)u_{yy} - 4(2uvu_y + u^2v_y)v_y - 4u^2vv_{yy} = 0. \end{cases}$$

Sustituyendo de nuevo $(x, y, z, u, v) = (1, 1, 1, 1, 1)$ y $(u_y, v_y) = (-1, 1)$ se obtiene

$$u_{yy}(1, 1, 1) = 0, \quad 8 - u_{yy}(1, 1, 1) - 4v_{yy}(1, 1, 1) = 0$$

de donde se sigue que

$$\boxed{u_{yy}(1, 1, 1) = 0, \quad v_{yy}(1, 1, 1) = 2}.$$

3.7 Extremos condicionados

En muchos problemas prácticos es necesario encontrar los extremos de una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ (que supondremos al menos de clase C^1) en un conjunto M de la forma

$$\boxed{M = \{x \in \mathbb{R}^n \mid g_k(x) = 0, \quad 1 \leq k \leq r\} \equiv g^{-1}(0)}, \quad (3.32)$$

con $g \equiv (g_1, \dots, g_r) : \mathbb{R}^n \rightarrow \mathbb{R}^r$ y $r < n$. Este es un problema de *extremos condicionados*, ya que buscamos los extremos de f con las r condiciones (o restricciones) sobre la variable independiente x dadas por las ecuaciones

$$g_1(x) = \dots = g_r(x) = 0. \quad (3.33)$$

Intuitivamente, el conjunto M definido por (3.32) debería ser una superficie de dimensión $n - r$, al menos si las funciones g_1, \dots, g_r son suficientemente regulares (por ejemplo, de clase C^1). En realidad, esto *no* tiene por qué ser así a menos que dichas funciones verifiquen alguna condición adicional que garantice la independencia de las ecuaciones (3.33). Por ejemplo, si $g_1 = \dots = g_r$ está claro que las r condiciones (3.33) no son independientes, sino que se reducen a la única condición $g_1 = 0$, y por tanto M debería tener dimensión $n - 1$. Como explicaremos a continuación, la condición clave para que las ecuaciones (3.33) definan un conjunto de dimensión $n - r$ resulta ser la siguiente:

$$\boxed{\text{rank } Dg(x) = r, \quad \forall x \in M}. \quad (3.34)$$

Nótese que la matriz $Dg(x)$ es de orden $r \times n$, y por tanto (al ser $r < n$) la condición anterior afirma que $Dg(x)$ tiene *rango máximo* en todos los puntos de $M = g^{-1}(0)$.

Más precisamente, supongamos que la función $g : \mathbb{R}^n \rightarrow \mathbb{R}^r$ es de clase C^1 en un abierto que contenga a M , y que se cumple la condición (3.34). Si $x_0 \in M$, en virtud de (3.34) hay un menor $r \times r$ de $Dg(x_0)$ que no se anula. Supongamos, por ejemplo, que

$$\det \frac{\partial (g_1, \dots, g_r)}{\partial (x_{n-r+1}, \dots, x_n)}(x_0) \neq 0.$$

Por el teorema de la función implícita, hay un entorno V de $x_0 \equiv (x_{01}, \dots, x_{0n})$ y un entorno U de $(x_{01}, \dots, x_{0, n-r})$ tales que

$$x \in V, g(x) = 0 \iff (x_1, \dots, x_{n-r}) \in U, (x_{n-r+1}, \dots, x_n) = f(x_1, \dots, x_{n-r}),$$

siendo $f : U \subset \mathbb{R}^{n-r} \rightarrow \mathbb{R}^r$ de clase C^1 en U . En otras palabras, el conjunto $M \cap V$ está descrito por las r ecuaciones

$$x_{n-r+1} = f_1(x_1, \dots, x_{n-r}), \dots, x_n = f_r(x_1, \dots, x_{n-r}),$$

donde (x_1, \dots, x_{n-r}) varía en el abierto $U \subset \mathbb{R}^{n-r}$ sin ninguna restricción. Por tanto, si se cumple la condición (3.34) el conjunto $M = g^{-1}(0)$ es efectivamente $(n-r)$ -dimensional, ya que está determinado por $n-r$ coordenadas independientes en un entorno de cada uno de sus puntos. Esto motiva la siguiente definición:

Definición 3.30. Diremos que el conjunto $M = g^{-1}(0)$ definido por (3.32) es una **variedad de dimensión $n-r$** si la función $g \equiv (g_1, \dots, g_r) : \mathbb{R}^n \rightarrow \mathbb{R}^r$ es de clase C^1 en un abierto que contiene a M y $\text{rank } Dg(x) = r$ para todo $x \in M$.

Si x_0 es un punto del conjunto M dado por (3.32), diremos que un vector $v \in \mathbb{R}^n$ es **tangente** a M en x_0 si v es el vector tangente en x_0 a alguna curva diferenciable $\gamma : \mathbb{R} \rightarrow \mathbb{R}^n$ contenida en M . El **espacio tangente** a M en $x_0 \in M$ es por definición el conjunto de todos los vectores tangentes a M en x_0 . El espacio tangente a una variedad M en uno de sus puntos admite la siguiente caracterización:

Proposición 3.31. Sea $M = g^{-1}(0)$ (con $g : \mathbb{R}^n \rightarrow \mathbb{R}^r$) una variedad de dimensión $n-r$ en \mathbb{R}^n , y sea x_0 un punto de M . Entonces el espacio tangente a M en x_0 es el núcleo $\ker Dg(x_0)$ de la aplicación lineal $Dg(x_0)$, es decir el conjunto de los vectores $v \in \mathbb{R}^n$ que verifican la condición

$$Dg(x_0) \cdot v = 0. \tag{3.35}$$

Demostración. Es fácil demostrar que si v es un vector tangente a M en x_0 entonces verifica la condición (3.35). En efecto, por definición $v = \gamma'(t_0)$, donde $\gamma : \mathbb{R} \rightarrow \mathbb{R}^n$ es una curva diferenciable tal que $\gamma(t) \in M$ para todo t y $\gamma(t_0) = x_0$. La condición $\gamma(t) \in M$ para todo t es equivalente a

$$g(\gamma(t)) = 0, \quad \forall t.$$

Por la regla de la cadena,

$$0 = \frac{d}{dt} g(\gamma(t)) \Big|_{t=t_0} = Dg(\gamma(t_0)) \cdot \gamma'(t_0) = Dg(x_0) \cdot v.$$

El recíproco de esta afirmación (es decir, que cualquier vector v que verifique (3.35) es un vector tangente a M en x_0) se demuestra utilizando el teorema de la función implícita (véase, por ejemplo, el libro de Fleming en la bibliografía adicional recomendada). \square

Como es sabido por el curso de Álgebra Lineal, el núcleo de cualquier aplicación lineal es un subespacio vectorial. Por tanto el espacio tangente a M en x_0 es un subespacio vectorial, de dimensión

$$\dim \ker Dg(x_0) = n - \text{rank } Dg(x_0) = n - r$$

igual a la dimensión de la variedad M . Nótese también que

$$Dg(x_0) \cdot v = (\nabla g_1(x_0) \cdot v, \dots, \nabla g_r(x_0) \cdot v),$$

y por tanto

$$Dg(x_0) \cdot v = 0 \iff \nabla g_k(x_0) \cdot v = 0, \quad 1 \leq k \leq r.$$

En otras palabras,

$$\ker Dg(x_0) = N^\perp,$$

siendo

$$N = \text{lin}\{\nabla g_1(x_0), \dots, \nabla g_r(x_0)\}$$

el subespacio generado por los r vectores $\nabla g_k(x_0)$ ($1 \leq k \leq r$) y N^\perp su *complemento ortogonal*².

Tras estos preliminares, resulta sencillo encontrar la condición *necesaria* que han de cumplir los extremos relativos de una función diferenciable $f : \mathbb{R}^n \rightarrow \mathbb{R}$ en la variedad $M = g^{-1}(0)$ de dimensión $n - r$. En efecto, sea x_0 un extremo relativo de f en M y sea v un vector tangente a M en x_0 . Por definición de vector tangente, hay una curva diferenciable $\gamma : \mathbb{R} \rightarrow \mathbb{R}^n$ contenida en M tal que $\gamma(t_0) = x_0$ y $v = \gamma'(t_0)$. La función $\phi(t) = f(\gamma(t))$ tiene entonces un extremo relativo en $t = t_0$. Si f es diferenciable en x_0 , por la regla de la cadena se tiene

$$\phi'(t_0) = Df(\gamma(t_0)) \cdot \gamma'(t_0) = Df(x_0) \cdot v \equiv \nabla f(x_0) \cdot v = 0.$$

En otras palabras, si x_0 es un extremo relativo de f en M entonces

$$\nabla f(x_0) \cdot v = 0, \quad \forall v \in \ker Dg(x_0),$$

o equivalentemente

$$\nabla f(x_0) \in (\ker Dg(x_0))^\perp = (N^\perp)^\perp = N. \tag{3.36}$$

Por consiguiente han de existir r constantes reales μ_1, \dots, μ_r tales que

$$\nabla f(x_0) = \mu_1 \nabla g_1(x_0) + \dots + \mu_r \nabla g_r(x_0) \equiv \nabla(\mu_1 g_1 + \dots + \mu_r g_r)(x_0),$$

o lo que es lo mismo (llamando $\lambda_i \equiv -\mu_i$)

$$\nabla(f + \lambda_1 g_1 + \dots + \lambda_m g_m)(x_0) = 0. \tag{3.37}$$

Hemos probado por tanto el siguiente resultado fundamental:

Proposición 3.32 (método de los multiplicadores de Lagrange). *Sea x_0 un extremo relativo de una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ en la variedad $M = g^{-1}(0)$ de dimensión $n - r$, donde $g : \mathbb{R}^n \rightarrow \mathbb{R}^r$. Si f es diferenciable en x_0 , entonces existen r números reales $\lambda_1, \dots, \lambda_r$ tales que x_0 es un punto crítico de la función $f + \lambda_1 g_1 + \dots + \lambda_r g_r$.*

Comentario. Las r constantes $\lambda_1, \dots, \lambda_r$ de la proposición anterior se denominan **multiplicadores de Lagrange**. □

Por lo que acabamos de ver, las n coordenadas de un extremo relativo x_0 de f en la variedad M han de satisfacer la condición (3.37) (n ecuaciones escalares), junto con $g(x_0) = 0$ (r ecuaciones escalares). Estas condiciones forman por tanto un sistema de $n + r$ ecuaciones escalares en las $n + r$ incógnitas x_0 (vector de n componentes), $\lambda_1, \dots, \lambda_r$, a saber

$$\begin{aligned} \frac{\partial f}{\partial x_i}(x_0) + \lambda_1 \frac{\partial g_1}{\partial x_i}(x_0) + \dots + \lambda_r \frac{\partial g_r}{\partial x_i}(x_0) &= 0, & 1 \leq i \leq n, \\ g_k(x_0) &= 0, & 1 \leq k \leq r. \end{aligned}$$

²Si W es un subespacio de un espacio euclidiano V (de dimensión finita), su *complemento ortogonal* es el conjunto

$$W^\perp = \{v \in V \mid v \cdot w = 0, \quad \forall w \in W\}.$$

Se demuestra que W^\perp es también un subespacio vectorial de V , y que $V = W \oplus W^\perp$. En particular, de esta última igualdad se sigue que $(W^\perp)^\perp = W$.

Definición 3.33. Diremos que $x_0 \in \mathbb{R}^n$ es un **punto crítico de f en la variedad M** si verifica las ecuaciones anteriores para ciertas constantes $\lambda_1, \dots, \lambda_r$.

Aunque hay algunos criterios para determinar si uno de estos puntos críticos es un extremo relativo de f en M , y averiguar de qué tipo, ninguno de ellos es tan sencillo como los vistos en la Sección 3.2 para extremos no condicionados. Por este motivo, en la mayor parte de los casos la naturaleza de un punto crítico de una función en una variedad se determina por inspección. Por ejemplo, como el conjunto (3.32) siempre es cerrado, si es acotado es también *compacto*, y por tanto está garantizado en este caso que la función f alcanzará en él sus valores máximo y mínimo. Para encontrar los extremos absolutos de f en una variedad *compacta* M basta entonces calcular sus puntos críticos en M y comparar los valores que la función alcanza en dichos puntos.

Comentario. El método de los multiplicadores de Lagrange se puede aplicar también para hallar los extremos de una función escalar (de clase C^1) en un conjunto de la forma $M \cap A$, donde $M = g^{-1}(0)$ es una variedad y A un conjunto *abierto*. La razón es que este método está basado en la Proposición 3.31, que sigue siendo válida en este caso. Por el contrario, si A es cerrado (es decir, $A = \overset{\circ}{A} \cup \partial A$) hay que buscar los extremos de f en $M \cap A$ por separado en los conjuntos $M \cap \overset{\circ}{A}$ y $M \cap \partial A$. En el primero de estos conjuntos podemos aplicar el método de los multiplicadores de Lagrange, en virtud del comentario anterior. Frecuentemente, también se puede aplicar dicho método en el segundo de estos conjuntos, ya que generalmente $M \cap \partial A$ es una variedad (de dimensión más baja que M).

Supongamos, por ejemplo, que es necesario hallar los extremos de $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ en el conjunto $M \cap A$, siendo M la esfera unidad y A el semiplano $z \geq 0$. Para ello hallaremos por separado los extremos de dicha función en los conjuntos $M_1 = M \cap \overset{\circ}{A} = M \cap \{(x, y, z) \in \mathbb{R}^3 \mid z > 0\}$ y $M_2 = M \cap \partial A = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 - 1 = 0, z = 0\}$. En el primer conjunto podemos aplicar el método de los multiplicadores con $g(x, y, z) = x^2 + y^2 + z^2 - 1$, rechazando los puntos críticos que no satisfagan la restricción adicional $z > 0$. Para hallar los extremos de f en el segundo de estos conjuntos también se puede aplicar el método de los multiplicadores de Lagrange, ya que es inmediato comprobar que $M_2 = g^{-1}(0)$, con $g(x, y, z) = (x^2 + y^2 - 1, z)$, es una variedad de dimensión 1. \square

Ejemplo 3.34. Calculemos la distancia al origen de la curva de ecuaciones

$$\boxed{z^2 + xy = 0, \quad x^2 + y^2 = 1}. \quad (3.38)$$

Por definición, dicha distancia es el valor mínimo de la función $\sqrt{x^2 + y^2 + z^2}$ en la curva dada por las ecuaciones anteriores o, equivalentemente, la raíz cuadrada del valor mínimo de la función $f(x, y, z) = x^2 + y^2 + z^2$ en la curva. La curva (3.38) es claramente un conjunto *compacto*. En efecto, es cerrada por ser la imagen inversa del punto $(0, 0)$ bajo la función continua (C^∞)

$$g(x, y, z) = (z^2 + xy, x^2 + y^2 - 1).$$

Es también acotada, ya que

$$x^2 + y^2 = 1 \quad \Rightarrow \quad |x| \leq 1, \quad |y| \leq 1 \quad \Rightarrow \quad |z| = \sqrt{|xy|} \leq 1.$$

Por tanto está garantizado que la función f alcance sus valores mínimo y máximo en la curva (3.38). Dicha curva es una variedad de dimensión 1, ya que la función g que la define es de clase C^∞ y

$$Dg(x, y, z) = \begin{pmatrix} y & x & 2z \\ 2x & 2y & 0 \end{pmatrix}$$

tiene rango 2 en todos los puntos de la curva. En efecto, los menores de orden 2 de $Dg(x, y, z)$ se anulan simultáneamente en los puntos (x, y, z) que verifican

$$y = \pm x, \quad xz = 0,$$

y ninguno de dichos puntos pertenece a la curva (3.38). Por tanto los extremos de f en la curva (3.38) han de estar entre los puntos críticos de f en dicha curva. Para determinar dichos puntos críticos introducimos dos multiplicadores de Lagrange λ_1, λ_2 y calculamos los puntos críticos de la función

$$f + \lambda_1 g_1 + \lambda_2 g_2 = x^2 + y^2 + z^2 + \lambda_1(z^2 + xy) + \lambda_2(x^2 + y^2 - 1).$$

Obtenemos así el sistema

$$2(1 + \lambda_2)x + \lambda_1 y = 0, \quad 2(1 + \lambda_2)y + \lambda_1 x = 0, \quad 2(1 + \lambda_1)z = 0, \quad (3.39)$$

que junto con las ecuaciones (3.38) de la curva forman un sistema de 5 ecuaciones en las 5 incógnitas $(x, y, z, \lambda_1, \lambda_2)$. Las dos primeras ecuaciones se pueden escribir en forma matricial como

$$\begin{pmatrix} 2(1 + \lambda_2) & \lambda_1 \\ \lambda_1 & 2(1 + \lambda_2) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0.$$

Como x e y no pueden anularse simultáneamente (ya que $x^2 + y^2 = 1$), de la esta ecuación se sigue que

$$\begin{vmatrix} 2(1 + \lambda_2) & \lambda_1 \\ \lambda_1 & 2(1 + \lambda_2) \end{vmatrix} = 4(1 + \lambda_2)^2 - \lambda_1^2 = 0 \iff 2(1 + \lambda_2) = \pm \lambda_1.$$

Sustituyendo en (3.39) queda

$$\lambda_1(x \pm y) = 0, \quad 2(1 + \lambda_1)z = 0.$$

Si $\lambda_1 = 0$ entonces $z = xy = 0$, y por tanto (al ser $x^2 + y^2 = 1$) o bien $x = 0, y = \pm 1$ o bien $y = 0, x = \pm 1$. Se obtienen de esta forma los cuatro puntos críticos

$$(0, \pm 1, 0), \quad (\pm 1, 0, 0). \quad (3.40)$$

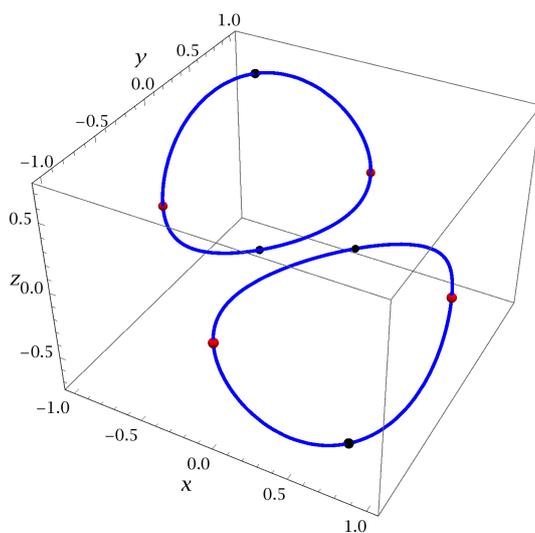


Figura 3.4: Curva de ecuaciones $z^2 + xy = 0, x^2 + y^2 = 1$. En rojo se han representado los cuatro puntos (3.40) de la curva en que la distancia al origen es mínima, en negro los cuatro puntos (3.41) en que dicha distancia es máxima.

La otra posibilidad es que $x \pm y = 0$ aunque, al ser $z^2 = -xy$, la solución $x = y$ está de hecho descartada. Por tanto en este caso $x = -y = \pm\sqrt{2}/2$, $z = \pm\sqrt{-xy} = \pm\sqrt{2}/2$ (donde el signo de z es independiente del de x), lo que proporciona los cuatro puntos críticos

$$\frac{\sqrt{2}}{2} (\varepsilon_1, -\varepsilon_1, \varepsilon_2), \quad (3.41)$$

donde $\varepsilon_{1,2} = \pm 1$ son dos signos independientes. En los cuatro primeros puntos críticos la función $x^2 + y^2 + z^2$ vale 1, mientras que sobre los últimos cuatro vale $3/2$. Por tanto la distancia de la curva (3.38) al origen es 1, y la distancia máxima de un punto de la curva al origen es $\sqrt{3/2}$. Además, los puntos de la curva a distancia mínima del origen son los cuatro puntos (3.40), y los puntos a distancia máxima son los cuatro puntos (3.41).

Capítulo 4

Integrales múltiples

4.1 Integral doble en un rectángulo

4.1.1 Definición de integral

Nuestro objetivo es definir la integral de una función de dos variables $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ en el rectángulo compacto $R = [a, b] \times [c, d]$ (con $a < b$ y $c < d$). Como en el caso de la integral (de Riemann) en una dimensión, la integral se definirá como el límite de una suma obtenida subdividiendo el rectángulo en un cierto número de subrectángulos, y aproximando en cada uno de ellos la función por el supremo o el ínfimo de sus valores.

Más precisamente, consideremos sendas particiones $P_1 = \{x_0, x_1, \dots, x_n\}$ y $P_2 = \{y_0, y_1, \dots, y_m\}$ de los lados $[a, b]$ y $[c, d]$ del rectángulo R , siendo

$$a = x_0 < x_1 < \dots < x_n = b, \quad c = y_0 < y_1 < \dots < y_m = d.$$

Dichas particiones definen una partición del rectángulo $R = [a, b] \times [c, d]$, que denotaremos por $P_1 \times P_2 \equiv P$, en nm subrectángulos de la forma

$$R_{ij} \equiv [x_{i-1}, x_i] \times [y_{j-1}, y_j], \quad 1 \leq i \leq n, \quad 1 \leq j \leq m,$$

cada uno de los cuales tiene área

$$a(R_{ij}) = (x_i - x_{i-1})(y_j - y_{j-1}) \equiv \Delta x_i \Delta y_j.$$

Si $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ está *acotada*¹ en el rectángulo R (y, por tanto, en cada uno de los subrectángulos R_{ij} de la partición), definimos las **sumas superior** e **inferior** $U(f; P)$ y $L(f; P)$ asociadas a la partición P por

$$U(f; P) = \sum_{i=1}^n \sum_{j=1}^m M_{ij}(f; P) \Delta x_i \Delta y_j, \quad L(f; P) = \sum_{i=1}^n \sum_{j=1}^m m_{ij}(f; P) \Delta x_i \Delta y_j,$$

siendo

$$M_{ij}(f; P) \equiv \sup\{f(x, y) \mid (x, y) \in R_{ij}\}, \quad m_{ij}(f; P) \equiv \inf\{f(x, y) \mid (x, y) \in R_{ij}\}$$

respectivamente el ínfimo y el supremo de f en el rectángulo R_{ij} . Es evidente que

$$L(f; P) \leq U(f; P), \tag{4.1}$$

por la definición de ínfimo y supremo. Intuitivamente, $L(f; P)$ debería ser una aproximación por

¹Recordemos que la función f está acotada en R si existe una constante M tal que $|f(x, y)| \leq M$, para todo $(x, y) \in R$. Nótese que una función acotada está acotada tanto inferiormente (por $-M$) como superiormente (por M). Recíprocamente, si f está acotada inferiormente por M_1 y superiormente por M_2 en R entonces f está acotada en R , ya que en tal caso $|f(x, y)| \leq M$ en R si (por ejemplo) $M = \max(|M_1|, |M_2|)$.

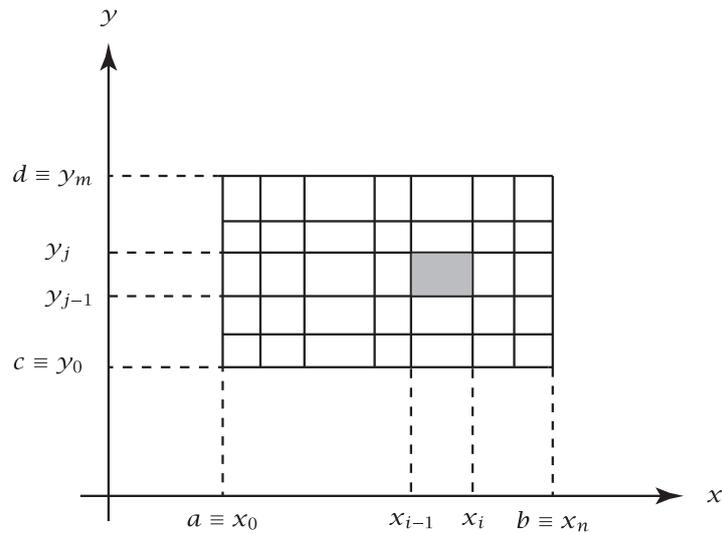


Figura 4.1: Partición del rectángulo $[a, b] \times [c, d]$ en 35 subrectángulos (en gris se ha representado uno de los rectángulos R_{ij} de la partición).

defecto a la integral de f en R , mientras que $U(f; P)$ debería ser una aproximación por exceso a dicha integral. De hecho, puede probarse que se cumple un resultado aún más fuerte que (4.1), a saber

$$L(f; P) \leq U(f; P'), \quad (4.2)$$

siendo P y P' dos particiones (no necesariamente iguales) de R . Esto implica que el conjunto de todas las sumas inferiores

$$\{L(f; P) \mid P \text{ partición de } R\}$$

está acotado superiormente por cualquier suma superior $U(f; P')$, y por tanto posee un supremo. Dicho supremo, que denotaremos por $\int_R f$, se denomina **integral inferior** de Riemann de la función f en el rectángulo R :

$$\int_R f \equiv \sup \{L(f; P) \mid P \text{ partición de } R\}.$$

Análogamente, en virtud de (4.2) el conjunto de las sumas superiores

$$\{U(f; P) \mid P \text{ partición de } R\}$$

está acotado inferiormente por cualquier suma inferior $L(f; P')$, y por tanto posee un ínfimo. Dicho ínfimo, que denotaremos por $\overline{\int}_R f$, se denomina **integral superior** de Riemann de la función f en el rectángulo R :

$$\overline{\int}_R f = \inf \{U(f; P) \mid P \text{ partición de } R\}.$$

Además, es evidente que

$$\int_R f \leq \overline{\int}_R f,$$

ya que de (4.2) se sigue que

$$\int_R f \equiv \sup \{L(f; P) \mid P \text{ partición de } R\} \leq U(f; P'), \quad \forall P' \text{ partición de } R,$$

y por tanto

$$\int_R f \leq \inf\{U(f;P) \mid P \text{ partición de } R\} \equiv \overline{\int}_R f.$$

Definición 4.1. Sea $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ una función definida y acotada en el rectángulo cerrado $R = [a, b] \times [c, d]$. Diremos que f es **integrable** Riemann en R si las integrales superior e inferior de Riemann de f en R coinciden. El valor común de ambas integrales se denomina **integral** de Riemann de f en R , y se denota por

$$\int_R f.$$

En otras palabras, f es integrable Riemann en R si

$$\int_R f = \overline{\int}_R f \equiv \int_R f.$$

Como en este curso elemental no utilizaremos otras definiciones de integral distintas de la de Riemann (por ejemplo, la integral de Lebesgue), a partir de ahora omitiremos el calificativo “de Riemann” al referirnos a la integral de f en el rectángulo R .

Notación. Es muy frecuente utilizar la notación tradicional (ligeramente redundante)

$$\int_R f(x, y) \, dx \, dy$$

para denotar la integral doble $\int_R f$.

Si la función f es no negativa en el rectángulo R , es intuitivamente evidente en virtud de su definición que las sumas superior e inferior $U(f;P)$ y $L(f;P)$ proporcionan una aproximación respectivamente por exceso y por defecto al *volumen* $v(\Omega)$ del sólido

$$\Omega = \{(x, y, z) \in \mathbb{R}^3 \mid (x, y) \in R, 0 \leq z \leq f(x, y)\}$$

limitado por el plano $z = 0$ y la gráfica de la función f en R . En efecto, el paralelepípedo de base R_{ij} y altura $m_{ij}(f;P)$ (resp. $M_{ij}(f;P)$) está contenido en (resp. contiene a) la región del sólido limitada por el plano $z = 0$ y la gráfica de f en R_{ij} . Por tanto deberían cumplirse las desigualdades

$$L(f;P) \leq v(\Omega) \leq U(f;P)$$

para toda partición P del rectángulo R , de donde se sigue que

$$\int_R f \leq v(\Omega) \leq \overline{\int}_R f.$$

En particular, si f es integrable en R el volumen de Ω está dado por

$$v(\Omega) = \int_R f.$$

De esta observación se sigue fácilmente que si $f_{1,2}$ son integrables en R y $f_1(x, y) \leq f_2(x, y)$ para todo $(x, y) \in R$, el volumen del sólido

$$\Omega = \{(x, y, z) \in \mathbb{R}^3 \mid (x, y) \in R, f_1(x, y) \leq z \leq f_2(x, y)\}$$

limitado por las gráficas de las funciones f_1 y f_2 en el rectángulo R está dado por

$$v(\Omega) = \int_R f_2 - \int_R f_1 = \int_R (f_2 - f_1),$$

donde la última igualdad es consecuencia de la propiedad de linealidad de la integral doble que veremos más adelante (cf. la pág. 113).

Ejemplo 4.2. Sea $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ la función constante dada por $f(x, y) = h$, para todo $(x, y) \in \mathbb{R}^2$. Entonces

$$m_{ij}(f; P) = M_{ij}(f; P) = h,$$

y por tanto

$$L(f; P) = U(f; P) = h \sum_{i=1}^n \sum_{j=1}^m \Delta x_i \Delta y_j = h \left(\sum_{i=1}^n (x_i - x_{i-1}) \right) \left(\sum_{j=1}^m (y_j - y_{j-1}) \right) = h(b-a)(d-c).$$

Luego f es integrable en R , y se cumple

$$\int_R f = h(b-a)(d-c) \equiv h a(R).$$

Si $h > 0$, el valor de la integral es el volumen del paralelepípedo de base R y altura h .

Ejemplo 4.3. Consideremos la función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por

$$f(x, y) = \begin{cases} 1, & (x, y) \in \mathbb{Q} \times \mathbb{Q} \\ 0, & (x, y) \notin \mathbb{Q} \times \mathbb{Q}. \end{cases}$$

Como cualquier rectángulo S contiene tanto puntos con ambas coordenadas racionales como puntos que no son de este tipo, es evidente que

$$\inf\{f(x, y) \mid (x, y) \in S\} = 0, \quad \sup\{f(x, y) \mid (x, y) \in S\} = 1.$$

Por tanto si $P = P_1 \times P_2$ es una partición cualquiera del rectángulo $R = [a, b] \times [c, d]$ se tiene

$$m_{ij}(f; P) = 0, \quad M_{ij}(f; P) = 1,$$

de donde se sigue que

$$L(f; P) = \sum_{i=1}^n \sum_{j=1}^m 0 = 0, \quad U(f; P) = \sum_{i=1}^n \sum_{j=1}^m \Delta x_i \Delta y_j = (b-a)(d-c) = a(R) > 0.$$

Luego en este caso se tiene

$$\int_{\underline{R}} f = 0 < \int_{\overline{R}} f = a(R),$$

y la función f no es integrable en ningún rectángulo R . □

4.1.2 Criterio de integrabilidad de Lebesgue

La función f del ejemplo anterior es discontinua en todo \mathbb{R}^2 . De hecho, enunciaremos a continuación un resultado clásico (debido a Lebesgue) que relaciona la integrabilidad de una función en un rectángulo R con el “tamaño” del conjunto de sus puntos de discontinuidad en R . Para ello necesitamos antes una definición y algunas consideraciones previas. En primer lugar, dado un n -paralelepípedo $H = [a_1, b_1] \times \cdots \times [a_n, b_n] \subset \mathbb{R}^n$, definimos de forma natural el volumen $v(H)$ de H como el producto de las longitudes de sus lados, es decir

$$v(H) = \prod_{i=1}^n (b_i - a_i). \quad (4.3)$$

Definición 4.4. Diremos que un conjunto $A \subset \mathbb{R}^n$ tiene **medida cero** si para todo $\varepsilon > 0$ existe una familia *numerable*² de n -paralelepípedos H_i (con $i \in \mathbb{N}$) tal que

$$A \subset \bigcup_{i \in \mathbb{N}} H_i, \quad \sum_{i=1}^{\infty} v(H_i) < \varepsilon.$$

Ejemplo 4.5. Cualquier subconjunto *numerable* $A \subset \mathbb{R}^n$ tiene medida cero. En efecto, si $A = \{z_i \mid i \in \mathbb{N}\}$ (con $z_i \in \mathbb{R}^n$), dado $\varepsilon > 0$ sea H_i un n -cubo centrado en el punto z_i de volumen menor que $\varepsilon/2^i$ (evidentemente, basta tomar el lado del n -cubo menor que $(\varepsilon/2^i)^{1/n}$). Entonces A está contenido en la unión de los H_i ($i \in \mathbb{N}$), y se tiene

$$\sum_{i=1}^{\infty} v(H_i) < \varepsilon \sum_{i=1}^{\infty} 2^{-i} = \varepsilon. \quad \square$$

También se demuestra, por un procedimiento análogo al del ejemplo anterior, que *la unión numerable de conjuntos de medida cero tiene medida cero*. Por último, es obvio que cualquier subconjunto de un conjunto de medida cero tiene medida cero.

Proposición 4.6 (criterio de integrabilidad de Lebesgue). Sea $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ una función definida y acotada en el rectángulo $R = [a, b] \times [c, d]$. Entonces f es integrable en R si y solo si el conjunto de los puntos $x \in R$ en que f es discontinua tiene medida cero.

Si $g : \mathbb{R} \rightarrow \mathbb{R}$ es una función *continua* en el intervalo compacto $[a, b]$, se demuestra que la gráfica de g

$$\mathcal{G}(g) = \{(x, g(x)) \mid x \in [a, b]\} \subset \mathbb{R}^2$$

tiene medida 0 (ver, por ejemplo, *Calculus* (vol. II), de Tom M. Apostol). En virtud del comentario al final del Ejemplo 4.5, la unión *numerable* de gráficas de funciones continuas en un intervalo compacto también tiene medida 0. De la proposición anterior se sigue entonces el siguiente corolario:

Corolario 4.7. Sea $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ una función definida y acotada en el rectángulo $R = [a, b] \times [c, d]$, y supongamos que el conjunto de los puntos de discontinuidad de f en R está contenido en la unión de las gráficas de un conjunto numerable de funciones de una variable continuas en un intervalo compacto. Entonces f es integrable en R .

En particular,

Si f es continua en el rectángulo R , entonces f es integrable en dicho rectángulo.

En efecto, si f es continua en R es automáticamente acotada en dicho conjunto (ya que R es compacto), y por tanto es integrable en él en virtud del criterio de Lebesgue.

La integral doble tiene las siguientes propiedades elementales, que se deducen fácilmente de su definición:

Sean $f, g : \mathbb{R}^2 \rightarrow \mathbb{R}$ funciones integrables en el rectángulo $R = [a, b] \times [c, d]$.

i) Si $\lambda, \mu \in \mathbb{R}$ entonces $\lambda f + \mu g$ es integrable en R , y se tiene:

$$\int_R (\lambda f + \mu g) = \lambda \int_R f + \mu \int_R g \quad (\text{linealidad}).$$

²Un conjunto A es **numerable** si hay una aplicación *biyectiva* $f : \mathbb{N} \rightarrow A$. En otras palabras, A es numerable si sus elementos se pueden *enumerar* como una sucesión $a_1, a_2, \dots, a_n, \dots$, o más formalmente si $A = \{a_i \mid i \in \mathbb{N}\}$ (obviamente, $a_i = f(i)$). Se demuestra que los conjuntos \mathbb{N} , \mathbb{Z} y \mathbb{Q} son numerables, mientras que \mathbb{R} o \mathbb{C} no lo son. Se demuestra también que la unión numerable de conjuntos numerables es numerable.

ii) Supongamos que $R = R_1 \cup \dots \cup R_k$, donde R_i ($1 \leq i \leq k$) es un rectángulo (con los lados paralelos a los ejes) y $R_i \cap R_j$ tiene interior vacío para todo $i \neq j$. Entonces f es integrable en R_i para todo $i = 1, \dots, k$, y se tiene

$$\int_R f = \sum_{i=1}^k \int_{R_i} f \quad (\text{aditividad}).$$

iii) Si $f(x, y) \leq g(x, y)$, para todo $(x, y) \in R$, entonces

$$\int_R f \leq \int_R g \quad (\text{monotonía}).$$

De la propiedad de monotonía se deducen varias consecuencias importantes. En primer lugar, nótese que si f es integrable en R también lo es $|f|$, ya que el conjunto de puntos de discontinuidad de $|f|$ está contenido en el de f . De las desigualdades

$$-|f(x, y)| \leq f(x, y) \leq |f(x, y)|$$

y la propiedad de monotonía se sigue entonces que

$$-\int_R |f| \leq \int_R f \leq \int_R |f|,$$

y por tanto

$$\left| \int_R f \right| \leq \int_R |f|. \quad (4.4)$$

En particular,

$$|f(x)| < M, \quad \forall x \in R \quad \Rightarrow \quad \left| \int_R f \right| \leq M a(R).$$

En segundo lugar,

$$m \leq f(x, y) \leq M, \quad \forall (x, y) \in R \quad \Rightarrow \quad m a(R) \leq \int_R f \leq M a(R). \quad (4.5)$$

En efecto, basta aplicar la propiedad iii) a las funciones m , f y M (de hecho, este resultado se deduce también fácilmente de la definición de integral). Por último, si la función f es *continua* en el rectángulo R ha de alcanzar necesariamente sus valores mínimo y máximo en dicho conjunto *compacto*. Existirán por tanto sendos puntos $(x_i, y_i) \in R$ ($i = 1, 2$) tales que

$$f(x_1, y_1) \leq f(x, y) \leq f(x_2, y_2), \quad \forall (x, y) \in R,$$

De (4.5) se sigue entonces que

$$f(x_1, y_1) \leq \frac{1}{a(R)} \int_R f \leq f(x_2, y_2).$$

Por el teorema de los valores intermedios (al ser el rectángulo *conexo*), existe $(x_0, y_0) \in R$ tal que

$$\frac{1}{a(R)} \int_R f = f(x_0, y_0).$$

Hemos demostrado por tanto el llamado **teorema del valor medio** para integrales dobles:

Si $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ es *continua* en el rectángulo R entonces

$$\int_R f = f(x_0, y_0) a(R), \quad \text{para algún } (x_0, y_0) \in R.$$

4.1.3 Teorema de Fubini

Si $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ es una función no negativa, hemos visto que la integral $\int_R f$ es el volumen del sólido limitado por el plano $z = 0$ y la gráfica de f en R . Intuitivamente, dicho volumen se puede calcular integrando el área de una sección del sólido por un plano $x = \text{const.}$, dada por

$$A(x) = \int_c^d f(x, y) dy,$$

respecto de la variable x en $[a, b]$, o bien integrando en $y \in [c, d]$ el área

$$B(y) = \int_a^b f(x, y) dx$$

de las secciones del sólido por planos $y = \text{const.}$ (cf. la Fig. 4.2). Esto sugiere que

$$\int_R f = \int_a^b A(x) dx = \int_c^d B(y) dy,$$

es decir

$$\int_R f = \int_a^b \left(\int_c^d f(x, y) dy \right) dx = \int_c^d \left(\int_a^b f(x, y) dx \right) dy.$$

El resultado anterior, conocido como teorema de Fubini, es cierto en situaciones mucho más generales:

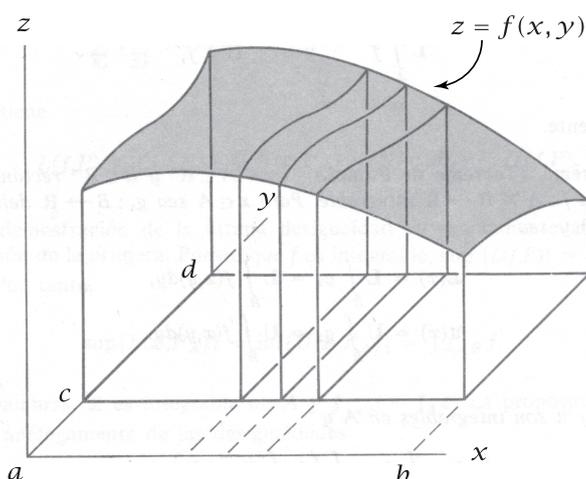


Figura 4.2: secciones del sólido de ecuaciones $0 \leq z \leq f(x, y)$, $(x, y) \in R$, por planos $x = \text{const.}$

Teorema de Fubini. Sea $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ una función integrable en el rectángulo $R = [a, b] \times [c, d]$, y supongamos que la integral

$$A(x) \equiv \int_c^d f(x, y) \, dy$$

existe para todo $x \in [a, b]$. Entonces A es integrable en $[a, b]$ y además $\int_R f = \int_a^b A(x) \, dx$, es decir

$$\int_R f(x, y) \, dx \, dy = \int_a^b \left(\int_c^d f(x, y) \, dy \right) dx. \tag{4.6}$$

Análogamente, si la integral

$$B(y) \equiv \int_a^b f(x, y) \, dx$$

existe para todo $y \in [c, d]$ entonces B es integrable en $[c, d]$ y se cumple $\int_R f = \int_c^d B(y) \, dy$, es decir

$$\int_R f(x, y) \, dx \, dy = \int_c^d \left(\int_a^b f(x, y) \, dx \right) dy. \tag{4.7}$$

Las integrales que aparecen en los miembros derechos de (4.6) y (4.7) se denominan **integrales iteradas** de f en el rectángulo R . En virtud del teorema de Fubini, si tanto $\int_c^d f(x, y) \, dy$ como $\int_a^b f(x, y) \, dx$ existen respectivamente para todo $x \in [a, b]$ y para todo $y \in [c, d]$, las integrales iteradas existen y son *iguales entre sí*, ya que ambas son iguales a la integral doble de f en el rectángulo:

$$\int_a^b \left(\int_c^d f(x, y) \, dy \right) dx = \int_c^d \left(\int_a^b f(x, y) \, dx \right) dy = \int f(x, y) \, dx \, dy. \tag{4.8}$$

Un caso particular importante de esta observación es la siguiente proposición:

Proposición 4.8. Si $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ es continua en el rectángulo R , entonces se verifica la igualdad (4.8).

Demostración. En efecto, si f es continua en el rectángulo R para cada $y \in [c, d]$ fijo la función de una variable dada por $g_y(x) = f(x, y)$ es continua en $[a, b]$ (¿por qué?), y por tanto es integrable en dicho intervalo. En otras palabras, la integral $B(y) \equiv \int_a^b f(x, y) \, dx = \int_a^b g_y(x) \, dx$ existe para todo $y \in [c, d]$. De la misma forma se demuestra que existe la integral $A(x) \equiv \int_c^d f(x, y) \, dy$ para todo $x \in [a, b]$. Por el teorema de Fubini, ambas integrales iteradas coinciden con la integral doble de f en el rectángulo. \square

Notación. Es muy frecuente, sobre todo en textos de Física, utilizar la siguiente notación más compacta para las integrales iteradas de la función f :

$$\int_c^d \left(\int_a^b f(x, y) \, dx \right) dy \equiv \int_c^d dy \int_a^b dx f(x, y), \quad \int_a^b \left(\int_c^d f(x, y) \, dy \right) dx \equiv \int_a^b dx \int_c^d dy f(x, y).$$

Utilizando este convenio, la integral doble de f en el rectángulo R se denota por

$$\int_R dx \, dy f(x, y).$$

Ejemplo 4.9. El teorema de Fubini proporciona un método práctico para evaluar integrales dobles en términos de integrales de funciones de una variable. Nótese sin embargo que, aunque las

integrales iteradas de la función f son ambas iguales a su integral doble en el rectángulo si se cumplen las condiciones de dicho teorema, es frecuente que una de ellas sea más fácil de evaluar que la otra. Consideremos, por ejemplo, la integral

$$I = \int_R |x - y| \, dx \, dy,$$

siendo R el rectángulo $[0, 1] \times [0, 2]$. Como $f(x, y) = |x - y|$ es continua, podemos calcular la integral doble evaluando cualquiera de las integrales iteradas. Integrando primero respecto de y y a continuación respecto de x se obtiene

$$\begin{aligned} I &= \int_0^1 dx \int_0^2 dy |x - y| = \int_0^1 dx \left[\int_0^x (x - y) \, dy + \int_x^2 (y - x) \, dy \right] \\ &= \frac{1}{2} \int_0^1 dx \left(-(x - y)^2 \Big|_{y=0}^{y=x} + (y - x)^2 \Big|_{y=x}^{y=2} \right) = \frac{1}{2} \int_0^1 [x^2 + (x - 2)^2] \, dx \\ &= \frac{1}{6} + \frac{1}{6}(-1 + 8) = \boxed{\frac{4}{3}}. \end{aligned}$$

Integrando en el orden inverso se obtiene, evidentemente, el mismo resultado, aunque de forma algo más laboriosa:

$$\begin{aligned} I &= \int_0^2 dy \int_0^1 dx |x - y| = \int_0^1 dy \left[\int_0^y (y - x) \, dx + \int_y^1 (x - y) \, dx \right] + \int_1^2 dy \int_0^1 (y - x) \, dx \\ &= \frac{1}{2} \int_0^1 dy \left(-(y - x)^2 \Big|_{x=0}^{x=y} + (x - y)^2 \Big|_{x=y}^{x=1} \right) - \frac{1}{2} \int_1^2 dy (y - x)^2 \Big|_{x=0}^{x=1} \\ &= \frac{1}{2} \int_0^1 dy [y^2 + (y - 1)^2] + \frac{1}{2} \int_1^2 dy [y^2 - (y - 1)^2] \\ &= \frac{1}{2} \int_0^2 y^2 \, dy + \frac{1}{2} \int_{-1}^0 t^2 \, dt - \frac{1}{2} \int_0^1 t^2 \, dt = \frac{1}{2} \int_0^2 y^2 \, dy = \frac{4}{3}. \end{aligned}$$

Ejemplo 4.10. Hallemos la integral de la función $f(x, y) = ye^{xy}$ en el cuadrado $[0, 1] \times [0, 1]$. En este caso, al ser

$$ye^{xy} = \frac{\partial}{\partial x} e^{xy},$$

es conveniente integrar primero respecto de x y a continuación respecto de y :

$$\begin{aligned} \int_{[0,1]^2} ye^{xy} \, dx \, dy &= \int_0^1 dy \int_0^1 dx \frac{\partial}{\partial x} (e^{xy}) = \int_0^1 dy e^{xy} \Big|_{x=0}^{x=1} = \int_0^1 (e^y - 1) \, dy = e^y \Big|_{y=0}^{y=1} - 1 \\ &= \boxed{e - 2}. \end{aligned}$$

De nuevo, el cálculo de las integrales iteradas en el orden inverso es bastante más engorroso. En efecto, si $x \neq 0$ integrando por partes se obtiene

$$A(x) \equiv \int_0^1 ye^{xy} \, dy = \frac{y}{x} e^{xy} \Big|_{y=0}^{y=1} - \frac{1}{x} \int_0^1 e^{xy} \, dy = \frac{e^x}{x} - \frac{1}{x^2} e^{xy} \Big|_{y=0}^{y=1} = \frac{1}{x^2} [(x - 1)e^x + 1],$$

mientras que

$$A(0) = \int_0^1 y \, dy = \frac{1}{2} = \lim_{x \rightarrow 0} A(x),$$

y por tanto

$$\int_{[0,1]^2} ye^{xy} \, dx \, dy = \int_0^1 \frac{1}{x^2} [(x - 1)e^x + 1] \, dx.$$

Si bien esta integral se puede evaluar integrando por partes (de hecho, el integrando es la derivada de $(e^x - 1)/x$), su cálculo es bastante más laborioso que el que resulta de integrar en el orden inverso.

4.2 Integral doble en recintos más generales

La integral de una función acotada de dos variables puede definirse sin dificultad en recintos mucho más generales que un rectángulo. Consideremos, en efecto, un conjunto *acotado* $D \subset \mathbb{R}^2$ tal que ∂D sea un conjunto de medida nula. Se dice en tal caso que el conjunto D es **medible Jordan**. Sea χ_D la **función característica** del conjunto D , definida por

$$\chi_D(x, y) = \begin{cases} 1, & (x, y) \in D \\ 0, & (x, y) \notin D. \end{cases}$$

Esta función es obviamente discontinua solamente en los puntos de la frontera de D (ejercicio). Al ser D acotado, existe un rectángulo cerrado $R = [a, b] \times [c, d]$ suficientemente grande tal que $D \subset R$. Diremos entonces que una función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida en D y acotada en dicho conjunto es **integrable** en D si la función $f\chi_D$, dada por

$$(f\chi_D)(x, y) = \begin{cases} f(x, y), & (x, y) \in D \\ 0, & (x, y) \notin D, \end{cases}$$

es integrable en el rectángulo R . (Nótese que $f\chi_D$ es acotada en R , al ser f acotada en D .) Si f es integrable en D , definimos su **integral** en dicho conjunto mediante la fórmula

$$\int_D f = \int_R f\chi_D.$$

Se demuestra fácilmente que la integral de f en el conjunto medible Jordan D no depende del rectángulo R utilizado, sino solo de los valores de f en D .

El conjunto S_R de los puntos de R en que $f\chi_D$ es discontinua está obviamente contenido en $\partial D \cup S_D$, siendo S_D el conjunto de los puntos de D en que f es discontinua (ya que $f\chi_D = 0$ en el exterior de D). Si S_D tiene medida nula, lo mismo ocurrirá con S_R (ya que la unión de dos conjuntos de medida nula tiene medida nula). El recíproco es también obvio, ya que si S_R tiene medida nula también la tendrá S_D , al ser $S_D \subset S_R$. En otras palabras, S_R tiene medida 0 si y solo si S_D tiene medida 0. Del criterio de integrabilidad de Lebesgue en un rectángulo se sigue entonces el resultado análogo para integrales en conjuntos medibles Jordan:

Proposición 4.11. Sea $D \subset \mathbb{R}^2$ un conjunto medible Jordan, y sea $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ una función definida y acotada en D . Entonces f es integrable en D si y solo si el conjunto de los puntos de D en que f es discontinua tiene medida 0.

De la proposición anterior se deduce el siguiente corolario:

Si f es continua en un conjunto medible Jordan cerrado D entonces f es integrable en dicho conjunto.

En efecto, si D es medible Jordan es por definición acotado, y al ser también cerrado es compacto. Obviamente, el resultado anterior sigue siendo válido si D no es necesariamente cerrado pero f es continua y *acotada* en D . En particular, la función constante 1 (al igual que cualquier función constante) es integrable en cualquier conjunto medible Jordan D , y su integral es el **área** $a(D)$ de dicho conjunto:

$$a(D) = \int_D dx dy.$$

Para justificar esta definición, nótese que si $R = [a, b] \times [c, d] \supset D$ y P es cualquier partición de R entonces

$$\begin{cases} L(\chi_D; P) = \text{suma de las áreas de los rectángulos de } P \text{ enteramente contenidos en } D, \\ U(\chi_D; P) = \text{suma de las áreas de los rectángulos de } P \text{ que cortan a } D. \end{cases}$$

Es por tanto intuitivamente obvio que el área de la región D debería cumplir $L(\chi_D; P) \leq a(D) \leq U(\chi_D; P)$, de donde se sigue que

$$\int_R \chi_D \leq a(D) \leq \int_R \chi_d.$$

Como la función 1 es integrable en D , de las desigualdades anteriores se deduce que

$$\int_D 1 = \int_R \chi_D = \int_R \chi_D = \int_R \chi_d = a(D).$$

La integral doble en conjuntos medibles Jordan tiene propiedades análogas a la integral en un rectángulo. Más precisamente, las propiedades i) y iii) de la integral doble vistas anteriormente (es decir, su linealidad y monotonía) siguen siendo ciertas en este caso cambiando el rectángulo R por la región D . También es cierta la propiedad de aditividad, reemplazando los rectángulos R_i por conjuntos medibles Jordan D_i tales que $D_i \cap D_j$ tiene medida 0 para todo $i \neq j$. Por último, son también válidas todas las consecuencias de la propiedad de monotonía de la integral doble deducidas anteriormente (cf. las ecs. (4.4) y (4.5)), sin más que cambiar el rectángulo R por el conjunto medible Jordan D . Si, además, D es cerrado y conexo, sigue siendo válido el *teorema del valor medio* (con la misma demostración vista anteriormente).

Comentario. Si $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ es integrable y no negativa en D , el volumen del sólido

$$\Omega = \{(x, y, z) \in \mathbb{R}^3 \mid (x, y) \in D, 0 \leq z \leq f(x, y)\}$$

está dado por

$$v(\Omega) = \int_D f.$$

En efecto, esta integral es igual al volumen del sólido limitado por el plano $z = 0$ y la gráfica de la función $f\chi_D$ en un rectángulo $R \supset D$, y el volumen de la porción de dicho sólido correspondiente a puntos $(x, y) \notin D$ es nulo al ser la altura del sólido nula en tales puntos. Más generalmente, el volumen del sólido

$$\Omega = \{(x, y, z) \in \mathbb{R}^3 \mid (x, y) \in D, f_1(x, y) \leq z \leq f_2(x, y)\}$$

donde las funciones $f_{1,2}$ son integrables en D y $f_1(x, y) \leq f_2(x, y)$ para todo $(x, y) \in D$, está dado por

$$v(\Omega) = \int_D (f_2 - f_1). \quad (4.9)$$

La mayoría de los conjuntos medibles Jordan en \mathbb{R}^2 que se presentan en la práctica son de alguno de los tres tipos sencillos que discutiremos a continuación. Como veremos, estos conjuntos tienen en común que su frontera está formada por la unión de un número finito de *gráficas* de funciones continuas (tanto de la forma $y = \phi(x)$ como de la forma $x = \psi(y)$) en intervalos compactos. Un conjunto acotado con esta propiedad es automáticamente medible Jordan, ya que la gráfica de una función continua en un intervalo compacto tiene medida 0.

Más concretamente, diremos que un subconjunto $D \subset \mathbb{R}^2$ es **de tipo I** si está dado por las ecuaciones

$$a \leq x \leq b, \quad \phi_1(x) \leq y \leq \phi_2(x),$$

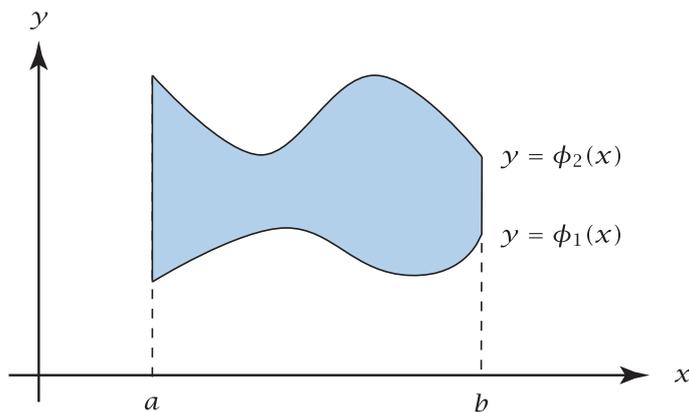


Figura 4.3: Región de tipo I.

siendo $\phi_{1,2} : \mathbb{R} \rightarrow \mathbb{R}$ funciones de una variable continuas en el intervalo compacto $[a, b]$ tales que $\phi_1(x) \leq \phi_2(x)$ para todo $x \in [a, b]$ (cf. la Fig. 4.3). Evidentemente, un conjunto de este tipo es *acotado*, ya que $a \leq x \leq b$ y $m \leq y \leq M$, siendo m y M respectivamente el mínimo de $\phi_1(x)$ y el máximo de $\phi_2(x)$ en el intervalo compacto $[a, b]$. Por otra parte, la frontera de una región de tipo I es la unión de los segmentos verticales

$$\{(a, y) \mid \phi_1(a) \leq y \leq \phi_2(a)\}, \quad \{(b, y) \mid \phi_1(b) \leq y \leq \phi_2(b)\}$$

junto con las dos gráficas

$$\{(x, \phi_i(x)) \mid a \leq x \leq b\}, \quad i = 1, 2,$$

y es por tanto un conjunto de medida nula (nótese que los segmentos verticales son las gráficas de las funciones constantes, y por tanto continuas, $x = a$ y $x = b$, respectivamente en los intervalos $[\phi_1(a), \phi_2(a)]$ y $[\phi_1(b), \phi_2(b)]$). Por tanto *todo conjunto de tipo I es medible Jordan*. La integral de una función integrable $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ en un conjunto de tipo I se puede calcular aplicando el teorema de Fubini a la función $f\chi_D$. Para ello basta tener en cuenta que si $R = [a, b] \times [c, d]$ es un rectángulo que contiene a D entonces

$$A(x) \equiv \int_c^d f(x, y)\chi_D(x, y) dy = \int_{\phi_1(x)}^{\phi_2(x)} f(x, y) dy,$$

ya que $\chi_D(x, y) = 0$ fuera de D . Si la integral del miembro derecho de la ecuación anterior existe para todo $x \in [a, b]$ (por ejemplo, si f es continua en D), entonces del teorema de Fubini se sigue que

$$\int_D f = \int_a^b dx \int_{\phi_1(x)}^{\phi_2(x)} f(x, y) dy. \tag{4.10}$$

La fórmula anterior proporciona un procedimiento práctico para calcular la integral doble de f en recintos de tipo I. En particular, el área de una región de este tipo está dada por

$$a(D) = \int_D 1 = \int_a^b dx \int_{\phi_1(x)}^{\phi_2(x)} dy = \int_a^b [\phi_2(x) - \phi_1(x)] dx,$$

resultado conocido del curso de Análisis de funciones de una variable.

Ejemplo 4.12. Como ejemplo, calcularemos el volumen del elipsoide Ω de semiejes a, b y c integrando una función adecuada en una región de tipo I. La ecuación del elipsoide puede tomarse sin pérdida de generalidad como

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leq 1,$$

o equivalentemente

$$(x, y) \in D, \quad -f(x, y) \leq z \leq f(x, y),$$

donde D es la elipse de ecuación

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1$$

y

$$f(x, y) = c\sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}}.$$

En virtud de la ec. (4.9), el volumen de Ω está dado por

$$v(\Omega) = 2 \int_D f(x, y) \, dx \, dy \equiv 2c \int_D \sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}} \, dx \, dy.$$

Para evaluar la integral, observamos que la región D es de tipo I, ya que se puede parametrizar como

$$-a \leq x \leq a, \quad -b\sqrt{1 - \frac{x^2}{a^2}} \leq y \leq b\sqrt{1 - \frac{x^2}{a^2}}.$$

Utilizando la ec. (4.10) con

$$\phi_2(x) = -\phi_1(x) = b\sqrt{1 - \frac{x^2}{a^2}}$$

se obtiene

$$v(\Omega) = 2c \int_{-a}^a dx \int_{-b\sqrt{1 - \frac{x^2}{a^2}}}^{b\sqrt{1 - \frac{x^2}{a^2}}} dy \sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}} = 8c \int_0^a dx \int_0^{b\sqrt{1 - \frac{x^2}{a^2}}} dy \sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}}.$$

Para evaluar la integral en y , efectuamos primero el cambio de variable

$$y = b\sqrt{1 - \frac{x^2}{a^2}} t, \quad 0 \leq t \leq 1,$$

en términos del cual

$$\int_0^{b\sqrt{1 - \frac{x^2}{a^2}}} \sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}} \, dy = b \left(1 - \frac{x^2}{a^2}\right) \int_0^1 \sqrt{1 - t^2} \, dt = \frac{\pi}{4} b \left(1 - \frac{x^2}{a^2}\right),$$

donde la integral de $\sqrt{1 - t^2}$ se ha calculado utilizando el cambio de variable $t = \sin \theta$:

$$\int_0^1 \sqrt{1 - t^2} \, dt = \int_0^{\pi/2} \cos^2 \theta \, d\theta = \frac{1}{2} \int_0^{\pi/2} (1 + \cos(2\theta)) \, d\theta = \frac{\pi}{4}.$$

De esta forma obtenemos

$$v(\Omega) = 2\pi bc \int_0^a \left(1 - \frac{x^2}{a^2}\right) dx = 2\pi abc \int_0^1 (1 - t^2) dt = \boxed{\frac{4}{3} \pi abc}. \quad \square$$

Análogamente, diremos que un conjunto $D \subset \mathbb{R}^2$ es **de tipo II** si es de tipo I intercambiando el papel de las coordenadas x e y . Más precisamente, una región de tipo II está descrita por ecuaciones de la forma

$$\boxed{c \leq y \leq d, \quad \phi_1(y) \leq x \leq \phi_2(y)},$$

donde las funciones de una variable $\phi_{1,2}$ son continuas en el intervalo $[c, d]$ y verifican $\phi_1(y) \leq \phi_2(y)$ para todo $y \in [c, d]$ (cf. la Fig. 4.4). Al igual que antes, un conjunto de este tipo es *acotado*, ya que $m \leq x \leq M$ y $c \leq y \leq d$, siendo m y M respectivamente el mínimo de $\phi_1(y)$ y el máximo

de $\phi_2(y)$ en el intervalo compacto $[c, d]$. Además, la frontera de D tiene medida 0, ya que es la unión de los segmentos horizontales

$$\{(x, c) \mid \phi_1(c) \leq x \leq \phi_2(c)\}, \quad \{(x, d) \mid \phi_1(d) \leq x \leq \phi_2(d)\}$$

junto con las gráficas de las funciones continuas $x = \phi_1(y)$ y $x = \phi_2(y)$ en el intervalo compacto $c \leq y \leq d$. Por tanto *una región de tipo II es automáticamente medible Jordan*. De nuevo, la integral de una función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ en un conjunto de tipo II puede calcularse aplicando el teorema de Fubini a la función $f\chi_D$. En efecto, si f es integrable en D y $R = [a, b] \times [c, d]$ es un rectángulo que contiene a D entonces

$$B(y) = \int_a^b f(x, y)\chi_D(x, y) dx = \int_{\phi_1(y)}^{\phi_2(y)} f(x, y) dx.$$

Si esta integral existe para todo $y \in [c, d]$ (por ejemplo, si f es continua en D), el teorema de Fubini garantiza que

$$\int_D f = \int_c^d dy \int_{\phi_1(y)}^{\phi_2(y)} f(x, y) dx. \tag{4.11}$$

En particular, el área de una región D de tipo II está dada por

$$a(D) = \int_D 1 = \int_c^d dy \int_{\phi_1(y)}^{\phi_2(y)} f(x, y) dx = \int_c^d [\phi_2(y) - \phi_1(y)] dy,$$

fórmula que de nuevo coincide con el resultado elemental conocido.

Finalmente, diremos que una región D es **de tipo III** si es a la vez de tipos I y II; en particular, toda región de tipo III es automáticamente medible Jordan. Por definición, una región de tipo III se puede describir simultáneamente por ecuaciones del tipo

$$a \leq x \leq b, \quad \phi_1(x) \leq y \leq \phi_2(x)$$

o

$$c \leq y \leq d, \quad \psi_1(y) \leq x \leq \psi_2(y),$$

con $\phi_{1,2}, \psi_{1,2}$ funciones de una variable continuas respectivamente en $[a, b]$ y $[c, d]$ tales que $\phi_1(x) \leq \phi_2(x)$ para todo $x \in [a, b]$ y $\psi_1(y) \leq \psi_2(y)$ para todo $y \in [c, d]$. Por ejemplo, la elipse D de ecuación

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 \leq 1$$

es una región de tipo III, ya que

$$\begin{aligned} D &= \left\{ (x, y) \in \mathbb{R}^2 \mid -a \leq x \leq a, -b\sqrt{1 - \frac{x^2}{a^2}} \leq y \leq b\sqrt{1 - \frac{x^2}{a^2}} \right\} \\ &= \left\{ (x, y) \in \mathbb{R}^2 \mid -b \leq y \leq b, -a\sqrt{1 - \frac{y^2}{b^2}} \leq x \leq a\sqrt{1 - \frac{y^2}{b^2}} \right\}. \end{aligned}$$

Si f es integrable en una región D de tipo III, y están definidas ambas integrales

$$A(x) = \int_{\phi_1(x)}^{\phi_2(x)} f(x, y) dy, \quad B(y) = \int_{\psi_1(y)}^{\psi_2(y)} f(x, y) dx$$

respectivamente para todo $x \in [a, b]$ y para todo $y \in [c, d]$ (esto ocurrirá, por ejemplo, si f es continua en D), el teorema de Fubini asegura que

$$\int_D f = \int_a^b dx \int_{\phi_1(x)}^{\phi_2(x)} dy f(x, y) = \int_c^d dy \int_{\psi_1(y)}^{\psi_2(y)} dx f(x, y).$$

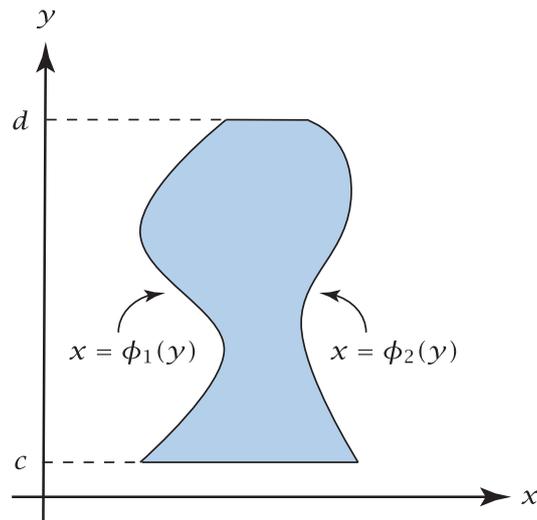
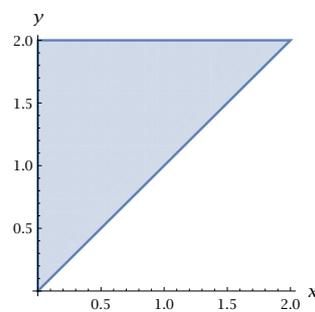


Figura 4.4: Región de tipo II.

Figura 4.5: Triángulo D del Ejemplo 4.13.

Nótese, sin embargo, que es frecuente en la práctica que una de las dos integrales iteradas sea mucho más difícil de calcular que la otra, o incluso que una de dichas integrables sea calculable elementalmente y la otra no lo sea.

Ejemplo 4.13. Sea $A(x)$ la función de una variable dada por

$$A(x) = \int_x^2 e^{y^2} dy.$$

Esta función no se puede expresar en términos de funciones elementales, ya que puede probarse que la primitiva de e^{y^2} *no* es una función elemental. Sin embargo, la integral $\int_0^2 A(x) dx$ se calcula fácilmente teniendo en cuenta que por el teorema de Fubini

$$\int_0^2 A(x) dx \equiv \int_0^2 dx \int_x^2 dy e^{y^2} = \int_D e^{y^2} dx dy,$$

siendo D el conjunto (triángulo) de ecuaciones

$$0 \leq x \leq 2, \quad x \leq y \leq 2$$

(cf. la Fig. 4.5). El conjunto D es de tipo III, ya que obviamente puede también describirse mediante las ecuaciones de tipo II

$$0 \leq y \leq 2, \quad 0 \leq x \leq y.$$

Calculando la integral doble considerando a D como un conjunto de tipo II se obtiene

$$\int_0^2 A(x) dx = \int_0^2 dy \int_0^y dx e^{y^2} = \int_0^2 y e^{y^2} dy = \frac{1}{2} e^{y^2} \Big|_{y=0}^{y=2} = \boxed{\frac{1}{2} (e^4 - 1)}.$$

Ejemplo 4.14. Calculemos la integral de la función $f(x, y) = x^3 y$ en el conjunto D acotado por el eje y y la parábola $x = 3 - 4y^2$. esta región es de tipo III, ya que puede describirse por las ecuaciones

$$0 \leq x \leq 3, \quad -\frac{1}{2}\sqrt{3-x} \leq y \leq \frac{1}{2}\sqrt{3-x}$$

(de tipo I), o bien

$$-\frac{\sqrt{3}}{2} \leq y \leq \frac{\sqrt{3}}{2}, \quad 0 \leq x \leq 3 - 4y^2$$

(de tipo II). Escogiendo la parametrización de tipo II y aplicando el teorema de Fubini obtenemos

$$\int_D x^3 y \, dx \, dy = \int_{-\frac{\sqrt{3}}{2}}^{\frac{\sqrt{3}}{2}} dy \int_0^{3-4y^2} x^3 y \, dx = \frac{1}{4} \int_{-\frac{\sqrt{3}}{2}}^{\frac{\sqrt{3}}{2}} y(3-4y^2)^4 \, dy = 0,$$

ya que el integrando es *impar*. En general, si D es *invariante* bajo la transformación $(x, y) \mapsto (x, -y)$ y $f(x, -y) = -f(x, y)$, utilizando la fórmula del cambio de variables en integrales múltiples (cf. la Sección 4.5) se demuestra que $\int_D f = 0$.

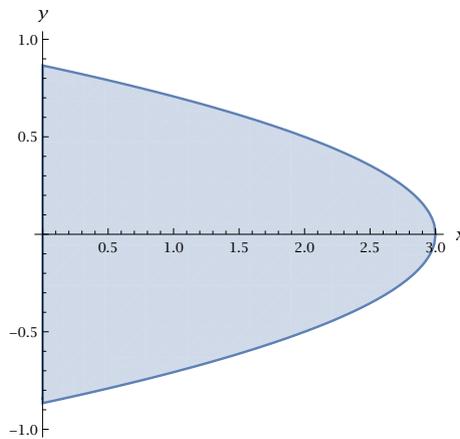


Figura 4.6: Región D limitada por el eje y y la parábola $x = 3 - 4y^2$.

4.3 Integrales múltiples en \mathbb{R}^n

4.3.1 Integral múltiple en un n -paralelepípedo

La integral de una función escalar *acotada* $f : \mathbb{R}^n \rightarrow \mathbb{R}$ se define de forma totalmente análoga a la integral doble que acabamos de discutir. En primer lugar, definimos la integral de la función f en un n -paralelepípedo $R \equiv [a_1, b_1] \times \cdots \times [a_n, b_n]$. Para ello, suponemos que f está definida y acotada en R y definimos las sumas superior e inferior asociadas a una partición $P = P_1 \times \cdots \times P_n$ de R , donde cada P_i es una partición del correspondiente intervalo $[a_i, b_i]$ en m_i subintervalos, al igual que en el caso bidimensional. Más precisamente,

$$L(f; P) = \sum_{S \in P} m(f; S) v(S), \quad U(f; P) = \sum_{S \in P} M(f; S) v(S),$$

siendo S uno de los $m_1 \cdots m_n$ sub-paralelepípedos de la partición P , $v(S)$ su volumen y

$$m(f; S) = \inf\{f(x) \mid x \in S\}, \quad M(f; S) = \sup\{f(x) \mid x \in S\}.$$

Al igual que en el caso $n = 2$, se demuestra que

$$L(f; P) \leq U(f; P')$$

para todo par de particiones P, P' de R . Por tanto podemos definir las integrales superior e inferior de f en R respectivamente mediante

$$\int_R f = \sup\{L(f;P) \mid P \text{ partición de } R\}, \quad \overline{\int}_R f = \inf\{U(f;P) \mid P \text{ partición de } R\},$$

y se tiene

$$\int_R f \leq \overline{\int}_R f.$$

Por definición, f es **integrable** en R si la integral superior de f en dicho conjunto coincide con su integral inferior. En tal caso, definimos la **integral** de f en el n -paralelepípedo R como el valor común de ambas integrales, es decir

$$\int_R f = \int_R f = \overline{\int}_R f.$$

Notación. La integral de f en R se suele denotar de distintas formas alternativas, como por ejemplo

$$\int_R f(x) dx_1 \cdots dx_n, \quad \int_R dx_1 \cdots dx_n f(x), \quad \int_R f(x) d^n x, \quad \int_R d^n x f(x). \quad \square$$

Al igual que en el caso bidimensional, es válido el *criterio de integrabilidad de Lebesgue*:

Una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ definida y acotada en el n -paralelepípedo R es *integrable* en dicho conjunto si y solo si el conjunto de los puntos de R en que f es discontinua tiene *medida nula*.

En particular:

Si $f : \mathbb{R}^n \rightarrow \mathbb{R}$ es *continua* en el n -paralelepípedo R , entonces f es *integrable* en dicho conjunto.

Como en dos dimensiones, la *gráfica* de una función continua $g : \mathbb{R}^{n-1} \rightarrow \mathbb{R}$ en un conjunto compacto K tiene medida 0 en \mathbb{R}^n , y por tanto la unión *numerable* de tales gráficas tiene también medida nula. Por último, las propiedades elementales de la integral múltiple son idénticas a las del caso bidimensional, sin más que cambiar el rectángulo R por el n -paralelepípedo R (o los rectángulos R_i en la propiedad de aditividad por n -paralelepípedos). En otras palabras, las integrales múltiples tienen las propiedades de linealidad, aditividad y monotonía, y se verifican las desigualdades (4.4) y (4.5) (con $a(R)$ reemplazado por $v(R)$) y el teorema del valor medio (si f es continua):

$$\int_R f = f(x_0) v(R), \quad \text{para algún } x_0 \in R.$$

También el *teorema de Fubini* admite una generalización relativamente sencilla a integrales múltiples n -dimensionales. En efecto, sean

$$R_m \equiv [a_1, b_1] \times \cdots \times [a_m, b_m] \subset \mathbb{R}^m, \quad R_{n-m} \equiv [a_{m+1}, b_{m+1}] \times \cdots \times [a_n, b_n] \subset \mathbb{R}^{n-m},$$

de modo que $R = R_m \times R_{n-m}$, y denotemos por (x, y) (con $x \in R_m$ e $y \in R_{n-m}$) un punto genérico del n -paralelepípedo R . Si f es integrable en R y existe la integral

$$A(x) \equiv \int_{R_{n-m}} f(x, y) d^{n-m} y$$

para todo $x \in R_m$ (por ejemplo, si f es continua en R) entonces A es integrable en R_m y además $\int_R f = \int_{R_m} A$, es decir

$$\int_R f = \int_{R_m} d^m x \int_{R_{n-m}} d^{n-m} y f(x, y).$$

Análogamente, si existe

$$B(\gamma) \equiv \int_{R_m} f(x, \gamma) d^m x$$

para todo $\gamma \in R_{n-m}$ entonces B es integrable en R_{n-m} y además $\int_R f = \int_{R_{n-m}} B$, es decir

$$\int_R f = \int_{R_{n-m}} d^{n-m} \gamma \int_{R_m} d^m x f(x, \gamma).$$

Las integrales del miembro derecho de las igualdades anteriores se llaman **integrales iteradas** de la función f . Evidentemente, si se cumplen las hipótesis del teorema de Fubini tanto para $A(x)$ como para $B(\gamma)$ (por ejemplo, si f es continua en R) entonces ambas integrales iteradas son iguales entre sí, ya que cualquiera de ellas es igual a la integral de f en R :

$$\int_R f = \int_{R_{n-m}} d^{n-m} \gamma \int_{R_m} d^m x f(x, \gamma) = \int_{R_m} d^m x \int_{R_{n-m}} d^{n-m} \gamma f(x, \gamma).$$

Aplicando el teorema de Fubini repetidamente (suponiendo que se cumplen en cada paso la hipótesis de dicho teorema, lo que ocurrirá por ejemplo si f es continua en R) se demuestra que la integral múltiple de f en el n -paralelepípedo $R = [a_1, b_1] \times \cdots \times [a_n, b_n]$ se puede calcular evaluando sucesivamente n integrales unidimensionales:

$$\int_R f = \int_{a_1}^{b_1} dx_1 \int_{a_2}^{b_2} dx_2 \cdots \int_{a_n}^{b_n} dx_n f(x_1, \dots, x_n).$$

De hecho, si f es continua en R podemos evaluar las integrales iteradas *en cualquier orden*, es decir

$$\int_R f = \int_{a_{i_1}}^{b_{i_1}} dx_{i_1} \int_{a_{i_2}}^{b_{i_2}} dx_{i_2} \cdots \int_{a_{i_n}}^{b_{i_n}} dx_{i_n} f(x_1, \dots, x_n).$$

siendo (i_1, \dots, i_n) cualquier permutación de $(1, \dots, n)$. En particular, en este caso las $n!$ integrales iteradas unidimensionales de f son iguales entre sí.

4.3.2 Integral múltiple en un conjunto medible Jordan

A continuación se define la integral de una función escalar acotada $f : \mathbb{R}^n \rightarrow \mathbb{R}$ en un conjunto **medible Jordan** $\Omega \subset \mathbb{R}^n$, que por definición es un conjunto *acotado* cuya frontera tiene *medida nula*. Como en el caso bidimensional, se dice que una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ definida y acotada en un conjunto medible Jordan Ω es **integrable** en dicho conjunto si $f\chi_\Omega$ es integrable en un n -paralelepípedo R que contiene al conjunto Ω . En tal caso, la **integral** de f en Ω se define de nuevo mediante

$$\int_\Omega f = \int_R f\chi_\Omega.$$

Al igual que en el caso bidimensional (y esencialmente con la misma demostración), se verifica el *criterio de integrabilidad de Lebesgue*:

Una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ definida y acotada en un conjunto medible Jordan Ω es *integrable* en dicho conjunto si y solo si el conjunto de los puntos de Ω en que f es discontinua tiene *medida nula*.

En particular,

Si $f : \mathbb{R}^n \rightarrow \mathbb{R}$ es continua en un conjunto medible Jordan cerrado Ω (o si f es continua y acotada en un conjunto medible Jordan Ω arbitrario) entonces f es integrable en dicho conjunto.

En particular, las funciones constantes son integrables en cualquier conjunto medible Jordan Ω . Por definición (análogamente a lo que ocurre en el caso bidimensional), se define el **volumen** $v(\Omega)$ de un conjunto medible Jordan Ω mediante la fórmula

$$v(\Omega) = \int_{\Omega} 1 \equiv \int_{\Omega} d^n x.$$

De nuevo, las propiedades de la integral múltiple de una función en un conjunto medible Jordan Ω (linealidad, aditividad, monotonía, teorema del valor medio, etc.) son esencialmente las mismas que en el caso bidimensional.

Ejemplo 4.15. Calculemos, como ejemplo, el volumen V_n del simplex n -dimensional de ecuaciones

$$x_1 + \dots + x_n \leq 1, \quad x_i \geq 0 \quad (i = 1, \dots, n).$$

Este conjunto es evidentemente acotado (en efecto, $0 \leq x_i \leq 1$ para todo $i = 1, \dots, n$), y su frontera está contenida en la unión de los hiperplanos $x_i = 0$ ($1 \leq i \leq n$) y $x_1 + \dots + x_n = 1$, cada uno de los cuales tiene medida cero. Por tanto el simplex es un conjunto medible Jordan. En dos dimensiones, dicho simplex es un triángulo rectángulo con ambos catetos iguales a 1, y por tanto $V_2 = 1/2$. Si $n = 3$, el 3-simplex es la pirámide Ω de ecuaciones

$$x + y + z \leq 1, \quad x, y, z \geq 0 \tag{4.12}$$

(cf. la Fig. 4.7). El volumen de esta pirámide se calcula fácilmente aplicando el teorema de Fubini al cubo $[0, 1]^3$, que claramente contiene a la pirámide. En efecto,

$$v(\Omega) = \int_{\Omega} dx dy dz = \int_{[0,1]^3} \chi_{\Omega} = \int_0^1 dx \int_0^1 dy \int_0^1 dz \chi_{\Omega}(x, y, z) = \int_0^1 dx \int_0^{1-x} dy \int_0^{1-x-y} dz,$$

ya que $\chi_{\Omega}(x, y, z) = 0$ a menos que (x, y, z) verifique las ecuaciones (4.12), que evidentemente son equivalentes a

$$0 \leq x \leq 1, \quad 0 \leq y \leq 1 - x, \quad 0 \leq z \leq 1 - x - y.$$

Nótese que podemos aplicar el teorema de Fubini, ya que todas las integrales iteradas que aparecen en la fórmula anterior obviamente existen. Realizando las integraciones se obtiene

$$V_3 \equiv v(\Omega) = \int_0^1 dx \int_0^{1-x} dy (1 - x - y) = \frac{1}{2} \int_0^1 (1 - x)^2 dx = \frac{1}{2} \int_0^1 x^2 dx = \frac{1}{6}.$$

Esto coincide con el conocido resultado de geometría elemental, según el cual el volumen de una pirámide es $1/3$ de la altura por el área de la base. Procediendo de la misma forma en el caso general n -dimensional y utilizando la identidad

$$\int_0^a (a - t)^k dt = \frac{a^{k+1}}{k + 1}, \quad k \in \mathbb{N},$$

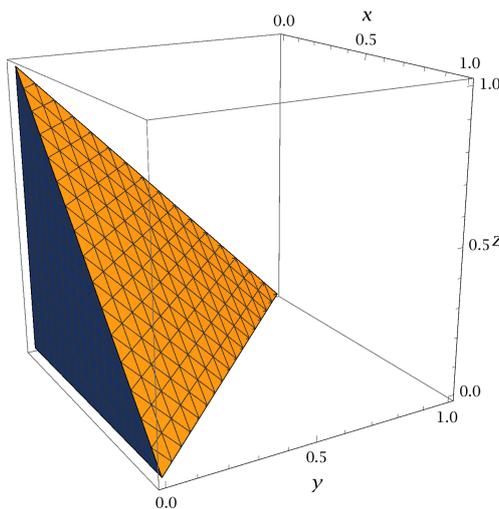


Figura 4.7: Pirámide del Ejemplo 4.15.

se obtiene

$$\begin{aligned}
 V_n &= \int_0^1 dx_1 \int_0^{1-x_1} dx_2 \cdots \int_0^{1-x_1-\cdots-x_{n-2}} dx_{n-1} \int_0^{1-x_1-\cdots-x_{n-1}} dx_n \\
 &= \int_0^1 dx_1 \int_0^{1-x_1} dx_2 \cdots \int_0^{1-x_1-\cdots-x_{n-2}} dx_{n-1} (1-x_1-\cdots-x_{n-1}) \\
 &= \frac{1}{2} \int_0^1 dx_1 \int_0^{1-x_1} dx_2 \cdots \int_0^{1-x_1-\cdots-x_{n-3}} dx_{n-2} (1-x_1-\cdots-x_{n-2})^2 \\
 &= \frac{1}{2 \cdot 3} \int_0^1 dx_1 \int_0^{1-x_1} dx_2 \cdots \int_0^{1-x_1-\cdots-x_{n-4}} dx_{n-3} (1-x_1-\cdots-x_{n-3})^3 \\
 &= \cdots = \frac{1}{(n-1)!} \int_0^1 dx_1 (1-x_1)^{n-1} = \boxed{\frac{1}{n!}}.
 \end{aligned}$$

4.4 Integrales triples

Estudiaremos a continuación en más detalle el caso de las integrales múltiples en \mathbb{R}^3 , llamadas también *integrales triples*. De forma análoga al caso bidimensional, se pueden definir cuatro tipos de regiones sencillas que se presentan frecuentemente en la práctica, a las cuales puede aplicarse fácilmente el teorema de Fubini.

Más precisamente, diremos que un conjunto $\Omega \subset \mathbb{R}^3$ es **de tipo I** si está dado por ecuaciones del tipo

$$\boxed{(x, y) \in D, \quad \phi_1(x, y) \leq z \leq \phi_2(x, y)}$$

siendo $\phi_{1,2}$ funciones de dos variables continuas en el conjunto medible Jordan compacto $D \subset \mathbb{R}^2$ tales que $\phi_1(x, y) \leq \phi_2(x, y)$ para todo $(x, y) \in D$. Una región de este tipo es automáticamente *medible Jordan*. En efecto, Ω es acotado, ya que las variables (x, y) están acotadas (al ser D compacto por hipótesis), mientras que si m y M denotan respectivamente el mínimo de ϕ_1 y el máximo de ϕ_2 en el conjunto compacto D se tiene $m \leq z \leq M$. Por otra parte, la frontera de Ω es la unión de las gráficas de las funciones $\phi_{1,2}$ en el conjunto compacto D , que tienen medida 0, con el conjunto

$$\{(x, y, z) \mid (x, y) \in \partial D, \phi_1(x, y) \leq z \leq \phi_2(x, y)\} \subset \partial D \times [m, M],$$

que claramente tiene medida 0 al ser por hipótesis ∂D un conjunto de medida nula en \mathbb{R}^2 (ejercicio). Sea $R = [a, b] \times [c, d] \times [m, M]$ un paralelepípedo que contiene a Ω (en particular, esto

implica que $D \subset [a, b] \times [c, d]$). Por el teorema de Fubini, si existe la integral

$$A(x, y) = \int_m^M f(x, y, z) \chi_\Omega(x, y, z) dz = \chi_D(x, y) \int_{\phi_1(x, y)}^{\phi_2(x, y)} f(x, y, z) dz$$

(por ejemplo, si f es continua) entonces

$$\int_\Omega f = \int_R f \chi_\Omega = \int_{[a, b] \times [c, d]} A(x, y) dx dy = \int_{[a, b] \times [c, d]} dx dy \chi_D(x, y) \int_{\phi_1(x, y)}^{\phi_2(x, y)} dz f(x, y, z),$$

es decir

$$\int_\Omega f = \int_D dx dy \int_{\phi_1(x, y)}^{\phi_2(x, y)} dz f(x, y, z). \tag{4.13}$$

En particular, el volumen de una región Ω de tipo I está dado por

$$v(\Omega) = \int_D dx dy \int_{\phi_1(x, y)}^{\phi_2(x, y)} dz = \int_D [\phi_2(x, y) - \phi_1(x, y)] dx dy,$$

como ya vimos en la sección anterior (cf. la ec. (4.9)). Evidentemente, si la región D es a su vez un conjunto de tipo I o II en \mathbb{R}^2 la integral (4.13) se puede expresar en términos de tres integrales sucesivas de funciones de una variable. Por ejemplo, si D es de tipo I y está determinada por las ecuaciones

$$a \leq x \leq b, \quad \psi_1(x) \leq y \leq \psi_2(x)$$

entonces se tiene

$$\int_\Omega f = \int_a^b dx \int_{\psi_1(x)}^{\psi_2(x)} dy \int_{\phi_1(x, y)}^{\phi_2(x, y)} dz f(x, y, z). \tag{4.14}$$

Análogamente, diremos que $\Omega \subset \mathbb{R}^3$ es **de tipo II** si es de tipo I intercambiando el papel de las variables x y z , es decir si está descrita por ecuaciones de la forma

$$(y, z) \in D, \quad \phi_1(y, z) \leq x \leq \phi_2(y, z)$$

siendo $\phi_{1,2}$ funciones de dos variables continuas en el conjunto medible Jordan compacto $D \subset \mathbb{R}^2$ tales que $\phi_1(y, z) \leq \phi_2(y, z)$ para todo $(y, z) \in D$. Evidentemente, en este caso la ecuación (4.13) ha de ser reemplazada por

$$\int_\Omega f = \int_D dy dz \int_{\phi_1(y, z)}^{\phi_2(y, z)} dx f(x, y, z). \tag{4.15}$$

En particular, en este caso el volumen de Ω está dado por

$$v(\Omega) = \int_D [\phi_2(y, z) - \phi_1(y, z)] dy dz.$$

Del mismo modo, diremos que $\Omega \subset \mathbb{R}^3$ es **de tipo III** si es de tipo I intercambiando el papel de las variables y y z , es decir si está descrita por ecuaciones de la forma

$$(x, z) \in D, \quad \phi_1(x, z) \leq y \leq \phi_2(x, z),$$

siendo como antes $\phi_{1,2}$ funciones de dos variables continuas en el conjunto medible Jordan compacto $D \subset \mathbb{R}^2$ tales que $\phi_1(x, z) \leq \phi_2(x, z)$ para todo $(x, z) \in D$. En este caso

$$\int_\Omega f = \int_D dx dz \int_{\phi_1(x, z)}^{\phi_2(x, z)} dy f(x, y, z), \tag{4.16}$$

y en particular

$$v(\Omega) = \int_D [\phi_2(x, z) - \phi_1(x, z)] dx dz.$$

Por último, diremos que una región $\Omega \subset \mathbb{R}^3$ es **de tipo IV** si es simultáneamente de los tres tipos I-III. Evidentemente, la integral de una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ en una región de este tipo se puede calcular por cualquiera de las tres expresiones (4.13)-(4.16) (si se cumplen las hipótesis del teorema de Fubini).

Ejemplo 4.16. Calculemos la integral de la función $f(x, y, z) = z$ en la región Ω del primer octante ($x, y, z \geq 0$) acotada por los planos $x = 0, y = 0, z = 0, x + y = 2, x + 2y = 6$ y el cilindro $y^2 + z^2 = 4$. Dicha región ha de estar en el interior $y^2 + z^2 \leq 4$ del cilindro, ya que su exterior no es acotado. Por tanto Ω es la región de tipo I

$$(x, y) \in D, \quad 0 \leq z \leq \sqrt{4 - y^2},$$

(cf. la Fig. 4.8, drcha.) siendo $D \subset \mathbb{R}^2$ el conjunto de ecuaciones

$$x \geq 0, \quad 0 \leq y \leq 2, \quad x + y \geq 2, \quad x + 2y \leq 6$$

(cf. la Fig. 4.8, izda.). Nótese que la restricción $y \leq 2$ proviene de la ecuación del cilindro, ya que $|y| \leq 2$ si $y^2 + z^2 \leq 4$. Podemos considerar a D como una región de tipo II, de ecuaciones

$$0 \leq y \leq 2, \quad 2 - y \leq x \leq 6 - 2y.$$

Por tanto

$$\begin{aligned} \int_{\Omega} z dx dy dz &= \int_0^2 dy \int_{2-y}^{6-2y} dx \int_0^{\sqrt{4-y^2}} dz z = \frac{1}{2} \int_0^2 dy \int_{2-y}^{6-2y} dx (4 - y^2) \\ &= \frac{1}{2} \int_0^2 (4 - y)(4 - y^2) dy = \frac{1}{2} \int_0^2 (16 - 4y - 4y^2 + y^3) dy \\ &= 16 - 4 - \frac{16}{3} + 2 = \boxed{\frac{26}{3}}. \end{aligned} \quad \square$$

Ejemplo 4.17 (principio de Cavalieri). Sea Ω el sólido de ecuaciones

$$h_1 \leq z \leq h_2, \quad (x, y) \in D(z), \tag{4.17}$$

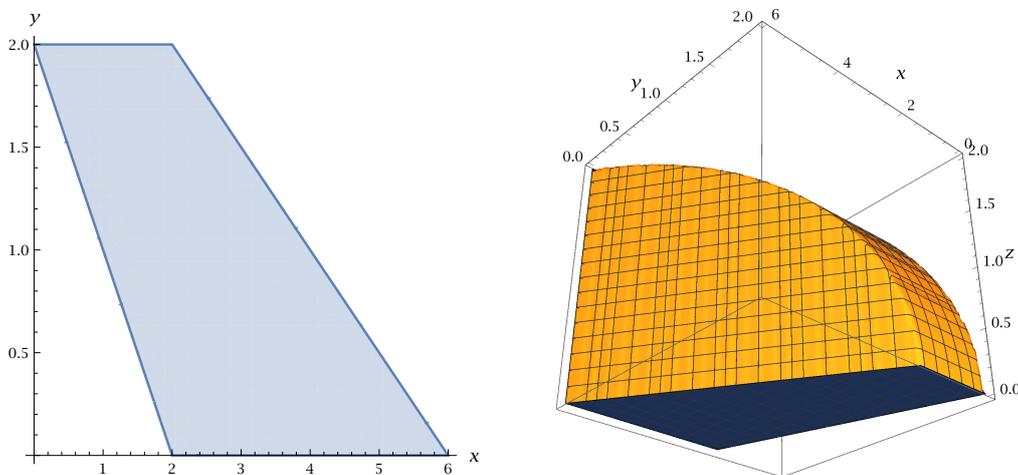


Figura 4.8: Región D (izda.) y sólido Ω (drcha.) en el Ejemplo 4.16.

donde $D(z) \subset \mathbb{R}^2$ es un conjunto compacto medible Jordan para todo $z \in [h_1, h_2]$ y $D(z) \subset [a, b] \times [c, d]$ para todo $z \in [h_1, h_2]$, de modo que Ω está contenido en el paralelepípedo $R = [a, b] \times [c, d] \times [h_1, h_2]$. Aplicando el teorema de Fubini se obtiene

$$\begin{aligned} v(\Omega) &= \int_R \chi_\Omega(x, y, z) \, dx \, dy \, dz = \int_{h_1}^{h_2} dz \int_{[a,b] \times [c,d]} dx \, dy \chi_\Omega(x, y, z) \\ &= \int_{h_1}^{h_2} dz \int_{[a,b] \times [c,d]} dx \, dy \chi_{D(z)}(x, y) = \int_{h_1}^{h_2} dz \int_{D(z)} dx \, dy, \end{aligned}$$

y por tanto

$$v(\Omega) = \int_{h_1}^{h_2} A(z) \, dz, \tag{4.18}$$

siendo $A(h) \equiv a(D(h))$ el área de la sección transversal de Ω por el plano $z = h$. Sea a continuación $\tilde{\Omega}$ otro sólido del mismo tipo que Ω , de ecuaciones

$$h_1 \leq z \leq h_2, \quad (x, y) \in \tilde{D}(z).$$

Si el área $\tilde{A}(h) \equiv a(\tilde{D}(h))$ de la sección de $\tilde{\Omega}$ por el plano $z = h$ es igual a $A(h)$ para todo $h \in [h_1, h_2]$, entonces Ω y $\tilde{\Omega}$ tienen obviamente el mismo volumen, ya que

$$v(\tilde{\Omega}) = \int_{h_1}^{h_2} \tilde{A}(z) \, dz = \int_{h_1}^{h_2} A(z) \, dz = v(\Omega).$$

Este es el llamado **principio de Cavalieri**.

Consideremos, por ejemplo, el *sólido de revolución* obtenido girando la gráfica de la función continua y no negativa

$$x = f(z), \quad h_1 \leq z \leq h_2,$$

alrededor del eje z . En este caso la sección del sólido por un plano $z = h$ es obviamente un disco de radio $f(h)$, y por tanto $A(h) = \pi f(h)^2$. Luego el volumen del sólido de revolución Ω está dado por

$$v(\Omega) = \pi \int_{h_1}^{h_2} f(z)^2 \, dz. \tag{4.19}$$

Ejemplo 4.18. Por ejemplo, hallemos el volumen del sólido de revolución Ω acotado por el cono $z = \sqrt{x^2 + y^2}$ y la esfera $x^2 + y^2 + z^2 = 1$ (cf. la Fig. 4.9, dcha.). En primer lugar, la sección de ambas superficies por el semiplano $y = 0, x \geq 0$ son las curvas $x = z$ y $x = \sqrt{1 - z^2}$, que se cortan en el punto $x = 1/\sqrt{2}$ (cf. la Fig. 4.9, izda.). Por tanto en este caso la función $f(z)$ está dada por

$$f(z) = \begin{cases} z, & 0 \leq z \leq 1/\sqrt{2} \\ \sqrt{1 - z^2}, & 1/\sqrt{2} \leq z \leq 1, \end{cases}$$

y el volumen pedido es igual a

$$\begin{aligned} v(\Omega) &= \pi \int_0^{1/\sqrt{2}} z^2 \, dz + \pi \int_{1/\sqrt{2}}^1 (1 - z^2) \, dz = \frac{\pi}{6\sqrt{2}} + \pi \left(1 - \frac{1}{\sqrt{2}} - \frac{1}{3} + \frac{1}{6\sqrt{2}} \right) \\ &= \frac{\pi}{3\sqrt{2}} + \frac{\pi}{3\sqrt{2}} (3\sqrt{2} - 3 - \sqrt{2}) = \frac{\pi}{3\sqrt{2}} \cdot 2(\sqrt{2} - 1) = \boxed{\frac{\pi}{3} (2 - \sqrt{2})}. \end{aligned}$$

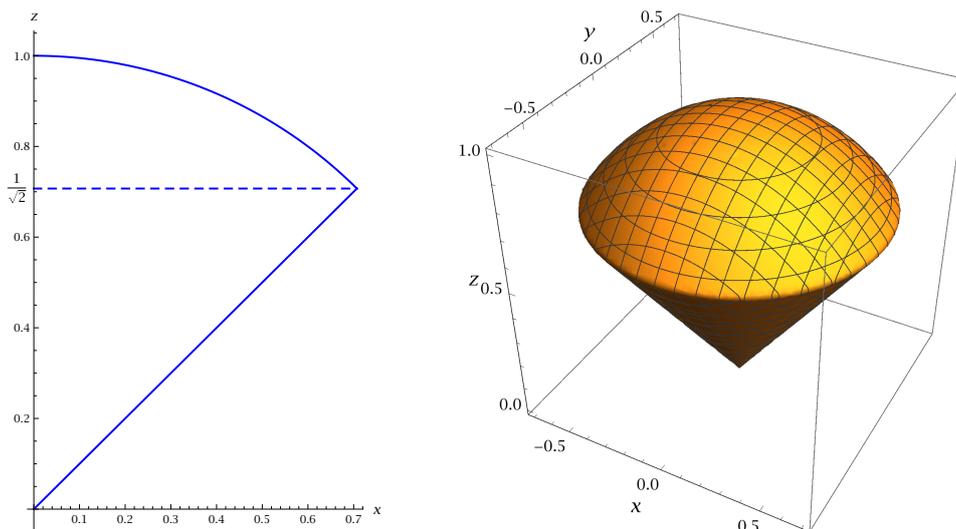


Figura 4.9: Gráfica de la función $x = f(z)$ (izda.) y sólido de revolución Ω (drcha.) en el Ejemplo 4.18.

Ejemplo 4.19. El volumen de la esfera de radio $a > 0$ se calcula fácilmente utilizando la ec. (4.19). En efecto, en este caso $f(z) = \sqrt{a^2 - z^2}$, con $-a \leq z \leq a$, y por tanto

$$v(B_a(0)) = \pi \int_{-a}^a (a^2 - z^2) dz = 2\pi \int_0^a (a^2 - z^2) dz = 2\pi a^3 \int_0^1 (1 - t^2) dt = \boxed{\frac{4}{3} \pi a^3}.$$

Ejercicio 35. Probar que la coordenada z del centro de masas de un sólido homogéneo de ecuación (4.17) está dada por

$$Z = \frac{\int_{h_1}^{h_2} zA(z) dz}{\int_{h_1}^{h_2} A(z) dz}.$$

En particular, para un sólido de revolución

$$Z = \frac{\int_{h_1}^{h_2} z f(z)^2 dz}{\int_{h_1}^{h_2} f(z)^2 dz}.$$

(En este último caso, las coordenadas x e y del centro de masas se anulan por simetría.)

4.5 Cambios de variables en integrales múltiples

Al igual que en el caso unidimensional, uno de los métodos más efectivos para evaluar una integral múltiple es efectuar un cambio de variables apropiado. Evidentemente, para poder utilizar este método necesitamos primero conocer cómo se transforman las integrales múltiples bajo cambios de variables. En una dimensión, dicha transformación es muy sencilla: en efecto, si $g : [a, b] \rightarrow \mathbb{R}$ es de clase C^1 y $f : \mathbb{R} \rightarrow \mathbb{R}$ es continua entonces

$$\int_{g(a)}^{g(b)} f(x) dx = \int_a^b f(g(t)) g'(t) dt.$$

Si g es *inyectiva*, es fácil convencerse de que esta fórmula se puede escribir

$$\boxed{\int_{g([a,b])} f(x) dx = \int_a^b f(g(t)) |g'(t)| dt.} \tag{4.20}$$

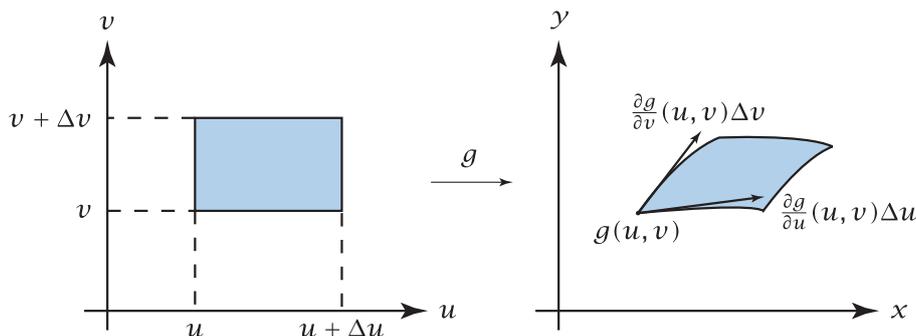


Figura 4.10: Rectángulo R en el plano (u, v) y su transformado $g(R)$ en el plano (x, y) .

En efecto, si g es monótona creciente en $[a, b]$ entonces $g'(t) \geq 0$ para todo $t \in [a, b]$, $g([a, b]) = [g(a), g(b)]$, y la fórmula anterior coincide con la usual, mientras que si g es monótona decreciente entonces $g'(t) \leq 0$ en $[a, b]$, $|g'(t)| = -g'(t)$ pero $g([a, b]) = [g(b), g(a)]$.

Para ver cuál es el análogo n -dimensional de la ecuación (4.20), consideremos en primer lugar el caso bidimensional. Sea $g : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ un difeomorfismo de clase C^1 , y supongamos que \tilde{D} y $D \equiv g(\tilde{D})$ son ambos medibles Jordan. Nótese que el difeomorfismo g determina un sistema de coordenadas curvilíneas en el abierto $g(\mathbb{R}^2)$, y por tanto los puntos del conjunto \tilde{D} se pueden interpretar como las coordenadas curvilíneas de los puntos del conjunto D . Consideremos un rectángulo $R \equiv [u, u + \Delta u] \times [v, v + \Delta v] \subset \tilde{D}$ de lados Δu y Δv . Si Δu y Δv son suficientemente pequeños, el conjunto $g(R)$ puede aproximarse por el paralelogramo generado por los vectores $\frac{\partial g}{\partial u}(u, v)\Delta u$ y $\frac{\partial g}{\partial v}(u, v)\Delta v$, con origen en el punto $g(u, v)$ (cf. la Fig. 4.10). El área de este paralelogramo es igual a

$$\left\| \frac{\partial g}{\partial u}(u, v) \right\| \left\| \frac{\partial g}{\partial v}(u, v) \right\| \sin \theta \Delta u \Delta v,$$

siendo θ el ángulo formado por los vectores $\frac{\partial g}{\partial u}(u, v)$ y $\frac{\partial g}{\partial v}(u, v)$. Nótese que

$$\begin{aligned} \left\| \frac{\partial g}{\partial u}(u, v) \right\| \left\| \frac{\partial g}{\partial v}(u, v) \right\| \sin \theta &= \left\| \frac{\partial g}{\partial u}(u, v) \times \frac{\partial g}{\partial v}(u, v) \right\| \\ &= \left\| \left(\frac{\partial g_1}{\partial u}(u, v) \frac{\partial g_2}{\partial v}(u, v) - \frac{\partial g_2}{\partial u}(u, v) \frac{\partial g_1}{\partial v}(u, v) \right) \mathbf{k} \right\| = | \det Dg(u, v) |, \end{aligned}$$

donde al calcular el producto vectorial se sobreentiende que la tercera componente de los vectores $\frac{\partial g}{\partial u}(u, v)$ y $\frac{\partial g}{\partial v}(u, v)$ es nula, y por tanto

$$a(g(R)) \simeq | \det Dg(u, v) | \Delta u \Delta v.$$

Consideremos a continuación una familia de pequeños subrectángulos $R_\alpha \equiv [u_\alpha, u_\alpha + \Delta u_\alpha] \times [v_\alpha, v_\alpha + \Delta v_\alpha]$ (con $1 \leq \alpha \leq p$) que no se solapan entre sí, de modo que

$$\tilde{D} \simeq \bigcup_{\alpha=1}^p R_\alpha.$$

Entonces

$$D = g(\tilde{D}) \simeq \bigcup_{\alpha=1}^p g(R_\alpha),$$

donde los conjuntos $g(R_\alpha)$ ($1 \leq \alpha \leq p$) no se solapan entre sí. Por la propiedad de aditividad, la

integral en el conjunto D de una función $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ está aproximadamente dada por

$$\begin{aligned} \int_D f &\simeq \sum_{\alpha=1}^p \int_{g(R_\alpha)} f \simeq \sum_{\alpha=1}^p f(g(u_\alpha, v_\alpha)) a(g(R_\alpha)) \simeq \sum_{\alpha=1}^p f(g(u_\alpha, v_\alpha)) |\det Dg(u_\alpha, v_\alpha)| \Delta u_\alpha \Delta v_\alpha \\ &\simeq \sum_{\alpha=1}^p \int_{R_\alpha} f(g(u, v)) |\det Dg(u, v)| du dv \simeq \int_{\tilde{D}} (f \circ g) |\det Dg|. \end{aligned} \quad (4.21)$$

En el límite en que los lados Δu_α y Δv_α de los rectángulos R_α tienden a cero, las aproximaciones anteriores deberían convertirse en igualdades. Este razonamiento heurístico sugiere que la igualdad

$$\boxed{\int_D f = \int_{\tilde{D}} (f \circ g) |\det Dg|} \quad (4.22)$$

debería ser el análogo bidimensional de la ec. (4.20). Nótese que la ecuación anterior se puede escribir de varias formas equivalentes, como por ejemplo

$$\int_D f(x, y) dx dy = \int_{\tilde{D}} f(g(u, v)) |\det Dg(u, v)| du dv,$$

o de forma más mnemotécnica (aunque con un ligero abuso de notación)

$$\boxed{\int_D f(x, y) dx dy = \int_{\tilde{D}} f(x(u, v), y(u, v)) \left| \det \frac{\partial(x, y)}{\partial(u, v)}(u, v) \right| du dv.} \quad (4.23)$$

Formalmente, por tanto, para efectuar el cambio de variable en la integral de f sobre la región D se sustituyen las variables x e y por las dos funciones $x(u, v)$ e $y(u, v)$ que definen el cambio de variables, se reemplaza el elemento de área $dx dy$ por

$$\left| \det \frac{\partial(x, y)}{\partial(u, v)}(u, v) \right| du dv,$$

y por supuesto se integra sobre el conjunto $\tilde{D} = g^{-1}(D)$, donde

$$g(u, v) = (x(u, v), y(u, v)).$$

Veremos a continuación que la ecuación (4.22) es esencialmente correcta también para integrales en $n > 2$ dimensiones, bajo condiciones muy generales sobre la función cambio de variables g .

El razonamiento anterior (cf. la ec. (4.21)) sugiere claramente que en cualquier número de dimensiones la fórmula del cambio de variables debería ser

$$\int_{\Omega} f = \int_{\tilde{\Omega}} (f \circ g) J,$$

donde

$$J(u) = \lim_{\Delta u_1, \dots, \Delta u_n \rightarrow 0} \frac{v(g(R))}{v(R)} = \lim_{\Delta u_1, \dots, \Delta u_n \rightarrow 0} \frac{v(g(R))}{\Delta u_1 \cdots \Delta u_n}$$

y R es el n -paralelepípedo recto con origen en el punto u y lados $\Delta u_1, \dots, \Delta u_n$. En el caso tridimensional, R es simplemente el paralelepípedo $[u, u + \Delta u] \times [v, v + \Delta v] \times [w, w + \Delta w]$. Si Δu , Δv y Δw son suficientemente pequeños, la imagen de dicho paralelepípedo bajo un difeomorfismo $g: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ de clase C^1 es un sólido que puede aproximarse por el paralelepípedo (no recto) generado por los vectores

$$\frac{\partial g}{\partial u}(u, v, w) \Delta u, \quad \frac{\partial g}{\partial v}(u, v, w) \Delta v, \quad \frac{\partial g}{\partial w}(u, v, w) \Delta w$$

con origen en el punto $g(u, v, w)$. Como es sabido, el volumen de este paralelepípedo está dado por

$$\left| \frac{\partial g}{\partial u}(u, v, w) \cdot \left(\frac{\partial g}{\partial v}(u, v, w) \times \frac{\partial g}{\partial w}(u, v, w) \right) \right| \Delta u \Delta v \Delta w = |\det Dg(u, v, w)| \Delta u \Delta v \Delta w,$$

donde se ha utilizado la siguiente identidad para el *producto triple* $a \cdot (b \times c)$ de tres vectores $a \equiv (a_1, a_2, a_3)$, $b \equiv (b_1, b_2, b_3)$, $c \equiv (c_1, c_2, c_3)$:

$$a \cdot (b \times c) = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

Razonando como en el caso bidimensional se llega a la fórmula

$$\int_{\Omega} f = \int_{\tilde{\Omega}} (f \circ g) |\det Dg|, \quad (4.24)$$

donde $\tilde{\Omega}$ y $\Omega \equiv g(\tilde{\Omega})$ son conjuntos medibles Jordan. Al igual antes, con un ligero abuso de notación esta fórmula se suele escribir de forma más explícita como

$$\int_{\Omega} f(x, y, z) dx dy dz = \int_{\tilde{\Omega}} f(x(u, v, w), y(u, v, w), z(u, v, w)) \left| \det \frac{\partial(x, y, z)}{\partial(u, v, w)}(u, v, w) \right| du dv dw. \quad (4.25)$$

En el caso n -dimensional (con $n > 3$ arbitrario), el razonamiento se basa de nuevo en que el n -paralelepípedo recto

$$R = [u_1, u_1 + \Delta u_1] \times \cdots \times [u_n, u_n + \Delta u_n]$$

se transforma bajo un difeomorfismo $g: \mathbb{R}^n \rightarrow \mathbb{R}^n$ de clase C^1 en un sólido que puede aproximarse (si Δu_i es suficientemente pequeño para todo i) por el paralelepípedo n -dimensional (no recto) generado por los vectores

$$\frac{\partial g}{\partial u_i}(u) \Delta u_i, \quad 1 \leq i \leq n,$$

con origen en el punto $g(u)$. El volumen de este paralelepípedo es el valor absoluto del determinante cuyas columnas son las componentes de los vectores anteriores, es decir

$$|\det Dg(u)| \Delta u_1 \cdots \Delta u_n,$$

lo que conduce de nuevo a la fórmula (4.24). Vemos pues que el análogo de la fórmula del cambio de variables (4.20) en cualquier número de dimensiones n debería ser la ec. (4.24). Esta afirmación es esencialmente correcta, como afirma de forma más precisa la siguiente proposición:

Proposición 4.20 (fórmula del cambio de variables). Sea $g: \mathbb{R}^n \rightarrow \mathbb{R}^n$ una función biyectiva de clase C^1 en un abierto $U \subset \mathbb{R}^n$ tal que $\det Dg(u) \neq 0$ para todo $u \in U$, salvo a lo sumo en un conjunto de medida 0, y sean $\tilde{\Omega} \subset U$ y $\Omega \equiv g(\tilde{\Omega})$ conjuntos medibles Jordan. Si $f: \mathbb{R}^n \rightarrow \mathbb{R}$ es integrable en Ω entonces se verifica la fórmula del cambio de variables (4.24).

Comentario. Como vimos en el capítulo anterior (cf. el Corolario 3.22), si el jacobiano de g no se anula en *todo* el abierto U entonces g es un difeomorfismo de clase C^1 en U , y por lo tanto en este caso g determina un sistema de coordenadas curvilíneas en el abierto $g(U)$. La condición sobre g en el enunciado de la proposición anterior es algo más débil, ya que permite que $\det Dg$ pueda anularse en un subconjunto de medida nula de U . \square

Al igual en los casos $n = 2$ y $n = 3$, la fórmula del cambio de variables (4.24) se suele escribir en la forma más explícita

$$\int_{\Omega} f(x_1, \dots, x_n) d^n x = \int_{\tilde{\Omega}} f(x_1(u), \dots, x_n(u)) \left| \det \frac{\partial(x_1, \dots, x_n)}{\partial(u_1, \dots, u_n)}(u) \right| d^n u, \quad (4.26)$$

donde la función $x(u) \equiv (x_1(u), \dots, x_n(u))$ que aparece en el miembro derecho es la función cambio de variables. Formalmente, la ecuación anterior se obtiene efectuando en el miembro izquierdo los cambios

$$x \rightarrow x(u), \quad d^n x \rightarrow \left| \det \frac{\partial(x_1, \dots, x_n)}{\partial(u_1, \dots, u_n)}(u) \right| d^n u \quad (4.27)$$

e integrando sobre el conjunto $\tilde{\Omega} = g^{-1}(\Omega)$ (donde $g(u) = (x_1(u), \dots, x_n(u))$ es el cambio de variable). En particular, la expresión

$$\left| \det \frac{\partial(x_1, \dots, x_n)}{\partial(u_1, \dots, u_n)}(u) \right| d^n u \equiv | \det Dg(u) | d^n u \quad (4.28)$$

se puede interpretar como el **elemento de volumen** infinitesimal en las coordenadas curvilíneas (u_1, \dots, u_n) . En otras palabras, el jacobiano de la transformación g es el *factor de dilatación local* del volumen bajo dicha transformación. En particular, el volumen de la región Ω está dado por

$$v(\Omega) = \int_{\tilde{\Omega}} \left| \det \frac{\partial(x_1, \dots, x_n)}{\partial(u_1, \dots, u_n)}(u) \right| d^n u.$$

Ejemplo 4.21. Consideremos, en primer lugar, las *coordenadas polares* en el plano \mathbb{R}^2 , definidas por

$$g(r, \theta) = r(\cos \theta, \sin \theta).$$

La función g es un difeomorfismo de clase C^∞ del abierto $U = (0, \infty) \times (0, 2\pi)$ en el abierto

$$g(U) = \mathbb{R}^2 - \{(x, 0) \mid x \geq 0\},$$

es decir en \mathbb{R}^2 menos el semieje $y = 0$, $x \geq 0$ (que tiene medida 0 en \mathbb{R}^2). El jacobiano de la transformación g está dado por

$$\det Dg(r, \theta) = \begin{vmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} = r > 0 \quad \text{en } U.$$

Por tanto la fórmula del cambio de variables se escribe en este caso

$$\int_D f(x, y) dx dy = \int_{\tilde{D}} f(r \cos \theta, r \sin \theta) r dr d\theta.$$

En particular, el área del conjunto D en coordenadas polares está dado por

$$a(D) = \int_{\tilde{D}} r dr d\theta.$$

Análogamente, las *coordenadas cilíndricas* en \mathbb{R}^3 están definidas por

$$g(\rho, \varphi, z) = (\rho \cos \varphi, \rho \sin \varphi, z),$$

que es un difeomorfismo de clase C^∞ del abierto $U = (0, \infty) \times (0, 2\pi) \times \mathbb{R}$ en \mathbb{R}^3 menos el semiplano $\{y = 0, x \geq 0\}$ (de nuevo, dicho semiplano tiene medida 0 en \mathbb{R}^3). Como vimos en el Ejemplo 3.26, el jacobiano de esta transformación es

$$\det Dg(\rho, \varphi, z) = \rho > 0 \quad \text{en } U.$$

Por tanto en este caso se tiene

$$\int_{\Omega} f(x, y, z) \, dx \, dy \, dz = \int_{\tilde{\Omega}} f(\rho \cos \varphi, \rho \sin \varphi, z) \rho \, d\rho \, d\varphi \, dz,$$

de donde se sigue que

$$v(\Omega) = \int_{\tilde{\Omega}} \rho \, d\rho \, d\varphi \, dz.$$

Finalmente, las *coordenadas esféricas* definidas por

$$g(r, \theta, \varphi) = r(\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta)$$

son un difeomorfismo de clase C^∞ de $U \equiv (0, \infty) \times (0, \pi) \times (0, 2\pi)$ en \mathbb{R}^3 menos el semiplano $\{y = 0, x \geq 0\}$, con jacobiano

$$\det Dg(r, \theta, \varphi) = r^2 \sin \theta > 0 \quad \text{en } U.$$

Por tanto en este caso la fórmula del cambio de variables se escribe

$$\int_{\Omega} f(x, y, z) \, dx \, dy \, dz = \int_{\tilde{\Omega}} f(r \sin \theta \cos \varphi, r \sin \theta \sin \varphi, r \cos \theta) r^2 \sin \theta \, dr \, d\theta \, d\varphi,$$

y en particular

$$v(\Omega) = \int_{\tilde{\Omega}} r^2 \sin \theta \, dr \, d\theta \, d\varphi.$$

Por ejemplo, el volumen de la esfera de radio $a > 0$ se calcula de forma muy sencilla en coordenadas esféricas, ya que en dichas coordenadas $B_a(0)$ es el paralelepípedo

$$0 \leq r \leq a, \quad 0 \leq \theta \leq \pi, \quad 0 \leq \varphi \leq 2\pi.$$

Aplicando la fórmula del cambio de variables se obtiene³

$$v(B_a(0)) = \int_0^a dr \int_0^\pi d\theta \int_0^{2\pi} d\varphi r^2 \sin \theta = 2\pi \left(\int_0^a r^2 dr \right) \left(\int_0^\pi \sin \theta d\theta \right) = \boxed{\frac{4}{3} \pi a^3}.$$

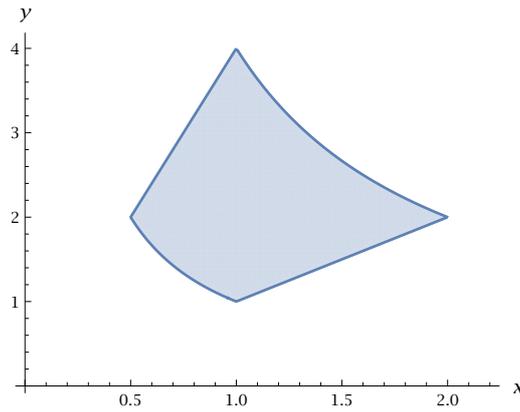
Ejemplo 4.22. Sea $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ una transformación lineal invertible. En este caso

$$DA(x) = A, \quad \forall x \in \mathbb{R}^n,$$

y por lo tanto el jacobiano de la transformación es el determinante de la matriz A (que es no nulo, al ser A invertible). La fórmula del cambio de variables proporciona en este caso (intercambiando los papeles de Ω y $\tilde{\Omega}$ en la ec. (4.24))

$$\int_{A(\Omega)} f(x) \, d^n x = |\det A| \int_{\Omega} f(Au) \, d^n u.$$

³En rigor, el cambio de variable a coordenadas esféricas solo puede aplicarse a la esfera menos su intersección con el semiplano $\{y = 0, x \geq 0\}$, pero este último conjunto tiene medida 0.

Figura 4.11: Región D del Ejemplo 4.23.

En particular, tomando $f = 1$ obtenemos la importante identidad

$$\boxed{v(A(\Omega)) = |\det A| \cdot v(\Omega)}.$$

Por ejemplo, si A es una transformación *ortogonal* entonces

$$A^T A = I \implies (\det A)^2 = 1 \implies |\det A| = 1,$$

y por tanto *las transformaciones ortogonales preservan el volumen*. Por otra parte, si A es la dilatación inhomogénea definida por $A(x) = (\lambda_1 x_1, \dots, \lambda_n x_n)$, entonces

$$v(A(\Omega)) = |\lambda_1 \cdots \lambda_n| v(\Omega).$$

Consideremos, por ejemplo, el elipsoide E de ecuación

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leq 1.$$

Aplicando la fórmula anterior con $\lambda_1 = a$, $\lambda_2 = b$, $\lambda_3 = c$ a la esfera unidad Ω se deduce que

$$v(E) = abc \cdot v(\Omega) = \boxed{\frac{4}{3} \pi abc}.$$

Ejemplo 4.23. Calculemos la integral de la función $f(x, y) = x^2 + y^2$ en la región D del primer cuadrante ($x, y \geq 0$) de ecuaciones

$$1 \leq xy \leq 4, \quad x \leq y \leq 4x \tag{4.29}$$

(cf. la Fig. 4.23) utilizando el cambio de variables $u = xy$, $v = y/x$. La función $(x, y) \mapsto (u, v) \equiv g^{-1}(x, y)$ es invertible si (x, y) pertenece al primer cuadrante, y su inversa está dada por

$$x = \sqrt{\frac{u}{v}}, \quad y = \sqrt{uv},$$

donde (u, v) también pertenece al primer cuadrante. Es inmediato comprobar utilizando las ecs. (4.29) junto con las ecuaciones del cambio de variables que en este caso el conjunto \tilde{D} es el cuadrado $[1, 4] \times [1, 4]$. El jacobiano de la transformación $(u, v) \mapsto (x, y) \equiv g(u, v)$ (que claramente es de clase C^∞ si (u, v) pertenece al primer cuadrante) está dado por

$$\det \frac{\partial(x, y)}{\partial(u, v)} = \left(\det \frac{\partial(u, v)}{\partial(x, y)} \right)^{-1} = \left| \begin{array}{cc} y & x \\ -\frac{y}{x^2} & \frac{1}{x} \end{array} \right|^{-1} = \frac{x}{2y} = \frac{1}{2v}.$$

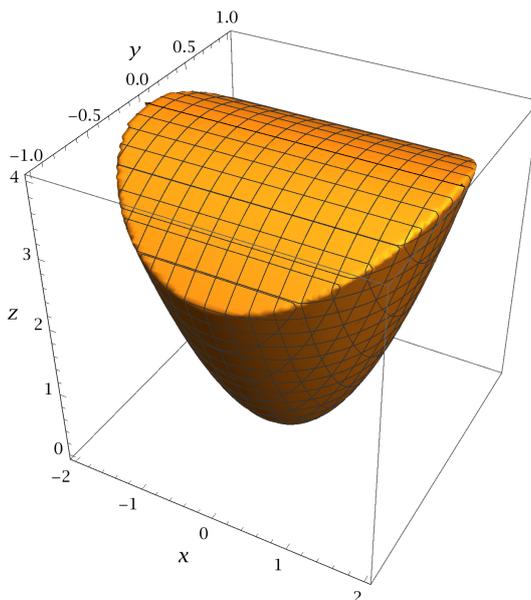


Figura 4.12: Sólido Ω del Ejemplo 4.24.

Vemos por tanto que dicho jacobiano es positivo en el primer cuadrante, y en particular en el cuadrado \tilde{D} . Aplicando la fórmula del cambio de variables se obtiene

$$\begin{aligned} \int_D (x^2 + y^2) \, dx \, dy &= \int_{[1,4]^2} \frac{1}{2v} \left(\frac{u}{v} + uv \right) \, du \, dv = \frac{1}{2} \int_{[1,4]^2} \left(\frac{u}{v^2} + u \right) \, du \, dv \\ &= \frac{1}{2} \left(\int_1^4 u \, du \right) \left(\int_1^4 v^{-2} \, dv \right) + \frac{3}{2} \int_1^4 u \, du = \frac{15}{4} \cdot \left(1 - \frac{1}{4} \right) + \frac{45}{4} = \boxed{\frac{225}{16}}. \end{aligned}$$

Ejemplo 4.24. Calculemos el volumen del sólido Ω acotado superiormente por el cilindro parabólico $z = 4 - y^2$ e inferiormente por el paraboloides elíptico $z = x^2 + 3y^2$ (cf. la Fig. 4.12). El sólido en cuestión es de tipo I, ya que

$$x^2 + 3y^2 \leq z \leq 4 - y^2 \implies x^2 + 4y^2 \leq 4,$$

y por consiguiente

$$\Omega = \{(x, y, z) \in \mathbb{R}^3 \mid (x, y) \in D, x^2 + 3y^2 \leq z \leq 4 - y^2\},$$

siendo D la elipse $(x/2)^2 + y^2 \leq 1$. Por tanto el volumen de Ω está dado por

$$v(\Omega) = \int_D (4 - y^2 - x^2 - 3y^2) \, dx \, dy = \int_D (4 - x^2 - 4y^2) \, dx \, dy.$$

Para evaluar esta integral efectuamos el cambio de variables⁴

$$x = 2r \cos \theta, \quad y = r \sin \theta,$$

que (al igual que las coordenadas polares) claramente es biyectivo si $r > 0$ y $0 < \theta < 2\pi$. El jacobiano de esta transformación está dado por

$$\frac{\partial(x, y)}{\partial(r, \theta)} = \begin{vmatrix} 2 \cos \theta & -2r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} = 2r > 0$$

⁴Este cambio de variables es la composición del paso a coordenadas polares con la dilatación $(x, y) \mapsto (2x, y)$.

en la región considerada. En las coordenadas (r, θ) , la elipse está descrita por las ecuaciones $0 \leq r \leq 1$, $0 \leq \theta < 2\pi$. Aplicando la fórmula del cambio de variables (sin tener en cuenta conjuntos de medida nula) se obtiene

$$v(\Omega) = 2 \int_0^{2\pi} d\theta \int_0^1 dr r \cdot 4(1-r^2) = 16\pi \int_0^1 (r-r^3) dr = 16\pi \left(\frac{1}{2} - \frac{1}{4} \right) = \boxed{4\pi}.$$

Ejemplo 4.25. Las integrales múltiples se pueden utilizar a veces para evaluar integrales de funciones de una variable. Uno de los ejemplos más conocidos de esta afirmación es el de la integral

$$\int_{-\infty}^{\infty} e^{-x^2} dx, \quad (4.30)$$

que calcularemos a continuación. Esta integral es de hecho una *integral impropia*, que (al ser e^{-x^2} una función par) se define más precisamente como

$$\lim_{a \rightarrow \infty} \int_{-a}^a e^{-x^2} dx.$$

Para calcular la integral (4.30), consideremos las dos integrales dobles

$$I(a) = \int_{C_a} e^{-(x^2+y^2)} dx dy, \quad J(a) = \int_{D_a} e^{-(x^2+y^2)} dx dy,$$

donde $a > 0$, $C_a = [-a, a]^2$ y D_a es el disco $x^2 + y^2 \leq a^2$. Aplicando el teorema de Fubini a la primera integral se obtiene

$$I(a) = \int_{-a}^a dx \int_{-a}^a dy e^{-x^2} e^{-y^2} = \left(\int_{-a}^a e^{-x^2} dx \right) \left(\int_{-a}^a e^{-y^2} dy \right) \equiv \left(\int_{-a}^a e^{-x^2} dx \right)^2,$$

y por tanto

$$\lim_{a \rightarrow \infty} I(a) = \left(\int_{-\infty}^{\infty} e^{-x^2} dx \right)^2.$$

La segunda integral se calcula fácilmente en coordenadas polares:

$$J(a) = \int_0^{2\pi} d\theta \int_0^a dr r e^{-r^2} = -2\pi \cdot \frac{1}{2} e^{-r^2} \Big|_{r=0}^{r=a} = \pi(1 - e^{-a^2}),$$

de donde se sigue que

$$\lim_{a \rightarrow \infty} J(a) = \pi.$$

Por otra parte, está claro que $D_a \subset C_a \subset D_{\sqrt{2}a}$ para todo $a > 0$, y por tanto

$$I(a) \leq J(a) \leq I(\sqrt{2}a) \implies \lim_{a \rightarrow \infty} I(a) = \lim_{a \rightarrow \infty} J(a),$$

de donde se sigue que

$$\left(\int_{-\infty}^{\infty} e^{-x^2} dx \right)^2 = \pi \implies \int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}.$$

En general, si $\alpha > 0$ se tiene

$$\int_{-\infty}^{\infty} e^{-\alpha x^2} dx = \frac{1}{\sqrt{\alpha}} \int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\frac{\pi}{\alpha}}.$$

En particular, la familia de funciones $\{f_\alpha\}_{\alpha \in (0, \infty)}$ definida por

$$f_\alpha(x) \equiv \sqrt{\frac{\alpha}{\pi}} e^{-\alpha x^2}$$

cumple

$$\int_{-\infty}^{\infty} f_{\alpha}(x) dx = 1$$

y

$$\lim_{\alpha \rightarrow \infty} f_{\alpha}(x) = \begin{cases} 0, & x \neq 0 \\ \infty, & x = 0. \end{cases}$$

De hecho, puede probarse que si $g : \mathbb{R} \rightarrow \mathbb{R}$ es continua tiende a cero suficientemente rápido cuando $|x| > R$ entonces

$$\lim_{\alpha \rightarrow \infty} \int_{-\infty}^{\infty} f_{\alpha}(x) g(x) dx = g(0).$$

Por este motivo, las funciones f_{α} con $\alpha \gg 1$ constituyen una aproximación de la *delta de Dirac*.

Capítulo 5

Integrales de línea y superficie

5.1 Integral de una función escalar a lo largo de una curva

5.1.1 Curvas parametrizadas

En este capítulo estudiaremos dos nuevos tipos de integral, a lo largo de *curvas* en \mathbb{R}^n y sobre *superficies* en \mathbb{R}^3 , tanto de funciones escalares como vectoriales. Comenzaremos definiendo el primero de estos dos tipos de integrales, para lo cual necesitamos dar una definición más precisa del concepto de curva.

En primer lugar, diremos que una función $\gamma : [a, b] \rightarrow \mathbb{R}^n$ es **de clase C^1 a trozos** si es continua en $[a, b]$, y el intervalo $[a, b]$ se puede subdividir en un número finito de intervalos en que γ es C^1 . Evidentemente, la función $\gamma \equiv (\gamma_1, \dots, \gamma_n) : [a, b] \rightarrow \mathbb{R}^n$ es C^1 a trozos si y solo si los son cada una de sus componentes $\gamma_i : [a, b] \rightarrow \mathbb{R}$.

Ejemplo 5.1. La función $f(t) = |t|$ (cf. la Fig. 5.1, izda.) es C^1 a trozos en $[-1, 1]$, ya que es continua en dicho intervalo y, aunque no es derivable en $t = 0$, es de clase C^∞ en cada uno de los dos intervalos $[-1, 0]$ y $[0, 1]$. La función $f(t) = \operatorname{sgn} t$ (donde $\operatorname{sgn}(0)$ se define como 0) no es de clase C^1 a trozos en $[-1, 1]$, ya que no es continua en $t = 0$ (cf. la Fig. 5.1, drcha.).

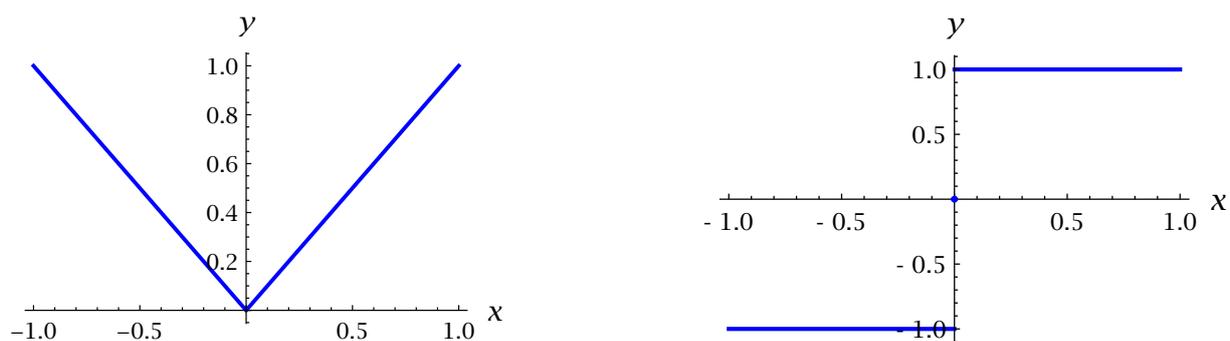


Figura 5.1: Gráfica de las funciones $|t|$ (izda.) y $\operatorname{sgn} t$ (drcha.) en el intervalo $-1 \leq t \leq 1$.

Definición 5.2.

- Una **curva parametrizada** (o **trayectoria**) en \mathbb{R}^n es una función $\gamma : [a, b] \rightarrow \mathbb{R}^n$. Una **curva** en \mathbb{R}^n es la imagen $\gamma([a, b]) \subset \mathbb{R}^n$ de una curva parametrizada $\gamma : [a, b] \rightarrow \mathbb{R}^n$.
- Una curva parametrizada es **suave** si la función γ es de clase C^1 a trozos en $[a, b]$.
- Una curva suave es **regular** si $\gamma'(t) \neq 0$ en $[a, b]$, salvo a lo sumo en un número *finito* de

puntos.

- Una curva parametrizada es **cerrada** si $\gamma(a) = \gamma(b)$.

Nótese que las curvas parametrizadas tienen un **sentido de recorrido** (u **orientación**) bien definido (*comienzan* en el punto $\gamma(a)$ y *terminan* en el punto $\gamma(b)$).

Ejemplo 5.3. La función $\gamma : [-2, 2] \rightarrow \mathbb{R}^2$ dada por

$$\gamma(t) = \begin{cases} (t, 0), & t \in [-2, -1] \cup [1, 2], \\ (\cos(\pi(1-t)/2), \sin(\pi(1-t)/2)), & t \in (-1, 1) \end{cases}$$

(cf. la Fig. 5.2, izda.) es una curva regular. En efecto, es claro que γ es continua en $[-2, 2]$, ya que

$$\lim_{t \rightarrow \pm 1} (\cos(\pi(1-t)/2), \sin(\pi(1-t)/2)) = \lim_{t \rightarrow \pm 1} (t, 0) = (\pm 1, 0) = \gamma(\pm 1).$$

Además, γ es de clase C^∞ en cada uno de los tres intervalos $[-2, -1]$, $[-1, 1]$ y $[1, 2]$, y por tanto es de clase C^1 a trozos en $[-2, 2]$ (aunque γ' no existe en los puntos $t = \pm 1$). Por último,

$$\gamma'(t) = \begin{cases} (1, 0), & t \in [-1, 1) \cup (1, 2], \\ \frac{\pi}{2} (\sin(\pi(1-t)/2), -\cos(\pi(1-t)/2)), & t \in (-1, 1), \end{cases}$$

existe y es no nulo en $[-2, 2] - \{-1, 1\}$. □

Si $\gamma : [a, b] \rightarrow \mathbb{R}^n$ es una curva regular, el vector $\gamma'(t)$ está definido y es no nulo en todo $[a, b]$ salvo a lo sumo en un número finito de puntos (los puntos de $[a, b]$ en que γ' es discontinua o se anula). En los puntos en que $\gamma'(t)$ no se anula, ya hemos visto que dicho vector es tangente a la curva γ . Si $\gamma'(t) \neq 0$, definimos entonces el **vector unitario tangente** a la curva γ en el punto $\gamma(t)$ mediante

$$T(t) = \frac{\gamma'(t)}{\|\gamma'(t)\|}. \quad (5.1)$$

Evidentemente, el vector $T(t)$ es tangente a la curva γ en el punto $\gamma(t)$ (ya que es proporcional a $\gamma'(t)$), es por definición unitario y su sentido es el de recorrido de la curva (ya que lleva la dirección límite de la secante a la curva que une $\gamma(t)$ con $\gamma(t + \Delta t)$ cuando $\Delta t > 0$ tiende a cero). Por ejemplo, el vector unitario tangente a la curva del ejemplo anterior está dado por

$$T(t) = \begin{cases} (1, 0), & t \in [-1, 1) \cup (1, 2], \\ (\sin(\pi(1-t)/2), -\cos(\pi(1-t)/2)), & t \in (-1, 1). \end{cases}$$

Ejemplo 5.4. Consideremos la curva plana (*cicloide*¹) dada por

$$\gamma(t) = (t - \sin t, 1 - \cos t), \quad t \in [-\pi, \pi]$$

(cf. la Fig. 5.2, dcha.). Dicha curva es C^1 en $[-\pi, \pi]$ y regular, ya que $\gamma'(t) = (1 - \cos t, \sin t)$ existe y es continua (C^∞) en $[-\pi, \pi]$ y solo se anula si $t = 0$, al ser

$$\cos t = 1, \quad \sin t = 0 \quad \Leftrightarrow \quad t = 2k\pi, \quad k \in \mathbb{Z}.$$

Nótese que en el punto $\gamma(0) = (0, 0)$ en que γ' se anula la curva tiene un punto anguloso² o

¹ Como es sabido, la cicloide es la trayectoria recorrida por un punto de una circunferencia (en este caso de radio 1) que rueda sin deslizar sobre un plano perpendicular al de dicha circunferencia.

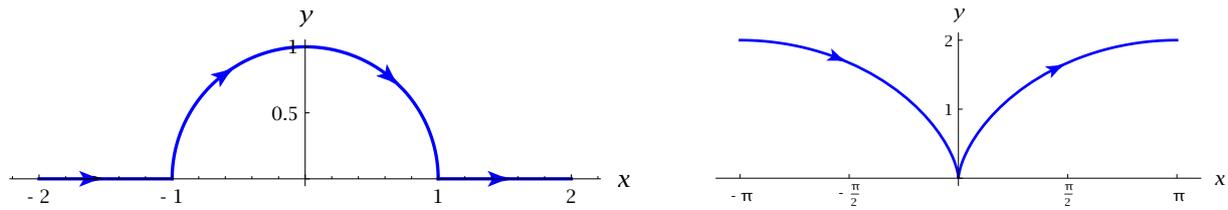


Figura 5.2: Izda: curva regular del Ejemplo 5.3. Dcha.: cicloide del Ejemplo 5.4.

“pico”. En cualquier otro punto, el vector unitario tangente a la curva está dado por

$$\begin{aligned} T(t) &= \frac{(1 - \cos t, \sin t)}{\|(1 - \cos t, \sin t)\|} = \frac{(1 - \cos t, \sin t)}{\sqrt{2(1 - \cos t)}} = \frac{(2 \sin^2(t/2), 2 \sin(t/2) \cos(t/2))}{2|\sin(t/2)|} \\ &= (|\sin(t/2)|, \operatorname{sgn} t \cdot \cos(t/2)) = \operatorname{sgn} t \cdot (\sin(t/2), \cos(t/2)), \quad t \neq 0. \end{aligned}$$

Ejemplo 5.5. La gráfica de una función $f : [a, b] \rightarrow \mathbb{R}$ determina una curva parametrizada, dada por $\gamma(t) = (t, f(t))$, que es suave siempre y cuando la función f sea de clase C^1 a trozos en el intervalo $[a, b]$. Nótese que la gráfica de una función C^1 a trozos siempre es *regular*, ya que su vector tangente $\gamma'(t) = (1, f'(t))$ no se anula en ningún punto en que exista f' . \square

5.1.2 Longitud de una curva

Sea $\gamma : [a, b] \rightarrow \mathbb{R}^n$ una curva suave, y sean $\gamma(t)$ y $\gamma(t + \Delta t)$ (con $\Delta t > 0$) dos puntos de dicha curva tales que $\gamma'(t) \neq 0$. Si Δt es muy pequeño, la longitud Δs del arco de curva entre ambos puntos se puede aproximar por la del segmento que une dichos puntos, es decir por la longitud del vector $\gamma(t + \Delta t) - \gamma(t)$, y por tanto

$$\Delta s \simeq \|\gamma(t + \Delta t) - \gamma(t)\| \simeq \|\gamma'(t)\Delta t\| = \|\gamma'(t)\| \Delta t.$$

Esta observación motiva la siguiente definición:

Definición 5.6. Si $\gamma : [a, b] \rightarrow \mathbb{R}^n$ es una curva suave, se define su **longitud** mediante

$$l(\gamma) = \int_a^b \|\gamma'(t)\| dt.$$

Se dice que

$$ds \equiv \|\gamma'(t)\| dt$$

es el **elemento de longitud de arco** a lo largo de la curva γ .

Nótese que la integral está bien definida, debido a la continuidad a trozos de γ' en $[a, b]$. De hecho, si γ' es continua en $[a, b] - \{c_1, \dots, c_k\}$, con $a \equiv c_0 < c_1 < \dots < c_k \equiv b$, la integral anterior está dada por

$$\int_a^b \|\gamma'(t)\| dt = \sum_{i=1}^k \int_{c_{i-1}}^{c_i} \|\gamma'(t)\| dt.$$

Ejemplo 5.7. La circunferencia de centro (a, b) y radio r se puede parametrizar mediante las ecuaciones

$$x = a + r \cos t, \quad y = b + r \sin t, \quad 0 \leq t \leq 2\pi,$$

²En efecto, cerca de $t = 0$ se tiene $\gamma(t) \simeq (t^3/6, t^2/2)$, y por tanto la cicloide coincide aproximadamente con la gráfica de la función $y = \frac{1}{2}(6|x|)^{2/3}$.

o equivalentemente $\gamma(t) = (a + r \cos t, b + r \sin t)$. Esta curva es suave y regular, ya que

$$\gamma'(t) = r(-\sin t, \cos t)$$

es continua y no se anula en ningún punto del intervalo $[0, 2\pi]$. La longitud de la circunferencia está dada por

$$\int_0^{2\pi} \|r(-\sin t, \cos t)\| dt = r \int_0^{2\pi} dt = 2\pi r.$$

Ejemplo 5.8. Calculemos la longitud del arco de cicloide del Ejemplo 5.4. En este caso se tiene

$$\gamma'(t) = (1 - \cos t, \sin t) \implies \|\gamma'(t)\| = \sqrt{(1 - \cos t)^2 + \sin^2 t} = \sqrt{2(1 - \cos t)} = 2|\sin(t/2)|.$$

Por tanto

$$l(\gamma) = 2 \int_{-\pi}^{\pi} |\sin(t/2)| dt = 4 \int_0^{\pi} \sin(t/2) dt = -8 \cos(t/2) \Big|_{t=0}^{t=\pi} = 8.$$

Ejemplo 5.9. La longitud de la gráfica de la función C^1 a trozos $f : [a, b] \rightarrow \mathbb{R}$ está dada por

$$\int_a^b \|(1, f'(x))\| dx = \int_a^b \sqrt{1 + f'(x)^2} dx. \quad \square$$

5.1.3 Reparametrizaciones

Es claro que una misma curva se puede parametrizar de infinitas formas distintas. Por ejemplo, el arco de parábola $y = x^2$ entre $x = 0$ y $x = 1/2$ se puede parametrizar como la gráfica de la función $f(x) = x^2$ en $[0, 1/2]$, es decir por $\gamma(t) = (t, t^2)$, pero también en la forma

$$\tilde{\gamma}(u) = (\sin u, \sin^2 u), \quad 0 \leq u \leq \pi/6.$$

La relación entre estas dos parametrizaciones está dada por

$$\tilde{\gamma}(u) = \gamma(\sin u).$$

En otras palabras, $\tilde{\gamma} = \gamma \circ \tau$, siendo $\tau : [0, \pi/6] \rightarrow [0, 1/2]$ la función dada por $\tau(u) = \sin u$. Nótese que τ es un difeomorfismo de clase C^∞ del intervalo $[0, \pi/6]$ en el intervalo $[0, 1/2]$, cuya inversa está dada por $\tau^{-1}(t) = \arcsin t$.

Definición 5.10. Una **reparametrización** de una curva suave $\gamma : [a, b] \rightarrow \mathbb{R}^n$ es cualquier curva parametrizada $\tilde{\gamma} : [c, d] \rightarrow \mathbb{R}^n$ tal que $\tilde{\gamma} = \gamma \circ \tau$, siendo $\tau : [c, d] \rightarrow [a, b]$ un difeomorfismo de clase C^1 . Se dice que la función τ es un **cambio de parámetro**.

Comentarios.

- Una curva suave γ y cualquier reparametrización suya $\tilde{\gamma} = \gamma \circ \tau$ determinan la misma curva. En efecto, al ser τ suprayectiva (por ser un cambio de parámetro) se tiene

$$\tilde{\gamma}([c, d]) = \gamma(\tau([c, d])) = \gamma([a, b]).$$

- De la definición de cambio de parámetro se sigue inmediatamente que si $\tilde{\gamma} = \gamma \circ \tau$ es una reparametrización de una curva suave γ entonces $\tilde{\gamma}$ es también una curva suave, ya que por la regla de la cadena

$$\tilde{\gamma}'(u) = \gamma'(\tau(u)) \tau'(u)$$

es continua en $u \in [c, d]$ si γ' es continua en $\tau(u) \in [a, b]$. Por el mismo motivo, si γ es regular también lo será su reparametrización $\tilde{\gamma}$ (ya que $\tau'(u) \neq 0$ para todo $u \in [c, d]$ por ser τ un difeomorfismo).

- Es evidente que si $\tau : [c, d] \rightarrow [a, b]$ es un cambio de parámetro también lo es $\tau^{-1} : [a, b] \rightarrow [c, d]$. En particular, la curva suave γ es una reparametrización de $\tilde{\gamma} = \gamma \circ \tau$, ya que $\gamma = \tilde{\gamma} \circ \tau^{-1}$.
- Es fácil ver que $\tau : [c, d] \rightarrow [a, b]$ es un cambio de parámetro si y solo si τ' es continua y no nula en $[c, d]$, y $\tau(\{c, d\}) = \{a, b\}$. En efecto, es evidente que si τ es un difeomorfismo de clase C^1 entonces τ' es continua y no nula en $[c, d]$ (ya que si $\tau'(u) = 0$ entonces τ^{-1} no es diferenciable en $\tau(u)$), y por tanto o bien $\tau' > 0$ en $[c, d]$ o bien $\tau' < 0$ en dicho intervalo. En el primer caso τ es monótona creciente en $[c, d]$ y en consecuencia $\tau(c) = a$, $\tau(d) = b$, mientras que en el segundo caso τ es monótona decreciente y $\tau(c) = b$, $\tau(d) = a$. Recíprocamente, si τ' es continua y no nula en $[c, d]$ entonces τ es monótona en dicho intervalo, y por tanto inyectiva. Si, además, $\tau(\{c, d\}) = \{a, b\}$ entonces $\tau([c, d]) = [a, b]$ por el teorema de los valores intermedios, y por tanto $\tau : [c, d] \rightarrow [a, b]$ es biyectiva. Por último, la función inversa $\tau^{-1} : [a, b] \rightarrow [c, d]$ es de clase C^1 en $[a, b]$ por el teorema de la función inversa (al ser $\tau' \neq 0$).

Por lo que acabamos de ver, un cambio de parámetro $\tau : [c, d] \rightarrow [a, b]$ es o bien monótono creciente, en cuyo caso

$$\tau(c) = a, \quad \tau(d) = b,$$

o monótono decreciente, en cuyo caso

$$\tau(c) = b, \quad \tau(d) = a.$$

En el primer caso diremos que τ **preserva la orientación** de la curva γ , ya que

$$\tilde{\gamma}(c) = \gamma(\tau(c)) = \gamma(a), \quad \tilde{\gamma}(d) = \gamma(\tau(d)) = \gamma(b).$$

Análogamente, si el cambio de parámetro τ es monótono decreciente diremos que **invierte la orientación** de γ , ya que en este caso

$$\tilde{\gamma}(c) = \gamma(\tau(c)) = \gamma(b), \quad \tilde{\gamma}(d) = \gamma(\tau(d)) = \gamma(a).$$

Ejemplo 5.11. La función $\tau : [0, \pi/6] \rightarrow [0, 1/2]$ definida por $\tau(u) = \sin u$ es claramente un cambio de parámetro que preserva la orientación, ya que $\tau'(u) = \cos u$ es continua (C^∞) y positiva en $[0, \pi/6]$, y $\tau(0) = 0$, $\tau(\pi/6) = 1/2$. Sin embargo, la función $\tau : [\pi/3, \pi/2] \rightarrow [0, 1/2]$ dada por $\tau(u) = \cos u$ es un cambio de parámetro que invierte la orientación, ya que $\tau'(u) = -\sin u$ es continua (C^∞) y negativa en $[\pi/3, \pi/2]$, siendo $\tau(\pi/3) = 1/2$ y $\tau(\pi/2) = 0$. Por ejemplo, la curva parametrizada $\gamma(u) = (\cos u, \cos^2 u)$ (con $\pi/3 \leq u \leq \pi/2$) es el arco de parábola $y = x^2$ entre $x = 0$ y $x = 1/2$, recorrida comenzando en $\gamma(\pi/3) = (1/2, 1/4)$ y terminando en $\gamma(\pi/2) = (0, 0)$. \square

En general, si $\gamma : [a, b] \rightarrow \mathbb{R}^n$ es una curva parametrizada entonces la curva $\tilde{\gamma} : [a, b] \rightarrow \mathbb{R}^n$ dada por $\tilde{\gamma}(u) = \gamma(a + b - u)$ es una reparametrización de γ , con orientación opuesta a la de dicha curva. En efecto, es claro que $\tau(u) = a + b - u$ es un cambio de parámetro de $[a, b]$ en $[a, b]$ con $\tau'(u) = -1 < 0$. Denotaremos a partir de ahora esta curva, denominada **opuesta** de la curva γ , por $-\gamma$.

Es evidente que la longitud de una curva suave no puede depender de su parametrización. En otras palabras, si $\tilde{\gamma}$ es una reparametrización de una curva suave γ entonces ha de cumplirse que

$$l(\tilde{\gamma}) = l(\gamma).$$

Para comprobar analíticamente este resultado, sea $\gamma : [a, b] \rightarrow \mathbb{R}^n$ y sea $\tilde{\gamma} = \gamma \circ \tau$, siendo $\tau : [c, d] \rightarrow [a, b]$ un cambio de parámetro. Supongamos, por sencillez, que la curva suave γ es de clase C^1 en $[a, b]$. Entonces se tiene

$$\begin{aligned} l(\tilde{\gamma}) &\equiv \int_c^d \|\tilde{\gamma}'(u)\| \, du = \int_c^d \|\gamma'(\tau(u))\tau'(u)\| \, du = \int_c^d \|\gamma'(\tau(u))\| |\tau'(u)| \, du \\ &= \int_a^b \|\gamma'(t)\| \, dt \equiv l(\gamma), \end{aligned}$$

donde hemos aplicado la fórmula del cambio de variable en integrales unidimensionales (4.20).

5.1.4 Integral de trayectoria de una función escalar

Si $\gamma : [a, b] \rightarrow \mathbb{R}^n$ es una curva suave y $f : \mathbb{R}^n \rightarrow \mathbb{R}$ es una función escalar continua en $\gamma([a, b])$, definimos la **integral de f a lo largo de γ** mediante

$$\int_{\gamma} f \, ds \equiv \int_a^b f(\gamma(t)) \|\gamma'(t)\| \, dt. \quad (5.2)$$

En particular,

$$l(\gamma) = \int_{\gamma} ds.$$

Nótese que la integral del miembro derecho de (5.2) está bien definida, debido a la continuidad de f en $\gamma([a, b])$ y a la continuidad a trozos de γ' en $[a, b]$. Formalmente, $\int_{\gamma} f \, ds$ se obtiene sustituyendo f por $f(\gamma(t))$ (el valor de f a lo largo de la curva γ) y el elemento de longitud de arco ds por su definición $\|\gamma'(t)\| \, dt$, e integrando en el intervalo $[a, b]$ en que varía el parámetro de la curva.

Ejemplo 5.12. Consideremos un alambre en el espacio \mathbb{R}^3 parametrizado por la curva suave $\gamma : [a, b] \rightarrow \mathbb{R}^3$. Si $f(x, y, z) \geq 0$ es la *densidad* (masa por unidad de longitud) del alambre en el punto (x, y, z) , la masa Δm del tramo de alambre comprendido entre dos de sus puntos $\gamma(t)$ y $\gamma(t + \Delta t)$ (con $\Delta t > 0$) es aproximadamente igual a la densidad del alambre en el punto $\gamma(t)$ por la longitud del tramo de alambre entre dichos puntos, que se puede aproximar por $\|\gamma'(t)\| \Delta t$. Por tanto

$$\Delta m \simeq f(\gamma(t)) \cdot \|\gamma'(t)\| \Delta t,$$

y la *masa total* del alambre está dada por

$$m(\gamma) = \int_a^b f(\gamma(t)) \|\gamma'(t)\| \, dt \equiv \int_{\gamma} f \, ds.$$

Ejemplo 5.13. Si $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ es una función continua no negativa, la integral $\int_{\gamma} f \, ds$ puede interpretarse como el área de la *superficie reglada*³

$$S = \{(\gamma(t), z) \mid a \leq t \leq b, 0 \leq z \leq f(\gamma(t))\} \subset \mathbb{R}^3.$$

En efecto, la pequeña región de la superficie correspondiente a $t_0 \leq t \leq t_0 + \Delta t$ (con $\Delta t > 0$) se puede aproximar por un rectángulo de base $\Delta s \simeq \|\gamma'(t_0)\| \Delta t$ y altura $f(\gamma(t_0))$. \square

De nuevo, de las interpretaciones anteriores es intuitivamente claro que la integral $\int_{\gamma} f \, ds$ ha de ser *independiente de la parametrización*. Para demostrar esta afirmación analíticamente, consideremos de nuevo una curva suave $\gamma : [a, b] \rightarrow \mathbb{R}^n$ y su reparametrización $\tilde{\gamma} = \gamma \circ \tau$, donde $\tau : [c, d] \rightarrow [a, b]$ es un cambio de parámetro. La integral de $f : \mathbb{R}^n \rightarrow \mathbb{R}$ a lo largo de $\tilde{\gamma}$ está dada entonces por

$$\begin{aligned} \int_{\tilde{\gamma}} f \, ds &\equiv \int_c^d f(\tilde{\gamma}(u)) \|\tilde{\gamma}'(u)\| \, du = \int_c^d f(\gamma(\tau(u))) \|\gamma'(\tau(u))\| |\tau'(u)| \, du \\ &= \int_a^b f(\gamma(t)) \|\gamma'(t)\| \, dt \equiv \int_{\gamma} f \, ds, \end{aligned}$$

donde de nuevo hemos aplicado la fórmula del cambio de variable en integrales unidimensionales (4.20).

³Por definición, una superficie S se dice *reglada* si por cada punto $p \in S$ pasa una recta contenida en S . Por ejemplo, un cono o un cilindro (de cualquier tipo) es una superficie reglada, mientras que una esfera no lo es.

Ejemplo 5.14. Hallemos la masa de un alambre de ecuaciones $x^2 + y^2 + z^2 = 1$, $x + y + z = 0$ y densidad $\rho(x, y, z) = x^2$ (en unidades apropiadas). Para ello, es necesario primero parametrizar la curva dada, que claramente es un círculo de radio 1 centrado en el origen y contenido en el plano $x + y + z = 0$ (círculo máximo en la esfera unidad). Nótese que podemos utilizar *cualquier* parametrización de la curva, ya que por lo que acabamos de ver la integral de una función a lo largo de una curva suave no depende de su parametrización. Para hallar una parametrización del alambre, sustituimos $z = -(x + y)$ en la ecuación de la esfera, lo que proporciona

$$2(x^2 + xy + y^2) = 1 \iff \left(x + \frac{y}{2}\right)^2 + \frac{3}{4}y^2 = \frac{1}{2}.$$

Esta relación sugiere la parametrización

$$x + \frac{y}{2} = \frac{1}{\sqrt{2}} \cos t, \quad \frac{\sqrt{3}}{2} y = \frac{1}{\sqrt{2}} \sin t, \quad z = -x - y,$$

con $t \in [0, 2\pi]$. Despejando (x, y, z) de las ecuaciones anteriores se obtiene la siguiente expresión para la parametrización de la curva γ que determina el alambre:

$$\gamma(t) = \left(\frac{1}{\sqrt{2}} \cos t - \frac{1}{\sqrt{6}} \sin t, \sqrt{\frac{2}{3}} \sin t, -\frac{1}{\sqrt{2}} \cos t - \frac{1}{\sqrt{6}} \sin t\right), \quad 0 \leq t \leq 2\pi.$$

Con esta parametrización es fácil ver que

$$\gamma'(t) = \left(-\frac{1}{\sqrt{2}} \sin t - \frac{1}{\sqrt{6}} \cos t, \sqrt{\frac{2}{3}} \cos t, \frac{1}{\sqrt{2}} \sin t - \frac{1}{\sqrt{6}} \cos t\right) \implies \|\gamma'(t)\|^2 = 1.$$

Por tanto $ds = dt$, y la masa del alambre es igual a

$$m(\gamma) = \int_{\gamma} x^2 ds = \int_0^{2\pi} \left(\frac{1}{\sqrt{2}} \cos t - \frac{1}{\sqrt{6}} \sin t\right)^2 dt = \pi \left(\frac{1}{2} + \frac{1}{6}\right) = \boxed{\frac{2\pi}{3}}.$$

5.2 Integral de línea de un campo vectorial

5.2.1 Integrales de línea

Supongamos que la *fuerza* que actúa sobre una partícula situada en el punto $(x, y, z) \in \mathbb{R}^3$ está dada por el valor $F(x, y, z)$ en dicho punto de un campo vectorial continuo $F: \mathbb{R}^3 \rightarrow \mathbb{R}^3$. Si la partícula se mueve a lo largo de la trayectoria $\gamma: [a, b] \rightarrow \mathbb{R}^3$, el *trabajo* realizado por la fuerza cuando la partícula se desplaza del punto $\gamma(t)$ de la trayectoria a un punto próximo $\gamma(t + \Delta t)$ está aproximadamente dado por

$$\Delta w \simeq F(\gamma(t)) \cdot (\gamma(t + \Delta t) - \gamma(t)) \simeq F(\gamma(t)) \cdot \gamma'(t) \Delta t.$$

Por tanto el trabajo total realizado por la fuerza cuando la partícula se desplaza del punto inicial $\gamma(a)$ al punto final $\gamma(b)$ a lo largo de la trayectoria γ está dado por

$$w = \int_a^b F(\gamma(t)) \cdot \gamma'(t) dt.$$

Este hecho motiva la siguiente definición:

Definición 5.15. Sea $\gamma: [a, b] \rightarrow \mathbb{R}^n$ una curva suave, y sea $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ un campo vectorial continuo en $\gamma([a, b])$. La **integral de línea** del campo F a lo largo de la trayectoria γ se define por

$$\int_{\gamma} F(x) \cdot dx \equiv \int_{\gamma} (F_1(x) dx_1 + \cdots + F_n(x) dx_n) = \int_a^b F(\gamma(t)) \cdot \gamma'(t) dt.$$

Nótese que la integral del miembro derecho está bien definida, ya que el integrando es una función continua a trozos en $[a, b]$ por la continuidad de F en $\gamma([a, b])$ y la continuidad a trozos de γ' en dicho intervalo. Formalmente, la integral $\int_{\gamma} F(x) \cdot dx$ se evalúa *parametrizándola*, es decir sustituyendo x por $\gamma(t)$ y $dx \equiv (dx_1, \dots, dx_n)$ por el **elemento de línea** $\gamma'(t) dt$. Nótese que el elemento de línea es un *vector* cuya norma es el elemento de longitud de arco $ds = \|\gamma'(t)\| dt$.

Notación. En \mathbb{R}^2 y \mathbb{R}^3 , es frecuente utilizar la notación $\int_{\gamma} \mathbf{F}(\mathbf{r}) \cdot d\mathbf{r}$ para la integral de línea de \mathbf{F} a lo largo de γ . También es habitual suprimir el argumento x o \mathbf{r} del campo \mathbf{F} , escribiendo simplemente $\int_{\gamma} F \cdot dx$ o $\int_{\gamma} \mathbf{F} \cdot d\mathbf{r}$.

La integral de línea del campo vectorial F a lo largo de la trayectoria $\gamma : [a, b] \rightarrow \mathbb{R}^n$ se puede expresar en términos de la **componente tangencial** del campo a lo largo de dicha trayectoria, definida por

$$F_T(\gamma(t)) = F(\gamma(t)) \cdot T(t).$$

En efecto, utilizando la definición (5.1) del vector tangente $T(t)$ se obtiene

$$F_T(\gamma(t)) = \frac{F(\gamma(t)) \cdot \gamma'(t)}{\|\gamma'(t)\|},$$

y por tanto

$$\int_{\gamma} F(x) \cdot dx = \int_a^b F_T(\gamma(t)) \|\gamma'(t)\| dt \equiv \int_{\gamma} F_T ds. \quad (5.3)$$

En otras palabras, la integral de línea del campo F a lo largo de la trayectoria γ es igual a la integral a lo largo de γ de su componente tangencial.

¿Cómo se comporta la integral de línea bajo reparametrizaciones de la trayectoria γ ? Para estudiar esta cuestión, sea $\tau : [c, d] \rightarrow [a, b]$ un cambio de parámetro y consideremos la reparametrización $\tilde{\gamma} = \gamma \circ \tau$ de γ . La integral de línea del campo vectorial continuo $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ a lo largo de la curva $\tilde{\gamma}$ está dada por

$$\int_{\tilde{\gamma}} F(x) \cdot dx = \int_c^d F(\tilde{\gamma}(u)) \cdot \tilde{\gamma}'(u) du = \int_c^d F(\gamma(\tau(u))) \cdot \gamma'(\tau(u)) \tau'(u) du.$$

Como el signo de τ' es *constante* en $[c, d]$, podemos escribir

$$\begin{aligned} \int_{\tilde{\gamma}} F(x) \cdot dx &= \operatorname{sgn} \tau' \cdot \int_c^d F(\gamma(\tau(u))) \cdot \gamma'(\tau(u)) |\tau'(u)| du = \operatorname{sgn} \tau' \cdot \int_a^b F(\gamma(t)) \cdot \gamma'(t) dt \\ &\equiv \operatorname{sgn} \tau' \cdot \int_{\gamma} F(x) \cdot dx, \end{aligned}$$

donde en la última igualdad hemos aplicado la fórmula del cambio de variable en integrales unidimensionales. En definitiva:

La integral de línea de un campo vectorial es *invariante* bajo reparametrizaciones de la trayectoria que *preservan la orientación*, y *cambia de signo* bajo reparametrizaciones que *invierten la orientación*.

Esta propiedad de las integrales de línea es intuitivamente clara desde el punto de vista físico, ya que el trabajo realizado por una fuerza cambia de signo cuando se invierte el sentido de recorrido de la trayectoria. Nótese también que no hay ninguna contradicción entre esta propiedad y la ec. (5.3), ya que al cambiar la orientación de la curva cambia también el sentido del vector tangente, y por tanto el signo de la componente transversal F_T del campo F .

Comentario. Nótese, en particular, que de la propiedad que acabamos de deducir de las integrales de línea se sigue que

$$\int_{-y} F(x) \cdot dx = - \int_y F(x) \cdot dx, \quad (5.4)$$

ya que y y $-y$ están relacionadas por el cambio de parámetro $\tau(u) = a + b - u$, que invierte la orientación.

Ejemplo 5.16. Calculemos la integral de línea del campo $\mathbf{F}(x, y) = (3x^2 + 2y) \mathbf{i} - (x + 3 \cos y) \mathbf{j}$ a lo largo de la trayectoria cerrada compuesta por segmentos de recta que unen los puntos $(0, 0)$, $(2, 0)$, $(3, 1)$, $(1, 1)$ y $(0, 0)$. En primer lugar,

$$\int_y \mathbf{F} \cdot d\mathbf{r} = \sum_{i=1}^4 \int_{y_i} \mathbf{F} \cdot d\mathbf{r},$$

siendo y_i cada uno de los cuatro segmentos orientados de que consta la trayectoria. En general, si y_i es un segmento con origen en el punto \mathbf{r}_1 y final en el punto \mathbf{r}_2 entonces una parametrización de y_i está dada por

$$y_i(t) = \mathbf{r}_1 + t(\mathbf{r}_2 - \mathbf{r}_1), \quad 0 \leq t \leq 1,$$

y por tanto

$$\int_{y_i} \mathbf{F} \cdot d\mathbf{r} = \int_0^1 \mathbf{F}(\mathbf{r}_1 + t(\mathbf{r}_2 - \mathbf{r}_1)) \cdot (\mathbf{r}_2 - \mathbf{r}_1) dt = (\mathbf{r}_2 - \mathbf{r}_1) \cdot \int_0^1 \mathbf{F}(\mathbf{r}_1 + t(\mathbf{r}_2 - \mathbf{r}_1)) dt.$$

Utilizando esta fórmula se obtiene

$$\begin{aligned} \int_{y_1} \mathbf{F} \cdot d\mathbf{r} &= 2 \int_0^1 F_1(2t, 0) dt = \int_0^1 24t^2 dt = 8, \\ \int_{y_2} \mathbf{F} \cdot d\mathbf{r} &= \int_0^1 (1, 1) \cdot \mathbf{F}((2, 0) + t(1, 1)) dt = \int_0^1 [F_1(2+t, t) + F_2(2+t, t)] dt \\ &= \int_0^1 [3(2+t)^2 + 2t - (2+t + 3 \cos t)] dt = \int_0^1 (3t^2 + 13t + 10) dt - 3 \sin 1 \\ &= 1 + \frac{13}{2} + 10 - 3 \sin 1 = \frac{35}{2} - 3 \sin 1, \\ \int_{y_3} \mathbf{F} \cdot d\mathbf{r} &= -2 \int_0^1 F_1((3, 1) - 2t(1, 0)) dt = -2 \int_0^1 F_1(3-2t, 1) dt = - \int_1^3 F_1(u, 1) du \\ &= - \int_1^3 (3u^2 + 2) du = -26 - 4 = -30, \\ \int_{y_4} \mathbf{F} \cdot d\mathbf{r} &= - \int_0^1 (1, 1) \cdot \mathbf{F}((1, 1) - t(1, 1)) dt = - \int_0^1 [F_1(1-t, 1-t) + F_2(1-t, 1-t)] dt \\ &= - \int_0^1 [F_1(t, t) + F_2(t, t)] dt = - \int_0^1 (3t^2 + 2t - t - 3 \cos t) dt = -\frac{3}{2} + 3 \sin 1. \end{aligned}$$

Por tanto

$$\int_y \mathbf{F} \cdot d\mathbf{r} = 8 + \frac{35}{2} - 3 \sin 1 - 30 - \frac{3}{2} + 3 \sin 1 = \boxed{-6}.$$

Veremos más adelante cómo evaluar esta integral de una forma mucho más sencilla utilizando el teorema de Green.

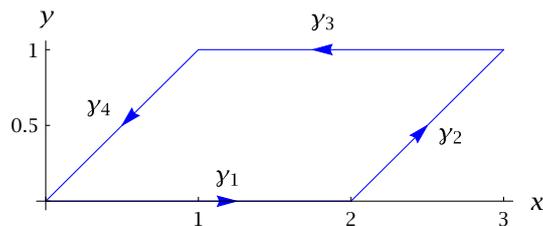


Figura 5.3: Curva cerrada γ del Ejemplo 5.16.

5.2.2 Campos conservativos

En general, la integral de línea de un campo de vectores a lo largo de trayectorias con los mismos extremos depende de la trayectoria considerada. En efecto, sea, por ejemplo, $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ el campo vectorial dado por $F(x, y) = y \mathbf{i} - x \mathbf{j}$, y consideremos las trayectorias

$$\gamma_1(t) = (t, t), \quad \gamma_2(t) = (t, t^2); \quad 0 \leq t \leq 1,$$

que unen el origen con el punto $(1, 1)$. La integral de línea del campo F a lo largo de la primera trayectoria (segmento) está dada por

$$\int_{\gamma_1} \mathbf{F}(x, y) \cdot d\mathbf{r} = \int_0^1 (t, -t) \cdot (1, 1) dt = 0,$$

mientras que la integral a lo largo de la segunda curva (arco de parábola) es igual a

$$\int_{\gamma_2} \mathbf{F}(x, y) \cdot d\mathbf{r} = \int_0^1 (t^2, -t) \cdot (1, 2t) dt = - \int_0^1 t^2 dt = -\frac{1}{3} \neq 0.$$

Por tanto en este caso

$$\int_{\gamma_1} \mathbf{F} \cdot d\mathbf{r} \neq \int_{\gamma_2} \mathbf{F} \cdot d\mathbf{r},$$

aún cuando γ_1 y γ_2 tienen los mismos extremos. Nótese que este hecho es bien conocido en Física, ya que es sabido que el trabajo realizado por una fuerza depende en general de la trayectoria recorrida por la partícula, y no solo de sus posiciones inicial y final.

Definición 5.17. Diremos que un campo vectorial $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ continuo en un abierto $U \subset \mathbb{R}^n$ es **conservativo** en U si

$$\int_{\gamma_1} F(x) \cdot dx = \int_{\gamma_2} F(x) \cdot dx \tag{5.5}$$

para todo par de curvas suaves $\gamma_{1,2}$ con los mismos extremos cuyas imágenes estén contenidas en U .

Notación. A partir de ahora, diremos abreviadamente que una curva parametrizada $\gamma : [a, b] \rightarrow \mathbb{R}^n$ está contenida en un conjunto $U \subset \mathbb{R}^n$ para indicar que su imagen $\gamma([a, b])$ está contenida en dicho conjunto. □

Por ejemplo, es claro que el campo vectorial del ejemplo anterior *no* es conservativo en \mathbb{R}^3 . Una primera observación interesante es que el carácter conservativo de un campo de vectores se puede determinar utilizando únicamente curvas *cerradas*:

Proposición 5.18. Un campo vectorial $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ continuo en un abierto $U \subset \mathbb{R}^n$ es conservativo en U si y solo

$$\int_{\gamma} F(x) \cdot dx = 0, \tag{5.6}$$

para toda curva suave cerrada γ contenida en U .

Demostración.

\Rightarrow) Si $\gamma : [a, b] \rightarrow \mathbb{R}^n$ es una curva cerrada contenida en U , la curva constante $\tilde{\gamma}(t) = \gamma(a)$, para todo $t \in [a, b]$, tiene los mismos extremos que γ . Al ser F conservativo se tiene

$$\int_{\gamma} F(x) \cdot dx = \int_{\tilde{\gamma}} F(x) \cdot dx = 0,$$

ya que $\tilde{\gamma}' = 0$.

\Leftarrow) Sean $\gamma_1 : [a, b] \rightarrow U \subset \mathbb{R}^n$ y $\gamma_2 : [c, d] \rightarrow U \subset \mathbb{R}^n$ dos curvas suaves contenidas en U con los mismos extremos (cf. la Fig. 5.4). Definimos entonces la curva $\gamma : [a, b + d - c] \rightarrow U \subset \mathbb{R}^n$ mediante

$$\gamma(t) = \begin{cases} \gamma_1(t), & a \leq t \leq b, \\ \gamma_2(d + b - t), & b \leq t \leq b + d - c. \end{cases}$$

Es evidente que γ es una curva suave (ya que es continua en $t = b$, al ser $\gamma_1(b) = \gamma_2(d)$) y cerrada (ya que $\gamma(b + d - c) = \gamma_2(c) = \gamma_1(a) \equiv \gamma(a)$). Al ser $\gamma \subset U$ cerrada por hipótesis se verifica (5.6), y por tanto

$$\int_{\gamma} F(x) \cdot dx = 0.$$

Pero la restricción de la curva γ al intervalo $[b, b + d - c]$ difiere de γ_2 en la reparametrización $t \mapsto d + b - t$, que invierte la orientación. Por tanto

$$0 = \int_{\gamma} F(x) \cdot dx = \int_{\gamma_1} F(x) \cdot dx + \int_{-\gamma_2} F(x) \cdot dx = \int_{\gamma_1} F(x) \cdot dx - \int_{\gamma_2} F(x) \cdot dx,$$

y el campo F es conservativo en U . □

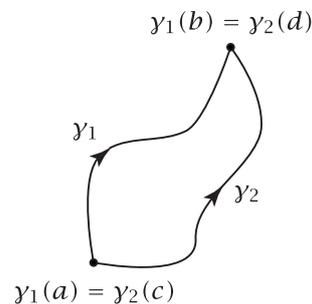


Figura 5.4: Curvas γ_1 y γ_2 en la demostración de la Proposición 5.18.

Por ejemplo, el campo vectorial del Ejemplo 5.16 *no* es conservativo. Nuestro siguiente objetivo es obtener una caracterización más directa de los campos conservativos. Para ello es de fundamental importancia el siguiente resultado, que se puede considerar como una generalización del teorema fundamental del Cálculo a integrales de línea:

Proposición 5.19 (teorema fundamental del Cálculo). Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función de clase C^1 en el abierto $U \subset \mathbb{R}^n$. Si $\gamma : [a, b] \rightarrow U$ es una curva suave de extremos $\gamma(a) = p$, $\gamma(b) = q$ entonces

$$\int_{\gamma} \nabla f(x) \cdot dx = f(q) - f(p).$$

En particular, el campo vectorial ∇f es conservativo en U .

Demostración. En primer lugar, nótese que el campo vectorial ∇f es continuo (y por tanto integrable a lo largo de γ) en U al ser por hipótesis $f \in C^1(U)$. Por otra parte, de la regla de la

cadena y del teorema fundamental del Cálculo para integrales unidimensionales se sigue que

$$\int_{\gamma} \nabla f(x) \cdot dx = \int_a^b \nabla f(\gamma(t)) \cdot \gamma'(t) dt = \int_a^b \frac{d}{dt} f(\gamma(t)) dt = f(\gamma(t)) \Big|_{t=a}^{t=b} = f(\gamma(b)) - f(\gamma(a)) \equiv f(q) - f(p).$$

En particular, el campo ∇f es conservativo, ya que $\int_{\gamma} \nabla f(x) \cdot dx$ solo depende de los extremos p y q de la trayectoria γ . □

Notación. Si $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ es un campo *conservativo* en un abierto $U \subset \mathbb{R}^n$ la integral de F a lo largo de cualquier curva suave $\gamma \subset U$ solo depende de los extremos de γ . Por tanto si F es conservativo en U utilizaremos a partir de ahora la notación

$$\boxed{\int_p^q F(x) \cdot dx}$$

para denotar la integral $\int_{\gamma} F(x) \cdot dx$, donde γ es *cualquier* curva suave contenida en U con extremos p y q .

Ejemplo 5.20. Consideremos la integral de línea $\int_{\gamma} \mathbf{F} \cdot d\mathbf{r}$, siendo $\mathbf{F} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ el campo vectorial definido por

$$\mathbf{F}(x, y, z) = e^x \sin y \mathbf{i} + e^x \cos y \mathbf{j} + z^2 \mathbf{k}$$

y $\gamma : [0, 1] \rightarrow \mathbb{R}^3$ la curva dada por

$$\gamma(t) = (t^{3/2}, t^3, e^{t\sqrt{t}}).$$

Nótese que γ es de clase C^1 (aunque no de clase C^2). El campo \mathbf{F} es claramente el gradiente de una función escalar, ya que

$$e^x \sin y \mathbf{i} + e^x \cos y \mathbf{j} = \nabla(e^x \sin y), \quad z^2 \mathbf{k} = \nabla\left(\frac{z^3}{3}\right) \implies \mathbf{F}(x, y, z) = \nabla\left(e^x \sin y + \frac{z^3}{3}\right).$$

Por tanto la integral considerada es igual a

$$\int_{\gamma} \mathbf{F} \cdot d\mathbf{r} = e^x \sin y + \frac{z^3}{3} \Big|_{\gamma(0)}^{\gamma(1)} = e^x \sin y + \frac{z^3}{3} \Big|_{(0,0,1)}^{(1,1,e)} = \boxed{e \sin 1 + \frac{1}{3}(e^3 - 1)}. \quad \square$$

Como acabamos de ver, el gradiente de una función escalar de clase C^1 es siempre un campo conservativo. Es natural preguntarse si estos son los *únicos* campos conservativos o, equivalentemente, si *todo* campo conservativo es el gradiente de una función escalar. Para responder a esta pregunta, supongamos en primer lugar que $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ es un campo conservativo en un abierto *conexo* $U \subset \mathbb{R}^n$, y que F es igual al gradiente de una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ de clase C^1 en U . Si fijamos un punto $x_0 \in U$ y x es cualquier punto de U , al ser U *conexo* hay una curva suave γ contenida en U cuyos extremos son los puntos x_0 y x . Por el teorema fundamental del Cálculo,

$$\int_{\gamma} F(\gamma) \cdot d\gamma \equiv \int_{x_0}^x F(\gamma) \cdot d\gamma = \int_{x_0}^x \nabla f(\gamma) \cdot d\gamma = f(x) - f(x_0).$$

De esta fórmula se deduce que si el campo F es el gradiente de una función escalar f , dicha función tienen que estar *necesariamente* dada por la fórmula

$$\boxed{f(x) = \int_{x_0}^x F(\gamma) \cdot d\gamma + c,} \tag{5.7}$$

donde $c = f(x_0)$ es una constante arbitraria. Nótese que este resultado concuerda con lo visto en la Sección 2.9 (cf. el Ejercicio 28), donde ya observamos que la función f está determinada por el campo F a menos de una constante arbitraria. Veremos a continuación que si F es un campo conservativo en un abierto *conexo* U entonces F es el gradiente de la función f definida por la ec. (5.7).

Proposición 5.21. Si $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ es un campo vectorial conservativo en el abierto conexo $U \subset \mathbb{R}^n$, entonces existe una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ de clase C^1 en U tal que $F = \nabla f$ en U .

Demostración. Si F es conservativo en un abierto conexo U , definimos la función $f : U \rightarrow \mathbb{R}$ por la ec. (5.7). Esta definición es correcta, ya que en este caso la integral de F a lo largo de cualquier curva suave que una x_0 con x (que ha de existir, al ser U conexo) solo depende de los puntos x_0 (que está fijo) y x . Probaremos a continuación que la función f definida por (5.7) efectivamente satisface

$$\frac{\partial f}{\partial x_i}(x) = F_i(x), \quad \forall x \in U, \quad 1 \leq i \leq n. \quad (5.8)$$

Nótese que esto implica automáticamente que f es de clase C^1 en U , ya que las componentes F_i del campo F son continuas en U al ser dicho campo conservativo (cf. la Definición 5.17). Para demostrar las igualdades (5.8), basta notar que

$$\begin{aligned} \frac{\partial f}{\partial x_i}(x) &= \lim_{h \rightarrow 0} \frac{1}{h} [f(x + he_i) - f(x)] \equiv \lim_{h \rightarrow 0} \frac{1}{h} \left[\int_{x_0}^{x+he_i} F(y) \cdot dy - \int_{x_0}^x F(y) \cdot dy \right] \\ &= \lim_{h \rightarrow 0} \frac{1}{h} \left[\int_x^{x_0} F(y) \cdot dy + \int_{x_0}^{x+he_i} F(y) \cdot dy \right] = \lim_{h \rightarrow 0} \frac{1}{h} \int_x^{x+he_i} F(y) \cdot dy. \end{aligned}$$

Como U es abierto, existe $r > 0$ tal que la bola de centro x y radio r está contenida en U . Si $|h| < r$ entonces el segmento $\gamma(t) = x + te_i$, $0 \leq t \leq h$, que une los puntos x y $x + he_i$ está contenido en $B_r(x) \subset U$. Utilizando este segmento para evaluar la integral del miembro derecho de la fórmula anterior se obtiene

$$\frac{\partial f}{\partial x_i}(x) = \lim_{h \rightarrow 0} \frac{1}{h} \int_0^h F(x + te_i) \cdot e_i dt = \lim_{h \rightarrow 0} \frac{1}{h} \int_0^h F_i(x + te_i) dt.$$

Como la función $g(t) \equiv F_i(x + te_i)$ es continua en t por la continuidad de F en U (en virtud de la regla de la cadena), aplicando el teorema fundamental del Cálculo para funciones de una variable se obtiene

$$\frac{\partial f}{\partial x_i}(x) = \lim_{h \rightarrow 0} \frac{1}{h} \int_0^h g(t) dt = \frac{d}{dh} \int_0^h g(t) dt \Big|_{h=0} = g(0) = F_i(x). \quad \square$$

De la proposición anterior y el teorema fundamental del Cálculo (Proposición 5.19) se sigue el siguiente corolario importante:

Corolario 5.22. Un campo vectorial $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ es conservativo en un abierto conexo $U \subset \mathbb{R}^n$ si y solo si existe una función escalar $f : \mathbb{R}^n \rightarrow \mathbb{R}$ de clase C^1 en U tal que $F = \nabla f$ en U .

A la función f del corolario anterior, que como sabemos está determinada por el campo vectorial F a menos de una constante arbitraria, se le denomina **potencial** del campo F .

Como adelantamos en la Sección 2.9, si $U \subset \mathbb{R}^n$ es un abierto *simplemente conexo* existe un criterio sencillo para determinar si un campo vectorial $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ de clase C^1 en U es el gradiente de una función escalar y, por tanto, es conservativo en dicho conjunto. (Recuérdese que un abierto conexo $U \subset \mathbb{R}^n$ es simplemente conexo si toda curva cerrada continua contenida en U puede deformarse de forma continua a un punto sin salirse de U .) Más precisamente:

Proposición 5.23. Un campo vectorial $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ de clase C^1 en un abierto simplemente conexo U es conservativo en dicho conjunto si y solo si

$$\frac{\partial F_i}{\partial x_j}(x) = \frac{\partial F_j}{\partial x_i}(x), \quad \forall x \in U, \quad 1 \leq i \neq j \leq n. \quad (5.9)$$

En particular, si $n = 3$ las ecuaciones anteriores son equivalentes a

$$\nabla \times \mathbf{F}(x, y, z) = 0, \quad \forall (x, y, z) \in U. \tag{5.10}$$

De hecho, las condiciones (5.9) o (5.10) son *necesarias* para que el campo vectorial F de clase C^1 sea conservativo en un abierto conexo U (¡incluso si U no es simplemente conexo!). En efecto, si $F = \nabla f$ en U entonces f es de clase C^2 en U , y del lema de Schwarz se sigue que

$$\frac{\partial F_i}{\partial x_j}(x) - \frac{\partial F_j}{\partial x_i}(x) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x) - \frac{\partial^2 f}{\partial x_i \partial x_j}(x) = 0, \quad \forall x \in U.$$

Ejemplo 5.24. Consideremos el campo vectorial $\mathbf{F} : U \equiv \mathbb{R}^2 - \{0\} \rightarrow \mathbb{R}^2$ dado por

$$\mathbf{F}(x, y) = \frac{1}{x^2 + y^2} (-y \mathbf{i} + x \mathbf{j}).$$

Este campo satisface la condición (5.9), ya que

$$\frac{\partial F_1}{\partial y} = \frac{\partial}{\partial y} \left(\frac{-y}{x^2 + y^2} \right) = \frac{y^2 - x^2}{(x^2 + y^2)^2}, \quad \frac{\partial F_2}{\partial x} = \frac{\partial}{\partial x} \left(\frac{x}{x^2 + y^2} \right) = \frac{y^2 - x^2}{(x^2 + y^2)^2}.$$

Sin embargo, el conjunto U no es simplemente conexo, ya que (por ejemplo) la circunferencia unidad $x^2 + y^2 = 1$ es una curva cerrada continua que no se puede deformar a un punto sin salirse de U . Por lo tanto, en este caso la condición (5.9) no garantiza que el campo \mathbf{F} sea conservativo. De hecho, no lo es, ya que la integral de \mathbf{F} a lo largo de la circunferencia unidad recorrida en sentido antihorario $\gamma(t) = (\cos t, \sin t)$ ($0 \leq t \leq 2\pi$) es no nula:

$$\int_{\gamma} \mathbf{F}(x, y) \cdot d\mathbf{r} = \int_0^{2\pi} (-\sin t, \cos t) \cdot (-\sin t, \cos t) dt = \int_0^{2\pi} dt = 2\pi.$$

El campo \mathbf{F} es por supuesto conservativo en cualquier abierto simplemente conexo que no contenga el origen, en virtud de la Proposición 5.23. Por ejemplo, si U es el semiplano $x > 0$ (o $x < 0$) es fácil comprobar que

$$\mathbf{F}(x, y) = \nabla \left(\arctan \left(\frac{y}{x} \right) \right). \quad \square$$

Si $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ es un campo vectorial de clase C^1 que cumple las condiciones (5.9) en un abierto simplemente conexo $U \subset \mathbb{R}^n$, está garantizado que F es conservativo en U . ¿Cómo se calcula, en tal caso, el potencial f de dicho campo? En primer lugar, podemos aplicar la ec. (5.7) evaluando la integral del miembro derecho utilizando *cualquier* curva que una un punto fijo $x_0 \in U$ con el punto variable $x \in U$. Por ejemplo, si $U = \mathbb{R}^n$ la fórmula (2.54) de la Sección 2.9 se obtiene tomando $x_0 = (c_1, \dots, c_n)$ y evaluando la integral $\int_{x_0}^x F(y) \cdot dy$ a lo largo de la línea quebrada que une los puntos

$$x_0, (c_1, \dots, c_{n-1}, x_n), \dots, (c_1, \dots, c_i, x_{i+1}, \dots, x_n), \dots, (c_1, x_2, \dots, x_n), x.$$

Si en lugar de esta curva utilizamos el segmento que une el punto fijo $x_0 \in \mathbb{R}^n$ con el punto variable $x \in \mathbb{R}^n$ (suponiendo que dicho segmento esté enteramente contenido en el abierto U) obtenemos la fórmula alternativa

$$f(x) = (x - x_0) \cdot \int_0^1 F(x_0 + t(x - x_0)) dt + c.$$

De hecho, la fórmula anterior es válida en cualquier abierto *convexo* U , ya que por definición un conjunto de este tipo contiene el segmento que une dos puntos cualesquiera de dicho conjunto.

Sin embargo, en la mayoría de los problemas prácticos el método más sencillo para calcular el potencial f es integrar sucesivamente respecto de cada variables las n ecuaciones

$$\frac{\partial f}{\partial x_i}(x) = F_i(x), \quad 1 \leq i \leq n$$

que determinan el potencial.

Ejemplo 5.25. Consideremos el campo vectorial $F: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ dado por

$$F(x, y, z) = (ayz + bx + cz)\mathbf{i} + (axz + bz)\mathbf{j} + (axy + by)\mathbf{k},$$

con $a, b, c \in \mathbb{R}$. El campo F es de clase C^∞ en \mathbb{R}^3 , y su rotacional está dado por

$$\nabla \times F(x, y, z) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ ayz + bx + cz & axz + bz & axy + by \end{vmatrix} = c\mathbf{j}.$$

Como \mathbb{R}^3 es simplemente conexo, F es conservativo en \mathbb{R}^3 si y solo si $c = 0$. En tal caso, su potencial $f(x, y, z)$ se calcula resolviendo el sistema

$$\frac{\partial f}{\partial x} = ayz + bx, \quad \frac{\partial f}{\partial y} = axz + bz, \quad \frac{\partial f}{\partial z} = axy + by.$$

Integrando la primera ecuación respecto de x se obtiene

$$f(x, y, z) = axyz + \frac{1}{2}bx^2 + g(y, z).$$

Sustituyendo en la segunda ecuación e integrando en y se obtiene

$$axz + \frac{\partial g}{\partial y} = axz + bz \implies g(y, z) = byz + h(z).$$

Finalmente, sustituyendo en la última ecuación e integrando respecto de z queda

$$axy + by + h'(z) = axy + by \implies h(z) = k = \text{const.}$$

Por tanto

$$f(x, y, z) = axyz + \frac{1}{2}bx^2 + byz + k,$$

con $k \in \mathbb{R}$ constante.

5.3 Integral de una función escalar sobre una superficie de \mathbb{R}^3

5.3.1 Superficies parametrizadas en \mathbb{R}^3

Intuitivamente, una superficie en \mathbb{R}^3 es un subconjunto bidimensional $S \subset \mathbb{R}^3$, donde el término “bidimensional” significa que los puntos de S están determinados por dos parámetros independientes. Más precisamente:

Definición 5.26.

- Una **superficie parametrizada** en \mathbb{R}^3 es una aplicación $\sigma : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$, donde D es un conjunto con interior no vacío. Una **superficie** en \mathbb{R}^3 es la imagen $\sigma(D)$ de una superficie parametrizada.
- Una superficie parametrizada es **suave** si la aplicación σ es de clase C^1 en D .
- Una superficie suave es **regular** si $\text{rank } D\sigma = 2$ en $\overset{\circ}{D}$, salvo a lo sumo en un conjunto de medida cero. Un punto $(u, v) \in D$ tal que $\text{rank } D\sigma(u, v) = 2$ se denomina **regular**.
- Una superficie parametrizada es **simple** si la aplicación σ es *inyectiva*. En tal caso, la imagen bajo σ de $\partial D \cap D$ se denomina el **borde** de la superficie.

Comentarios.

- Por definición, diremos que $\sigma : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ es de clase C^1 en el conjunto D (no necesariamente abierto) si es de clase C^1 en un *abierto* que contiene a D .
- Si $(u_0, v_0) \in \overset{\circ}{D}$ es un punto *regular*, hay un entorno U de (u_0, v_0) tal que $\sigma(U)$ es la *gráfica* de una función de dos variables. En efecto, supongamos que, por ejemplo,

$$\det \frac{\partial(\sigma_1, \sigma_2)}{\partial(u, v)}(u_0, v_0) \neq 0.$$

Por el teorema de la función inversa aplicado a la función $(\sigma_1, \sigma_2) : D \rightarrow \mathbb{R}^2$, en un entorno U del punto (u_0, v_0) las ecuaciones

$$x = \sigma_1(u, v), \quad y = \sigma_2(u, v)$$

permiten expresar las variables u y v como sendas funciones $u(x, y)$, $v(x, y)$ de las coordenadas $(x, y) \in V$, donde V es la imagen de U bajo (σ_1, σ_2) (es decir, la proyección de $\sigma(U)$ sobre el plano (x, y)). (Nótese que, de nuevo en virtud del teorema de la función inversa, V es abierto.) Los puntos de $\sigma(U)$ están por tanto dados por tanto por la ecuación

$$z = \sigma_3(u(x, y), v(x, y)), \quad (x, y) \in V.$$

Notación. Con un ligero abuso de notación, denotaremos frecuentemente a partir de ahora por $x(u, v)$, $y(u, v)$, $z(u, v)$ a las tres componentes del vector $\sigma(u, v) \in \mathbb{R}^3$. También utilizaremos a menudo la notación $\mathbf{r}(u, v)$ en lugar de $\sigma(u, v)$. En particular, denotaremos las derivadas parciales $\frac{\partial \sigma}{\partial u}(u, v)$ y $\frac{\partial \sigma}{\partial v}(u, v)$ por $\frac{\partial \mathbf{r}}{\partial u}(u, v)$ y $\frac{\partial \mathbf{r}}{\partial v}(u, v)$, o abreviadamente por $\mathbf{r}_u(u, v)$ y $\mathbf{r}_v(u, v)$.

Ejemplo 5.27. La gráfica de una función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ en un conjunto D (de interior no vacío) es una superficie parametrizada, dada por

$$\sigma(u, v) = (u, v, f(u, v)), \quad (u, v) \in D.$$

Esta superficie es suave si y solo si la función f es de clase C^1 en D . Nótese que la gráfica de una función $f : D \rightarrow \mathbb{R}^2$ es siempre *regular* (si es suave), ya que en este caso

$$D\sigma(u, v) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \frac{\partial f}{\partial u}(u, v) & \frac{\partial f}{\partial v}(u, v) \end{pmatrix}$$

tiene rango 2 en cualquier punto de D .

Ejemplo 5.28. La esfera de centro (x_0, y_0, z_0) y radio $a > 0$ se puede parametrizar como

$$\sigma(\theta, \varphi) = (x_0 + a \sin \theta \cos \varphi, y_0 + a \sin \theta \sin \varphi, z_0 + a \cos \theta), \quad (5.11)$$

con $(\theta, \varphi) \in [0, \pi] \times [0, 2\pi)$. Esta parametrización es evidentemente de clase C^∞ , pero no es simple (ya que $\sigma(0, \varphi) = (x_0, y_0, z_0 + a)$ y $\sigma(\pi, \varphi) = (x_0, y_0, z_0 - a)$ para todo $\varphi \in [0, 2\pi)$), y su derivada $D\sigma$ está dada por

$$D\sigma(\theta, \varphi) = \begin{pmatrix} a \cos \theta \cos \varphi & -a \sin \theta \sin \varphi \\ a \cos \theta \sin \varphi & a \sin \theta \cos \varphi \\ -a \sin \theta & 0 \end{pmatrix}.$$

Los tres menores de orden 2 de $D\sigma$ son

$$\det \frac{\partial(x, y)}{\partial(\theta, \varphi)} = a^2 \sin \theta \cos \theta, \quad \det \frac{\partial(x, z)}{\partial(\theta, \varphi)} = -a^2 \sin^2 \theta \sin \varphi, \quad \det \frac{\partial(y, z)}{\partial(\theta, \varphi)} = a^2 \sin^2 \theta \cos \varphi,$$

y por tanto

$$\text{rank } D\sigma(\theta, \varphi) = 2 \iff \theta \neq 0, \pi.$$

En particular, $\text{rank } D\sigma = 2$ en $\overset{\circ}{D} = (0, \pi) \times (0, 2\pi)$, por lo que la superficie es regular. Una parametrización semejante se puede utilizar para el *elipsoide* de semiejes $a, b, c > 0$ y centro (x_0, y_0, z_0) , a saber

$$\sigma(\theta, \varphi) = (x_0 + a \sin \theta \cos \varphi, y_0 + b \sin \theta \sin \varphi, z_0 + c \cos \theta), \quad (\theta, \varphi) \in [0, \pi] \times [0, 2\pi).$$

Ejemplo 5.29. La *superficie de revolución* obtenida girando la gráfica de la curva de ecuación $x = f(z)$ en el intervalo $[h_1, h_2]$ alrededor del eje z (donde $f(z) \geq 0$ para todo $z \in [h_1, h_2]$) está descrita en coordenadas cilíndricas por la ecuación $\rho = f(z)$, y por tanto se puede parametrizar como

$$\sigma(\varphi, z) = (f(z) \cos \varphi, f(z) \sin \varphi, z), \quad (\varphi, z) \in [0, 2\pi) \times [h_1, h_2]. \quad (5.12)$$

Esta parametrización es suave si la función f es de clase C^1 en $[h_1, h_2]$. La matriz jacobiana de σ está dada por

$$D\sigma(\varphi, z) = \begin{pmatrix} -f(z) \sin \varphi & f'(z) \cos \varphi \\ f(z) \cos \varphi & f'(z) \sin \varphi \\ 0 & 1 \end{pmatrix},$$

y por tanto

$$\det \frac{\partial(x, y)}{\partial(\varphi, z)} = -f(z)f'(z), \quad \det \frac{\partial(x, z)}{\partial(\varphi, z)} = -f(z) \sin \varphi, \quad \det \frac{\partial(y, z)}{\partial(\varphi, z)} = f(z) \cos \varphi.$$

En particular, el rango de $D\sigma$ es igual a 2 si y solo si $f(z) > 0$ (ya que f es no negativa). Por tanto la superficie de revolución (5.12) es automáticamente *regular* si y solo si f se anula a lo sumo en un conjunto de medida 0 (por ejemplo, en un número finito de puntos). También es inmediato comprobar que la superficie parametrizada (5.12) es *simple* si la función f es *positiva* en el intervalo $[h_1, h_2]$ (ejercicio).

Ejemplo 5.30. Consideremos el *hiperboloide de una hoja* de ecuación $x^2 + y^2 - z^2 = a^2$ (con $a > 0$). Esta es una superficie de revolución, obtenida girando la gráfica de la curva $x = \sqrt{a^2 + z^2}$ (con $z \in \mathbb{R}$) alrededor del eje z . Como hemos visto en el ejemplo anterior, podemos parametrizar esta superficie por

$$x = \sqrt{a^2 + z^2} \cos \varphi, \quad y = \sqrt{a^2 + z^2} \sin \varphi, \quad z = z; \quad (\varphi, z) \in [0, 2\pi) \times \mathbb{R}.$$

Otra parametrización de esta superficie se obtiene haciendo $z = a \sinh u$, con $u \in \mathbb{R}$, es decir

$$x = a \cosh u \cos \varphi, \quad y = a \cosh u \sin \varphi, \quad z = a \sinh u; \quad (\varphi, u) \in [0, 2\pi) \times \mathbb{R}. \quad (5.13)$$

Como

$$\begin{cases} \det \frac{\partial(x, y)}{\partial(\varphi, u)} = -a^2 \sinh u \cosh u, & \det \frac{\partial(x, z)}{\partial(\varphi, u)} = -a^2 \cosh^2 u \sin \varphi, \\ \det \frac{\partial(y, z)}{\partial(\varphi, u)} = a^2 \cosh^2 u \cos \varphi, \end{cases}$$

la parametrización (5.13) es *regular*, y es también inmediato comprobar que es *simple*.

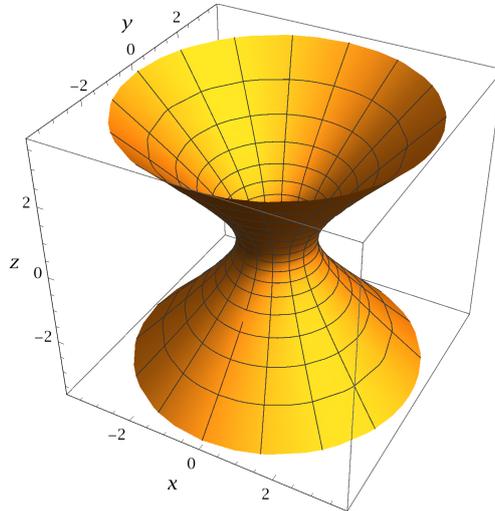


Figura 5.5: Hiperboloide $x^2 + y^2 - z^2 = 1$.

5.3.2 Producto vectorial fundamental

Sea $\sigma : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ una superficie parametrizada, y sea $(u_0, v_0) \in \overset{\circ}{D}$ un punto regular (es decir, $\text{rank } D\sigma(u_0, v_0) = 2$). Los dos segmentos $v = v_0$ (horizontal) y $u = u_0$ (vertical) en el plano (u, v) dan lugar a sendas curvas $\gamma_1(u) = \sigma(u, v_0)$ y $\gamma_2(v) = \sigma(u_0, v)$ contenidas en la superficie $S = \sigma(D)$ que se cortan en el punto $\gamma_1(u_0) = \gamma_2(v_0) = \sigma(u_0, v_0)$. Los vectores tangentes a estas curvas en dicho punto están dados por

$$\gamma_1'(u_0) = \frac{\partial \sigma}{\partial u}(u_0, v_0), \quad \gamma_2'(v_0) = \frac{\partial \sigma}{\partial v}(u_0, v_0).$$

Estos vectores son tangentes a la superficie S en el punto $\sigma(u_0, v_0)$, ya que son tangentes en dicho punto a sendas curvas contenidas en S . Son, además, *linealmente independientes* en virtud de la condición $\text{rank } D\sigma(u_0, v_0) = 2$, ya que

$$\begin{aligned} \frac{\partial \sigma}{\partial u} \times \frac{\partial \sigma}{\partial v} &\equiv \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} = \left(\frac{\partial y}{\partial u} \frac{\partial z}{\partial v} - \frac{\partial z}{\partial u} \frac{\partial y}{\partial v}, \frac{\partial z}{\partial u} \frac{\partial x}{\partial v} - \frac{\partial x}{\partial u} \frac{\partial z}{\partial v}, \frac{\partial x}{\partial u} \frac{\partial y}{\partial v} - \frac{\partial y}{\partial u} \frac{\partial x}{\partial v} \right) \\ &\equiv \left(\det \frac{\partial(y, z)}{\partial(u, v)}, \det \frac{\partial(z, x)}{\partial(u, v)}, \det \frac{\partial(x, y)}{\partial(u, v)} \right) \end{aligned}$$

tiene por componentes los tres menores de orden 2 de la matriz $D\sigma$. Por tanto:

Si (u_0, v_0) es un punto regular, los dos vectores $\frac{\partial \mathbf{r}}{\partial u}(u_0, v_0)$ y $\frac{\partial \mathbf{r}}{\partial v}(u_0, v_0)$ forman una *base* del *plano tangente* en el punto $\mathbf{r}(u_0, v_0)$ a la superficie S parametrizada por la función $\mathbf{r}(u, v)$.

Otra consecuencia importante de las consideraciones anteriores es que el vector

$$\mathbf{r}_u(u_0, v_0) \times \mathbf{r}_v(u_0, v_0),$$

llamado **producto vectorial fundamental**, es *normal* a la superficie S en el punto $\mathbf{r}(u_0, v_0)$, ya que es normal a los dos vectores que generan el plano tangente a la superficie en dicho punto. El **vector unitario normal** a la superficie en un punto $\mathbf{r}(u_0, v_0)$ (tal que (u_0, v_0) es regular) se define por

$$\mathbf{n}(u_0, v_0) = \frac{\mathbf{r}_u(u_0, v_0) \times \mathbf{r}_v(u_0, v_0)}{\|\mathbf{r}_u(u_0, v_0) \times \mathbf{r}_v(u_0, v_0)\|}. \quad (5.14)$$

En particular, la ecuación del plano tangente a la superficie en el punto $\mathbf{r}(u_0, v_0) \equiv (x_0, y_0, z_0)$ es

$$(x - x_0, y - y_0, z - z_0) \cdot (\mathbf{r}_u(u_0, v_0) \times \mathbf{r}_v(u_0, v_0)) = 0,$$

o equivalentemente

$$\begin{vmatrix} x - x_0 & y - y_0 & z - z_0 \\ x_u(u_0, v_0) & y_u(u_0, v_0) & z_u(u_0, v_0) \\ x_v(u_0, v_0) & y_v(u_0, v_0) & z_v(u_0, v_0) \end{vmatrix} = 0.$$

Ejemplo 5.31. Sea S la gráfica de la función $f : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}$, parametrizada por

$$\mathbf{r}(x, y) = (x, y, f(x, y)).$$

Si f es de clase C^1 en D , los vectores tangentes \mathbf{r}_x y \mathbf{r}_y están dados por

$$\mathbf{r}_x = (1, 0, f_x(x, y)), \quad \mathbf{r}_y = (0, 1, f_y(x, y)),$$

y por tanto

$$\mathbf{r}_x \times \mathbf{r}_y = (-f_x(x, y), -f_y(x, y), 1)$$

no se anula en ningún punto, como ya sabíamos (cf. el Ejemplo 5.27). De la fórmula anterior se sigue que el vector unitario tangente a la gráfica de f en el punto $(x_0, y_0, f(x_0, y_0))$ está dado por

$$\mathbf{n}(x_0, y_0) = \frac{(-f_x(x_0, y_0), -f_y(x_0, y_0), 1)}{\sqrt{1 + f_x^2(x_0, y_0) + f_y^2(x_0, y_0)}}.$$

La ecuación del plano tangente a la superficie en dicho punto es

$$(x - x_0, y - y_0, z - f(x_0, y_0)) \cdot (-f_x(x_0, y_0), -f_y(x_0, y_0), 1) = 0,$$

o equivalentemente

$$z = f(x_0, y_0) + \nabla f(x_0, y_0) \cdot (x - x_0, y - y_0),$$

como ya vimos en la Sección 2.8 (cf. la ec. (2.38)).

Ejemplo 5.32. Consideremos la superficie de revolución $\rho = f(z)$, que podemos parametrizar por

$$\mathbf{r}(\varphi, z) = (f(z) \cos \varphi, f(z) \sin \varphi, z). \quad (5.15)$$

En este caso los vectores \mathbf{r}_φ y \mathbf{r}_z están dados por

$$\mathbf{r}_\varphi = (-f(z) \sin \varphi, f(z) \cos \varphi, 0), \quad \mathbf{r}_z = (f'(z) \cos \varphi, f'(z) \sin \varphi, 1), \quad (5.16)$$

y por tanto

$$\mathbf{r}_\varphi \times \mathbf{r}_z = f(z)(\cos \varphi, \sin \varphi, -f'(z))$$

solo se anula si $f(z) = 0$. Si $f(z) \neq 0$, el vector unitario normal a la superficie es igual a

$$\mathbf{n}(\varphi, z) = \frac{(\cos \varphi, \sin \varphi, -f'(z))}{\sqrt{1 + f'(z)^2}},$$

y la ecuación del plano tangente a la superficie en el punto $(f(z_0) \cos \varphi_0, f(z_0) \sin \varphi_0, z_0)$ es

$$\begin{aligned} (x - f(z_0) \cos \varphi_0, y - f(z_0) \sin \varphi_0, z - z_0) \cdot (\cos \varphi_0, \sin \varphi_0, -f'(z_0)) &= 0 \\ \Leftrightarrow \cos \varphi_0 x + \sin \varphi_0 y - f'(z_0) z &= f(z_0) - z_0 f'(z_0). \end{aligned}$$

Multiplicando la ecuación anterior por $f(z_0) \neq 0$ se obtiene la ecuación equivalente

$$x_0 x + y_0 y - f(z_0) f'(z_0) z = f(z_0)^2 - z_0 f(z_0) f'(z_0).$$

A esta misma ecuación se llega teniendo en cuenta que S es la superficie de nivel $L_0(F)$ de la función $F(x, y, z) = x^2 + y^2 - f(z)^2$. En efecto, por lo visto en la Sección 2.8 la ecuación del plano tangente a S en el punto (x_0, y_0, z_0) es

$$\begin{aligned} (x - x_0, y - y_0, z - z_0) \cdot \nabla F(x_0, y_0, z_0) &= 2(x - x_0, y - y_0, z - z_0)(x_0, y_0, -f(z_0)f'(z_0)) = 0 \\ \Leftrightarrow x_0 x + y_0 y - f(z_0) f'(z_0) z &= f(z_0)^2 - z_0 f(z_0) f'(z_0), \end{aligned}$$

donde se ha tenido en cuenta que $x_0^2 + y_0^2 = f(z_0)^2 = 0$. Por ejemplo, el cono de ecuación $z = \sqrt{x^2 + y^2}$ se obtiene de (5.15) con $f(z) = z$ (con $z \geq 0$). Por tanto en este caso el vector unitario normal está dado por

$$\mathbf{n}(\varphi, z) = \frac{1}{\sqrt{2}} (\cos \varphi, \sin \varphi, -1)$$

(si $z > 0$), y la ecuación del plano tangente al cono en el punto (x_0, y_0, z_0) (con $z_0 = \sqrt{x_0^2 + y_0^2} > 0$) es

$$x_0 x + y_0 y - z_0 z = 0.$$

5.3.3 Área de una superficie parametrizada

Sea $\sigma : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ una superficie suave, y consideremos un punto regular $(u, v) \in \overset{\circ}{D}$ en que $D\sigma$ tiene rango 2, o equivalentemente

$$\frac{\partial \sigma}{\partial u}(u, v) \times \frac{\partial \sigma}{\partial v}(u, v) \neq 0.$$

La imagen de un pequeño rectángulo

$$R = [u, u + \Delta u] \times [v, v + \Delta v] \subset D$$

de lados $\Delta u, \Delta v > 0$ se puede aproximar por el paralelogramo generado por los vectores

$$\frac{\partial \sigma}{\partial u}(u, v) \Delta u, \quad \frac{\partial \sigma}{\partial v}(u, v) \Delta v$$

con origen en el punto $\sigma(u, v)$, cuya área está dada por

$$\left\| \frac{\partial \sigma}{\partial u}(u, v) \times \frac{\partial \sigma}{\partial v}(u, v) \right\| \Delta u \Delta v.$$

Este hecho motiva la siguiente definición:

Definición 5.33. Sea $\sigma : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ una superficie suave, y supongamos que $D \subset \mathbb{R}^2$ es un conjunto compacto medible Jordan. El **área** de la superficie σ se define por

$$a(\sigma) = \int_D \left\| \frac{\partial \sigma}{\partial u}(u, v) \times \frac{\partial \sigma}{\partial v}(u, v) \right\| du dv.$$

Comentario. Nótese que no es necesario pedir que D sea compacto (o, equivalentemente, cerrado) si $\left\| \frac{\partial \sigma}{\partial u}(u, v) \times \frac{\partial \sigma}{\partial v}(u, v) \right\|$ está acotada en D . En efecto, en tal caso la continuidad del integrando en la integral que define $a(\sigma)$ basta para garantizar la existencia de dicha integral. \square

Con un ligero abuso de notación, la fórmula anterior se suele escribir

$$a(\sigma) = \int_D \left\| \frac{\partial \mathbf{r}}{\partial u}(u, v) \times \frac{\partial \mathbf{r}}{\partial v}(u, v) \right\| du dv, \quad (5.17)$$

o equivalentemente

$$a(\sigma) = \int_D \left[\left(\det \frac{\partial(x, y)}{\partial(u, v)} \right)^2 + \left(\det \frac{\partial(y, z)}{\partial(u, v)} \right)^2 + \left(\det \frac{\partial(z, x)}{\partial(u, v)} \right)^2 \right]^{1/2} du dv.$$

Otra expresión muy útil para el área de la superficie σ se obtiene teniendo en cuenta que

$$\left\| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right\|^2 = \left(\frac{\partial \mathbf{r}}{\partial u} \right)^2 \left(\frac{\partial \mathbf{r}}{\partial v} \right)^2 - \left(\frac{\partial \mathbf{r}}{\partial u} \cdot \frac{\partial \mathbf{r}}{\partial v} \right)^2 \equiv EG - F^2,$$

donde hemos utilizado la notación habitual

$$E \equiv \left(\frac{\partial \mathbf{r}}{\partial u} \right)^2, \quad F \equiv \frac{\partial \mathbf{r}}{\partial u} \cdot \frac{\partial \mathbf{r}}{\partial v}, \quad G \equiv \left(\frac{\partial \mathbf{r}}{\partial v} \right)^2.$$

Sustituyendo en (5.17) se obtiene

$$a(\sigma) = \int_D \sqrt{E(u, v)G(u, v) - F(u, v)^2} du dv.$$

Sea $g : \tilde{D} \subset \mathbb{R}^2 \rightarrow D$ un **cambio de parámetros**, es decir un difeomorfismo de clase C^1 en \tilde{D} . Diremos que $\tilde{\sigma} : \tilde{D} \rightarrow \mathbb{R}^3$ es una **reparametrización** de la superficie parametrizada $\sigma : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ si $\tilde{\sigma} = \sigma \circ g$, siendo $g : \tilde{D} \rightarrow D$ un cambio de parámetros. Nótese que, por definición de cambio de parámetros,

$$\tilde{\sigma}(\tilde{D}) = \sigma(g(\tilde{D})) = \sigma(D),$$

y por tanto σ y su reparametrización $\tilde{\sigma}$ determinan la misma superficie. Además, de la regla de la cadena se sigue que

$$D\tilde{\sigma}(s, t) = D\sigma(g(s, t))Dg(s, t),$$

y por tanto (al ser Dg invertible en todo punto, ya que g es un difeomorfismo) la superficie $\tilde{\sigma}$ es suave o regular si y solo si lo es σ . Intuitivamente, es claro que el área de una superficie no debe depender de su parametrización. En otras palabras, debería ocurrir que

$$a(\sigma) = a(\tilde{\sigma}), \quad (5.18)$$

siendo $\tilde{\sigma} : \tilde{D} \rightarrow \mathbb{R}^3$ cualquier reparametrización de σ . Para probar esta identidad, nótese en primer lugar que aplicando la regla de la cadena se obtiene

$$\frac{\partial \tilde{\sigma}}{\partial s} = \frac{\partial \sigma}{\partial u}(u, v) \frac{\partial u}{\partial s} + \frac{\partial \sigma}{\partial v}(u, v) \frac{\partial v}{\partial s}, \quad \frac{\partial \tilde{\sigma}}{\partial t} = \frac{\partial \sigma}{\partial u}(u, v) \frac{\partial u}{\partial t} + \frac{\partial \sigma}{\partial v}(u, v) \frac{\partial v}{\partial t},$$

donde las derivadas parciales de σ en el miembro derecho están evaluadas en $(u, v) = g(s, t)$, y hemos denotado por $(u(s, t), v(s, t))$ las componentes del cambio de parámetros $g(s, t)$. Por tanto

$$\begin{aligned} \frac{\partial \tilde{\sigma}}{\partial s} \times \frac{\partial \tilde{\sigma}}{\partial t} &= \left(\frac{\partial \sigma}{\partial u}(u, v) \times \frac{\partial \sigma}{\partial v}(u, v) \right) \left(\frac{\partial u}{\partial s} \frac{\partial v}{\partial t} - \frac{\partial u}{\partial t} \frac{\partial v}{\partial s} \right) \\ &\equiv \left(\frac{\partial \sigma}{\partial u}(u, v) \times \frac{\partial \sigma}{\partial v}(u, v) \right) \det \frac{\partial(u, v)}{\partial(s, t)}. \end{aligned} \quad (5.19)$$

Sustituyendo en la definición de área y aplicando la fórmula del cambio de variables para integrales dobles se obtiene

$$\begin{aligned} a(\tilde{\sigma}) &\equiv \int_{\tilde{D}} \left\| \frac{\partial \tilde{\sigma}}{\partial s} \times \frac{\partial \tilde{\sigma}}{\partial t} \right\| ds dt = \int_{\tilde{D}} \left\| \frac{\partial \sigma}{\partial u}(u, v) \times \frac{\partial \sigma}{\partial v}(u, v) \right\| \left| \det \frac{\partial(u, v)}{\partial(s, t)} \right| ds dt \\ &= \int_D \left\| \frac{\partial \sigma}{\partial u} \times \frac{\partial \sigma}{\partial v} \right\| du dv \equiv a(\sigma), \end{aligned}$$

como habíamos afirmado.

Por lo que acabamos de probar, podemos definir el área de una superficie S mediante la fórmula

$$a(S) = \int_D \left\| \frac{\partial \sigma}{\partial u}(u, v) \times \frac{\partial \sigma}{\partial v}(u, v) \right\| du dv,$$

donde $\sigma : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ es *cualquier* parametrización de S . Formalmente, el área de la superficie $S = \sigma(D)$ es la integral sobre la región $D \subset \mathbb{R}^2$ del **elemento de área**

$$dS \equiv \left\| \frac{\partial \mathbf{r}}{\partial u}(u, v) \times \frac{\partial \mathbf{r}}{\partial v}(u, v) \right\| du dv = \sqrt{E(u, v)G(u, v) - F(u, v)^2} du dv. \quad (5.20)$$

Ejemplo 5.34. Calculemos el área del sector de la esfera de radio a de ecuaciones

$$\theta_1 \leq \theta \leq \theta_2, \quad \varphi_1 \leq \varphi \leq \varphi_2$$

(cf la fig. 5.6). Utilizando la parametrización (5.11) se obtiene

$$\frac{\partial \mathbf{r}}{\partial \theta} = a(\cos \theta \cos \varphi, \cos \theta \sin \varphi, -\sin \theta), \quad \frac{\partial \mathbf{r}}{\partial \varphi} = a \sin \theta (-\sin \varphi, \cos \varphi, 0),$$

y por tanto

$$E = \left(\frac{\partial \mathbf{r}}{\partial \theta} \right)^2 = a^2, \quad F = \frac{\partial \mathbf{r}}{\partial \theta} \cdot \frac{\partial \mathbf{r}}{\partial \varphi} = 0, \quad G = \left(\frac{\partial \mathbf{r}}{\partial \varphi} \right)^2 = a^2 \sin^2 \theta.$$

Luego el elemento de área en la esfera está dado por

$$dS = a^2 \sin \theta d\theta d\varphi,$$

y el área del sector S considerado es igual a

$$a(S) = a^2 \int_{\theta_1}^{\theta_2} d\theta \int_{\varphi_1}^{\varphi_2} d\varphi \sin \theta = a^2 (\varphi_2 - \varphi_1) (\cos \theta_1 - \cos \theta_2).$$

En particular, tomando

$$\varphi_1 = 0, \quad \varphi_2 = 2\pi, \quad \theta_1 = 0, \quad \theta_2 = \pi$$

se deduce que el área de la esfera es $4\pi a^2$.

Ejemplo 5.35. Sea S la gráfica de la función (de clase C^1) $f : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$, dada por

$$\mathbf{r}(x, y) = (x, y, f(x, y)), \quad (x, y) \in D.$$

En este caso

$$\left\| \frac{\partial \mathbf{r}}{\partial x} \times \frac{\partial \mathbf{r}}{\partial y} \right\| = \|(-f_x, -f_y, 1)\| = \sqrt{1 + f_x^2 + f_y^2}.$$

Luego el elemento de área de esta superficie es

$$dS = \sqrt{1 + f_x^2 + f_y^2} \, dx \, dy. \quad (5.21)$$

Calculemos, por ejemplo, el área del paraboloides de revolución S de ecuación

$$z = x^2 + y^2, \quad 0 \leq z \leq h.$$

En este caso $f(x, y) = x^2 + y^2$, y de la condición $0 \leq z \leq h$ se sigue que D es el disco $x^2 + y^2 \leq h$ de centro el origen y radio \sqrt{h} . Por tanto

$$\begin{aligned} a(S) &= \int_D \sqrt{1 + 4x^2 + 4y^2} \, dx \, dy = \int_0^{\sqrt{h}} dr \int_0^{2\pi} d\theta \cdot r \sqrt{1 + 4r^2} \\ &= 2\pi \cdot \frac{1}{8} \cdot \frac{2}{3} (1 + 4r^2)^{3/2} \Big|_{r=0}^{r=\sqrt{h}} = \frac{\pi}{6} [(1 + 4h)^{3/2} - 1]. \end{aligned}$$

Ejemplo 5.36. Consideremos la superficie de revolución de ecuación $\rho = f(z)$, $h_1 \leq z \leq h_2$, obtenida rotando la gráfica de la función no negativa $x = f(z)$ alrededor del eje z . Utilizando las ecuaciones (5.16) se obtiene

$$E = f(z)^2, \quad F = 0, \quad G = 1 + f'(z)^2,$$

y por tanto

$$dS = f(z) \sqrt{1 + f'(z)^2} \, d\varphi \, dz. \quad (5.22)$$

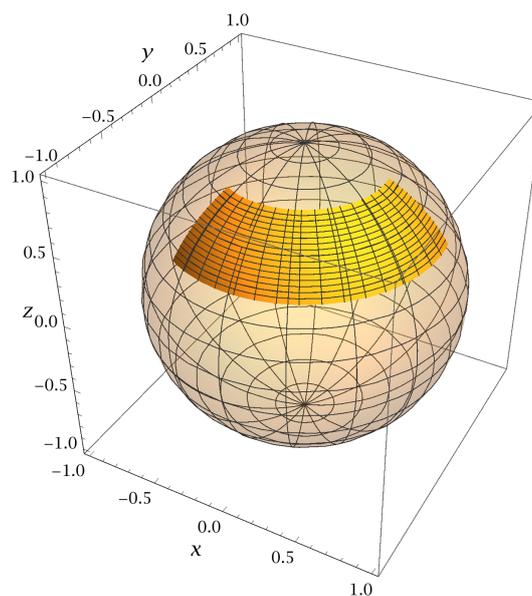


Figura 5.6: Sector de la esfera de radio a de ecuaciones $\theta_1 \leq \theta \leq \theta_2$, $\varphi_1 \leq \varphi \leq \varphi_2$.

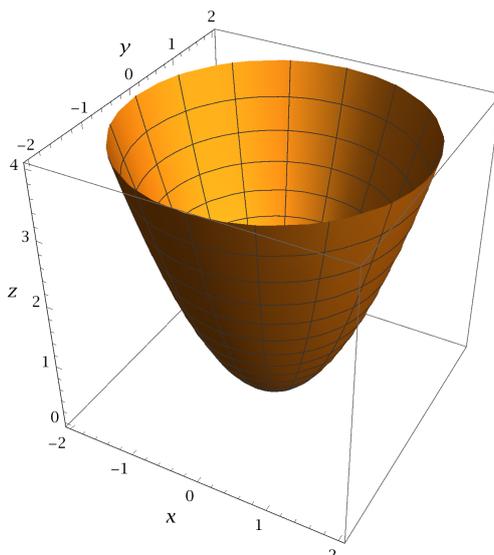


Figura 5.7: Paraboloides $z = x^2 + y^2$.

Luego el área de la superficie de revolución (5.15) está dada por

$$a(S) = 2\pi \int_{h_1}^{h_2} f(z) \sqrt{1 + f'(z)^2} dz. \tag{5.23}$$

Nótese que $\sqrt{1 + f'(z)^2} dz$ es la longitud de arco ds a lo largo de la curva plana $x = f(z)$, por lo que el integrando de (5.23) es formalmente el área de un cilindro infinitesimal de altura ds y radio $\rho = f(z)$. La fórmula anterior es por tanto equivalente a

$$a(S) = 2\pi \int_{\gamma} x ds,$$

siendo γ la curva de ecuación $x = f(z)$ ($h_1 \leq z \leq h_2$). Introduciendo la coordenada x del centro de masas de la curva γ , definida por

$$X = \frac{1}{l(\gamma)} \int_{\gamma} x ds$$

se obtiene la relación

$$a(S) = 2\pi X l(\gamma).$$

Este es el llamado *teorema de Pappus* para superficies de revolución. Consideremos, por ejemplo, el *toro*

$$(\rho - b)^2 + z^2 = a^2,$$

con $0 < a < b$, que es la unión de las dos superficies de revolución

$$\rho = b \pm \sqrt{a^2 - z^2} \equiv f_{\pm}(z).$$

Es fácil ver que podemos aplicar el teorema de Pappus a la circunferencia completa γ , ya que si llamamos γ_{\pm} a las semicircunferencias $x = f_{\pm}(z)$ se tiene

$$a(S) = 2\pi \int_{\gamma_-} x ds + 2\pi \int_{\gamma_+} x ds = 2\pi \int_{\gamma} x ds \equiv 2\pi X l(\gamma).$$

Por simetría, el centro de masas de la circunferencia $(x - b)^2 + z^2 = a^2$ está situado en su centro $(b, 0)$. Por tanto en este caso $X = b$, $l(\gamma) = 2\pi a$ y

$$a(S) = 4\pi^2 ab.$$

Ejemplo 5.37. Consideremos el conjunto $S \subset \mathbb{R}^3$ de ecuación $F(x, y, z) = 0$, y supongamos que se puede despejar z como una función $f(x, y)$ en una región $D \subset \mathbb{R}^2$. En tal caso, el conjunto S coincide con la gráfica de la función f en D y (si la función f es de clase C^1) su elemento de área está dado por la ec. (5.21). Derivando implícitamente se obtiene

$$F_x + F_z \frac{\partial f}{\partial x} = 0, \quad F_y + F_z \frac{\partial f}{\partial y} = 0,$$

donde z debe sustituirse por $f(x, y)$ en los argumentos de las derivadas parciales de F . Utilizando estas fórmulas (y suponiendo que F_z no se anula) se obtiene:

$$dS = \sqrt{1 + \frac{F_x^2}{F_z^2} + \frac{F_y^2}{F_z^2}} dx dy = \frac{\sqrt{F_x^2 + F_y^2 + F_z^2}}{|F_z|} dx dy \equiv \boxed{\frac{\|\nabla F\|}{|F_z|} dx dy}. \quad (5.24)$$

Ejemplo 5.38. Como aplicación de la fórmula anterior, hallemos el área de la región de la esfera $x^2 + y^2 + z^2 = 1$ interior al cilindro de ecuación $x^2 + y^2 - x = 0$ (cf. la Fig. 5.8, izda.). La superficie dada es simétrica bajo $z \mapsto -z$, $y \mapsto -y$, y por tanto el área pedida es cuatro veces el área de la intersección de la superficie con el cuadrante $y > 0$, $z > 0$, que denotaremos por S_{++} . En tal caso podemos tomar

$$F(x, y, z) = \frac{1}{2}(x^2 + y^2 + z^2 - 1),$$

ya que si $z > 0$ la relación anterior determina z en función de (x, y) a través de la ecuación $z = \sqrt{1 - x^2 - y^2}$. Por tanto el elemento de área en S_{++} está dado por

$$dS = \frac{dx dy}{|z|} = \frac{dx dy}{\sqrt{1 - x^2 - y^2}}. \quad (5.25)$$

La región de integración D es el semicírculo $x^2 + y^2 - x \leq 0$, $y \geq 0$, que en coordenadas polares (r, θ) se convierte en el conjunto \tilde{D} de ecuaciones

$$r \leq \cos \theta, \quad 0 \leq \theta \leq \frac{\pi}{2},$$

y el elemento de área está dado por

$$dS = \frac{r dr d\theta}{\sqrt{1 - r^2}}.$$

Nótese que el denominador de dS se anula en el punto $(r, \theta) = (1, 0) \in \tilde{D}$ (como ya ocurría en coordenadas cartesianas con el denominador de (5.25), que se anula en el punto $(1, 0) \in D$). Para evitar este problema, integramos en la región \tilde{D}_ε de ecuación

$$r \leq \cos \theta, \quad \varepsilon \leq \theta \leq \frac{\pi}{2}, \quad (5.26)$$

con $0 < \varepsilon < \pi/2$ (cf. la Fig. 5.8, dcha.), y a continuación hacemos ε tender a 0. La integral de dS en el conjunto \tilde{D}_ε está dada por

$$\begin{aligned} \int_\varepsilon^{\pi/2} d\theta \int_0^{\cos \theta} dr \frac{r}{\sqrt{1 - r^2}} &= - \int_\varepsilon^{\pi/2} d\theta \sqrt{1 - r^2} \Big|_{r=0}^{r=\cos \theta} = \int_\varepsilon^{\pi/2} (1 - \sin \theta) d\theta \\ &= \frac{\pi}{2} - \varepsilon - \cos \varepsilon. \end{aligned}$$

Por tanto

$$a(S_{++}) = \lim_{\varepsilon \rightarrow 0^+} \left(\frac{\pi}{2} - \varepsilon - \cos \varepsilon \right) = \frac{\pi}{2} - 1 \quad \Rightarrow \quad \boxed{a(S) = 2(\pi - 2)}.$$

Este problema también se puede resolver en coordenadas esféricas. En estas coordenadas, la

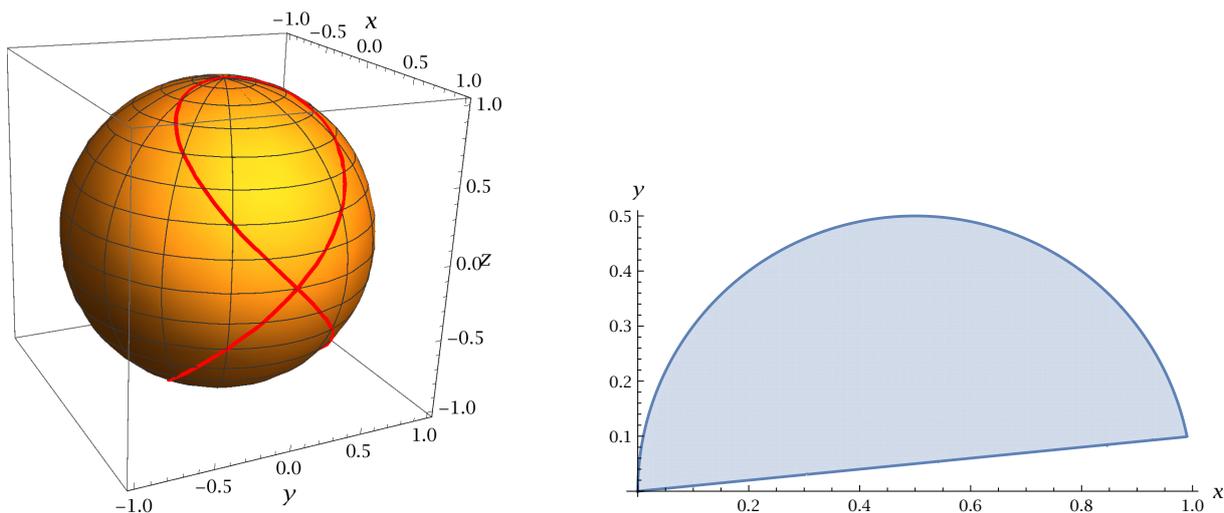


Figura 5.8: Izda: intersección de la esfera $x^2 + y^2 + z^2 = 1$ con el cilindro $x^2 + y^2 = x$ (en rojo). Drcha.: recinto de integración (5.26) en el Ejemplo 5.38.

ecuación del cilindro se escribe

$$\sin^2 \theta \leq \sin \theta \cos \varphi \iff \sin \theta \leq \cos \varphi \iff \cos \left(\frac{\pi}{2} - \theta \right) \leq \cos \varphi .$$

En la región S_{++} el ángulo θ varía entre 0 y $\pi/2$ (ya que $z \geq 0$), y lo mismo ocurre con φ (al ser $y > 0, x \geq x^2 + y^2 \geq 0$). De las desigualdades anteriores se sigue por tanto que la ecuación del cilindro es equivalente a

$$0 \leq \varphi \leq \frac{\pi}{2} - \theta .$$

Utilizando el elemento de área de la esfera en coordenadas esféricas se obtiene

$$\begin{aligned} a(S_{++}) &= \int_0^{\pi/2} d\theta \int_0^{\frac{\pi}{2}-\theta} d\varphi \sin \theta = \int_0^{\pi/2} \left(\frac{\pi}{2} - \theta \right) \sin \theta d\theta = \left(\theta - \frac{\pi}{2} \right) \cos \theta \Big|_0^{\pi/2} - \int_0^{\pi/2} \cos \theta d\theta \\ &= \frac{\pi}{2} - 1 , \end{aligned}$$

como antes.

Por último, una tercera forma de resolver el problema es utilizando coordenadas cilíndricas. En efecto, en estas coordenadas la esfera tiene por ecuación $\rho = \sqrt{1 - z^2}$, y el interior del cilindro está dado por $\rho^2 \leq \rho \cos \varphi$. Por tanto

$$\sqrt{1 - z^2} \leq \cos \varphi \iff 1 - z^2 \leq \cos^2 \varphi \iff z^2 \geq \sin^2 \varphi \iff z \geq \sin \varphi$$

(ya que en S_{++} la coordenada z es positiva). Utilizando la ecuación (5.22) para el elemento de área de una superficie de revolución alrededor del eje z se obtiene

$$dS = \sqrt{1 - z^2} \sqrt{1 + \frac{z^2}{1 - z^2}} d\varphi dz = d\varphi dz .$$

Como en S_{++} el ángulo φ varía entre 0 y $\pi/2$, el área de S_{++} está dada por

$$a(S_{++}) = \int_0^{\pi/2} d\varphi \int_{\sin \varphi}^1 dz = \int_0^{\pi/2} (1 - \sin \varphi) d\varphi = \varphi + \cos \varphi \Big|_0^{\pi/2} = \frac{\pi}{2} - 1 .$$

5.3.4 Integral de superficie de una función escalar

Una vez que hemos definido el elemento de área de una superficie (cf. la ec. (5.20)), es natural definir la integral de una función escalar sobre la superficie como la integral de dicha función respecto del elemento de área. Más precisamente:

Definición 5.39. Sea $\sigma : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ una superficie suave tal que D es un conjunto compacto medible Jordan, y sea $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ una función escalar continua en $\sigma(D)$. La **integral** de f sobre la superficie σ se define por

$$\int_{\sigma} f \, dS = \int_D f(\sigma(u, v)) \left\| \frac{\partial \sigma}{\partial u}(u, v) \times \frac{\partial \sigma}{\partial v}(u, v) \right\| \, du \, dv.$$

Nótese que la integral doble del miembro derecho está bien definida, ya que el conjunto D es por hipótesis compacto y medible Jordan, y el integrando es continuo en D al ser f continua en $\sigma(D)$ y σ de clase C^1 en D . Normalmente utilizaremos la notación más descriptiva

$$\begin{aligned} \int_{\sigma} f \, dS &= \int_D f(\mathbf{r}(u, v)) \left\| \frac{\partial \mathbf{r}}{\partial u}(u, v) \times \frac{\partial \mathbf{r}}{\partial v}(u, v) \right\| \, du \, dv \\ &\equiv \int_D f(\mathbf{r}(u, v)) \sqrt{(E(u, v)G(u, v) - F^2(u, v))} \, du \, dv \end{aligned}$$

para denotar la integral $\int_{\sigma} f \, dS$.

Con el mismo argumento que utilizamos para probar la ec. (5.18) se demuestra la siguiente propiedad importante de las integrales de superficie:

La integral de superficie $\int_{\sigma} f \, dS$ es *invariante bajo reparametrizaciones*, es decir

$$\int_{\sigma} f \, dS = \int_{\tilde{\sigma}} f \, dS,$$

donde $\tilde{\sigma}$ es cualquier reparametrización de σ .

Por este motivo, si $S \subset \mathbb{R}^3$ es una superficie es posible definir la integral en dicha superficie de una función f continua en S mediante la fórmula

$$\int_S f \, dS = \int_{\sigma} f \, dS,$$

donde $\sigma : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ es *cualquier* parametrización de S .

Ejemplo 5.40. Evaluemos la integral $\int_S z \, dS$, donde S es la semiesfera $x^2 + y^2 + z^2 = a^2$, $z \geq 0$. Utilizando el elemento de área en coordenadas esféricas obtenemos inmediatamente

$$\int_S z \, dS = \int_0^{\pi/2} d\theta \int_0^{2\pi} d\varphi \, a^2 \sin \theta \cdot a \cos \theta = 2\pi a^3 \int_0^{\pi/2} \sin \theta \cos \theta \, d\theta = \pi a^3 \sin^2 \theta \Big|_{\theta=0}^{\theta=\pi/2} = \boxed{\pi a^3}.$$

Como la semiesfera tiene área $2\pi a^2$, la coordenada z del centro de masas de la semiesfera (supuesta homogénea) es igual a

$$Z = \frac{\pi a^3}{a(S)} = \frac{\pi a^3}{2\pi a^2} = \frac{a}{2}.$$

Está claro, por simetría, que las otras dos coordenadas del centro de masas de S son $X = Y = 0$. Por ejemplo,

$$\int_S x \, dS = \int_0^{\pi/2} d\theta \int_0^{2\pi} d\varphi \, a^2 \sin \theta \cdot a \sin \theta \cos \varphi = 0,$$

ya que $\int_0^{2\pi} \cos \varphi \, d\varphi = 0$. Luego el centro de masas de la semiesfera es el punto $(0, 0, a/2)$.

Ejemplo 5.41. Calculemos la integral $\int_S z \, dS$, donde ahora S es la región del hiperboloide $x^2 + y^2 - z^2 = a^2$ comprendida entre los planos $z = 0$ y $z = h$. Para ello necesitamos primero calcular el elemento de área del hiperboloide, para lo cual podemos utilizar *cualquier* parametrización de dicha superficie. Por ejemplo, considerando el hiperboloide como la superficie de revolución alrededor del eje z de ecuación $\rho = \sqrt{a^2 + z^2} \equiv f(z)$ y utilizando la fórmula general (5.22) para el elemento de área de dicha superficie obtenemos

$$dS = \sqrt{a^2 + z^2} \sqrt{1 + \frac{z^2}{a^2 + z^2}} \, d\varphi \, dz = \sqrt{a^2 + 2z^2} \, d\varphi \, dz.$$

Por tanto

$$\begin{aligned} \int_S z \, dS &= \int_0^h dz \int_0^{2\pi} d\varphi \cdot z \sqrt{a^2 + 2z^2} = 2\pi \int_0^h z \sqrt{a^2 + 2z^2} \, dz = 2\pi \cdot \frac{1}{4} \cdot \frac{2}{3} (a^2 + 2z^2)^{3/2} \Big|_{z=0}^{z=h} \\ &= \frac{\pi}{3} \left[(a^2 + 2h^2)^{3/2} - a^3 \right]. \end{aligned}$$

5.4 Integral de superficie de un campo vectorial

Sea $\sigma : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ una superficie *regular* (con D conjunto compacto medible Jordan), y sea $\mathbf{F} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ un campo vectorial *continuo* en $\sigma(D)$. La integral de la componente del campo en la dirección del vector unitario normal a la superficie

$$\mathbf{n}(u, v) = \frac{\mathbf{r}_u(u, v) \times \mathbf{r}_v(u, v)}{\|\mathbf{r}_u(u, v) \times \mathbf{r}_v(u, v)\|} \quad (5.27)$$

determinado por la parametrización $\sigma(u, v) \equiv \mathbf{r}(u, v)$ se denomina **integral de superficie** de \mathbf{F} sobre la superficie parametrizada σ , y se denota por $\int_\sigma \mathbf{F}(\mathbf{r}) \cdot d\mathbf{S}$. En otras palabras,

$$\int_\sigma \mathbf{F}(\mathbf{r}) \cdot d\mathbf{S} = \int_\sigma \mathbf{F}(\mathbf{r}) \cdot \mathbf{n} \, dS.$$

Nótese que la integral del miembro derecho está bien definida si la superficie σ es regular, ya que el campo vectorial \mathbf{F} es continuo sobre la superficie y el vector normal \mathbf{n} está bien definido y es continuo en D salvo a lo sumo en un conjunto de medida nula. Utilizando las definiciones de \mathbf{n} y del elemento de superficie dS se obtiene la siguiente fórmula explícita para la integral de superficie $\int_\sigma \mathbf{F}(\mathbf{r}) \cdot d\mathbf{S}$:

$$\begin{aligned} \int_\sigma \mathbf{F}(\mathbf{r}) \cdot d\mathbf{S} &= \int_D F(\mathbf{r}(u, v)) \cdot \frac{\mathbf{r}_u(u, v) \times \mathbf{r}_v(u, v)}{\|\mathbf{r}_u(u, v) \times \mathbf{r}_v(u, v)\|} \|\mathbf{r}_u(u, v) \times \mathbf{r}_v(u, v)\| \, du \, dv \\ &= \int_D F(\mathbf{r}(u, v)) \cdot (\mathbf{r}_u(u, v) \times \mathbf{r}_v(u, v)) \, du \, dv. \end{aligned} \quad (5.28)$$

Equivalentemente,

$$\int_\sigma \mathbf{F}(\mathbf{r}) \cdot d\mathbf{S} = \int_D \begin{vmatrix} F_1(\mathbf{r}(u, v)) & F_2(\mathbf{r}(u, v)) & F_3(\mathbf{r}(u, v)) \\ x_u(u, v) & y_u(u, v) & z_u(u, v) \\ x_v(u, v) & y_v(u, v) & z_v(u, v) \end{vmatrix} \, du \, dv.$$

Nótese que la integral (5.28) está de hecho bien definida para cualquier superficie *suave* σ . Formalmente, la integral $\int_\sigma \mathbf{F}(\mathbf{r}) \cdot d\mathbf{S}$ es el resultado de integrar en el conjunto D el producto escalar del campo \mathbf{F} evaluado en un punto $\mathbf{r}(u, v)$ de la superficie por el **elemento de superficie**

$$d\mathbf{S} = \mathbf{r}_u(u, v) \times \mathbf{r}_v(u, v) \, du \, dv \equiv \mathbf{n}(u, v) \, dS.$$

¿Cómo varía la integral de superficie de un campo vectorial bajo reparametrizaciones de la superficie? Para responder a esta pregunta, consideremos una reparametrización $\tilde{\sigma} = \sigma \circ g : \tilde{D} \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$, donde $g : \tilde{D} \rightarrow D$ es un cambio de parámetros (es decir, un difeomorfismo de clase C^1). Supondremos a partir de ahora que el conjunto D (y, por tanto, $\tilde{D} = g^{-1}(D)$) es *conexo*. Si utilizamos la notación abreviada

$$g(s, t) \equiv (u(s, t), v(s, t)), \quad \tilde{\sigma}(s, t) \equiv \tilde{\mathbf{r}}(s, t) \equiv \mathbf{r}(u(s, t), v(s, t)),$$

de la ec. (5.19), que podemos escribir

$$\tilde{\mathbf{r}}_s(s, t) \times \tilde{\mathbf{r}}_t(s, t) = \mathbf{r}_u(u, v) \times \mathbf{r}_v(u, v) \det \frac{\partial(u, v)}{\partial(s, t)} \quad (5.29)$$

(con $(u, v) \equiv (u(s, t), v(s, t))$) se sigue que

$$\int_{\tilde{\sigma}} \mathbf{F}(\mathbf{r}) \cdot d\mathbf{S} \equiv \int_{\tilde{D}} \mathbf{F}(\mathbf{r}(u, v)) \cdot (\mathbf{r}_u(u, v) \times \mathbf{r}_v(u, v)) \det \frac{\partial(u, v)}{\partial(s, t)} ds dt.$$

Como la aplicación $g : \tilde{D} \rightarrow D$ es por hipótesis un *difeomorfismo*, el jacobiano que aparece en la ecuación anterior no se anula en \tilde{D} . Al ser este conjunto *conexo*, por el teorema de los valores intermedios dicho jacobiano ha de ser o bien positivo o bien negativo en \tilde{D} . Por tanto

$$\det \frac{\partial(u, v)}{\partial(s, t)} = \varepsilon \left| \det \frac{\partial(u, v)}{\partial(s, t)} \right|,$$

con $\varepsilon \equiv \operatorname{sgn} \left(\det \frac{\partial(u, v)}{\partial(s, t)} \right)$ constante en \tilde{D} . Aplicando la fórmula del cambio de variables a la ecuación anterior se obtiene entonces la relación

$$\int_{\tilde{\sigma}} \mathbf{F}(\mathbf{r}) \cdot d\mathbf{S} = \varepsilon \int_{\sigma} \mathbf{F}(\mathbf{r}) \cdot d\mathbf{S}, \quad \varepsilon \equiv \operatorname{sgn} \left(\det \frac{\partial(u, v)}{\partial(s, t)} \right).$$

Por tanto:

La integral de superficie $\int_{\sigma} \mathbf{F}(\mathbf{r}) \cdot d\mathbf{S}$ es *invariante* bajo reparametrizaciones cuyo jacobiano es *positivo*, y *cambia de signo* bajo reparametrizaciones cuyo jacobiano es *negativo*.

El resultado anterior sugiere que las reparametrizaciones con jacobiano positivo preservan la *orientación* de la superficie, mientras que las que tienen jacobiano negativo la cambian. Para precisar esta afirmación, es interesante analizar cómo se transforma el vector unitario normal $\mathbf{n}(u, v)$ bajo una reparametrización. Utilizando la ec. (5.29) se obtiene fácilmente

$$\tilde{\mathbf{n}}(s, t) \equiv \frac{\tilde{\mathbf{r}}_s(s, t) \times \tilde{\mathbf{r}}_t(s, t)}{\|\tilde{\mathbf{r}}_s(s, t) \times \tilde{\mathbf{r}}_t(s, t)\|} = \frac{\mathbf{r}_u(u, v) \times \mathbf{r}_v(u, v) \det \frac{\partial(u, v)}{\partial(s, t)}}{\|\mathbf{r}_u(u, v) \times \mathbf{r}_v(u, v)\| \left| \det \frac{\partial(u, v)}{\partial(s, t)} \right|} = \varepsilon \cdot \mathbf{n}(u, v),$$

donde ε es el signo del jacobiano $\det \frac{\partial(u, v)}{\partial(s, t)}$. Por tanto, las reparametrizaciones que dejan invariante la integral de superficie $\int_{\sigma} \mathbf{F} \cdot d\mathbf{S}$ son aquellas que preservan el vector unitario normal $\mathbf{n}(u, v)$, mientras que las que cambian su signo son aquellas que invierten la orientación de dicho vector.

Las consideraciones anteriores motivan la siguiente definición:

Definición 5.42. Diremos que una superficie $S \subset \mathbb{R}^3$ es **orientable** si se puede definir un campo de vectores $\mathbf{N} : S \rightarrow \mathbb{R}^3$ *continuo* en S tal que $\mathbf{N}(\mathbf{r})$ es unitario y normal a la superficie en cada punto $\mathbf{r} \in S$.

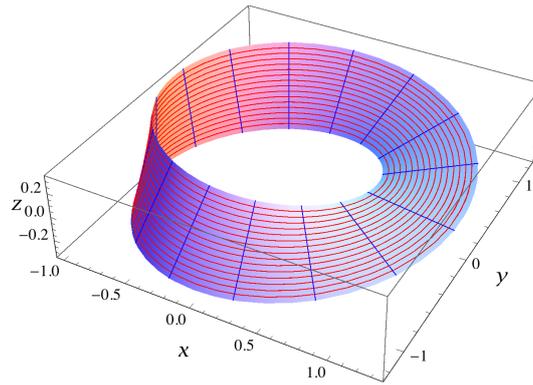


Figura 5.9: Banda de Möbius (con $a = 1$ y $h = 1/3$). En azul se han representado las curvas (segmentos de longitud $2h$) $v = \text{const.}$, en rojo las curvas $u = \text{const.}$

- Es fácil probar⁴ que en una superficie orientable *conexa* hay exactamente dos campos de vectores unitarios normales continuos, a saber el campo \mathbf{N} de la definición y el campo opuesto $-\mathbf{N}$. Cada uno de estos dos campos se denomina una **orientación** de la superficie S .
- Por definición, **orientar** una superficie orientable S significa escoger uno de estos dos campos como el vector unitario normal a S . Intuitivamente, esto significa escoger una de las dos “caras” de la superficie (aquella sobre la que “apoya” el vector unitario normal \mathbf{N} escogido). Una **superficie orientada** es una superficie S junto con una elección de orientación \mathbf{N} .

Sea $S = \sigma(D)$, con $\sigma : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$, y supongamos que todos los puntos de D son *regulares*. Podría pensarse que S es automáticamente orientable, ya que en tal caso el vector $\mathbf{n}(u, v)$ dado por la ec. (5.27) es un vector unitario normal a S para todo $(u, v) \in D$. Sin embargo, esta afirmación es, en general, *incorrecta*. En efecto, el campo $\mathbf{n}(u, v)$ está definido y es continuo en D , no en la superficie S . Para obtener a partir de \mathbf{n} un campo de vectores \mathbf{N} definido en S , es necesario que σ sea *inyectiva* (es decir, que la parametrización σ sea *simple*), ya que solo en tal caso podemos definir

$$\mathbf{N}(\mathbf{r}) = \mathbf{n}(\sigma^{-1}(\mathbf{r})). \quad (5.30)$$

Incluso en tal caso, el campo \mathbf{N} definido de esta forma puede no ser continuo, ya que σ^{-1} no siempre es continua. Esto es lo que ocurre, por ejemplo, con la banda de Möbius, como veremos en el ejemplo siguiente.

Ejemplo 5.43. La *banda de Möbius* es la superficie obtenida girando los extremos de una cinta de papel antes de unirlos (cf. la Fig. 5.9). Más precisamente, una parametrización sencilla de esta superficie está dada por

$$\sigma(u, v) = \left((a + u \cos(\frac{v}{2})) \cos v, (a + u \cos(\frac{v}{2})) \sin v, u \sin(\frac{v}{2}) \right), \quad (u, v) \in [-h, h] \times [0, 2\pi),$$

con $0 < h < a$. Se puede probar que esta parametrización es simple (ejercicio), y que todos sus puntos son regulares. Sin embargo, en este caso el campo \mathbf{N} definido por la ecuación (5.30) es discontinuo en los puntos de la forma $\mathbf{r}(u, 0)$. En efecto, un cálculo sencillo demuestra que

$$\lim_{v \rightarrow 2\pi} \mathbf{r}(-u, v) = \mathbf{r}(u, 0), \quad (5.31)$$

⁴En efecto, supongamos que $\mathbf{N}' : S \rightarrow \mathbb{R}^3$ es otro campo de vectores unitarios normales a S y continuo en dicha superficie. Entonces en cada punto $\mathbf{r} \in S$ se tiene $\mathbf{N}' = \varepsilon(\mathbf{r}) \mathbf{N}$, con $\varepsilon(\mathbf{r}) = \mathbf{N}(\mathbf{r}) \cdot \mathbf{N}'(\mathbf{r}) \in \{\pm 1\}$. Como S es conexa y ε es continua en S (al serlo los campos \mathbf{N} y \mathbf{N}'), la imagen de S bajo ε es un conjunto conexo. Al ser $\varepsilon(S) \subset \{\pm 1\}$, esto significa que o bien $\varepsilon(S) = 1$ (es decir, $\mathbf{N} = \mathbf{N}'$) o bien $\varepsilon(S) = \{-1\}$ (es decir, $\mathbf{N} = -\mathbf{N}'$).

y sin embargo (ejercicio)

$$\lim_{v \rightarrow 2\pi} \mathbf{n}(-u, v) = -\mathbf{n}(u, 0). \quad (5.32)$$

De hecho, la banda de Möbius es un ejemplo de superficie *no orientable*, en que *no* es posible definir un campo continuo de vectores unitarios normales a la superficie. Esto se debe, esencialmente, a que la banda de Möbius *solo tiene una cara*. Por ejemplo, consideremos el vector normal \mathbf{n} en el punto $\mathbf{r}(0, 0)$, dado por $\mathbf{n}(0, 0) = \mathbf{k}$ (ejercicio), y desplacémoslo de forma continua sobre la superficie a lo largo de la curva $\mathbf{r}(0, v)$ ($0 \leq v \leq 2\pi$), que es *cerrada* en virtud de la ec. (5.31). Cuando regresamos de nuevo al punto de partida (es decir, para $v = 2\pi$) el vector \mathbf{n} pasa a ser (cf. la ec. (5.32))

$$\lim_{v \rightarrow 2\pi} \mathbf{n}(0, v) = -\mathbf{n}(0, 0) = -\mathbf{k}.$$

En otras palabras, hemos pasado de forma continua de una cara de la superficie en el punto $\mathbf{r}(0, 0)$ a la opuesta.

Comentarios.

- Si la superficie suave $\sigma : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ es *simple*, D es *abierto* y todos los puntos de D son *regulares*, el teorema de la función inversa garantiza que $\sigma^{-1} : \sigma(D) \rightarrow D$ es *continua* (ejercicio). Por tanto este tipo de superficies son siempre *orientables*.
- La práctica totalidad de las superficies de interés son orientables, aún cuando no sean simples o dejen de ser regulares en algún punto. Este es, por ejemplo, el caso de la esfera, el cilindro o el toro. Por ejemplo, en la esfera de centro el origen y radio $a > 0$ las dos orientaciones son

$$\mathbf{N}(\mathbf{r}) = \pm \frac{\mathbf{r}}{r} = \pm \frac{\mathbf{r}}{a},$$

donde el signo “+” (resp. “-”) corresponde a la elección de la *normal exterior* (resp. *interior*). En este caso, la normal exterior coincide con el vector unitario $\mathbf{n}(\theta, \varphi)$ correspondiente a la parametrización usual de la esfera (salvo en los polos norte y sur, en que $\mathbf{n}(\theta, \varphi)$ no está definido), ya que

$$\begin{aligned} \mathbf{r}_\theta \times \mathbf{r}_\varphi &= a^2 \sin \theta (\cos \theta \cos \varphi \mathbf{i} + \cos \theta \sin \varphi \mathbf{j} - \sin \theta \mathbf{k}) \times (-\sin \varphi \mathbf{i} + \cos \varphi \mathbf{j}) \\ &= a^2 \sin \theta (\cos \theta \cos^2 \varphi \mathbf{k} + \cos \theta \sin^2 \varphi \mathbf{k} + \sin \theta \sin \varphi \mathbf{j} + \sin \theta \cos \varphi \mathbf{i}) \\ &= a \sin \theta \mathbf{r} \quad \Rightarrow \quad \mathbf{n}(\theta, \varphi) = \frac{\mathbf{r}_\theta \times \mathbf{r}_\varphi}{\|\mathbf{r}_\theta \times \mathbf{r}_\varphi\|} = \frac{\mathbf{r}}{a}. \quad \square \end{aligned}$$

Si S es una superficie orientable y $\mathbf{N} : S \rightarrow \mathbb{R}^3$ es la orientación escogida en S , se define la integral de superficie del campo vectorial $\mathbf{F} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ sobre S como

$$\int_S \mathbf{F}(\mathbf{r}) \cdot d\mathbf{S} \equiv \int_S \mathbf{F}(\mathbf{r}) \cdot \mathbf{N}(\mathbf{r}) \, dS = \int_\sigma \mathbf{F}(\mathbf{r}) \cdot \mathbf{N}(\mathbf{r}) \, dS,$$

donde $\sigma : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ es *cualquier* parametrización de S . Supongamos que todos los puntos de D son *regulares*. Si $(u, v) \in D$ entonces

$$\mathbf{N}(\mathbf{r}(u, v)) = \varepsilon(u, v) \mathbf{n}(u, v), \quad \forall (u, v) \in D,$$

con $\varepsilon(u, v) \in \{\pm 1\}$. Si (como es habitual) D es *conexo*, en virtud de la *continuidad* del campo vectorial \mathbf{N} el signo $\varepsilon(u, v)$ es de hecho *constante*⁵ en todo D . Por tanto en este caso podemos escribir

$$\int_S \mathbf{F}(\mathbf{r}) \cdot d\mathbf{S} = \varepsilon \int_\sigma \mathbf{F}(\mathbf{r}) \cdot d\mathbf{S}.$$

⁵La demostración de esta afirmación es muy sencilla. En efecto, al ser $\varepsilon(u, v) = \mathbf{N}(\mathbf{r}(u, v)) \cdot \mathbf{n}(u, v)$ continua en el conjunto conexo D , la imagen de D bajo ε ha de ser conexa. Como $\varepsilon(D) \subset \{\pm 1\}$, necesariamente $\varepsilon(D) = \{1\}$ o $\varepsilon(D) = \{-1\}$.

Ejercicio 36. Calcular $\int_S \mathbf{F} \cdot d\mathbf{S}$, siendo $\mathbf{F}(x, y, z) = 18z\mathbf{i} - 12y\mathbf{j} + 3y\mathbf{k}$ y S la región del plano $z = (12 - 2x - 3y)/6$ situada en el primer octante.

Solución. En general, consideremos una superficie S definida en forma implícita por una ecuación de la forma $\Phi(x, y, z) = 0$ (con Φ de clase C^1), y supongamos que (por ejemplo) de esta ecuación puede despejarse z en función de (x, y) . Como vimos anteriormente (cf. la ec. (5.24)), si $\Phi_z \neq 0$ el elemento de área de S está dado por

$$dS = \frac{\|\nabla\Phi\|}{|\Phi_z|} dx dy,$$

mientras que un vector unitario normal a la superficie es

$$\mathbf{N} = \pm \frac{\nabla\Phi}{\|\nabla\Phi\|}.$$

Por tanto el elemento de superficie en este caso está dado por

$$d\mathbf{S} = \pm \frac{\nabla\Phi}{|\Phi_z|} dx dy.$$

Al ser $\Phi_z \neq 0$ por hipótesis, si (como estamos suponiendo) la superficie es conexa el signo de F_z es *constante*, y por tanto podemos escribir la ecuación anterior en la forma más sencilla

$$d\mathbf{S} = \pm \frac{\nabla\Phi}{\Phi_z} dx dy.$$

En este caso la superficie S es la gráfica de la función $f(x, y) = (12 - 2x - 3y)/6$ en el conjunto D del plano descrito por las ecuaciones

$$x \geq 0, \quad y \geq 0, \quad 2x + 3y \leq 12.$$

(En otras palabras, S es el triángulo de vértices $(6, 0, 0)$, $(0, 4, 0)$ y $(0, 0, 2)$.) Nótese que la última desigualdad se sigue de la condición $z \geq 0$, ya que solo pertenecen a S los puntos del plano $z = f(x, y)$ que están en el primer octante. En este caso podemos tomar $F(x, y, z) = 2x + 3y + 6z - 12$, y por tanto el elemento de superficie es

$$d\mathbf{S} = \pm \frac{(2, 3, 6)}{6} dx dy = \pm \left(\frac{1}{3}, \frac{1}{2}, 1\right) dx dy$$

y la integral pedida está dada por

$$\begin{aligned} \int_S \mathbf{F} \cdot d\mathbf{S} &= \pm \int_D \mathbf{F}(x, y, f(x, y)) \cdot \left(\frac{1}{3}, \frac{1}{2}, 1\right) dx dy \\ &= \pm \int_D (12 - 2x - 3y - 6 + 3y) dx dy = \pm 2 \int_D (3 - x) dx dy, \end{aligned}$$

donde el signo “ \pm ” depende de la elección del vector unitario normal a la superficie. Evaluando la integral doble se obtiene finalmente

$$\begin{aligned} \int_S \mathbf{F} \cdot d\mathbf{S} &= \pm 2 \int_0^4 dy \int_0^{6-\frac{3}{2}y} dx (3 - x) = \mp \int_0^4 dy (3 - x)^2 \Big|_{x=0}^{x=6-\frac{3}{2}y} = \mp \int_0^4 \left[\left(\frac{3y}{2} - 3\right)^2 - 9 \right] dy \\ &= \mp 9 \int_0^4 \left(\frac{1}{4}y^2 - y\right) dy = \mp (48 - 72) = \boxed{\pm 24}. \quad \square \end{aligned}$$

Ejemplo 5.44. La integral de superficie de un campo vectorial sobre una superficie orientada S se denomina también **flujo** del campo vectorial a través de la superficie. Para entender el origen de esta terminología, sea $\mathbf{v} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ el campo de velocidades de un fluido, sea S una superficie orientable, y sea \mathbf{N} la orientación escogida de S . El volumen ΔV de fluido que atraviesa un elemento de superficie ΔS centrado en el punto \mathbf{r} en el tiempo Δt y en la dirección del vector \mathbf{N} es aproximadamente igual al volumen de un paralelepípedo de base ΔS y altura $\|\mathbf{v}(\mathbf{r})\| \Delta t \cdot \cos \theta$, donde θ es el ángulo formado por los vectores $\mathbf{N}(\mathbf{r})$ y $\mathbf{v}(\mathbf{r})$. Como $\mathbf{N}(\mathbf{r})$ es unitario,

$$\|\mathbf{v}(\mathbf{r})\| \cos \theta = \mathbf{v}(\mathbf{r}) \cdot \mathbf{N}(\mathbf{r}),$$

y por tanto el volumen de fluido que atraviesa el elemento de superficie ΔS por unidad de tiempo está dado por⁶

$$\frac{\Delta V}{\Delta t} = \mathbf{v}(\mathbf{r}) \cdot \mathbf{N}(\mathbf{r}) \Delta S.$$

Por consiguiente, el volumen total de fluido que atraviesa la superficie S por unidad de tiempo es igual a la integral de superficie

$$\boxed{\int_S \mathbf{v}(\mathbf{r}) \cdot \mathbf{N}(\mathbf{r}) \, dS \equiv \int_S \mathbf{v}(\mathbf{r}) \cdot d\mathbf{S}}.$$

Supongamos, por ejemplo, que

$$\mathbf{v}(\mathbf{r}) = \mathbf{i} + x\mathbf{j} + z\mathbf{k} \quad \text{m s}^{-1},$$

y que S es la semiesfera $x^2 + y^2 + z^2 = 1, z \geq 0$, con la orientación $\mathbf{N}(\mathbf{r}) = \mathbf{r}$ (obsérvese que $\|\mathbf{N}\| = 1$, ya que $\|\mathbf{r}\| = 1$ en la semiesfera). En coordenadas esféricas,

$$dS = \sin \theta \, d\theta \, d\varphi,$$

(ya que el radio de la esfera es igual a 1), y

$$\mathbf{v} \cdot \mathbf{N} = (\mathbf{i} + x\mathbf{j} + z\mathbf{k}) \cdot (x\mathbf{i} + y\mathbf{j} + z\mathbf{k}) = x + xy + z^2 = \sin \theta \cos \varphi + \sin^2 \theta \sin \varphi \cos \varphi + \cos^2 \theta.$$

Por tanto el volumen de fluido por unidad de tiempo (en metros cúbicos por segundo) que atraviesa la semiesfera S está dado por

$$\begin{aligned} \int_S \mathbf{v} \cdot \mathbf{N} \, dS &= \int_0^{\pi/2} d\theta \int_0^{2\pi} d\varphi \sin \theta (\sin \theta \cos \varphi + \sin^2 \theta \sin \varphi \cos \varphi + \cos^2 \theta) \\ &= 2\pi \int_0^{\pi/2} \sin \theta \cos^2 \theta \, d\theta = -\frac{2\pi}{3} \cos^3 \theta \Big|_{\theta=0}^{\theta=\pi/2} = \boxed{\frac{2\pi}{3}} \text{ m}^3 \text{ s}^{-1}. \end{aligned}$$

Nótese que el resultado es positivo, y por tanto el fluido atraviesa la superficie en la dirección del vector \mathbf{N} (y no en la de $-\mathbf{N}$).

⁶Nótese que si el ángulo θ es obtuso, es decir si $\mathbf{v}(\mathbf{r}) \cdot \mathbf{N}(\mathbf{r}) < 0$, el valor de ΔV es *negativo*. En tal caso el fluido atraviesa el elemento de superficie ΔS en la dirección de $-\mathbf{N}(\mathbf{r})$.

Capítulo 6

Teoremas de Green, Stokes y Gauss

6.1 Teorema de Green

En este capítulo deduciremos los tres teoremas clásicos del Cálculo integral en \mathbb{R}^2 y \mathbb{R}^3 , que relacionan las integrales de línea, superficie y volumen. Comenzaremos con el teorema de Green, el más sencillo de los tres, ya que se aplica a la integral línea de un campo vectorial en el plano. Para enunciar este teorema, necesitamos la siguiente definición previa:

Definición 6.1. Una curva suave *cerrada* $\gamma : [a, b] \rightarrow \mathbb{R}^2$ es **simple** si la aplicación γ es inyectiva en el intervalo $[a, b)$. Una **curva de Jordan** es una curva cerrada simple.

Intuitivamente, una curva cerrada es simple si no tiene “autointersecciones”. Por ejemplo, una circunferencia es claramente una curva de Jordan, mientras que la figura “8” no lo es.

Por definición, el **interior** de una curva de Jordan γ es la región *acotada* del plano limitada por dicha curva. Este conjunto es *abierto*, y su frontera es la curva γ . Sea $D \subset \mathbb{R}^2$ el conjunto formado por la unión de los puntos de una curva de Jordan γ con su interior. Por lo que acabamos de ver, D es un conjunto cerrado con frontera

$$\partial D = \gamma,$$

donde (como es habitual) hemos confundido la aplicación γ con su imagen $\gamma([a, b])$. Nótese que el conjunto D es *medible Jordan*, ya que es acotado (¿por qué?) y su frontera es la curva γ , que tiene medida cero. A partir de ahora, si no se dice nada en contrario entenderemos que una curva de Jordan γ está orientada en sentido **antihorario** (llamado también **positivo**). Con esta orientación, al recorrer la curva γ el conjunto D queda siempre a la *izquierda*. Por ejemplo, la circunferencia de centro el origen y radio $a > 0$ con la parametrización habitual $\gamma(t) = a(\cos t, \sin t)$, con $0 \leq t \leq 2\pi$, está orientada positivamente.

Tras estos preliminares, estamos ya en condiciones de enunciar y demostrar el teorema de Green:

Teorema de Green. Sea $\gamma : [a, b] \rightarrow \mathbb{R}^2$ una curva de Jordan, y sea D el conjunto del plano formado por la curva γ y su interior. Si $\mathbf{F} \equiv (A, B) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ es un campo vectorial de clase C^1 en D entonces

$$\int_{\gamma} \mathbf{F}(\mathbf{r}) \cdot d\mathbf{r} \equiv \int_{\gamma} (A dx + B dy) = \int_D \left(\frac{\partial B}{\partial x} - \frac{\partial A}{\partial y} \right) dx dy. \quad (6.1)$$

Comentario. Recuérdese que $\mathbf{F} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ es de clase C^1 en un conjunto arbitrario $D \subset \mathbb{R}^2$ si \mathbf{F} (o, equivalentemente, sus componentes A y B) es de clase C^1 en un abierto que contiene a D .

Demostración. Demostraremos el teorema suponiendo, por sencillez, que D es una región de tipo III (es decir, simultáneamente de tipos I y II). Para ello probaremos por separado que

$$\int_y A dx = - \int_D \frac{\partial A}{\partial y} dx dy \tag{6.2}$$

si D es de tipo I, mientras que

$$\int_y B dy = \int_D \frac{\partial B}{\partial x} dx dy \tag{6.3}$$

si D es de tipo II. Si D es de tipo III, se verifican *simultáneamente* las dos identidades anteriores. Sumando dichas identidades se obtiene entonces el teorema de Green (6.1).

Supongamos por tanto, en primer lugar, que D de tipo I, es decir que está descrita por ecuaciones de la forma

$$a \leq x \leq b, \quad \phi_1(x) \leq y \leq \phi_2(x),$$

con $\phi_{1,2}$ de clase C^1 a trozos en $[a, b]$ (cf. la Fig. 6.1, izda.). En tal caso, la curva γ (que es, por

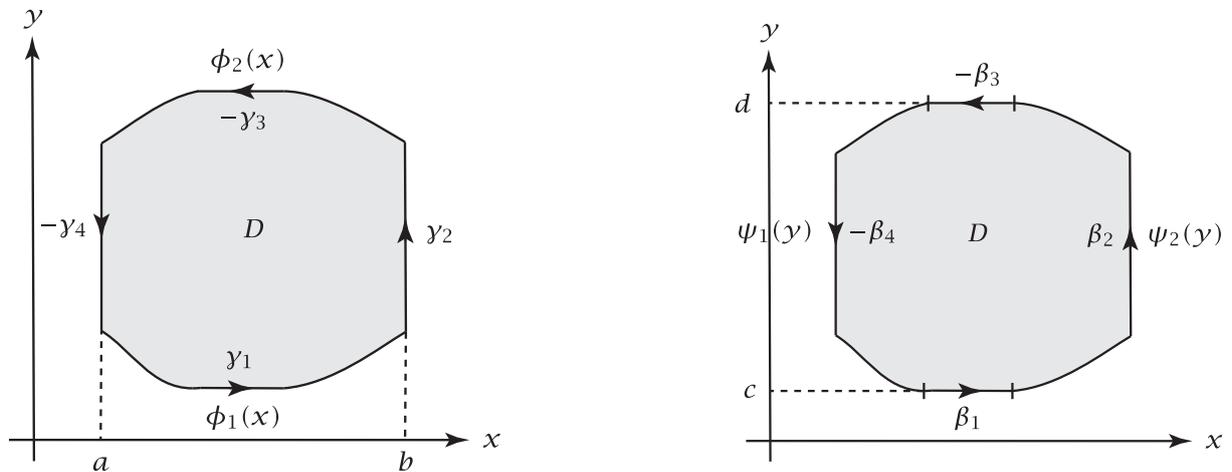


Figura 6.1: Región D en la demostración del teorema de Green vista como un conjunto de tipo I (izda.) y de tipo II (drcha.)

construcción, la frontera de D) está formada por las cuatro curvas $\gamma_1, \gamma_2, -\gamma_3$ y $-\gamma_4$, donde γ_2 y γ_4 son los dos segmentos verticales (¡alguno de los cuales puede reducirse a un punto!)

$$\gamma_2(t) = (b, t), \quad t \in [\phi_1(b), \phi_2(b)]; \quad \gamma_4(t) = (a, t), \quad t \in [\phi_1(a), \phi_2(a)]$$

y $\gamma_{1,3}$ son las gráficas de las funciones $\phi_{1,2}$. Teniendo esto en cuenta y aplicando el teorema de Fubini a la integral en D de la derivada parcial $-A_y$ se obtiene

$$\begin{aligned} - \int_D \frac{\partial A}{\partial y}(x, y) dx dy &= - \int_a^b dx \int_{\phi_1(x)}^{\phi_2(x)} dy \frac{\partial A}{\partial y}(x, y) \\ &= \int_a^b [A(x, \phi_1(x)) - A(x, \phi_2(x))] dx. \end{aligned} \tag{6.4}$$

Por otra parte, por lo que acabamos de ver

$$\begin{aligned} \int_y A(x, y) dx &\equiv \int_y (A(x, y), 0) \cdot dr \\ &= \int_{\gamma_1} A(x, y) dx + \int_{\gamma_2} A(x, y) dx - \int_{\gamma_3} A(x, y) dx - \int_{\gamma_4} A(x, y) dx. \end{aligned}$$

La integral del campo $(A, 0)$ se anula claramente en los segmentos verticales $\gamma_{2,4}$, ya que en ellos $x(t)$ es constante, y por tanto $dx = x'(t) dt = 0$. Análogamente, en las curvas $\gamma_{1,3}$ se tiene

$$\gamma_{1,3}(x) = (x, \phi_{1,2}(x)) \implies A(x, y) dx = A(x, \phi_{1,2}(x)) dx$$

y por tanto

$$\int_{\gamma_{1,3}} A(x, y) dx = \int_a^b A(x, \phi_{1,2}(x)) dx.$$

En definitiva,

$$\begin{aligned} \int_{\gamma} A(x, y) dx &= \int_{\gamma_1} A(x, y) dx - \int_{\gamma_3} A(x, y) dx = \int_a^b [A(x, \phi_1(x)) - A(x, \phi_2(x))] dx \\ &= - \int_D \frac{\partial A}{\partial y}(x, y) dx dy, \end{aligned}$$

en virtud de (6.4). Esto demuestra la ec. (6.2) si D es de tipo I.

Supongamos a continuación que el conjunto D es de tipo II, y está por tanto descrito por ecuaciones de la forma

$$c \leq y \leq d, \quad \psi_1(y) \leq x \leq \psi_2(y)$$

(cf. la Fig. 6.1, drcha.). Aplicando el teorema de Fubini se obtiene

$$\int_D \frac{\partial B}{\partial x}(x, y) dx dy = \int_c^d dy \int_{\psi_1(y)}^{\psi_2(y)} dx \frac{\partial B}{\partial x}(x, y) = \int_c^d [B(\psi_2(y), y) - B(\psi_1(y), y)] dy. \quad (6.5)$$

Al igual que antes, la curva γ está formada por las curvas $\beta_1, \beta_2, -\beta_3$ y $-\beta_4$, donde β_1 y β_3 son los segmentos horizontales (que, de nuevo, pueden reducirse a un punto)

$$\beta_1(t) = (t, c), \quad t \in [\psi_1(c), \psi_2(c)]; \quad \beta_3(t) = (t, d), \quad t \in [\psi_1(d), \psi_2(d)],$$

y $\beta_{2,4}$ son las gráficas de las funciones ψ_2 y ψ_1 en el intervalo $[c, d]$, respectivamente. En este caso,

$$\int_{\beta_{1,3}} B(x, y) dy = 0,$$

ya que en los segmentos verticales $\gamma(t)$ es constante, y por tanto $dy = \gamma'(t) dt = 0$. Por otra parte, en las curvas $\beta_{2,4}$ se cumple

$$\beta_{2,4}(y) = (\psi_{2,1}(y), y) \implies B(x, y) dy = B(\psi_{2,1}(y), y) dy.$$

Utilizando (6.5) se obtiene

$$\begin{aligned} \int_{\gamma} B(x, y) dy &= \int_{\beta_2} B(x, y) dy - \int_{\beta_4} B(x, y) dy = \int_c^d [B(\psi_2(y), y) - B(\psi_1(y), y)] dy \\ &= \int_D \frac{\partial B}{\partial x}(x, y) dx dy, \end{aligned}$$

que es la ec. (6.3). □

La demostración anterior puede extenderse a regiones muy generales, que no son ni siquiera de tipos I o II, subdividiendo dichas regiones en conjuntos de tipo III y aplicando el teorema de Green a cada uno de dichos conjuntos. Consideremos, por ejemplo, el conjunto D de la Fig. 6.2. Este conjunto es la unión de las regiones de tipo III D_1 y D_2 , cuya intersección es un conjunto de medida 0. Por tanto

$$\int_D (B_x - A_y) dx dy = \int_{D_1} (B_x - A_y) dx dy + \int_{D_2} (B_x - A_y) dx dy.$$

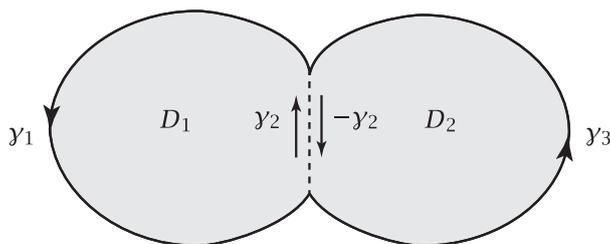


Figura 6.2: Conjunto D que no es de tipo III al que se puede aplicar la demostración del teorema de Green.

Aplicando el teorema de Green a cada una de estas integrales se obtiene

$$\int_D (B_x - A_y) \, dx \, dy = \int_{\partial D_1} (A \, dx + B \, dy) + \int_{\partial D_2} (A \, dx + B \, dy). \tag{6.6}$$

Claramente,

$$\partial D_1 = \gamma_1 + \gamma_2, \quad \partial D_2 = \gamma_3 + (-\gamma_2) \equiv \gamma_3 - \gamma_2,$$

donde la suma de dos curvas α y β (tales que el extremo final de α coincide con el inicial de β) es la curva obtenida recorriendo primero α y a continuación β . Por tanto, el miembro derecho de (6.6) está dado por

$$\begin{aligned} \int_{\gamma_1} (A \, dx + B \, dy) + \int_{\gamma_2} (A \, dx + B \, dy) + \int_{\gamma_3} (A \, dx + B \, dy) - \int_{\gamma_2} (A \, dx + B \, dy) \\ = \int_{\gamma_1} (A \, dx + B \, dy) + \int_{\gamma_3} (A \, dx + B \, dy) \equiv \int_{\gamma} (A \, dx + B \, dy), \end{aligned}$$

siendo $\gamma \equiv \gamma_1 + \gamma_3$ la frontera de D . □

Ejemplo 6.2 (*teorema de Green en regiones múltiplemente conexas*). De hecho, el teorema de Green se puede extender a regiones que no son ni siquiera simplemente conexas. Consideremos, por ejemplo, el conjunto D de la Fig. 6.3, limitado por las dos curvas de Jordan $\gamma_1 = \gamma_1^+ + \gamma_1^-$ y $\gamma_2 = \gamma_2^+ + \gamma_2^-$.

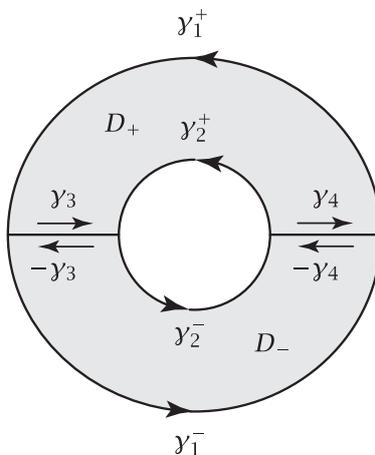


Figura 6.3: Región no simplemente conexa a la que se puede aplicar el teorema de Green.

Este conjunto es la unión de las regiones D_- y D_+ , cuya intersección es un conjunto de medida 0, a cada una de la cuales se le puede aplicar el teorema de Green (cf. el comentario anterior). Por

tanto

$$\begin{aligned} \int_D (B_x - A_y) dx dy &= \int_{D_-} (B_x - A_y) dx dy + \int_{D_+} (B_x - A_y) dx dy \\ &= \int_{\partial D_-} (A dx + B dy) + \int_{\partial D_+} (A dx + B dy). \end{aligned} \quad (6.7)$$

Claramente, las fronteras de las regiones D_{\pm} son las curvas de Jordan

$$\partial D_- = \gamma_1^- - \gamma_4 - \gamma_2^- - \gamma_3, \quad \partial D_+ = \gamma_1^+ + \gamma_3 - \gamma_2^+ + \gamma_4.$$

Por tanto, el miembro derecho de (6.7) es igual a

$$\begin{aligned} &\left(\int_{\gamma_1^-} - \int_{\gamma_4} - \int_{\gamma_2^-} - \int_{\gamma_3} + \int_{\gamma_1^+} + \int_{\gamma_3} - \int_{\gamma_2^+} + \int_{\gamma_4} \right) (A dx + B dy) \\ &= \left(\int_{\gamma_1^-} + \int_{\gamma_1^+} - \int_{\gamma_2^-} - \int_{\gamma_2^+} \right) (A dx + B dy) = \int_{\gamma_1} (A dx + B dy) - \int_{\gamma_2} (A dx + B dy). \end{aligned}$$

En definitiva, hemos demostrado que en este caso

$$\int_D \left(\frac{\partial B}{\partial x} - \frac{\partial A}{\partial y} \right) dx dy = \int_{\gamma_1} (A dx + B dy) - \int_{\gamma_2} (A dx + B dy).$$

Nótese que la igualdad anterior se puede escribir en la forma

$$\int_D \left(\frac{\partial B}{\partial x} - \frac{\partial A}{\partial y} \right) dx dy = \int_{\gamma_1} (A dx + B dy) + \int_{-\gamma_2} (A dx + B dy),$$

donde las curvas γ_1 y $-\gamma_2$, cuya unión es la frontera de D , están ambas recorridas de modo que dicho conjunto quede a la *izquierda*. Podemos entonces escribir también en este caso

$$\int_D \left(\frac{\partial B}{\partial x} - \frac{\partial A}{\partial y} \right) dx dy = \int_{\gamma} (A dx + B dy),$$

donde $\gamma = \gamma_1 - \gamma_2$ es la llamada **frontera orientada** de la región D .

Ejemplo 6.3 (*área de una figura plana en términos de una integral de línea*). Mediante el teorema de Green es posible expresar el área de una región plana por medio de una integral de línea. En efecto, si D es el conjunto cerrado acotado por una curva de Jordan y entonces

$$\int_{\gamma} [(ax + by) dx + (cx + ey) dy] = \int_D (c - b) dx dy = (c - b)a(D). \quad (6.8)$$

(Nótese que el resultado es independiente de a y e , como era de esperar, ya que

$$\int_{\gamma} (ax dx + ey dy) = \frac{1}{2} \int_{\gamma} \nabla(ax^2 + ey^2) \cdot d\mathbf{r} = 0,$$

por el teorema fundamental del Cálculo.) Sustituyendo $c = -b = 1/2$, $c = 1$, $b = 0$ y $c = 0$, $b = -1$ en (6.8) se obtienen las fórmulas

$$a(D) = \frac{1}{2} \int_{\gamma} (x dy - y dx) = \int_{\gamma} x dy = - \int_{\gamma} y dx.$$

Por ejemplo, calculemos el área de la región D limitada por la elipse γ de ecuación $(x/a)^2 + (y/b)^2 = 1$. Utilizando la parametrización

$$x = a \cos t, \quad y = b \sin t$$

con $0 \leq t \leq 2\pi$ (con la cual la orientación de γ es claramente *positiva*) se obtiene inmediatamente

$$a(D) = \frac{1}{2} \int_0^{2\pi} [a \cos t \cdot b \cos t - b \sin t \cdot (-a \sin t)] dt = \frac{1}{2} \int_0^{2\pi} ab dt = \boxed{\pi ab}.$$

Ejercicio 37. Calcular la integral del campo vectorial $\mathbf{F}(x, y) = (3x^2 + 2y)\mathbf{i} - (x + 3 \cos y)\mathbf{j}$ a lo largo de la trayectoria cerrada compuesta por segmentos de recta que unen los puntos $(0, 0)$, $(2, 0)$, $(3, 1)$, $(1, 1)$ y $(0, 0)$.

Solución. La curva γ es una curva de Jordan positivamente orientada que delimita el paralelogramo D generado por los vectores $(2, 0)$ y $(1, 1)$. Aplicando el teorema de Green se obtiene

$$\int_{\gamma} [(3x^2 + 2y) dx - (x + 3 \cos y) dy] = \int_D (-2 - 1) dx dy = -3a(D) = -3 \cdot 2 = \boxed{-6}$$

(cf. el Ejemplo 5.16).

Ejemplo 6.4 (teorema de la divergencia en el plano). Sea, de nuevo, D el conjunto formado por la unión de una curva de Jordan $\gamma : [a, b] \rightarrow \mathbb{R}^2$ y su interior, y sea $\mathbf{F} = (A, B) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ un campo vectorial de clase C^1 en D . Si $\mathbf{y}'(t) \equiv (x'(t), y'(t)) \neq 0$, el vector unitario normal *exterior* a γ está dado por

$$\mathbf{n} = \frac{\mathbf{y}'(t) \times \mathbf{k}}{\|\mathbf{y}'(t) \times \mathbf{k}\|} = \frac{(y'(t), -x'(t))}{\sqrt{x'(t)^2 + y'(t)^2}},$$

de donde se sigue que

$$\mathbf{n} ds = (y'(t), -x'(t)) dt.$$

Por tanto

$$\int_{\gamma} \mathbf{F} \cdot \mathbf{n} ds = \int_a^b [A(x(t), y(t))y'(t) - B(x(t), y(t))x'(t)] dt \equiv \int_{\gamma} A dy - B dx,$$

y aplicando el teorema de Green se obtiene

$$\boxed{\int_{\gamma} \mathbf{F} \cdot \mathbf{n} ds = \int_D (A_x + B_y) dx dy \equiv \int_D \nabla \cdot \mathbf{F} dx dy}. \quad (6.9)$$

Este es el llamado *teorema de la divergencia en el plano*.

6.2 Teorema de Stokes

El teorema de Green se expresa fácilmente en notación vectorial. En efecto, si $\mathbf{F} \equiv A\mathbf{i} + B\mathbf{j}$ es un campo vectorial en \mathbb{R}^2 , podemos considerarlo como un campo vectorial en \mathbb{R}^3 (independiente de z) cuya componente z es nula. Con este convenio, el rotacional de \mathbf{F} está dado por

$$\nabla \times \mathbf{F} = \left(\frac{\partial B}{\partial x} - \frac{\partial A}{\partial y} \right) \mathbf{k}.$$

Análogamente, si D es la región del plano (x, y) formada por la unión de una curva de Jordan γ con su interior, podemos considerar a D como un subconjunto de \mathbb{R}^3 contenido en el plano $z = 0$. En tal caso,

$$\frac{\partial B}{\partial x} - \frac{\partial A}{\partial y} = \nabla \times \mathbf{F} \cdot \mathbf{N},$$

siendo $\mathbf{N} = \mathbf{k}$ un vector unitario normal a $D \subset \mathbb{R}^3$. Por tanto, el teorema de Green se puede escribir en la forma

$$\boxed{\int_D \nabla \times \mathbf{F} \cdot \mathbf{N} \, dS = \int_\gamma \mathbf{F} \cdot d\mathbf{r}}, \quad (6.10)$$

ya que $dS = dx \, dy$ es el elemento de área en el plano $z = 0$. Nótese que la normal $\mathbf{N} = \mathbf{k}$ que aparece en la fórmula anterior está relacionada con la orientación (antihoraria) de la curva γ mediante la regla siguiente: al recorrer la curva γ , la cara de la superficie D determinada por la normal \mathbf{N} (intuitivamente, la cara en la que “se apoya” el vector \mathbf{N}) queda siempre a la *izquierda*. La ec. (6.10) se puede extender esencialmente a cualquier superficie *orientable* de \mathbb{R}^3 . Más precisamente, se verifica el siguiente resultado:

Teorema de Stokes. Sea $D \subset \mathbb{R}^2$ el conjunto formado por la unión de una curva de Jordan y su interior, y sea $\sigma : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ una superficie parametrizada de clase C^2 . Si $\mathbf{F} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ es un campo vectorial de clase C^1 en $\sigma(D)$ entonces se verifica

$$\int_\sigma \nabla \times \mathbf{F} \cdot d\mathbf{S} = \int_C \mathbf{F} \cdot d\mathbf{r}, \quad (6.11)$$

donde C es la imagen de γ bajo la parametrización σ , es decir $C(t) = \sigma(\gamma(t))$.

Demostración. Si

$$\gamma(t) = (u(t), v(t)), \quad a \leq t \leq b,$$

es la parametrización de la frontera ∂D de D entonces

$$C(t) = \mathbf{r}(\gamma(t)) = \mathbf{r}(u(t), v(t)), \quad a \leq t \leq b,$$

donde $\mathbf{r}(u, v)$ (con $(u, v) \in D$) es la parametrización de la superficie $\sigma(D)$. Por tanto

$$\begin{aligned} \int_C \mathbf{F} \cdot d\mathbf{r} &= \int_a^b \mathbf{F}(\mathbf{r}(\gamma(t))) \cdot \frac{d}{dt} \mathbf{r}(\gamma(t)) \, dt = \int_a^b \mathbf{F}(\mathbf{r}(\gamma(t))) \cdot \left[\frac{\partial \mathbf{r}}{\partial u}(\gamma(t)) u'(t) + \frac{\partial \mathbf{r}}{\partial v}(\gamma(t)) v'(t) \right] dt \\ &\equiv \int_\gamma \left[\mathbf{F}(\mathbf{r}(u, v)) \cdot \frac{\partial \mathbf{r}}{\partial u}(u, v) \, du + \mathbf{F}(\mathbf{r}(u, v)) \cdot \frac{\partial \mathbf{r}}{\partial v}(u, v) \, dv \right]. \end{aligned}$$

Aplicando el teorema de Green a esta última integral se obtiene

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_D f(u, v) \, du \, dv, \quad (6.12)$$

siendo

$$\begin{aligned} f(u, v) &= \frac{\partial}{\partial u} \left(\mathbf{F}(\mathbf{r}(u, v)) \cdot \frac{\partial \mathbf{r}}{\partial v} \right) - \frac{\partial}{\partial v} \left(\mathbf{F}(\mathbf{r}(u, v)) \cdot \frac{\partial \mathbf{r}}{\partial u} \right) \\ &= \frac{\partial}{\partial u} \left[\mathbf{F}(\mathbf{r}(u, v)) \right] \cdot \frac{\partial \mathbf{r}}{\partial v} - \frac{\partial}{\partial v} \left[\mathbf{F}(\mathbf{r}(u, v)) \right] \cdot \frac{\partial \mathbf{r}}{\partial u}. \end{aligned}$$

Llamando $\mathbf{r}(u, v) = (x_1(u, v), x_2(u, v), x_3(u, v))$ y abreviando $\mathbf{F}(\mathbf{r}(u, v))$ por \mathbf{F} se tiene:

$$\begin{aligned} f(u, v) &= \sum_{j=1}^3 \left(\frac{\partial F_j}{\partial u} \frac{\partial x_j}{\partial v} - \frac{\partial F_j}{\partial v} \frac{\partial x_j}{\partial u} \right) = \sum_{i,j=1}^3 \left(\frac{\partial F_j}{\partial x_i} \frac{\partial x_i}{\partial u} \frac{\partial x_j}{\partial v} - \frac{\partial F_j}{\partial x_i} \frac{\partial x_i}{\partial v} \frac{\partial x_j}{\partial u} \right) \\ &= \sum_{i,j=1}^3 \frac{\partial F_j}{\partial x_i} \det \frac{\partial (x_i, x_j)}{\partial (u, v)} = \frac{1}{2} \sum_{i,j=1}^3 \left(\frac{\partial F_j}{\partial x_i} - \frac{\partial F_i}{\partial x_j} \right) \det \frac{\partial (x_i, x_j)}{\partial (u, v)} \\ &= \sum_{1 \leq i < j \leq 3} \left(\frac{\partial F_j}{\partial x_i} - \frac{\partial F_i}{\partial x_j} \right) \det \frac{\partial (x_i, x_j)}{\partial (u, v)} = \nabla \times \mathbf{F} \cdot \left(\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right), \end{aligned}$$

donde hemos utilizado (en la cuarta igualdad) que $\det \frac{\partial(x_i, x_j)}{\partial(u, v)}$ es antisimétrico¹ bajo el intercambio de i y j . De la expresión anterior para $f(u, v)$ y la ec. (6.12) se sigue finalmente que

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_D \nabla \times \mathbf{F} \cdot \left(\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right) du dv \equiv \int_\sigma \nabla \times \mathbf{F} \cdot d\mathbf{S}. \quad \square$$

Nótese que la curva C es la *frontera* de la superficie $\sigma(D) = S$, con la orientación inducida por la de la curva de Jordan γ que constituye la frontera del conjunto D . Además, la integral de superficie del miembro izquierdo de (6.11) es por definición igual a

$$\int_\sigma \nabla \times \mathbf{F} \cdot \mathbf{n} dS,$$

donde \mathbf{n} es el vector unitario normal a $\sigma(D)$ inducido por la parametrización σ , es decir

$$\mathbf{n} = \frac{\mathbf{r}_u \times \mathbf{r}_v}{\|\mathbf{r}_u \times \mathbf{r}_v\|}$$

(con $\mathbf{r}(u, v) \equiv \sigma(u, v)$). Se demuestra que este vector normal está relacionado con la orientación de la frontera C de $\sigma(D)$ por la misma regla que en el caso bidimensional. En otras palabras, al recorrer C la cara de la superficie $\sigma(D)$ determinada por el vector normal \mathbf{n} queda siempre a la *izquierda*. Si $S = \sigma(D)$ podemos por tanto escribir la ec. (6.11) en la forma

$$\int_S \nabla \times \mathbf{F} \cdot \mathbf{N} dS = \int_{\partial S} \mathbf{F} \cdot d\mathbf{r},$$

siempre y cuando la orientación \mathbf{N} de S esté relacionada con el sentido de recorrido de ∂S en la forma que acabamos de indicar.

Ejemplo 6.5. Calculemos la integral $\int_S (\nabla \times \mathbf{F}) \cdot d\mathbf{S}$, donde $\mathbf{F}(x, y, z) = y\mathbf{i} - x\mathbf{j} + x^3y^2z\mathbf{k}$ y S es la región del elipsoide $x^2 + y^2 + 3z^2 = 1$ con $z \leq 0$. En primer lugar, como la superficie S satisface la relación implícita $F(x, y, z) = 0$, siendo $\Phi(x, y, z) = (x^2 + y^2 + 3z^2 - 1)/2$, podemos escoger como orientación de S la dada por el vector

$$\mathbf{N} = \frac{\nabla \Phi}{\|\nabla \Phi\|} = \frac{(x, y, 3z)}{\sqrt{x^2 + y^2 + 9z^2}}.$$

Al ser $z \leq 0$ en la superficie, la componente vertical de este vector es *negativa* en S , y por tanto \mathbf{N} es la normal *exterior* al elipsoide. Con esta orientación, la frontera de S es la circunferencia $x^2 + y^2 = 1$ en el plano $z = 0$, recorrida en sentido *horario*. Aplicando el teorema de Stokes obtenemos

$$\int_S \nabla \times \mathbf{F} \cdot d\mathbf{S} = - \int_C \mathbf{F} \cdot d\mathbf{r},$$

donde C es la circunferencia $x^2 + y^2 = 1$ en el plano $z = 0$ orientada *positivamente*. Por tanto

$$\int_S \nabla \times \mathbf{F} \cdot d\mathbf{S} = - \int_0^{2\pi} (\sin t, -\cos t, 0) \cdot (-\sin t, \cos t, 0) dt = \int_0^{2\pi} dt = \boxed{2\pi}.$$

Ejemplo 6.6. Sea S la región de la esfera $x^2 + y^2 + z^2 = 1$ comprendida entre los planos $z = 0$ y $z = 1/2$, y sea $\mathbf{F} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ el campo vectorial dado por $\mathbf{F}(x, y, z) = (z, x, y)$. En coordenadas esféricas, la superficie S está parametrizada por

$$\mathbf{r}(\theta, \varphi) = (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta), \quad (\theta, \varphi) \in D \equiv [\pi/3, \pi/2] \times [0, 2\pi]. \quad (6.13)$$

¹Si $a_{ij} = -a_{ji}$ para todo i, j (y por tanto, en particular, $a_{ii} = 0$ para todo i) entonces

$$\sum_{i,j} a_{ij} b_{ij} = \frac{1}{2} \sum_{i,j} a_{ij} b_{ij} - \frac{1}{2} \sum_{i,j} a_{ji} b_{ji} = \frac{1}{2} \sum_{i,j} a_{ij} (b_{ij} - b_{ji}) = \sum_{i < j} a_{ij} (b_{ij} - b_{ji}).$$

En este caso, hemos visto que

$$\mathbf{n}(\theta, \varphi) = \mathbf{r}$$

salvo en $\theta = 0, \pi$, en que \mathbf{n} no está definido porque el producto vectorial fundamental se anula. Por tanto el flujo del campo vectorial $\nabla \times \mathbf{F} = (1, 1, 1)$ que atraviesa la superficie en la dirección de la normal \mathbf{n} está dado por

$$\int_S \nabla \times \mathbf{F} \cdot d\mathbf{S} \equiv \int_S \nabla \times \mathbf{F} \cdot \mathbf{n} dS = \int_S (1, 1, 1) \cdot \mathbf{r} dS = \int_S (x + y + z) dS.$$

Evaluando esta última integral de superficie obtenemos

$$\begin{aligned} \int_{\pi/3}^{\pi/2} d\theta \int_0^{2\pi} d\varphi \sin \theta (\sin \theta \cos \varphi + \sin \theta \sin \varphi + \cos \theta) &= 2\pi \int_{\pi/3}^{\pi/2} \sin \theta \cos \theta d\theta \\ &= \pi \sin^2 \theta \Big|_{\pi/3}^{\pi/2} = \pi \left(1 - \frac{3}{4}\right) = \boxed{\frac{\pi}{4}}. \end{aligned}$$

Este resultado también se puede obtener aplicando el teorema de Stokes. Para ello, obsérvese

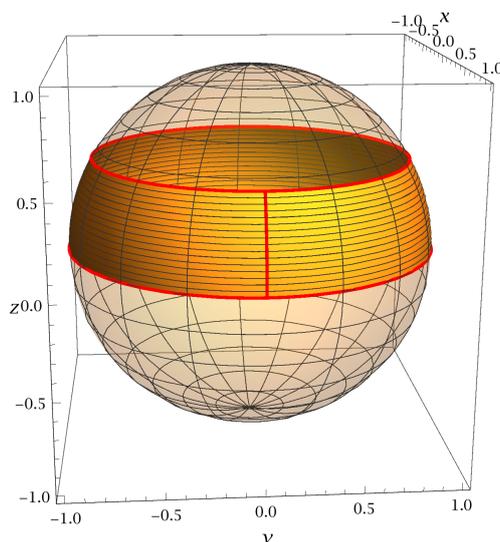


Figura 6.4: Superficie S del Ejemplo 6.6 (en rojo se ha representado la imagen de la frontera del conjunto D bajo la parametrización (6.13)).

que la frontera γ de D es el perímetro del rectángulo $[\pi/3, \pi/2] \times [0, 2\pi]$ orientado en sentido *antihorario*. La imagen de γ bajo la parametrización $\mathbf{r}(\theta, \varphi)$ de la esfera es el meridiano $\varphi = 0, \pi/3 \leq \theta \leq \pi/2$ recorrido *de arriba hacia abajo*, la circunferencia $x^2 + y^2 = 1, z = 0$ recorrida en sentido *antihorario*, a continuación el meridiano $\varphi = 0, \pi/3 \leq \theta \leq \pi/2$ recorrido *de abajo hacia arriba*, y finalmente la circunferencia $x^2 + y^2 = 3/4, z = 1/2$ recorrida en sentido *horario*. Como el meridiano $\varphi = 0, \pi/3 \leq \theta \leq \pi/2$ aparece dos veces, pero recorrido en sentidos opuestos, podemos considerar (a la hora de aplicar el teorema de Stokes) que la frontera de S está formada por la circunferencia $x^2 + y^2 = 1, z = 0$ recorrida en sentido *antihorario*, y la circunferencia $x^2 + y^2 = 3/4, z = 1/2$ recorrida en sentido *horario*. Nótese que este resultado es consistente con la regla que hemos dado anteriormente para relacionar la orientación de la superficie S con la de su frontera, ya que en ambos casos al recorrer la circunferencia la cara de la superficie determinada por la normal $\mathbf{n} = \mathbf{r}$ queda a la *izquierda*. Aplicando el teorema de Stokes se obtiene

$$\boxed{\int_S \nabla \times \mathbf{F} \cdot d\mathbf{S} = \int_{C_1} \mathbf{F} \cdot d\mathbf{r} - \int_{C_2} \mathbf{F} \cdot d\mathbf{r}},$$

donde C_1 y C_2 son las circunferencias anteriores (en los planos $z = 0$ y $z = 1/2$, respectivamente) recorridas en ambos casos en sentido *antihorario*. Parametrizando estas integrales se obtiene

$$\begin{aligned} \int_{C_1} \mathbf{F} \cdot d\mathbf{r} - \int_{C_2} \mathbf{F} \cdot d\mathbf{r} &= \int_0^{2\pi} (0, \cos t, \sin t) \cdot (-\sin t, \cos t, 0) dt \\ &\quad - \int_0^{2\pi} \left(\frac{1}{2}, \frac{\sqrt{3}}{2} \cos t, \frac{\sqrt{3}}{2} \sin t\right) \cdot \left(-\frac{\sqrt{3}}{2} \sin t, \frac{\sqrt{3}}{2} \cos t, 0\right) dt \\ &= \int_0^{2\pi} \cos^2 t dt - \frac{3}{4} \int_0^{2\pi} \cos^2 t dt = \pi - \frac{3\pi}{4} = \boxed{\frac{\pi}{4}}, \end{aligned}$$

como antes. □

Como ya vimos para el teorema de Green, el teorema de Stokes puede extenderse a superficies *múltiplemente conexas*, cuya frontera está formada por la unión de n curvas cerradas C_1, \dots, C_n . En este caso se tiene

$$\boxed{\int_S \nabla \times \mathbf{F} \cdot d\mathbf{S} = \sum_{i=1}^n \int_{C_i} \mathbf{F} \cdot d\mathbf{r}},$$

donde cada curva C_i ha de estar orientada de forma que al recorrer dicha curva la cara de la superficie S determinada por la normal escogida \mathbf{N} quede siempre a la *izquierda*. Este resultado se demuestra fácilmente, sin más que subdividir S en regiones simplemente conexas a las que se pueda aplicar el teorema de Stokes.

6.3 Teorema de Ostrogradsky–Gauss

El teorema de Ostrogradsky–Gauss (también llamado teorema de Gauss, o de la divergencia) es la generalización tridimensional del teorema de la divergencia en el plano (6.9). Más precisamente:

Teorema de Ostrogradsky–Gauss. Sea $\Omega \subset \mathbb{R}^3$ un conjunto medible Jordan cuya frontera $\partial\Omega$ es una superficie suave cerrada y orientable. Si $\mathbf{F} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ es un campo vectorial de clase C^1 en Ω , entonces

$$\int_{\Omega} \nabla \cdot \mathbf{F} d^3\mathbf{r} = \int_{\partial\Omega} \mathbf{F} \cdot \mathbf{N} dS, \tag{6.14}$$

siendo \mathbf{N} la normal exterior a $\partial\Omega$.

Demostración. Por sencillez, demostraremos este resultado suponiendo que Ω es un sólido de tipo IV. Consideremos, por ejemplo, el término

$$\int_{\Omega} \frac{\partial F_3}{\partial z} dx dy dz$$

que aparece en la integral del miembro izquierdo de (6.14). Al ser Ω de tipo I (por ser de tipo IV), está descrito por ecuaciones de la forma

$$(x, y) \in D, \quad \phi_1(x, y) \leq z \leq \phi_2(x, y),$$

con $\phi_{1,2} : D \rightarrow \mathbb{R}$ funciones de clase C^1 . Por tanto

$$\begin{aligned} \int_{\Omega} \frac{\partial F_3}{\partial z} dx dy dz &= \int_D dx dy \int_{\phi_1(x,y)}^{\phi_2(x,y)} dz \frac{\partial F_3}{\partial z}(x, y, z) \\ &= \int_D [F_3(x, y, \phi_2(x, y)) - F_3(x, y, \phi_1(x, y))] dx dy. \end{aligned} \tag{6.15}$$

Probaremos a continuación que

$$\int_{\Omega} \frac{\partial F_3}{\partial z} dx dy dz = \int_{\partial\Omega} F_3 N_3 dS. \quad (6.16)$$

En efecto, la frontera $\partial\Omega$ de Ω está formada por las dos gráficas

$$\partial\Omega_i = \{(x, y, z) \in \mathbb{R}^3 \mid (x, y) \in D, z = \phi_i(x, y)\}, \quad i = 1, 2.$$

junto con la superficie lateral

$$\partial\Omega_0 = \{(x, y, z) \in \mathbb{R}^3 \mid (x, y) \in \partial D, \phi_1(x, y) \leq z \leq \phi_2(x, y)\}.$$

La integral

$$\int_{\partial\Omega_0} F_3 N_3 dS$$

es nula, ya que la normal a $\partial\Omega_0$ no tiene componente vertical. En efecto, si $(x_0, y_0, z_0) \in \partial\Omega_0$ el segmento vertical $x = x_0, y = y_0, \phi_1(x_0, y_0) \leq z \leq \phi_2(x_0, y_0)$ está contenido en $\partial\Omega_0$, y por tanto el vector \mathbf{k} es tangente a $\partial\Omega_0$ en el punto (x_0, y_0, z_0) . Como el vector normal $\mathbf{N}(x_0, y_0, z_0)$ ha de ser perpendicular a cualquier vector tangente a $\partial\Omega_0$ en (x_0, y_0, z_0) ,

$$\mathbf{N}(x_0, y_0, z_0) \cdot \mathbf{k} \equiv N_3(x_0, y_0, z_0) = 0, \quad \forall (x_0, y_0, z_0) \in \partial\Omega_0.$$

Por otra parte, en las dos gráficas $z = \phi_i(x, y)$ ($i = 1, 2$) el vector normal unitario exterior a Ω está dado por

$$\mathbf{N} = \pm \frac{(-\frac{\partial\phi_i}{\partial x}, -\frac{\partial\phi_i}{\partial y}, 1)}{\sqrt{1 + \left(\frac{\partial\phi_i}{\partial x}\right)^2 + \left(\frac{\partial\phi_i}{\partial y}\right)^2}}$$

donde el signo “+” corresponde a $\partial\Omega_2$ y el signo “-” a $\partial\Omega_1$. Como

$$dS = \sqrt{1 + \left(\frac{\partial\phi_i}{\partial x}\right)^2 + \left(\frac{\partial\phi_i}{\partial y}\right)^2} dx dy$$

en ambas superficies, el producto $N_3 dS$ está dado por

$$N_3 dS = \pm dx dy,$$

donde de nuevo el signo “+” corresponde a $\partial\Omega_2$ y el “-” a $\partial\Omega_1$. Por tanto

$$\begin{aligned} \int_{\partial\Omega} F_3 N_3 dS &= \int_{\partial\Omega_1} F_3 N_3 dS + \int_{\partial\Omega_2} F_3 N_3 dS \\ &= \int_D F_3(x, y, \phi_2(x, y)) dx dy - \int_D F_3(x, y, \phi_1(x, y)) dx dy, \end{aligned}$$

que efectivamente coincide con la ec. (6.15). Procediendo de la misma forma (ya que, por hipótesis, el sólido Ω es también de tipos II y III) se demuestran las igualdades análogas

$$\int_{\Omega} \frac{\partial F_1}{\partial x} dx dy dz = \int_{\partial\Omega} F_1 N_1 dS, \quad \int_{\Omega} \frac{\partial F_2}{\partial y} dx dy dz = \int_{\partial\Omega} F_2 N_2 dS. \quad (6.17)$$

Sumando las ecs. (6.16) y (6.17) se obtiene el teorema de Gauss (6.14). \square

Comentario. El teorema de Gauss tiene numerosas aplicaciones en Física, especialmente en electromagnetismo, mecánica de fluidos y teoría del potencial. Por ejemplo, una de las ecuaciones de Maxwell afirma que

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0},$$

donde $\mathbf{E}(t, x, y, z)$ es el valor del *campo eléctrico* en el punto $(x, y, z) \in \mathbb{R}^3$ en el instante t , $\rho(t, x, y, z)$ es la *densidad de carga eléctrica* y ϵ_0 es una constante (la llamada *permitividad del vacío*), que en el sistema SI está dada por

$$\epsilon_0 = 8,854\,187\,817 \dots \cdot 10^{-12} \text{ F m}^{-1}.$$

Aplicando el teorema de Gauss se obtiene

$$\int_{\partial\Omega} \mathbf{E} \cdot \mathbf{N} \, dS = \frac{1}{\epsilon_0} \int_{\Omega} \rho \, d^3\mathbf{r} = \frac{Q}{\epsilon_0},$$

donde Q es la carga total encerrada por la región Ω . En otras palabras, *el flujo del campo eléctrico hacia el exterior de un volumen $\Omega \subset \mathbb{R}^3$ es igual a la carga eléctrica encerrada en Ω dividida por la constante ϵ_0* . Esta es la llamada *ley de Gauss*.

Ejemplo 6.7. Halleemos el flujo del campo vectorial $\mathbf{F}: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ dado por $\mathbf{F}(x, y, z) = (xz, xy, 2)$ hacia el exterior del cilindro S de ecuaciones $x^2 + y^2 = 4$, $0 \leq z \leq 3$ aplicando el teorema de Gauss. Para ello, hay que tener en cuenta en primer lugar que la superficie S *no* es cerrada, por lo que no se puede aplicar directamente dicho teorema. Considérese, sin embargo, el sólido Ω (interior del cilindro) de ecuaciones

$$x^2 + y^2 \leq 4, \quad 0 \leq z \leq 3,$$

cuya frontera es

$$\partial\Omega = S \cup \partial\Omega_1 \cup \partial\Omega_2,$$

siendo $\partial\Omega_1$ y $\partial\Omega_2$ los discos $x^2 + y^2 \leq 4$, $z = 0$ y $x^2 + y^2 \leq 4$, $z = 3$, respectivamente. Aplicando el teorema de Gauss se obtiene

$$\int_{\Omega} \nabla \cdot \mathbf{F} \, d^3\mathbf{r} = \int_S \mathbf{F} \cdot \mathbf{N} \, dS + \int_{\partial\Omega_1} \mathbf{F} \cdot \mathbf{N} \, dS + \int_{\partial\Omega_2} \mathbf{F} \cdot \mathbf{N} \, dS.$$

La integral del miembro izquierdo se evalúa fácilmente en coordenadas cilíndricas:

$$\begin{aligned} \int_{\Omega} \nabla \cdot \mathbf{F} \, d^3\mathbf{r} &= \int_{\Omega} (x + z) \, dx \, dy \, dz = \int_0^2 d\rho \int_0^3 dz \int_0^{2\pi} d\varphi \, \rho(\rho \cos \varphi + z) \\ &= 2\pi \left(\int_0^2 \rho \, d\rho \right) \left(\int_0^3 z \, dz \right) = 2\pi \cdot 2 \cdot \frac{9}{2} = 18\pi. \end{aligned}$$

Por otra parte,

$$\begin{aligned} \int_{\partial\Omega_1} \mathbf{F} \cdot \mathbf{N} \, dS + \int_{\partial\Omega_2} \mathbf{F} \cdot \mathbf{N} \, dS &= - \int_{\partial\Omega_1} \mathbf{F} \cdot \mathbf{k} \, dS + \int_{\partial\Omega_2} \mathbf{F} \cdot \mathbf{k} \, dS \\ &= 2 \int_{\partial\Omega_2} dS - 2 \int_{\partial\Omega_1} dS = 2a(\partial\Omega_2) - 2a(\partial\Omega_1) = 0. \end{aligned}$$

y por tanto

$$\int_S \mathbf{F} \cdot \mathbf{N} \, dS = \int_{\Omega} \nabla \cdot \mathbf{F} \, d^3\mathbf{r} = \boxed{18\pi}.$$

Ejercicio 38. Calcular el flujo del campo vectorial $\mathbf{F}(x, y, z) = (xy^2, x^2y, (x^2 + y^2)z)$ a través de la región del cono S de ecuaciones $z = \sqrt{x^2 + y^2}$, $z \leq 2$, directamente y utilizando el teorema de Gauss.

Solución. Consideremos el sólido Ω (interior del cono) de ecuaciones

$$\sqrt{x^2 + y^2} \leq z, \quad z \leq 2,$$

cuya frontera $\partial\Omega$ es la unión del cono S con el disco $\partial\Omega_1$ de ecuación $x^2 + y^2 \leq 4$, $z = 2$. Aplicando el teorema de Gauss se obtiene

$$\int_{\Omega} \nabla \cdot \mathbf{F} \, d^3\mathbf{r} = \int_S \mathbf{F} \cdot \mathbf{N} \, dS + \int_{\partial\Omega_1} \mathbf{F} \cdot \mathbf{N} \, dS,$$

siendo \mathbf{N} la normal exterior a Ω . La integral del miembro izquierdo se calcula fácilmente en coordenadas cilíndricas:

$$\int_{\Omega} \nabla \cdot \mathbf{F} \, d^3\mathbf{r} = 2 \int_{\Omega} (x^2 + y^2) \, dx \, dy \, dz = 2 \int_0^2 dz \int_0^z d\rho \int_0^{2\pi} d\varphi \, \rho \cdot \rho^2 = \pi \int_0^2 z^4 \, dz = \frac{32\pi}{5}.$$

Por otra parte, en el disco $\partial\Omega_1$ la normal exterior a Ω es \mathbf{k} y el elemento de área es $dx \, dy$, por lo que

$$\int_{\partial\Omega_1} \mathbf{F} \cdot \mathbf{N} \, dS = \int_{x^2+y^2 \leq 4} 2(x^2 + y^2) \, dx \, dy = 2 \int_0^2 dr \int_0^{2\pi} d\theta \, r \cdot r^2 = 4\pi \int_0^2 r^3 \, dr = 16\pi.$$

Por consiguiente,

$$\int_S \mathbf{F} \cdot \mathbf{N} \, dS = \int_{\Omega} \nabla \cdot \mathbf{F} \, d^3\mathbf{r} - \int_{\partial\Omega_1} \mathbf{F} \cdot \mathbf{N} \, dS = \frac{32\pi}{5} - 16\pi = \boxed{-\frac{48\pi}{5}}.$$

Este resultado también se puede obtener, aunque algo más laboriosamente, de forma directa. En efecto, una parametrización del cono S es la dada por

$$\mathbf{r}(\varphi, z) = z(\cos \varphi, \sin \varphi, 1), \quad (\varphi, z) \in D \equiv [0, 2\pi] \times [0, 2],$$

y por tanto

$$\mathbf{n} \, dS = \frac{\partial \mathbf{r}}{\partial \varphi} \times \frac{\partial \mathbf{r}}{\partial z} \, d\varphi \, dz = z(\cos \varphi, \sin \varphi, -1) \, d\varphi \, dz \equiv (x, y, -z) \, d\varphi \, dz.$$

Al ser la componente vertical del vector $(x, y, -z)$ negativa, el vector unitario normal \mathbf{n} inducido por esta parametrización es el *exterior* a la superficie. Por tanto

$$\begin{aligned} \int_S \mathbf{F} \cdot \mathbf{N} \, dS &= \int_S \mathbf{F} \cdot \mathbf{n} \, dS = \int_D \mathbf{F} \cdot (x, y, -z) \, d\varphi \, dz = \int_D [2x^2y^2 - z^2(x^2 + y^2)] \, d\varphi \, dz \\ &= \int_0^{2\pi} d\varphi \int_0^2 dz (2 \sin^2 \varphi \cos^2 \varphi - 1)z^4 = \frac{32}{5} \int_0^{2\pi} (2 \sin^2 \varphi \cos^2 \varphi - 1) \, d\varphi. \end{aligned}$$

La integral en φ se evalúa fácilmente:

$$\int_0^{2\pi} (2 \sin^2 \varphi \cos^2 \varphi - 1) \, d\varphi = \frac{1}{2} \int_0^{2\pi} \sin^2(2\varphi) \, d\varphi - 2\pi = \frac{\pi}{2} - 2\pi = -\frac{3\pi}{2},$$

lo que de nuevo proporciona

$$\int_S \mathbf{F} \cdot \mathbf{N} \, dS = -\frac{32}{5} \cdot \frac{3\pi}{2} = \boxed{-\frac{48\pi}{5}}.$$

Apéndice

En este apéndice recordaremos brevemente algunos conceptos elementales de la teoría de conjuntos que serán de utilidad a lo largo del curso. En primer lugar, aunque aceptaremos por sencillez la noción de **conjunto** como un *concepto primitivo* (es decir, no definido en términos de otros conceptos más fundamentales), la idea intuitiva de conjunto es la de una *colección de objetos* de cualquier tipo. Lo esencial es que la pertenencia de un objeto a un conjunto determinado sea una noción *bien definida* y no ambigua. Por ejemplo,

$$A = \{1, 2, \sqrt{7}, \text{España}\}, \quad B = \{x \mid x \text{ es ciudadano español}\}$$

son conjuntos (en el primer caso hemos enumerado explícitamente los elementos del conjunto, en el segundo hemos indicado de manera no ambigua cuál es la propiedad que caracteriza a sus elementos), mientras que

$$C = \{x \mid x \text{ es un país desarrollado}\}$$

no lo es (la definición de país desarrollado es ambigua). Usaremos la notación $x \in A$ (respectivamente $x \notin A$) para denotar que x es (respectivamente no es) un elemento del conjunto A . Denotaremos por \emptyset al **conjunto vacío**, definido como aquél conjunto que no posee ningún elemento. Por ejemplo,

$$\{x \in \mathbb{R} \mid x^2 < 0\} = \emptyset.$$

Dados dos conjuntos A y B , diremos que A está **contenido** en B (ó que A es un **subconjunto** de B), y escribiremos $A \subset B$, si todo elemento de A pertenece también a B . Simbólicamente²,

$$A \subset B \iff (x \in A \implies x \in B).$$

La notación $B \supset A$ significa lo mismo que $A \subset B$. Por definición, dos conjuntos A y B son **iguales** (lo que será denotado por $A = B$) si $A \subset B$ y $B \subset A$. Equivalentemente,

$$A = B \iff (x \in A \iff x \in B).$$

Utilizaremos a veces la notación $A \subsetneq B$ para indicar que $A \subset B$ y $A \neq B$.

Una parte fundamental del lenguaje matemático está relacionada con nociones elementales de lógica que repasaremos brevemente a continuación. En primer lugar, dadas dos proposiciones P_1 y P_2 diremos que P_1 **implica** P_2 , y escribiremos $P_1 \implies P_2$, si P_2 es verdadera siempre que P_1 lo sea. Por definición, $P_2 \impliedby P_1$ significa lo mismo que $P_1 \implies P_2$. Análogamente, diremos que las proposiciones P_1 y P_2 son **equivalentes**, y escribiremos $P_1 \iff P_2$, si $P_1 \implies P_2$ y $P_2 \implies P_1$ (es decir, P_2 es verdadera si y solo si lo es P_1). Por ejemplo, dadas las proposiciones $P_1 = "x \geq 5"$, $P_2 = "x^2 > 100"$, $P_3 = "x^2 \geq 25"$ y $P_4 = "e^x \geq e^5"$ entonces se verifica $P_1 \not\implies P_2$, $P_2 \implies P_3$, $P_1 \not\iff P_3$ y $P_1 \iff P_4$. Nótese, en particular, que si P es una proposición siempre falsa y Q es *cualquier* proposición entonces $P \implies Q$. También son muy utilizados en el lenguaje matemático

²Los símbolos " \implies " y " \iff " se definirán más abajo.

los **cuantificadores existencial** \exists (“existe”) y **universal** \forall (“para todo”). Nótese que al negar una proposición en la que figuren estos cuantificadores cada uno de ellos ha de ser reemplazado por el otro. Por ejemplo, la negación de la proposición

$$“\forall \varepsilon > 0, \exists \delta > 0 \text{ t.q. } |x| < \delta \Rightarrow |f(x) - f(0)| < \varepsilon”$$

(definición “ ε - δ ” de límite en 0 de la función $f : \mathbb{R} \rightarrow \mathbb{R}$) es

$$“\exists \varepsilon > 0 \text{ t.q. } \forall \delta > 0, |x| < \delta \not\Rightarrow |f(x) - f(0)| < \varepsilon”,$$

que a su vez es equivalente a

$$“\exists \varepsilon > 0 \text{ t.q. } \forall \delta > 0, \exists x \in \mathbb{R} \text{ t.q. } (|x| < \delta, |f(x) - f(0)| \geq \varepsilon)”.$$

Dado un conjunto A , denotaremos por $\mathcal{P}(A)$ al **conjunto de las partes de A** , definido como el conjunto de los subconjuntos de A :

$$\mathcal{P}(A) = \{B \mid B \subset A\}.$$

La relación de inclusión (\subset) es una **relación de orden parcial** en $\mathcal{P}(A)$, ya que goza de las siguientes propiedades:

- i) $\forall B \in \mathcal{P}(A), B \subset B$ (*propiedad reflexiva*)
- ii) $\forall B, C \in \mathcal{P}(A), B \subset C \text{ y } C \subset B \Rightarrow B = C$ (*propiedad antisimétrica*)
- iii) $\forall B, C, D \in \mathcal{P}(A), B \subset C \text{ y } C \subset D \Rightarrow B \subset D$ (*propiedad transitiva*)

Evidentemente, cualquiera que sea A los conjuntos \emptyset y A son subconjuntos de A (i.e., son elementos del conjunto $\mathcal{P}(A)$). A estos dos subconjuntos en cierto modo triviales se les llama subconjuntos **impropios** de A , mientras que un subconjunto **propio** de A es cualquier subconjunto de A distinto de \emptyset y de A .

Dados dos conjuntos A y B , se definen los conjuntos $A \cap B$ (**intersección** de A con B) y $A \cup B$ (**unión** de A y B) mediante

$$A \cap B = \{x \mid x \in A \text{ y } x \in B\}, \quad A \cup B = \{x \mid x \in A \text{ ó } x \in B\}.$$

Obviamente, $A \cap B \subset A, B \subset A \cup B$. Denotaremos por $A - B$ a la **diferencia** de los conjuntos A y B , definida por

$$A - B = \{x \in A \mid x \notin B\} = A \cap \{x \mid x \notin B\}.$$

La unión y la intersección de conjuntos gozan de las siguientes propiedades elementales:

- i) $A \cap B = B \cap A, (A \cap B) \cap C = A \cap (B \cap C)$
- ii) $A \cap B = A \Leftrightarrow A \subset B$
- iii) $A \cap \emptyset = \emptyset$
- iv) $A \cup B = B \cup A, (A \cup B) \cup C = A \cup (B \cup C)$
- v) $A \cup B = A \Leftrightarrow B \subset A$
- vi) $A \cup \emptyset = A$
- vii) $A \subset B \Rightarrow A \cap C \subset B \cap C, A \subset C, B \subset C \Rightarrow A \cup B \subset C$

Se verifican además las siguientes **propiedades distributivas**:

$$\text{i) } A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

$$\text{ii) } A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

Demostremos, por ejemplo, la primera de estas igualdades. En virtud de la definición de igualdad de dos conjuntos, basta probar que el miembro izquierdo está contenido en el miembro derecho y viceversa. En primer lugar,

$$B, C \subset B \cup C \Rightarrow A \cap B \subset A \cap (B \cup C), A \cap C \subset A \cap (B \cup C),$$

de donde se deduce que el miembro derecho está contenido en el miembro izquierdo. Recíprocamente, si $x \in A \cap (B \cup C)$ entonces $x \in A$ y $x \in B \cup C$, es decir $(x \in A \text{ y } x \in B)$ ó $(x \in A \text{ y } x \in C)$, de donde se sigue que $x \in (A \cap B) \cup (A \cap C)$. Esto prueba que el miembro izquierdo está contenido en el miembro derecho.

En general, si I es un conjunto de índices arbitrario (finito ó infinito) tal que a todo $i \in I$ le corresponde un conjunto A_i , se definen la unión y la intersección de la familia de conjuntos $\{A_i \mid i \in I\}$ mediante las fórmulas

$$\bigcup_{i \in I} A_i = \{x \mid \exists j \in I \text{ tal que } x \in A_j\} \quad \bigcap_{i \in I} A_i = \{x \mid x \in A_j, \forall j \in I\}.$$

El **producto cartesiano** de dos conjuntos A y B es el conjunto

$$A \times B = \{(x, y) \mid x \in A, y \in B\}.$$

Nótese que en esta definición (x, y) es un **par ordenado**. En otras palabras,

$$(x_1, y_1) = (x_2, y_2) \Leftrightarrow x_1 = x_2 \text{ e } y_1 = y_2.$$

De esto se deduce que $A \times B$ es en general distinto de $B \times A$. (*Ejercicio*: ¿cuándo es $A \times B = B \times A$?) Más generalmente, dados n conjuntos A_1, \dots, A_n se define su producto cartesiano mediante

$$A_1 \times A_2 \times \dots \times A_n = \{(x_1, x_2, \dots, x_n) \mid x_i \in A_i, \forall i = 1, 2, \dots, n\}.$$

En particular, si $A_1 = A_2 = \dots = A_n = A$ escribiremos

$$\underbrace{A \times \dots \times A}_{n \text{ veces}} = A^n.$$

En otras palabras,

$$A^n = \{(x_1, x_2, \dots, x_n) \mid x_i \in A, \forall i = 1, 2, \dots, n\}.$$

Ejercicio. ¿Son iguales los conjuntos $\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \emptyset\}, \{\emptyset, \{\emptyset\}\}$?

