

# Lecciones de Análisis Matemático II

Gabriel Vera

15 de julio de 2011

# Índice general

Prólogo	I
<b>1. Preliminares sobre funciones de varias variables</b>	<b>1</b>
1.1. Introducción . . . . .	1
1.2. Funciones de una variable . . . . .	2
1.3. Funciones de varias variables . . . . .	2
1.4. Coordenadas curvilíneas . . . . .	5
1.5. Ejercicios propuestos . . . . .	7
<b>2. Espacios métricos y espacios normados</b>	<b>9</b>
2.1. El espacio $\mathbb{R}^n$ . Espacios normados . . . . .	10
2.2. Sucesiones y conjuntos compactos . . . . .	17
2.3. Espacios completos . . . . .	21
2.4. Normas en $C[a, b]$ . . . . .	23
2.5. Ejercicios resueltos . . . . .	24
2.6. Ejercicios propuestos . . . . .	27
<b>3. Límites y continuidad</b>	<b>32</b>
3.1. Definiciones y resultados básicos . . . . .	33
3.2. Reglas para obtener el límite y la continuidad . . . . .	36
3.3. Funciones continuas en conjuntos compactos . . . . .	41
3.4. Espacios normados de dimensión finita . . . . .	42
3.5. Continuidad uniforme . . . . .	45
3.6. Convergencia uniforme . . . . .	50
3.7. Ejercicios resueltos . . . . .	53
3.8. Ejercicios propuestos . . . . .	57
<b>4. Funciones vectoriales de una variable</b>	<b>62</b>
4.1. Derivada de una función vectorial . . . . .	64
4.2. Desarrollo de Taylor . . . . .	70
4.3. Integral de una función vectorial . . . . .	72
4.4. Caminos rectificables . . . . .	75
4.5. Integral respecto al arco . . . . .	82
4.6. Ejercicios resueltos . . . . .	85
4.7. Ejercicios propuestos . . . . .	87

<b>5. Funciones diferenciables</b>	<b>91</b>
5.1. Derivada según un vector . . . . .	92
5.2. Aplicaciones diferenciables . . . . .	102
5.3. Las reglas del cálculo diferencial . . . . .	109
5.4. Gradiente . . . . .	116
5.5. Espacio tangente . . . . .	118
5.6. Ejercicios resueltos . . . . .	124
5.7. Ejercicios propuestos . . . . .	132
<b>6. Funciones dos veces diferenciables</b>	<b>139</b>
6.1. Funciones dos veces diferenciables. . . . .	140
6.2. Extremos relativos . . . . .	146
6.3. Funciones convexas . . . . .	150
6.4. Ejercicios resueltos . . . . .	154
6.5. Ejercicios propuestos . . . . .	160
<b>7. Desarrollo de Taylor</b>	<b>163</b>
7.1. Funciones diferenciables $m$ veces . . . . .	164
7.2. Desarrollo de Taylor . . . . .	168
7.3. Serie de Taylor de una función de clase $C^\infty$ . . . . .	173
7.4. Fórmula integral para el resto . . . . .	175
7.5. Ejercicios resueltos . . . . .	176
7.6. Ejercicios propuestos . . . . .	180
<b>8. Función inversa y función implícita</b>	<b>181</b>
8.1. Aplicaciones con inversa local . . . . .	182
8.2. Funciones implícitas . . . . .	189
8.3. Cálculo con funciones implícitas e inversas . . . . .	193
8.4. Cambio de variable en el cálculo diferencial . . . . .	196
8.5. Ejercicios resueltos . . . . .	199
8.6. Ejercicios propuestos . . . . .	204
<b>9. Extremos condicionados</b>	<b>208</b>
9.1. Subvariedades diferenciables . . . . .	209
9.2. Extremos condicionados . . . . .	214
9.3. Ejercicios resueltos . . . . .	221
9.4. Ejercicios propuestos . . . . .	234
<b>10. Integral de Riemann</b>	<b>238</b>
10.1. Funciones integrables Riemann . . . . .	238
10.2. Conjuntos medibles Jordan . . . . .	245
10.3. Caracterización de las funciones integrables . . . . .	256
10.4. Ejercicios resueltos . . . . .	261
10.5. Ejercicios propuestos . . . . .	265

<b>11. Técnicas de cálculo integral</b>	<b>267</b>
11.1. Integración iterada . . . . .	268
11.2. Utilización del cambio de variable . . . . .	274
11.3. Ejercicios resueltos . . . . .	282
11.4. Ejercicios propuestos . . . . .	292
<b>12. Integrales impropias. Integrales dependientes de un parámetro</b>	<b>298</b>
12.1. Integrales impropias . . . . .	298
12.2. Paso al límite bajo la integral . . . . .	301
12.3. Ejercicios resueltos . . . . .	307
12.4. Ejercicios propuestos . . . . .	312
<b>13. Integral curvilínea</b>	<b>314</b>
13.1. Formas diferenciales e integral curvilínea . . . . .	315
13.2. Formas diferenciales en el plano . . . . .	324
13.3. El teorema de Green . . . . .	332
13.4. Ejercicios resueltos . . . . .	340
13.5. Ejercicios propuestos . . . . .	344
<b>14. Integrales de superficie</b>	<b>348</b>
14.1. Preliminares geométricos . . . . .	349
14.2. Área de una superficie . . . . .	351
14.3. Integral respecto al elemento de área . . . . .	358
14.4. Flujo de un campo de vectores . . . . .	360
14.5. Integración sobre variedades paramétricas $k$ -dimensionales . . . . .	365
14.6. Ejercicios resueltos . . . . .	369
14.7. Ejercicios propuestos . . . . .	373
<b>A. Sucesiones y series de funciones</b>	<b>374</b>
A.1. Convergencia puntual y uniforme . . . . .	375
A.2. Continuidad, derivabilidad e integrabilidad del límite . . . . .	377
A.3. Series de funciones . . . . .	381
A.4. Ejercicios resueltos . . . . .	385
A.5. Ejercicios propuestos . . . . .	392
<b>B. Complementos al capítulo 2</b>	<b>397</b>
B.1. La recta real . . . . .	397
B.2. Completitud y compacidad . . . . .	399
B.3. Espacios de sucesiones . . . . .	400
B.4. Formas lineales y producto escalar . . . . .	402
B.5. Espacios complejos con producto interior . . . . .	403
<b>C. Complementos al capítulo 3</b>	<b>406</b>
C.1. Intercambio de límites . . . . .	406
C.2. Convergencia uniforme de series de funciones vectoriales . . . . .	411

<b>D. Integración de funciones vectoriales</b>	<b>414</b>
D.1. Integración de funciones regladas . . . . .	414
D.2. Definición general de la integral de Riemann . . . . .	417
<b>E. Complementos sobre diferenciabilidad</b>	<b>422</b>
E.1. Caracterización de las funciones de clase $C^1$ . . . . .	422
E.2. La definición general de diferencial segunda . . . . .	423
E.3. Teorema de Schwarz sobre la igualdad de las derivadas mixtas . . . . .	425
<b>F. Funciones convexas</b>	<b>427</b>
F.1. Caracterización de las funciones convexas de una variable . . . . .	427
F.2. Continuidad de las funciones convexas de varias variables . . . . .	433
<b>G. Funciones analíticas</b>	<b>437</b>
G.1. Funciones analíticas . . . . .	437
<b>H. Dependencia funcional. Subvariedades diferenciables</b>	<b>443</b>
H.1. Dependencia e independencia funcional . . . . .	443
H.2. Parametrizaciones regulares . . . . .	446
H.3. Subvariedades orientables . . . . .	449
<b>I. Extremos y formas cuadráticas</b>	<b>452</b>
I.1. Extremos y formas cuadráticas . . . . .	452
<b>J. Cambio de variable en la integral de Riemann</b>	<b>457</b>
J.1. Preliminares . . . . .	457
J.2. La demostración del teorema de cambio de variable . . . . .	463
<b>K. Formas diferenciales</b>	<b>472</b>
K.1. Producto mixto y producto vectorial . . . . .	472
K.2. Formas multilineales alternadas . . . . .	478
K.3. Formas diferenciales . . . . .	484

# Prólogo

El material que se ofrece en este texto es fruto de una larga experiencia docente enseñando esta materia en la Facultad de Matemáticas de la Universidad de Murcia. Contiene, además de los contenidos básicos de la asignatura, otro material complementario que en alguna ocasión ha sido expuesto o entregado por escrito a los alumnos. Por esta razón, el temario desarrollado en estas *Lecciones* está adaptado y cubre lo que habitualmente se enseña en la Facultad de Matemáticas de esta Universidad, aunque excede lo que se puede enseñar durante un curso académico. Para solventar esta dificultad aquellos temas que se pueden considerar de carácter complementario o más avanzado, han sido incluidos en apéndices independientes al final del texto. Allí el estudiante interesado podrá ampliar y estudiar con mayor profundidad algunos de los temas propios de la asignatura.

A lo largo del texto se exponen con detalle ejemplos que ilustran y aclaran los conceptos teóricos nuevos. Cada capítulo termina con un repertorio de problemas resueltos donde se analizan comentan y enseñan diferentes estrategias para abordarlos, seguido de un amplio repertorio de problemas propuestos.

Por su enfoque, por el amplio repertorio de problemas resueltos, y por los temas complementarios incluidos, estas *Lecciones* puedan interesar no sólo a los estudiantes de Matemáticas que quieran profundizar en los asuntos propios del Análisis Matemático II, sino a profesores jóvenes que comiencen a enseñar de esta materia. Esperamos que también sean útiles a estudiantes de otras titulaciones, de carácter científico, que estudien, en universidades de habla hispana, el cálculo diferencial e integral para funciones de varias variables.

Los conocimientos previos asumidos al redactar estas *Lecciones* han sido:

- Cálculo Diferencial y Cálculo Integral para funciones reales de una variable real.
- Nociones básicas de Álgebra Lineal (aplicaciones lineales, matrices y determinantes) y de Geometría Euclídea.
- El vocabulario y la terminología usual de la Teoría de Conjuntos y de la Topología en el ámbito del espacio euclídeo  $\mathbb{R}^n$  o de los espacios métricos: Conjuntos abiertos, cerrados, compactos. Frontera, interior y adherencia de un conjunto (esencialmente, el capítulo 2 y la primera parte del capítulo 3 del libro de Apostol [2]).

NOTA: La versión completa en formato `.pdf` de estas *Lecciones* permite navegar a lo largo de todo el texto, y acudir directamente a las referencias y citas bibliográficas.

# Capítulo 1

## Preliminares sobre funciones de varias variables

*Diversas formas de describir analíticamente curvas y superficies. Curvas y superficies de nivel. Introducción a los sistemas de coordenadas curvilíneas.*

En este capítulo se hace una breve introducción a la geometría analítica tridimensional con el fin de dar interpretaciones geométricas y físicas de las funciones de varias variables. Se consideran las diversas formas (explícita, implícita y parametrizada) de describir curvas y superficies, nociones que de momento se manejan en un sentido intuitivo, y también se introducen los sistemas de coordenadas curvilíneas usuales (polares en el plano; cilíndricas y esféricas en el espacio).

Esta introducción permitirá presentar desde un punto de vista geométrico algunos de los problemas que se abordan con el cálculo diferencial y el cálculo integral de funciones de varias variables: Existencia de planos tangentes a superficies, problemas de optimización (con y sin restricciones), existencia de inversas locales, definición implícita de funciones, cálculo de áreas, volúmenes y longitudes de curvas.

Para este capítulo introductorio se recomienda el manejo del programa DpGraph especialmente diseñado para ilustrar los diversos aspectos teóricos y prácticos de la materia: Visualización de curvas planas y alabeadas, curvas y superficies de nivel, recintos de integración, extremos relativos o absolutos de funciones sometidas a ligaduras, uso de parámetros en fórmulas, etc.

### 1.1. Introducción

El objetivo del curso es el estudio de las funciones vectoriales de varias variables reales, es decir, funciones  $\mathbf{f} : \Omega \rightarrow \mathbb{R}^m$  definidas en un abierto  $\Omega \subset \mathbb{R}^n$ . En lo que sigue  $\mathbf{x}$  denotará siempre un elemento genérico de  $\mathbb{R}^n$  de componentes  $(x_1, x_2, \dots, x_n)$ .

En bastantes cuestiones el hecho de que sea  $m > 1$  no involucra dificultades realmente significativas pues frecuentemente el estudio de la función

$$\mathbf{f}(\mathbf{x}) = (f_1(x_1, x_2, \dots, x_n), f_2(x_1, x_2, \dots, x_n), \dots, f_m(x_1, x_2, \dots, x_n))$$

se reduce al de sus componentes  $f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_m(\mathbf{x})$ . Si  $n = 2$ , (resp.  $n = 3$ ) en lugar de  $\mathbf{f}(x_1, x_2)$  (resp.  $\mathbf{f}(x_1, x_2, x_3)$ ) se suele escribir  $\mathbf{f}(x, y)$  (resp.  $\mathbf{f}(x, y, z)$ ).

Con el fin de motivar el estudio de las funciones vectoriales de varias variables conviene empezar comentando los diferentes tipos de representación geométrica que admiten estas funciones, según los valores de  $n$  y  $m$ , lo que permitirá interpretaciones geométricas ilustrativas de los conceptos que se vayan introduciendo. Con este fin conviene comenzar utilizando las nociones de curva y superficie en un sentido intuitivo, mostrando ejemplos concretos de estos objetos geométricos que más adelante se definirán de manera precisa. Uno de los objetivos de este curso es el de dar definiciones matemáticamente rigurosas de estas nociones. Mientras tanto utilizaremos los términos “curva” y “superficie” entre comillas para indicar que estamos considerando estos conceptos desde un punto de vista intuitivo completamente informal. Comenzamos con el caso  $n = 1$  donde las interpretaciones geométricas son de distinta naturaleza que en el caso  $n \geq 2$ .

## 1.2. Funciones de una variable

En el curso de Análisis I, que se refiere al caso  $n = 1, m = 1$ , al efectuar la representación gráfica de una función aparecen “curvas” planas de un tipo muy especial pues cada recta paralela al eje  $OY$  sólo las puede cortar a lo más en un punto. Estas curvas, que vienen dadas como la gráfica de una función real de una variable real  $G(f) = \{(x, y) \in \mathbb{R}^2 : y = f(x)\}$  diremos que admiten la representación *explícita*  $y = f(x)$ .

Más general es el caso de las “curvas” planas en forma *paramétrica* que corresponden al caso  $n = 1, m = 2$ . En este caso hay una interpretación geométrica y física natural: Si  $\mathbf{f}(t) = (f_1(t), f_2(t))$  se dice que  $x = f_1(t), y = f_2(t)$  son las ecuaciones paramétricas de una “curva” en el plano. Ahora la “curva” se puede interpretar físicamente como la trayectoria de una partícula que se mueve de modo que en el instante  $t$  se encuentra en la posición  $\mathbf{f}(t) = (f_1(t), f_2(t))$ . Esta clase de “curvas” incluye a las anteriores ya que toda “curva” dada en la forma explícita  $y = f(x)$  admite la parametrización canónica  $\mathbf{f}(t) = (t, f(t))$ . Un ejemplo muy sencillo lo proporciona la parametrización canónica de la circunferencia de centro  $(0, 0)$  y radio 1:  $\mathbf{f}(t) = (\cos t, \sin t), 0 \leq t \leq 2\pi$ .

Análogamente, el caso  $n = 1, m = 3$ , se considerará a la hora de estudiar “curvas” parametrizadas en el espacio euclídeo ordinario  $\mathbf{f}(t) = (f_1(t), f_2(t), f_3(t))$ . Interpretando que el parámetro  $t$  es el tiempo, con una función de este tipo se describe la trayectoria de una partícula que se mueve en el espacio.

## 1.3. Funciones de varias variables

Para estudiar las funciones de varias variables se utilizan con frecuencia los recursos del cálculo con funciones de una variable, considerando las funciones parciales que se obtienen fijando todas las variables menos una. Para estudiar una función  $f$



de  $n$  variables cerca de un punto  $\mathbf{a} = (a_1, a_2, \dots, a_n) \in \Omega$  es natural considerar las funciones parciales determinadas por ese punto, es decir, las funciones de variable real

$$x_1 \rightarrow f(x_1, a_2, \dots, a_n), \quad x_2 \rightarrow f(a_1, x_2, \dots, a_n) \quad \dots, \quad x_n \rightarrow f(a_1, a_2, \dots, a_{n-1}, x_n)$$

donde la primera función está definida en  $\Omega^1 = \{x_1 : (x_1, a_2, \dots, a_n) \in \Omega\}$  la segunda en  $\Omega^2 = \{x_2 : (a_1, x_2, \dots, a_n) \in \Omega\}$  etc.

**Funciones de dos variables** Comencemos con el caso  $n = 2$ ,  $m = 1$ . Para una función  $f : \Omega \rightarrow \mathbb{R}$  de dos variables reales  $(x, y)$  definida en un recinto  $\Omega \subset \mathbb{R}^2$ , su gráfica  $G(f) = \{(x, y, z) : (x, y) \in \Omega, z = f(x, y)\}$  suele ser una “superficie” con la que se pueden dar interpretaciones geométricas de los conceptos básicos del cálculo diferencial e integral análogas a las del caso  $n = 1$ ,  $m = 1$ .

La noción de función diferenciable en un punto  $(a, b) \in \Omega$  significará que la “superficie”  $G(f)$  tiene plano tangente en  $\mathbf{p} = (a, b, f(a, b))$ . Por otra parte, la noción de integral para funciones de dos variables permitirá definir y calcular volúmenes de recintos tridimensionales del tipo  $R(f, \Omega) = \{(x, y, z) : (x, y) \in \Omega, 0 \leq z \leq f(x, y)\}$ .

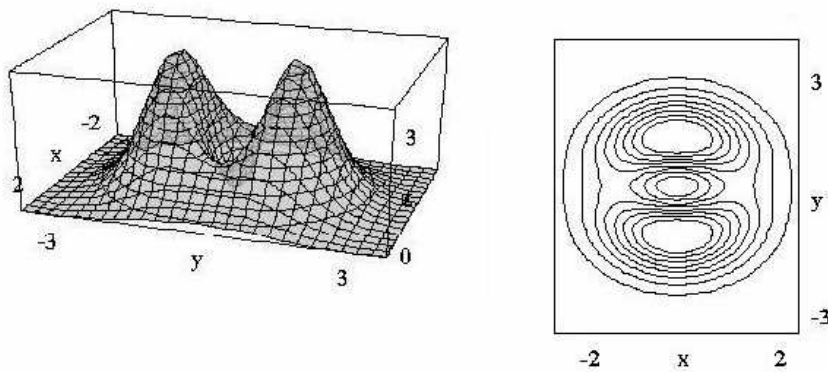
Una “superficie” de este tipo, que es la gráfica  $G(f)$  de una función real de dos variables reales, diremos que admite la representación *explícita*  $z = f(x, y)$ . Las “superficies” que admiten una representación explícita son muy particulares pues cada recta paralela al eje  $OZ$  sólo las corta, a lo más, en un punto.

Las funciones reales de dos variables también intervienen al considerar “curvas” planas definidas mediante una ecuación implícita de la forma  $g(x, y) = c$ , como ocurre con la circunferencia  $x^2 + y^2 = 1$ . Toda “curva” dada en forma explícita  $y = f(x)$  se puede representar en forma implícita  $g(x, y) = 0$ , usando la función  $g(x, y) = f(x) - y$ . Con la circunferencia se pone de manifiesto que hay “curvas” planas que admiten una ecuación implícita pero no admiten una representación explícita global. (El teorema de la función implícita servirá para estudiar cuando una curva dada en forma implícita admite representaciones explícitas locales).

En el caso de las funciones reales de dos variables reales, aunque es posible visualizar la gráfica de la función, también suele ser útil acudir a la técnica de las “curvas” de nivel que proporciona una representación gráfica bidimensional de la “superficie” tridimensional  $G(f)$ . Proyectando sobre el plano  $XY$  la intersección de la gráfica  $G(f)$  con los planos  $z = c$  se obtienen los conjuntos de nivel

$$N_c = \{(x, y) \in \Omega : f(x, y) = c\}$$

Estos conjuntos, si no son vacíos, son “curvas” definidas implícitamente. Dibujando estas “curvas” para distintos valores de  $c$  (variando en progresión aritmética) se obtiene un mapa topográfico de la “superficie”  $G(f)$ .



Gráfica y curvas de nivel de  $z = (x^2 + 3y^2)e^{1-x^2-y^2}$

Figura 1

Para motivar el estudio de funciones de dos variables con valores en  $\mathbb{R}^3$  (caso  $n = 2, m = 3$ ) se puede considerar la representación paramétrica de una “superficie”. Si  $\mathbf{f}(u, v) = (f_1(u, v), f_2(u, v), f_3(u, v))$ , se dice que

$$x = f_1(u, v), \quad y = f_2(u, v), \quad z = f_3(u, v)$$

son las ecuaciones paramétricas de una “superficie” en el espacio ordinario. Cuando  $(u, v)$  recorre el dominio  $\Omega \subset \mathbb{R}^2$  la imagen  $\mathbf{f}(u, v)$  recorre una “superficie”  $S = \mathbf{f}(\Omega)$  que se puede visualizar trazando en el espacio  $(x, y, z)$  las “curvas” imágenes de las rectas  $u = cte$ ,  $v = cte$ . Obsérvese que toda “superficie” dada en forma explícita  $z = f(x, y)$  admite la parametrización canónica  $\mathbf{f}(u, v) = (u, v, f(u, v))$ .

Un ejemplo estandar lo proporciona la parametrización usual de la esfera de centro  $(0, 0, 0)$  y radio  $R$ , usando los parámetros habituales, latitud  $\varphi$ , y longitud  $\theta$ :

$$x = R \cos \varphi \cos \theta, \quad y = R \cos \varphi \sin \theta, \quad z = R \sin \varphi$$

Según sea el dominio  $\Omega$  donde varían los parámetros se obtendrá como imagen un trozo de esfera, o toda la esfera. Así por ejemplo, el trozo de esfera que queda en  $\{(x, y, z) : y > 0, z > 0\}$  se parametriza con  $\Omega = \{(\varphi, \theta) : 0 < \varphi < \pi/2, 0 < \theta < \pi\}$ .

**Funciones de tres variables** Una motivación geométrica para el estudio de las funciones reales de tres variables reales es el de las “superficies” definidas mediante una ecuación de la forma  $g(x, y, z) = c$ , como es el caso de la esfera  $x^2 + y^2 + z^2 = 1$ . Estas “superficies” se dice que admiten una representación *implícita* mediante la ecuación  $g(x, y, z) = c$ . Es claro que cualquier “superficie” dada en forma explícita  $z = f(x, y)$  se puede representar implícitamente usando la ecuación  $g(x, y, z) = 0$  donde  $g(x, y, z) = f(x, y) - z$ . Con la esfera se pone de manifiesto que la clase de las “superficies” que admiten una ecuación implícita es estrictamente más amplia que la clase de las que admiten una representación explícita. (El teorema de la función implícita permitirá estudiar cuando una “superficie” dada en forma implícita admite representaciones explícitas locales).

La gráfica de una función real de tres variables reales (caso  $n = 3, m = 1$ ) es un subconjunto de  $\mathbb{R}^4$  y es imposible visualizarla. Una alternativa para visualizar

geométricamente la función y dar interpretaciones físicas de su comportamiento es considerar sus “superficies” de nivel

$$N_c = \{(x, y, z) \in \Omega : f(x, y, z) = c\}$$

Estas “superficies”, dadas en forma implícita, se pueden visualizar en el espacio ordinario usando un programa de ordenador como DpGraph. Cuando se interpreta que la función  $t = f(x, y, z)$  proporciona la temperatura  $t$  del punto  $(x, y, z) \in \Omega$ , entonces las “superficies” de nivel se llaman isotermas y su distribución en el espacio permite apreciar como varía la temperatura en el recinto  $\Omega \subset \mathbb{R}^3$ .

Una “curva” en el espacio tridimensional  $\mathbb{R}^3$  puede venir dada como intersección de dos “superficies” expresadas en forma implícita

$$g_1(x, y, z) = 0, \quad g_2(x, y, z) = 0$$

En el estudio de una curva de esta clase interviene una función de tres variables reales con valores en  $\mathbb{R}^2$ ,  $\mathbf{g}(x, y, z) = (g_1(x, y, z), g_2(x, y, z))$ . El teorema de la función implícita servirá para decidir cuando una “curva” de este tipo admite una representación paramétrica local.

## 1.4. Coordenadas curvilíneas

En el caso  $n = 2$  y  $m = 2$ , una función  $\mathbf{f}(u, v) = (f_1(u, v), f_2(u, v))$  se puede interpretar como una transformación entre dos planos: El plano  $(u, v)$  donde varían las variables independientes y el plano  $(x, y)$  donde toman valores las variables dependientes  $x = f_1(u, v)$ ,  $y = f_2(u, v)$ . La transformación se puede visualizar considerando las curvas, en el plano  $(x, y)$ , imágenes de las rectas  $u = cte$ ,  $v = cte$ . Estas transformaciones intervienen en los problemas de cambio de variable en cálculo diferencial e integral. En este asunto, un problema expresado en términos de las variables originales  $(x, y)$  mediante la sustitución  $x = f_1(u, v)$ ,  $y = f_2(u, v)$  se transforma en otro problema en términos de las nuevas variables  $(u, v)$ .

Un ejemplo notable lo proporciona el cambio de variable a *coordenadas polares*, asociado a la transformación

$$\mathbf{g}(r, \theta) = (r \cos \theta, r \sin \theta)$$

En este caso, si  $(x, y) = (r \cos \theta, r \sin \theta)$ , con  $r \geq 0$ , se dice que  $(r, \theta)$  son las coordenadas polares del punto  $(x, y)$ . Las coordenadas polares de un punto  $(x, y) \neq (0, 0)$  son únicas cuando se exige que  $\theta$  varíe en un intervalo  $(\alpha, \beta)$  con  $\beta - \alpha \leq 2\pi$ .

En el caso  $n = 3$ ,  $m = 3$ , una función  $\mathbf{f}(t, u, v) = (f_1(t, u, v), f_2(t, u, v), f_3(t, u, v))$  se puede interpretar como una transformación en el espacio  $\mathbb{R}^3$ . En una copia del espacio varían las variables independientes  $(t, u, v)$  y en la otra copia toman valores

las variables dependientes  $x = f_1(t, u, v)$ ,  $y = f_2(t, u, v)$ ,  $z = f_3(t, u, v)$ . Se puede visualizar considerando las superficies paramétricas que se obtienen manteniendo constante uno de los parámetros  $(t, u, v)$  y haciendo que varíen los otros dos. Igual que en el caso  $n = 2, m = 2$  estas transformaciones intervendrán en los problemas de cambio de variable, donde un problema expresado en términos de las variables originales  $(x, y, z)$  mediante la sustitución  $x = f_1(t, u, v)$ ,  $y = f_2(t, u, v)$ ,  $z = f_3(t, u, v)$  se transforma en otro problema en términos de las nuevas variables  $(t, u, v)$ .

Un ejemplo importante es el cambio de variable a *coordenadas esféricas*. Estas coordenadas son las asociadas a la transformación

$$\mathbf{g}(\rho, \varphi, \theta) = (\rho \cos \varphi \cos \theta, \rho \cos \varphi \sin \theta, \rho \sin \varphi)$$

En este caso, si  $(x, y, z) = (\rho \cos \varphi \cos \theta, \rho \cos \varphi \sin \theta, \rho \sin \varphi)$  con  $\rho \geq 0$ , se dice que  $(\rho, \varphi, \theta)$  son las coordenadas esféricas del punto  $(x, y, z)$ , a las variables  $\varphi, \theta$  se les llama latitud y longitud por su interpretación obvia como coordenadas geográficas). Las coordenadas esféricas de un punto  $(x, y, z)$ , con  $(x, y) \neq (0, 0)$ , son únicas cuando se exige que  $\theta$  varíe en un intervalo  $(\theta_0, \theta_1)$  con  $\theta_1 - \theta_0 \leq 2\pi$  y que  $\varphi$  varíe en un intervalo  $(\varphi_0, \varphi_1) \subset (-\pi/2, \pi/2)$ .

Otro ejemplo notable lo proporcionan las *coordenadas cilíndricas*, asociadas a la transformación

$$\mathbf{g}(r, \theta, t) = (r \cos \theta, r \sin \theta, t)$$

Si  $(x, y, z) = (r \cos \theta, r \sin \theta, t)$  con  $r \geq 0$ , se dice que  $(r, \theta, t)$  son las coordenadas cilíndricas del punto  $(x, y, z)$ . Si  $(x, y) \neq (0, 0)$  las coordenadas cilíndricas de  $(x, y, z)$  son únicas cuando se exige que  $\theta$  varíe en un intervalo  $(\theta_0, \theta_1)$  con  $\theta_1 - \theta_0 \leq 2\pi$ .

Cierto tipo de conjuntos  $M \subset \mathbb{R}^3$  se describen fácilmente usando coordenadas esféricas o cilíndricas. Como esta descripción se utilizará con frecuencia al efectuar cambios de variable en integrales triples, conviene adquirir destreza en el problema geométrico de describir subconjuntos de  $\mathbb{R}^3$  usando este tipo de coordenadas.

## 1.5. Ejercicios propuestos

◇ **1.5.1** Sea  $\mathbf{f}(t) = (f_1(t), f_2(t), \dots, f_n(t))$ ,  $t \in [a, b]$ , una trayectoria de clase  $C^1$ , que describe la curva  $C = \mathbf{f}([a, b])$  y cumple la condición

$$(f_1'(t), f_1'(t), \dots, f_n'(t)) \neq (0, 0, \dots, 0) \text{ para cada } t \in [a, b]$$

Demuestre que la trayectoria sólo pasa un número finito de veces por cada punto de la curva  $C$ .

◇ **1.5.2** Se sabe que la longitud de una trayectoria de clase  $C^1$ ,  $\mathbf{f}(t) = (f_1(t), f_2(t), \dots, f_n(t))$ ,  $t \in [a, b]$ , se calcula mediante la integral:

$$L = \int_a^b \sqrt{x_1'(t)^2 + x_2'(t)^2 + \dots + x_n'(t)^2} dt$$

Obtenga la interpretación física del número  $\sqrt{x_1'(t)^2 + x_2'(t)^2 + \dots + x_n'(t)^2}$ .

◇ **1.5.3** La cicloide es la curva plana que describe un punto de un aro circular que rueda sin deslizarse sobre una recta. Escriba sus ecuaciones paramétricas considerando una circunferencia de radio  $R$  que rueda sobre el eje de abscisas de modo que en instante inicial  $t = 0$  el punto toca el suelo en  $(0, 0)$ . Calcule la longitud del arco de cicloide entre dos pasos consecutivos por el suelo.

◇ **1.5.4** Escriba las ecuaciones paramétricas de la curva del espacio tridimensional que sigue el pasamanos de una escalera de caracol de radio  $R$ . Se sabe que la escalera sube tres pisos de altura  $h$ , y que emplea dos vueltas completas para subir desde un piso al siguiente. Utilice la fórmula dada en el problema 1.5.2 para calcular la longitud del pasamanos. (La curva que sigue el pasamanos, que tiene la forma de un muelle, se llama hélice.)

◇ **1.5.5** Utilice DpGraph para visualizar la superficie de ecuaciones paramétricas

$$x = r \cos t, \quad y = r \sin t, \quad z = t, \quad 0 < r < 2 \quad 0 < t < 4\pi$$

◇ **1.5.6** Sea considera la transformación  $\mathbf{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  definida por

$$\mathbf{f}(s, t) = (s^2 + t^2, 2st)$$

Compruebe que  $\mathbf{f}(\mathbb{R}^2) = \{(x, y) : 0 < x, |y| < x\}$  y que la restricción  $\mathbf{f}|_U$  al abierto  $U = \{(x, y) : |y| < x\}$  es inyectiva. Calcule  $V = \mathbf{f}(U)$  y obtenga las ecuaciones de la inversa  $\mathbf{g} = \mathbf{f}^{-1} : V \rightarrow U$ . Visualize la transformación con DpGraph. (Indic:  $\mathbf{f}(r \cos \varphi, r \sin \varphi) = (r^2, r^2 \sin 2\varphi)$ ).

◇ **1.5.7** En  $U = \{(s, t) : 0 < t < s\}$  se define  $\mathbf{f}(s, t) = (\log st, 1/(s^2 + t^2))$ . Calcule  $V = \mathbf{f}(U)$  y compruebe que  $f$  establece una biyección entre  $U$  y  $V$ . Visualize la transformación con DpGraph.

◇ **1.5.8** Sea  $U = \{(s, t) : s > 0\}$  y  $A = \{(s, t) : 0 < s, 0 < t < 2\pi\}$ . Se considera la transformación  $\mathbf{f} : U \rightarrow \mathbb{R}^2$  definida por  $\mathbf{f}(s, t) = (\operatorname{ch} s \cos t, \operatorname{sh} s \operatorname{sen} t)$ . Obtenga  $\mathbf{f}(U)$ ,  $\mathbf{f}(A)$ . (Indic: Determine las imágenes de las rectas  $I_\alpha = \{(\alpha, t) : t \in \mathbb{R}\}$ ,  $\alpha > 0$ , o las imágenes de las semirrectas  $L_\beta = \{(s, \beta) : s > 0\}$ ).

◇ **1.5.9** Demuestre que cada una de las siguientes aplicaciones establece una biyección entre su dominio y su imagen. Obtenga la imagen en cada caso.

a)  $\mathbf{f} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $\mathbf{f}(x, y, z) = (e^{2y} + e^{2z}, e^{2x} - e^{2z}, x - y)$ .

b)  $\mathbf{g} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $\mathbf{g}(x, y) = (e^x + e^y, e^x - e^y)$ .

◇ **1.5.10** La temperatura de un punto  $(x, y, z)$  de la esfera  $x^2 + y^2 + z^2 = 3$  viene dada por la función  $t = x^2 + y^2 + 8xy + 10z$ . Utilice *DpGraph* para visualizar el punto más caliente y el punto más frío de la esfera.

◇ **1.5.11** En cada uno de los siguientes casos considere las curvas de nivel  $N_t = \{(x, y) : f(x, y) = t\}$  y utilice *DpGraph* para visualizar los puntos de la curva  $C = \{(x, y) : g(x, y) = 0\}$  donde  $f|_C$  presenta extremos absolutos o relativos.

a)  $f(x, y) = x + y^2$   $g(x, y) = 2x^2 + y^2 - 1$   
 b)  $f(x, y) = x^2 + y^2 - 4xy + 20x + 20y$   $g(x, y) = x^2 + y^2 + xy - 12$ .

◇ **1.5.12** Se considera el polinomio  $Q(x, y) = Ax^2 + 2Bxy + Cy^2$ , donde  $A, B, C \in \mathbb{R}$ . Utilice *DpGraph* para visualizar las curvas de nivel  $Q(x, y) = cte$  en los casos

i)  $AC - B^2 < 0$ ; ii)  $AC - B^2 > 0, A > 0$ ; iii)  $AC - B^2 > 0; A < 0$ .

Estudie cuando estas curvas son elipses, hipérbolas o rectas.

◇ **1.5.13** Utilice *DpGraph* para visualizar los siguientes subconjuntos de  $\mathbb{R}^3$

i)  $A = \{(x, y, z) : 0 \leq x, 0 \leq y, 0 \leq z, \sqrt{x} + \sqrt{y} + \sqrt{z} \leq 1\}$ .

ii)  $B = \{(x, y, z) : 0 \leq x, 0 \leq y, 0 \leq z, x + y + z \leq a, az \leq xy\}; (a > 0)$ .

iii)  $C = \{(x, y, z) : 2z^2 \leq x^2 + y^2 \leq 1 + z^2\}$ .

iv)  $D = \{(x, y, z) : x^2 + z^2 \leq R^2, y^2 + z^2 \leq R^2\}$ .

v)  $E = \{(x, y, z) : x^2 + y^2 + z^2 \leq a^2, x^2 + y^2 \leq z^2\}$ .

vi)  $F = \{(x, y, z) : x^2 + y^2 + z^2 \leq a^2, x^2 + y^2 \leq ax; 0 \leq z\}$ .

vii)  $G = \{(x, y, z) : x^2/a^2 + y^2/b^2 \leq 1 + z^2/c^2, 0 \leq z \leq 1\}$ .

viii)  $H = \{(x, y, z) : x^2 + y^2 \leq 2y; x^2 + y^2 \leq 1; 0 \leq x; 0 \leq z \leq x^2 + y^2\}$ .

# Capítulo 2

## Espacios métricos y espacios normados

*Normas y distancias en  $\mathbb{R}^n$ . Espacios métricos y espacios normados. Topología de un espacio normado. Normas en  $\mathbb{R}^n$ . Normas equivalentes. Espacios completos. Conjuntos compactos. Normas en  $C[a, b]$ .*

En este capítulo se repasan los resultados de topología que intervienen en la teoría de funciones reales de varias variables reales. En el espacio  $\mathbb{R}^n$  ( $n > 1$ ) no hay un orden natural, como ocurre en  $\mathbb{R}$ , y para establecer los resultados básicos de su topología no sirven los métodos y técnicas basados en el orden que se suelen utilizar en la recta real (resumidos en el apéndice B.1). Por ello es preciso acudir a los métodos generales de la topología de los espacios métricos.

En  $\mathbb{R}^n$  la distancia euclídea se define en términos de la norma euclídea. Generalmente las distancias que intervienen en Análisis Matemático proceden, en forma similar, de una norma y por ello se introduce en este capítulo la noción general de norma en un espacio vectorial sobre el cuerpo de los números reales, haciendo énfasis en el caso especial de que la norma proceda de un producto escalar. La estrecha relación que hay entre la estructura algebraica y la estructura topológica de los espacios normados hace que su topología tenga propiedades especiales que en general no tienen las topologías de los espacios métricos.

Se supone que el lector conoce los conceptos básicos de topología en el contexto de los espacios métricos y sólo se insiste en algunos aspectos particulares de la topología de los espacios normados como la caracterización de las normas equivalentes y sus consecuencias en relación con la noción de conjunto acotado y de espacio completo, hechos que no tienen contrapartida en el contexto de las distancias equivalentes.

En este capítulo, al repasar algunos de los resultados generales de la topología de los espacios métricos se hace énfasis en el manejo de las sucesiones. En un espacio normado las sucesiones se pueden someter a operaciones algebraicas y esto hace que sean una herramienta teórica muy adecuada para establecer resultados donde intervienen simultáneamente la estructura algebraica y la topología del espacio. También se insiste en los dos ingredientes básicos que garantizan la convergencia de una sucesión: La condición de Cauchy cuando el espacio es completo, y la existencia de un

único punto de aglomeración cuando la sucesión está contenida en un compacto.

Los espacios métricos completos se caracterizan mediante la validez del principio de encaje métrico 2.14 que sirve para insistir en la técnica de las sucesiones, y en la noción de diámetro de un conjunto.

En relación con la compacidad se establece su caracterización por sucesiones, el principio de encaje, y la caracterización de los subconjuntos compactos de  $\mathbb{R}^n$  (teorema de Bolzano-Weierstrass). La caracterización similar de los subconjuntos compactos de un espacio métrico completo utilizando la noción de conjunto totalmente acotado o precompacto, se ofrece como material complementario en el apéndice B.2.

Los convergencia uniforme de sucesiones (véase el apéndice A), se toma como base para introducir la norma de la convergencia uniforme  $\| \cdot \|_\infty$  sobre el espacio de las funciones continuas  $C[a, b]$ , viendo luego que el espacio  $(C[a, b], \| \cdot \|_\infty)$  es completo. En los ejercicios final del capítulo se muestra que  $C[a, b]$  no es completo para las normas  $\| \cdot \|_1$ ,  $\| \cdot \|_2$ , y que en  $(C[a, b], \| \cdot \|_\infty)$  hay subconjuntos cerrados y acotados que no son compactos.

## 2.1. El espacio $\mathbb{R}^n$ . Espacios normados

Como introducción al estudio de los espacios normados y en particular del espacio  $\mathbb{R}^n$ , recordemos que en cuerpo de los números complejos  $\mathbb{C}$  el módulo o valor absoluto de  $z = x + iy \in \mathbb{C}$ ,  $|z| = \sqrt{x^2 + y^2}$  tiene propiedades análogas a las del valor absoluto de los números reales:

- i)  $|z| \geq 0$  y  $|z| = 0$  si y sólo si  $z = 0$ .
- ii)  $|z + w| \leq |z| + |w|$  si  $z, w \in \mathbb{C}$
- iii)  $|zw| = |z||w|$  si  $z, w \in \mathbb{C}$

y lleva asociada la distancia euclídea en el plano

$$d_2(z, w) = |z - w| = \sqrt{(x - u)^2 + (y - v)^2}, \quad \text{si } z = x + iy, \quad w = u + iv.$$

Identificando  $\mathbb{R}^2$  con  $\mathbb{C}$  en la forma natural  $(x_1, x_2) \leftrightarrow x_1 + ix_2$ , en  $\mathbb{R}^2$  obtenemos la distancia euclídea  $d_2$  cuya topología asociada es la usual.

Con las operaciones habituales,  $\mathbb{R}^n$  tiene estructura de espacio vectorial, sobre el cuerpo  $\mathbb{R}$ , de dimensión  $n$ . (Análogamente  $\mathbb{C}^n$  es espacio vectorial sobre  $\mathbb{C}$  de dimensión  $n$ , pero también se puede considerar como espacio vectorial sobre  $\mathbb{R}$  de dimensión  $2n$ , que se identifica con  $\mathbb{R}^{2n}$ .)

En lo que sigue denotaremos por  $\mathbf{x}$  un punto genérico de  $\mathbb{R}^n$  de coordenadas  $\mathbf{x} = (x_1, x_2, \dots, x_n)$  y por  $\mathbf{e}_j$ ,  $1 \leq j \leq n$ , a los elementos de la base canónica:

$$\mathbf{e}_1 = (1, 0, 0, \dots, 0), \quad \mathbf{e}_2 = (0, 1, 0, \dots, 0), \quad \dots \quad \mathbf{e}_n = (0, 0, \dots, 0, 1).$$

A los elementos de  $\mathbb{R}^n$  a veces les llamaremos puntos y a veces vectores, según la interpretación que sea más adecuada en cada caso. Así por ejemplo, se suele hablar



de la recta que pasa por un punto  $\mathbf{p}$  según la dirección de un vector  $\mathbf{v}$ , o del vector  $\mathbf{v}$  tangente a una curva  $C$  en un punto  $\mathbf{p}$  de la misma. Esta manera de hablar se basa en que  $\mathbb{R}^n$ , además la estructura de espacio vectorial sobre  $\mathbb{R}$  también tiene estructura canónica de espacio afín con el vector nulo  $\mathbf{0} = (0, 0, \dots, 0)$  como origen. Según que  $\mathbf{x} = (x_1, x_2, \dots, x_n)$  se considere como punto o como vector se dice que  $x_j$ ,  $1 \leq j \leq n$ , son las coordenadas del punto o las componentes del vector. A estas dos formas de designar los elementos de  $\mathbb{R}^n$  corresponden dos formas de representación geométrica, según la interpretación que convenga en cada caso. La representación geométrica de un punto  $\mathbf{p} = (a, b, c) \in \mathbb{R}^3$  es la usual, mediante un sistema de ejes cartesianos rectangulares con origen en  $\mathbf{0}$ . Por otra parte, la representación geométrica de un vector  $\mathbf{v} \in \mathbb{R}^3$  es la habitual, fijando un punto  $\mathbf{p}$  y dibujando una flecha con origen en  $\mathbf{p}$  y extremo en  $\mathbf{p} + \mathbf{v}$ .

En  $\mathbb{R}^n$  la distancia euclídea  $d_2(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_2$ , se define en términos de la norma euclídea

$$\|\mathbf{x}\|_2 = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$$

Generalmente las distancias que intervienen en Análisis Matemático proceden, en forma similar, de una norma y por ello comenzaremos dando la noción general de norma sobre un espacio vectorial real o complejo  $E$ , es decir, un espacio vectorial sobre el cuerpo de los números reales o el cuerpo de los números complejos.

En lo que sigue, cuando una propiedad o definición se refiera indistintamente al caso de espacios vectoriales reales o complejos la formularemos hablando de un espacio vectorial sobre el cuerpo  $\mathbb{K}$ , donde  $\mathbb{K}$  será siempre el cuerpo real o el complejo. Generalmente consideraremos espacios vectoriales reales de dimensión finita, cuyo modelo estandar es  $\mathbb{R}^n$ , pero de momento no nos restringiremos a esta situación particular porque también conviene considerar algunos ejemplos importantes de espacios de funciones que no son finito dimensionales.

**Definición 2.1** *Si  $E$  es un espacio vectorial sobre el cuerpo  $\mathbb{K}$ , una norma sobre  $E$  es una aplicación  $\|\cdot\| : E \rightarrow [0, +\infty)$  que cumple:*

- i)  $\|\mathbf{x}\| \geq 0$  y  $\|\mathbf{x}\| = 0$  si y sólo si  $\mathbf{x} = 0$ .
- ii)  $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$  si  $\mathbf{x}, \mathbf{y} \in E$ .
- iii)  $\|\mu\mathbf{x}\| = |\mu| \|\mathbf{x}\|$  si  $\mu \in \mathbb{K}$  y  $\mathbf{x} \in E$ .

*Un espacio normado es un par  $(E, \|\cdot\|)$  donde  $E$  es un espacio vectorial sobre  $\mathbb{K}$  y  $\|\cdot\|$  es una norma sobre  $E$ . Cuando  $\mathbb{K} = \mathbb{R}$  (resp.  $\mathbb{K} = \mathbb{C}$ ) se dice que  $(E, \|\cdot\|)$  es un espacio normado real (resp. complejo)*

$(\mathbb{C}, |\cdot|)$  es un espacio normado complejo de dimensión 1, que se puede considerar como espacio normado real de dimensión 2, que se identifica con  $\mathbb{R}^2$ , dotado de la norma euclídea  $\|(x, y)\|_2 = |x + iy| = \sqrt{x^2 + y^2}$ .

**Proposición 2.2** Sea  $E$  un espacio vectorial real dotado de un producto escalar

$$\langle \cdot | \cdot \rangle : E \times E \rightarrow \mathbb{R}, \quad (\mathbf{x}, \mathbf{y}) \rightarrow \langle \mathbf{x} | \mathbf{y} \rangle$$

(aplicación bilineal simétrica que verifica  $\langle \mathbf{x} | \mathbf{x} \rangle \geq 0$  y  $\langle \mathbf{x} | \mathbf{x} \rangle = 0$  si y sólo si  $\mathbf{x} = 0$ ). Entonces  $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x} | \mathbf{x} \rangle}$  define en  $E$  una norma que cumple la desigualdad de Cauchy-Schwarz:

$$|\langle \mathbf{x} | \mathbf{y} \rangle| \leq \|\mathbf{x}\| \|\mathbf{y}\| \quad \text{para cada } \mathbf{x}, \mathbf{y} \in E$$

DEM: Dados  $\mathbf{x}, \mathbf{y} \in E$ , para todo  $t \in \mathbb{R}$  es  $h(t) = \langle \mathbf{x} + t\mathbf{y} | \mathbf{x} + t\mathbf{y} \rangle \geq 0$ . Usando la bilinealidad del producto escalar se obtiene:

$$h(t) = \langle \mathbf{x} | \mathbf{x} \rangle + 2t\langle \mathbf{x} | \mathbf{y} \rangle + t^2\langle \mathbf{y} | \mathbf{y} \rangle = \|\mathbf{x}\|^2 + 2t\langle \mathbf{x} | \mathbf{y} \rangle + t^2\|\mathbf{y}\|^2$$

luego la gráfica de la función  $h$  es una parábola que queda por encima del eje de abscisas. Por lo tanto la ecuación de segundo grado

$$\|\mathbf{x}\|^2 + 2t\langle \mathbf{x} | \mathbf{y} \rangle + t^2\|\mathbf{y}\|^2 = 0$$

no tiene dos soluciones reales distintas, luego su discriminante

$$\Delta = 4\langle \mathbf{x} | \mathbf{y} \rangle^2 - 4\|\mathbf{x}\|^2\|\mathbf{y}\|^2$$

debe cumplir  $\Delta \leq 0$ , es decir  $|\langle \mathbf{x} | \mathbf{y} \rangle| \leq \|\mathbf{x}\| \|\mathbf{y}\|$ . Con esta desigualdad se comprueba fácilmente que  $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x} | \mathbf{x} \rangle}$  cumple la desigualdad triangular  $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$  (las otras propiedades de la norma son inmediatas):

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|^2 &= \langle \mathbf{x} + \mathbf{y} | \mathbf{x} + \mathbf{y} \rangle = \|\mathbf{x}\|^2 + 2\langle \mathbf{x} | \mathbf{y} \rangle + \|\mathbf{y}\|^2 \leq \\ &= \|\mathbf{x}\|^2 + 2\|\mathbf{x}\| \|\mathbf{y}\| + \|\mathbf{y}\|^2 = (\|\mathbf{x}\| + \|\mathbf{y}\|)^2 \end{aligned}$$

La norma euclídea  $\|\cdot\|_2$  en  $\mathbb{R}^n$  es la asociada al producto escalar ordinario:

$$\langle \mathbf{x} | \mathbf{y} \rangle = \sum_{i=1}^n x_i y_i, \quad \text{es decir:}$$

$$\|\mathbf{x}\|_2 = \sqrt{\langle \mathbf{x} | \mathbf{x} \rangle} = \sqrt{\sum_{i=1}^n x_i^2}$$

NOTA: La desigualdad de Cauchy-Schwarz, cuando se aplica en  $\mathbb{R}^n$ , se escribe en la forma  $\sum_{j=1}^n x_j y_j \leq \sqrt{\sum_{j=1}^n x_j^2} \sqrt{\sum_{j=1}^n y_j^2}$ . Aplicada a los vectores  $(|x_1|, |x_2|, \dots, |x_n|)$ ,  $(|y_1|, |y_2|, \dots, |y_n|)$ , se obtiene una desigualdad mejorada:

$$\sum_{j=1}^n |x_j y_j| \leq \sqrt{\sum_{j=1}^n x_j^2} \sqrt{\sum_{j=1}^n y_j^2}$$

Más adelante se verán ejemplos de espacios normados, de dimensión infinita, con una norma que procede de un producto escalar. Uno de ellos es el espacio  $l^2$ ,

prototipo estandar de los espacios de Hilbert que desempeñan un papel destacado en el Análisis Funcional. Otro es el espacio de las funciones continuas  $C[a, b]$  con la norma  $\|f\|_2$  asociada al producto escalar,

$$\langle f | g \rangle = \int_a^b f(t)g(t)dt$$

que interviene en los problemas de aproximación de funciones en el sentido de los mínimos cuadrados.

**Topología de un espacio normado.** Una *distancia*  $d$  en un conjunto  $E$  es una aplicación  $d : E \times E \rightarrow [0, +\infty)$  que para cada terna  $\mathbf{x}, \mathbf{y}, \mathbf{z} \in E$  verifica:

- a)  $d(\mathbf{x}, \mathbf{y}) = 0$  si y sólo si  $\mathbf{x} = \mathbf{y}$ ;
- b)  $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$ ;
- c)  $d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$ ;

Un *espacio métrico* es un par  $(E, d)$  donde  $E$  es un conjunto  $d$  una distancia en  $E$ . En un espacio métrico  $(E, d)$  se define la *bola abierta* de centro  $\mathbf{a} \in E$  y radio  $r > 0$  como  $B(\mathbf{a}, r) = \{\mathbf{x} \in E : d(\mathbf{x}, \mathbf{a}) < r\}$ . La *bola cerrada* del mismo centro y radio es el conjunto  $\{\mathbf{x} \in E : d(\mathbf{x}, \mathbf{a}) \leq r\}$ .

Si  $(E, \|\cdot\|)$  es un espacio normado, es inmediato que  $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$  define en  $E$  una distancia, y la topología del espacio normado  $(E, \|\cdot\|)$  es la de este espacio métrico, que tiene como base la familia de las bolas abiertas  $\{B(\mathbf{a}, r) : \mathbf{a} \in E, r > 0\}$  donde  $B(\mathbf{a}, r) = \{\mathbf{x} \in E : \|\mathbf{x} - \mathbf{a}\| < r\}$ . Un conjunto  $A \subset E$  es *abierto* si para cada  $\mathbf{a} \in A$  existe  $r > 0$  tal que  $B(\mathbf{a}, r) \subset A$ . La familia de los abiertos (la topología de  $E$ ) la denotaremos por  $\mathcal{G}_{\|\cdot\|}(E)$  para hacer explícito que depende de la norma. A veces, cuando esté claro por el contexto la norma que se está considerando escribiremos simplemente  $\mathcal{G}$ . Es bien conocido que la familia de los abiertos  $\mathcal{G}$  es estable frente a intersecciones finitas y uniones arbitrarias y que  $\{E, \emptyset\} \subset \mathcal{G}$ .

Recordemos que dos distancias  $d, d'$  definidas sobre un mismo conjunto  $T$  se dice que son equivalentes cuando definen en  $T$  la misma topología.

**Definición 2.3** *Dos normas  $\|\cdot\|, \|\cdot\|'$  sobre un mismo espacio vectorial  $E$  se dice que son equivalentes cuando las distancias asociadas*

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|, \quad d'(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|'$$

*son equivalentes, es decir, las topologías asociadas coinciden.*

**Proposición 2.4** *Sea  $E$  un espacio vectorial sobre  $\mathbb{K}$  y  $\|\cdot\|, \|\cdot\|'$ , normas sobre  $E$ . Una condición necesaria y suficiente para que las dos normas sean equivalentes es que existan constantes  $\alpha > 0, \beta > 0$  verificando*

$$\alpha \|\mathbf{x}\| \leq \|\mathbf{x}\|' \leq \beta \|\mathbf{x}\|, \quad \text{para todo } \mathbf{x} \in E$$

DEM: La condición es suficiente: La primera desigualdad implica que

$$\{\mathbf{x} : \|\mathbf{x} - \mathbf{a}\| < r\} \supset \{\mathbf{x} : \|\mathbf{x} - \mathbf{a}\|' < \alpha r\}$$

luego  $\mathcal{G}_{\|\cdot\|}(E) \subset \mathcal{G}_{\|\cdot\|'}(E)$ . Análogamente la segunda desigualdad implica que

$$\{\mathbf{x} : \|\mathbf{x} - \mathbf{a}\|' < r\} \supset \{\mathbf{x} : \|\mathbf{x} - \mathbf{a}\| < r/\beta\}$$

luego  $\mathcal{G}_{\|\cdot\|'}(E) \subset \mathcal{G}_{\|\cdot\|}(E)$  y queda demostrado que las dos normas son equivalentes.

Recíprocamente, si las dos normas son equivalentes la bola  $\{\mathbf{y} \in E : \|\mathbf{y}\| < 1\}$  es abierta para la norma  $\|\cdot\|'$ , luego debe existir  $\alpha > 0$  tal que

$$\{\mathbf{y} \in E : \|\mathbf{y}\|' < \alpha\} \subset \{\mathbf{y} \in E : \|\mathbf{y}\| < 1\}$$

Si  $0 \neq \mathbf{x} \in E$  es arbitrario y  $0 < r < \alpha/\|\mathbf{x}\|'$  se cumple  $\|r\mathbf{x}\|' = r\|\mathbf{x}\|' < \alpha$ , luego  $r\|\mathbf{x}\| = \|r\mathbf{x}\| < 1$ . De la implicación

$$0 < r < \alpha/\|\mathbf{x}\|' \Rightarrow 0 < r < 1/\|\mathbf{x}\|$$

se sigue que  $\alpha/\|\mathbf{x}\|' \leq 1/\|\mathbf{x}\|$ , y queda establecida la desigualdad  $\alpha\|\mathbf{x}\| \leq \|\mathbf{x}\|'$ . Análogamente se demuestra que existe  $\beta' > 0$  tal que  $\beta'\|\mathbf{x}\|' \leq \|\mathbf{x}\|$ , luego  $\beta = 1/\beta'$  hace que se cumpla la otra desigualdad. ■

**La topología de  $\mathbb{R}^n$ .** La topología usual de  $\mathbb{R}^n$  es la asociada a la norma euclídea

$$\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$$

Es fácil ver que las siguientes fórmulas también definen normas en  $\mathbb{R}^n$ :

$$\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|; \quad \|\mathbf{x}\|_\infty = \sup_{1 \leq i \leq n} |x_i|.$$

Las distancias asociadas a estas normas las denotaremos  $d_1$  y  $d_\infty$ , respectivamente. Las normas  $\|\cdot\|_1$ ,  $\|\cdot\|_\infty$ , aunque no proceden de un producto escalar (véase el problema 2.6.1), también definen la topología usual de  $\mathbb{R}^n$ . Para ver que  $\|\cdot\|_1$  y  $\|\cdot\|_\infty$  son equivalentes a la norma euclídea  $\|\cdot\|_2$  basta aplicar la proposición 2.4 teniendo en cuenta las desigualdades.

$$\text{a) } \|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_2 \leq \sqrt{n} \|\mathbf{x}\|_\infty; \quad \text{b) } \|\mathbf{x}\|_1 / \sqrt{n} \leq \|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_1.$$

que son inmediatas, excepto  $\|\mathbf{x}\|_1 \leq \sqrt{n} \|\mathbf{x}\|_2$  que se puede obtener aplicando la desigualdad de Cauchy-Schwarz a la pareja de vectores  $\mathbf{a}, \mathbf{x}$ , donde las coordenadas  $a_j \in \{-1, 1\}$  de  $\mathbf{a}$  se han elegido de modo que para cada  $1 \leq j \leq n$ , sea  $a_j x_j = |x_j|$ .

$$\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i| = \sum_{i=1}^n a_i x_i = \langle \mathbf{a} | \mathbf{x} \rangle \leq \|\mathbf{a}\|_2 \|\mathbf{x}\|_2 = \sqrt{n} \|\mathbf{x}\|_2$$

En lo que sigue las bolas abiertas las normas  $\|\cdot\|_1$ ,  $\|\cdot\|_2$ ,  $\|\cdot\|_\infty$ , se designarán usando un subíndice que indique la norma que se está considerando:

$$B_p(\mathbf{a}, r) = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{a}\|_p < r\}, \quad p \in \{1, 2, \infty\}$$

(Cuando sea indiferente la norma considerada no usaremos subíndices). Con esta notación, las desigualdades en a) y b) se traducen en las inclusiones

$$B_\infty(\mathbf{a}, r) \supset B_2(\mathbf{a}, r) \supset B_\infty(\mathbf{a}, r/\sqrt{n}); \quad B_2(\mathbf{a}, r) \supset B_1(\mathbf{a}, r) \supset B_2(\mathbf{a}, r/\sqrt{n})$$

Se deja al cuidado del lector la interpretación geométrica, en  $\mathbb{R}^2$  y en  $\mathbb{R}^3$ , de las distancias  $d_1$ ,  $d_2$ ,  $d_\infty$ , y de las correspondientes bolas.

Además de las tres normas que ya definidas en  $\mathbb{R}^n$ , también se verifica que para cada  $p \geq 1$ , la fórmula  $\|\mathbf{x}\|_p = (\sum_{k=1}^n |x_k|^p)^{1/p}$  define una norma en  $\mathbb{R}^n$  (véase B.5). En el capítulo 3 se demostrará que en  $\mathbb{R}^n$  todas las normas son equivalentes.

Un subconjunto  $M$  de un espacio métrico  $(E, d)$  se dice que es *acotado* si está contenido en alguna bola. Si  $d$  y  $d'$  son dos distancias equivalentes sobre un mismo conjunto  $E$ , no es cierto en general que las dos distancias definan los mismos conjuntos acotados (basta considerar en  $\mathbb{R}$  la distancia usual  $d$  y la distancia acotada  $d'(x, y) = \min\{1, d(x, y)\}$ ). Sin embargo, en virtud de la proposición 2.4, dos normas equivalentes sobre un espacio vectorial  $E$  producen los mismos conjuntos acotados. Un subconjunto de  $\mathbb{R}^n$  se dice que es acotado si lo es para la norma euclídea (o para cualquier norma equivalente, como  $\|\cdot\|_1$ ,  $\|\cdot\|_\infty$ ).

**Nociones básicas de topología.** A continuación hacemos un breve resumen de algunas nociones y resultados básicos de la topología de los espacios métricos, que se aplican en particular a la topología de los espacios normados. Al mismo tiempo fijamos la notación y terminología que se empleará en lo que sigue.

- Sea  $M$  un subconjunto del espacio métrico  $(E, d)$ . Se dice que  $\mathbf{a} \in E$  es *interior* al conjunto  $M$ , ( $\mathbf{a} \in M^\circ$ ) si existe  $r > 0$  tal que  $B(\mathbf{a}, r) \subset M$ , y se dice que  $\mathbf{a}$  es *adherente* al conjunto  $M$ , ( $\mathbf{a} \in \overline{M}$ ), si para cada  $r > 0$  es  $B(\mathbf{a}, r) \cap M \neq \emptyset$ . Al conjunto  $\overline{M}$  (resp.  $M^\circ$ ) se le llama *cierre o clausura* (resp. *interior* de  $M$ ).

- Un conjunto  $G \subset E$  es *abierto* si y sólo si  $G = G^\circ$  y un conjunto  $F \subset E$  es *cerrado* si y sólo si  $F = \overline{F}$ . El interior de  $M$  es el mayor abierto contenido en  $M$  y la clausura de  $M$  es el menor cerrado que contiene a  $M$ .

- La *frontera* de  $M$  denotada  $\partial M$  es el conjunto formado por los puntos adherentes a  $M$  que no son interiores a  $M$  y el *exterior* de  $M$  es el interior de su complemento, que coincide con el complemento de su adherencia.

- Si para cada  $r > 0$  el conjunto  $B(\mathbf{a}, r) \cap M$  tiene infinitos elementos se dice que  $\mathbf{a}$  es un *punto de acumulación* de  $M$ , y se escribe  $\mathbf{a} \in M'$ . Evidentemente  $M' \subset \overline{M}$  y  $\overline{M} \setminus M' \subset M$ . Se verifica que  $\overline{M} = M^\circ \cup \partial M = M \cup M'$  donde la primera unión es disjunta. El conjunto  $M$  es cerrado si y sólo si  $M' \subset M$ .

- Los puntos de  $\overline{M} \setminus M'$  se dice que son *puntos aislados* de  $M$ . Claramente  $\mathbf{a}$  es un punto aislado de  $M$  si para algún  $r > 0$  es  $B(\mathbf{a}, r) \cap M = \{\mathbf{a}\}$ .

- Si  $M$  es un subconjunto de  $E$ , y  $d_M$  la distancia que  $d$  induce en  $M$  (la restricción de  $d$  al subconjunto  $M \times M \subset E \times E$ ) entonces la topología relativa

de  $M$  es la asociada a la distancia  $d_M$ . Cuando es un  $M$  subespacio vectorial del espacio normado  $(E, \|\cdot\|)$ , la topología relativa de  $M$  es la asociada a la norma que se obtiene restringiendo a  $M$  la norma de  $E$ .

Recordemos que un subconjunto  $A$  de  $M$  es *abierto* (resp. *cerrado* relativo a  $M$  si  $A$  es la intersección de  $M$  con un subconjunto abierto (resp. cerrado) de  $E$ . Esto ocurre si y sólo si  $A$ , considerado como subconjunto de  $M$ , es abierto (resp. cerrado) en el espacio métrico  $(M, d_M)$ . Se comprueba fácilmente que  $(A^\circ) \cap M$ ,  $\overline{A} \cap M$  y  $A' \cap M$  son los subconjuntos de  $M$  formados, respectivamente, por los puntos de  $A$  que son interiores, adherentes y de acumulación de  $A$  relativos a  $M$ , es decir, en el espacio métrico  $(M, d_M)$ .

- La topología de un espacio normado, y en particular la de  $(\mathbb{R}^n, \|\cdot\|_2)$  tiene propiedades especiales que no tienen sentido en un espacio métrico general: En un espacio normado la distancia  $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$  es invariante por traslaciones:  $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{a} + \mathbf{x}, \mathbf{a} + \mathbf{y})$  y como consecuencia la bola  $B(\mathbf{a}, r) = \mathbf{a} + B(\mathbf{0}, r)$  es la trasladada, con el vector  $\mathbf{a}$ , de la bola  $B(\mathbf{0}, r)$ . La distancia también se comporta bien con las homotecias respecto al origen:  $d(\mu\mathbf{x}, \mu\mathbf{y}) = \mu d(\mathbf{x}, \mathbf{y})$  si  $\mu > 0$ .

La estrecha relación que hay entre la estructura algebraica y la estructura topológica de los espacios normados hace que su topología tenga propiedades especiales que en general no tienen las topologías de los espacios métricos. A título de ejemplo se puede señalar la propiedad de las bolas considerada en el ejercicio 2.6.11.

**Conjuntos conexos.** En un espacio métrico general y en particular en  $\mathbb{R}^n$  los subconjuntos conexos desempeñan un papel análogo al que desempeñan los intervalos en la recta real. Conviene empezar con la definición de espacio métrico conexo, para formular luego, en términos de ella, la definición de subconjunto conexo.

Un espacio métrico  $(E, d)$  se dice que es *conexo* cuando los únicos subconjuntos de  $E$  que son simultáneamente abiertos y cerrados son  $\emptyset$  y  $E$ . Un subconjunto  $M$  del espacio métrico  $(E, d)$  se dice que es *conexo* cuando el espacio métrico  $(M, d_M)$  es conexo.

La condición necesaria y suficiente para que  $M \subset E$  no sea conexo es que  $M$  pueda recubrir mediante dos abiertos de  $(E, d)$ ,  $U, V \subset E$ , tales que los conjuntos  $U \cap M, V \cap M$  sean no vacíos y disjuntos. En topología general se demuestran, entre otras, las siguientes propiedades:

- i) si  $M$  es conexo entonces su adherencia  $\overline{M}$  también lo es.
- ii) Si  $\{M_j : j \in J\}$  es una familia de subconjuntos conexos con intersección no vacía entonces su unión  $M = \bigcup_{j \in J} M_j$  también es conexa.

Obsérvese un conjunto  $C \subset M \subset E$  es conexo como subconjunto del espacio métrico  $(M, d_M)$  si y sólo si es conexo como subconjunto del espacio métrico  $(E, d)$ , es decir, la conexión de un conjunto  $M \subset E$  es una propiedad intrínseca del mismo.

Dos puntos  $\mathbf{x}, \mathbf{y}$  de un conjunto  $M \subset E$  se dice que están conectados en  $M$  si existe un conjunto conexo  $C$  tal que  $\{\mathbf{x}, \mathbf{y}\} \subset C \subset M$ . Así queda definida en  $M$  una relación de equivalencia. Las *componentes conexas* de  $M$  son las clases de equivalencia en que queda descompuesto  $M$  mediante esta relación.

Para cada  $\mathbf{x} \in M$  se llama *componente conexa* de  $\mathbf{x}$  en  $M$  a la clase de equivalencia de  $\mathbf{x}$ ; está formada por la unión de todos los subconjuntos conexos de  $M$  que

contienen a  $\mathbf{x}$ . En virtud de ii) la componente conexa de  $\mathbf{x}$  en  $M$  es un conjunto conexo, y por lo tanto es el mayor subconjunto conexo de  $M$  que contiene a  $\mathbf{x}$ . En virtud de i) se puede asegurar que las componentes conexas de  $M$  son subconjuntos cerrados relativos a  $M$ , es decir conjuntos cerrados del espacio métrico  $(M, d_M)$ .

Un subconjunto  $M$  de la recta real  $\mathbb{R}$  es conexo si y sólo si es un intervalo. Este hecho tiene como consecuencia otros resultados importantes en los que interviene la continuidad (véase la definición 3.4) que anticipamos a continuación con el fin de completar aquí los resultados básicos de conexión que interesan en Análisis Matemático.

- Si  $M$  es un subconjunto del espacio métrico  $(E, d)$  y  $\gamma : [a, b] \rightarrow E$  es una función continua con  $\gamma([a, b]) \subset M$ , se dice que  $\gamma$  es un *camino* en  $M$  de origen  $\gamma(a)$  y extremo  $\gamma(b)$ . Si para cada par de puntos  $\mathbf{x}, \mathbf{y}$  de  $M$  existe un camino en  $M$  con origen  $\mathbf{x}$  y extremo  $\mathbf{y}$  se dice que  $M$  es *conexo por caminos*. Todo subconjunto  $M \subset E$  conexo por caminos es conexo. El recíproco se cumple si el espacio métrico  $(E, d)$  tiene la propiedad de que sus bolas abiertas son conexas por caminos.

- En cualquier espacio normado  $(E, \|\cdot\|)$ , con la distancia asociada a la norma las bolas abiertas son conexas por caminos. Por lo tanto, en un espacio normado un subconjunto abierto es conexo si y sólo si es conexo por caminos. Como en este caso las bolas abiertas son conexas, se sigue que en los espacios normados las componentes conexas de los abiertos son abiertas.

- Si  $(E, \|\cdot\|)$  es un espacio normado, un segmento de origen  $\mathbf{a}$  y extremo  $\mathbf{b}$  es un camino de la forma  $\sigma(t) = \mathbf{a} + t(\mathbf{b} - \mathbf{a})$ ,  $0 \leq t \leq 1$ , y un camino poligonal es un camino obtenido concatenando un número finito de segmentos, de manera que el extremo de cada segmento es el origen del que le sigue. Un camino poligonal en  $\mathbb{R}^n$  se dice que es de lados paralelos a los ejes cuando cada segmento que lo forma tiene la dirección de alguno de los vectores de la base canónica.

Un subconjunto  $M$  de un espacio normado  $(E, \|\cdot\|)$  se dice que es *conexo por poligonales* si para cada  $\mathbf{x}, \mathbf{y} \in M$  existe un camino poligonal en  $M$  con origen  $\mathbf{x}$  y extremo  $\mathbf{y}$ . Un subconjunto  $M$  de  $\mathbb{R}^n$  se dice que es conexo por poligonales de lados paralelos a los ejes si para cada  $\mathbf{x}, \mathbf{y} \in M$  existe un camino poligonal en  $M$  de lados paralelos a los ejes con origen  $\mathbf{x}$  y extremo  $\mathbf{y}$ .

En un espacio normado (resp. en  $\mathbb{R}^n$ ) todo abierto conexo es conexo por poligonales (resp. por poligonales de lados paralelos a los ejes).

## 2.2. Sucesiones y conjuntos compactos

Una sucesión  $(\mathbf{x}_n)$  en el espacio métrico  $(E, d)$  se dice que es *convergente* hacia  $\mathbf{x} \in E$  si  $\lim_n d(\mathbf{x}, \mathbf{x}_n) = 0$ . En este caso el punto  $\mathbf{x}$ , necesariamente único, se dice que es el *límite* de la sucesión y se escribe  $\lim_n \mathbf{x}_n = \mathbf{x}$ .

- En los espacios métricos las sucesiones proporcionan caracterizaciones útiles de diversas nociones topológicas, como las a) y b) que siguen:

a) Un punto  $\mathbf{x} \in E$  es adherente a  $M \subseteq E$  (resp. es de acumulación de  $M$ ) si y sólo si es límite de alguna sucesión contenida en  $M$  (resp.  $M \setminus \{\mathbf{x}\}$ ).

b) Un conjunto  $M \subset E$  es cerrado si y sólo si toda sucesión convergente,

contenida en  $M$ , tiene su límite en  $M$ .

- Si  $\sigma : \mathbb{N} \rightarrow \mathbb{N}$  es estrictamente creciente y  $\sigma(k) = n_k$ , se dice que la sucesión  $(\mathbf{x}_{n_k})$  es una *subsucesión* de  $(\mathbf{x}_n)$ . Si la sucesión  $(\mathbf{x}_n)$  converge hacia  $\mathbf{x}$ , cada subsucesión de  $(\mathbf{x}_n)$  también converge hacia  $\mathbf{x}$ .

- Se dice que  $\mathbf{x} \in E$  es *punto de aglomeración* de la sucesión  $(\mathbf{x}_n)$  si es límite de alguna subsucesión de  $(\mathbf{x}_n)$ . Las sucesiones convergentes tienen un único punto de aglomeración (su límite).

Todo punto de acumulación de la imagen  $\mathbf{x}(\mathbb{N})$  es un punto de aglomeración de la sucesión  $(\mathbf{x}_n)$ , pero el recíproco no es cierto en general (los puntos 1 y  $-1$  son de aglomeración de la sucesión  $((-1)^n)$  cuya imagen no tiene puntos de acumulación porque es finita). El recíproco es cierto cuando la aplicación  $n \rightarrow \mathbf{x}_n$  es inyectiva (y más generalmente, cuando cada conjunto  $\{k \in \mathbb{N} : \mathbf{x}_k = \mathbf{x}_n\}$  es finito).

El siguiente lema proporciona una sencilla y útil caracterización del conjunto (que puede ser vacío) formado por los puntos de aglomeración de una sucesión :

**Lema 2.5** Si  $(\mathbf{x}_n)$  es una sucesión en el espacio métrico  $(E, d)$ , son equivalentes

- $\mathbf{x}$  es punto de aglomeración de la sucesión  $(\mathbf{x}_n)$ .
- Para cada  $\epsilon > 0$  el conjunto  $\{n \in \mathbb{N} : \mathbf{x}_n \in B(\mathbf{x}, \epsilon)\}$  es infinito.
- $\mathbf{x} \in \bigcap_{m=1}^{\infty} \overline{\{\mathbf{x}_n : n \geq m\}}$ .

DEM: Es una comprobación sencilla que se deja como ejercicio. ■

Un subconjunto  $K$  de un espacio métrico  $(E, d)$  se dice que es *compacto* si de cada recubrimiento abierto de  $K$  es posible extraer un subrecubrimiento finito. Los conjuntos compactos son cerrados y los subconjuntos cerrados de los conjuntos compactos también son compactos. La familia de los conjuntos compactos es estable por uniones finitas e intersecciones arbitrarias.

Una familia de conjuntos  $\{F_\alpha : \alpha \in A\}$  se dice que tiene la *propiedad de la intersección finita* cuando toda subfamilia finita tiene intersección no vacía. Es fácil ver que  $K \subset E$  es compacto si y sólo si toda familia de subconjuntos cerrados de  $K$ , con la propiedad de la intersección finita, tiene intersección no vacía. Una consecuencia inmediata es el siguiente *principio de encaje*

**Proposición 2.6** Toda sucesión decreciente de conjuntos cerrados no vacíos  $C_n$  contenidos en un subconjunto compacto  $K$  de un espacio métrico tiene intersección no vacía  $\bigcap_{n \in \mathbb{N}} C_n \neq \emptyset$ .

**Teorema 2.7** Para un subconjunto  $K$  de un espacio métrico  $(E, d)$  las siguientes propiedades son equivalentes:

- $K$  es compacto.
- De cada sucesión en  $K$  se pueda extraer una subsucesión que converge hacia un punto de  $K$ .
- Para cada conjunto infinito  $M \subset K$  se cumple  $M' \cap K \neq \emptyset$ .



DEM: a)  $\Rightarrow$  b): Si  $(\mathbf{x}_n)$  es una sucesión en  $K$ , aplicando la proposición 2.6 a la sucesión decreciente de cerrados  $C_m = \overline{\{\mathbf{x}_n : n \geq m\}} \subset K$  se obtiene que no es vacía la intersección

$$A = \bigcap_m^{\infty} \overline{\{\mathbf{x}_n : n \geq m\}}$$

y según el lema 2.5, para cada  $\mathbf{a} \in A \subset K$  hay una subsucesión de  $(\mathbf{x}_n)$  que converge hacia  $\mathbf{a}$ .

b)  $\Rightarrow$  a): Sea  $\{A_j : j \in J\}$  un recubrimiento abierto de  $K$ . Demostramos en primer lugar que la condición b) implica la siguiente propiedad:

(L) *Existe  $\rho > 0$  tal que para cada  $\mathbf{x} \in K$  la bola  $B(\mathbf{x}, \rho)$  está contenida en algún  $A_j$ .*

Basta razonar por reducción al absurdo: Si no se cumple (L), para cada  $n \in \mathbb{N}$  existe  $\mathbf{x}_n \in K$  tal que la bola  $B(\mathbf{x}_n, 1/n)$  no está contenida en ningún  $A_j$ . Por hipótesis, de la sucesión  $(\mathbf{x}_n)$  se puede extraer una subsucesión  $(\mathbf{x}_{n_k})$  convergente hacia un punto  $\mathbf{x} \in K$ . Entonces  $\mathbf{x} \in A_j$  para algún  $j \in J$ , y existirá  $r > 0$  tal que  $B(\mathbf{x}, r) \subset A_j$ . Por ser  $\mathbf{x}$  el límite de la subsucesión  $(\mathbf{x}_{n_k})$ , existirá  $k \in \mathbb{N}$  tal que  $1/n_k < r/2$  y  $\mathbf{x}_{n_k} \in B(\mathbf{x}, r/2)$ . Se llega así a la contradicción:

$$B(\mathbf{x}_{n_k}, 1/n_k) \subset B(\mathbf{x}_{n_k}, r/2) \subset B(\mathbf{x}, r) \subset A_j$$

y queda establecida la propiedad (L).

Para demostrar que de  $\{A_j : j \in J\}$  se puede extraer un recubrimiento finito de  $K$  procedemos en la forma siguiente:

Fijado un punto  $\mathbf{x}_1 \in K$ , si  $B(\mathbf{x}_1, \rho) \supset K$  hemos terminado en virtud de (L). En caso contrario existe  $\mathbf{x}_2 \in K$  tal que  $d(\mathbf{x}_1, \mathbf{x}_2) \geq \rho$ . Si  $B(\mathbf{x}_1, \rho) \cup B(\mathbf{x}_2, \rho) \supset K$ , la demostración ha concluido en virtud de (L). En caso contrario existe  $\mathbf{x}_3 \in K$  tal que  $d(\mathbf{x}_1, \mathbf{x}_3) \geq \rho$ ,  $d(\mathbf{x}_2, \mathbf{x}_3) \geq \rho$ . Si las tres bolas  $B(\mathbf{x}_1, \rho)$ ,  $B(\mathbf{x}_2, \rho)$ ,  $B(\mathbf{x}_3, \rho)$  recubren  $K$ , la demostración ha concluido en virtud de (L). En caso contrario el proceso continúa. Para terminar la demostración basta ver que el proceso se detiene en un número finito de pasos. En caso contrario obtendríamos una sucesión infinita  $\mathbf{x}_n \in K$  tal que  $d(\mathbf{x}_p, \mathbf{x}_q) \geq \rho$  si  $p \neq q$ , y es claro que una sucesión con esta propiedad no puede tener subsucesiones convergentes.

b)  $\Rightarrow$  c): Como  $M$  es infinito existe una sucesión  $\mathbf{x}_n \in M$  sin términos repetidos y todo punto de aglomeración de la sucesión es un punto de acumulación de  $M$  que está en  $K$ .

c)  $\Rightarrow$  b) es inmediato. ■

Como corolario se obtiene el siguiente criterio que suele ser útil para demostrar que una sucesión  $(\mathbf{x}_n)$  converge y obtener su límite:

**Corolario 2.8** *Una sucesión contenida en un subconjunto compacto de un espacio métrico es convergente si y sólo si tiene un único punto de aglomeración. En este caso el único punto de aglomeración es su límite.*

DEM: Es claro que el límite de una sucesión convergente es su único punto de aglomeración, por lo que basta demostrar que el recíproco es cierto cuando se supone que la sucesión está contenida en un compacto. Sea  $(\mathbf{x}_n)$  una sucesión con esta propiedad, tal que toda subsucesión convergente de  $(\mathbf{x}_n)$  converge hacia  $\mathbf{x}$ . Demostraremos por reducción al absurdo que  $(\mathbf{x}_n)$  converge hacia  $\mathbf{x}$ . Efectivamente, en caso contrario existiría  $\epsilon > 0$  tal que el conjunto  $P = \{p \in \mathbb{N} : d(\mathbf{x}_p, \mathbf{x}) \geq \epsilon\}$  es infinito. Ordenando sus elementos de modo creciente  $P = \{p_1 < p_2 < p_3 < \dots\}$  obtendríamos una subsucesión  $(\mathbf{x}_{p_j})_{j \in \mathbb{N}}$  la cual, por estar contenida en un compacto, tendría una subsucesión convergente  $(\mathbf{x}_{q_j})_{q_j \in Q}$ , donde  $Q = \{q_1 < q_2 < q_3 < \dots\}$  es un subconjunto infinito de  $P$ . Para cada  $j \in \mathbb{N}$  es  $d(\mathbf{x}_{q_j}, \mathbf{x}) \geq \epsilon$  luego el límite  $\mathbf{a} = \lim_j \mathbf{x}_{q_j}$  cumpliría  $d(\mathbf{a}, \mathbf{x}) = \lim_j d(\mathbf{x}_{q_j}, \mathbf{x}) \geq \epsilon$ . Se obtendría así una subsucesión  $(\mathbf{x}_{q_j})_{j \in \mathbb{N}}$  convergente hacia un punto  $\mathbf{a} \neq \mathbf{x}$ , lo que contradice la hipótesis. ■

En un espacio métrico los conjuntos compactos son cerrados y acotados, pero el recíproco no es cierto en general ( $\mathbb{R}$  con la distancia  $d'(x, y) = \min\{1, |x - y|\}$  es cerrado y acotado pero no es compacto).

El siguiente teorema establece que en  $\mathbb{R}^n$ , con la norma euclídea (o cualquier norma equivalente) vale el recíproco. En la demostración se utiliza que una sucesión  $\mathbf{x}_k = (x_k(1), x_k(2), \dots, x_k(n))$  en  $\mathbb{R}^n$ , converge hacia  $\mathbf{x} = (x(1), x(2), \dots, x(n))$  si y sólo si  $\lim_k x_k(j) = x(j)$ , para cada  $1 \leq j \leq n$ .

**Teorema 2.9 (Bolzano-Weierstrass)** *Un conjunto  $K \subset \mathbb{R}^n$  es compacto (para la topología usual) si y sólo si es cerrado y acotado.*

DEM: Basta demostrar que todo conjunto cerrado y acotado  $K \subset \mathbb{R}^n$  es compacto. Para ello conviene trabajar con la norma  $\|\cdot\|_\infty$  que tiene la propiedad de que la bola cerrada  $\overline{B_\infty(\mathbf{0}, r)} = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_\infty \leq r\}$  es un producto finito de intervalos compactos:

$$\overline{B_\infty(\mathbf{0}, r)} = [-r, r] \times [-r, r] \times \dots \times [-r, r]$$

Con esta propiedad es fácil demostrar que cada sucesión en  $\overline{B_\infty(\mathbf{0}, r)}$  tiene una subsucesión convergente luego, en virtud del teorema 2.7, la bola cerrada  $\overline{B_\infty(\mathbf{0}, r)}$  es compacta. Todo conjunto cerrado y acotado  $K \subset \mathbb{R}^n$  está contenido en alguna bola compacta  $\overline{B_\infty(\mathbf{0}, r)}$  y por lo tanto es compacto. ■

NOTA. Según el teorema anterior una sucesión de números reales  $(x_n)$  está contenida en un compacto si y sólo si es acotada. Es bien conocido que  $\underline{\lim}_n x_n$  y  $\overline{\lim}_n x_n$  son el menor y el mayor punto de aglomeración de la sucesión  $(x_n)$ , de modo que, en este caso, el corolario 2.8 se concreta en el resultado que afirma que una sucesión acotada de números reales tiene límite si y sólo si su límite inferior coincide con su límite superior.

## 2.3. Espacios completos

Una sucesión  $(\mathbf{x}_n)$  en un espacio métrico  $(E, d)$  se dice que es de Cauchy cuando cumple la condición

$$(C): \text{ Para cada } \epsilon > 0 \text{ existe } n(\epsilon) \in \mathbb{N} \text{ tal que } p > q \geq n(\epsilon) \Rightarrow d(\mathbf{x}_p, \mathbf{x}_q) < \epsilon$$

que se suele expresar en la forma abreviada  $\lim_{p,q} d(\mathbf{x}_p, \mathbf{x}_q) = 0$ .

Es fácil ver que toda sucesión convergente es de Cauchy y que toda sucesión de Cauchy  $(\mathbf{x}_n)$  con una subsucesión convergente (hacia  $\mathbf{x}$ ) es convergente (hacia  $\mathbf{x}$ ).

**Definición 2.10** *Un espacio métrico  $(E, d)$  se dice que es completo cuando toda sucesión de Cauchy es convergente. Un espacio normado, real o complejo,  $(E, \|\cdot\|)$ , se dice que es completo cuando el espacio métrico asociado es completo. A los espacios normados completos se les llama también espacios de Banach.*

Es bien conocido que los espacios métricos  $(\mathbb{R}, d)$ ,  $(\mathbb{C}, d)$ , con las distancias asociadas al valor absoluto, son completos.

- Si  $M$  es un subconjunto del espacio métrico  $(E, d)$  y el espacio métrico  $(M, d_M)$  es completo se dice que  $M$  es un *subconjunto completo* de  $E$  (recuérdese que  $d_M$  es la distancia que  $d$  induce en  $M$ ).

- Todo subconjunto completo es cerrado y todo subconjunto cerrado de un espacio métrico completo es completo. Así, en los espacios métricos completos la familia de los subconjuntos completos coincide con la familia de los subconjuntos cerrados.

- Las nociones de sucesión de Cauchy y de espacio métrico completo no son topológicas: Puede haber dos distancias equivalentes  $d, d'$  sobre un mismo conjunto  $E$  que no tengan las mismas sucesiones de Cauchy, de modo  $E$  sea completo con una distancia y no lo sea con la otra (véase el ejercicio 2.6.5). Según la siguiente proposición esta situación no se presenta cuando se consideran normas equivalentes.

**Proposición 2.11** *Sean  $\|\cdot\|, \|\cdot\|'$  normas equivalentes sobre el espacio vectorial  $E$  (real o complejo). Entonces  $(E, \|\cdot\|)$  es completo si y sólo si  $(E, \|\cdot\|')$  lo es.*

DEM: Según la proposición 2.4 existen  $\alpha > 0, \beta > 0$  tales que para todo  $\mathbf{x} \in E$

$$\alpha \|\mathbf{x}\| \leq \|\mathbf{x}\|' \leq \beta \|\mathbf{x}\|$$

Si  $(E, \|\cdot\|)$  es completo y  $(\mathbf{x}_n)$  es una sucesión de Cauchy para  $\|\cdot\|'$ , en virtud de la primera desigualdad, también lo es para la norma  $\|\cdot\|$ , luego es convergente para esta norma, y también lo es para la norma equivalente  $\|\cdot\|'$ . Esto demuestra que  $(E, \|\cdot\|')$  es completo. Análogamente se demuestra el recíproco. ■

**Teorema 2.12** *Si  $\|\cdot\|$  es una de las tres normas usuales de  $\mathbb{R}^n$ ,  $(\|\cdot\|_1, \|\cdot\|_2, \|\cdot\|_\infty)$  entonces  $(\mathbb{R}^n, \|\cdot\|)$  es completo.*

DEM: Como las tres normas son equivalentes, según la proposición 2.11 basta demostrar que  $(\mathbb{R}^n, \|\cdot\|_2)$  es completo. Si  $\mathbf{x}_k = (x_k(1), x_k(2), \dots, x_k(n))$  es una sucesión de Cauchy en  $(\mathbb{R}^n, \|\cdot\|_2)$  las primeras componentes  $(x_k(1))$  forman una sucesión de Cauchy de números reales ya que  $|x_p(1) - x_q(1)| \leq \|\mathbf{x}_p - \mathbf{x}_q\|_2$ , luego existe el límite  $\lim_k x_k(1) = x(1)$ . Razonando igual con las otras componentes se obtiene el vector  $\mathbf{x} = (x(1), x(2), \dots, x(n)) \in \mathbb{R}^n$  con  $x(j) = \lim_k x_k(j)$ ,  $1 \leq j \leq n$ , y es inmediato que  $\lim_k \|\mathbf{x}_k - \mathbf{x}\|_2 = 0$ . ■

- Para  $1 \leq p \leq +\infty$ , en el apéndice B se introducen los espacios  $(l^p, \|\cdot\|_p)$ , (versiones infinito dimensionales de  $(\mathbb{R}^n, \|\cdot\|_p)$  que también resultan completos.

- Tomando como base el teorema anterior en el siguiente capítulo (3.20) se demostrará que todo espacio normado de dimensión finita es completo

El *diámetro* de un subconjunto  $A$  del espacio métrico  $(E, d)$ , se define como el supremo (en la recta real ampliada  $\overline{\mathbb{R}}$ )

$$\text{diam}(A) = \sup\{d(\mathbf{x}, \mathbf{y}) : \mathbf{x} \in A, \mathbf{y} \in A\} \leq +\infty$$

Obsérvese que  $\text{diam}(A) < +\infty$  si y sólo si  $A$  es acotado (es decir, está contenido en alguna bola).

**Lema 2.13** *En un espacio métrico  $(E, d)$ , si  $A \subset E$  se verifica  $\text{diam}(A) = \text{diam}(\overline{A})$ .*

DEM: Es obvio que  $\text{diam}(A) \leq \text{diam}(\overline{A})$  y basta ver que  $\text{diam}(A) \geq \text{diam}(\overline{A})$ . Dados  $\mathbf{x}, \mathbf{y} \in \overline{A}$  existen sucesiones  $\mathbf{x}_n, \mathbf{y}_n$  en  $A$  tales que

$$\lim_n d(\mathbf{x}, \mathbf{x}_n) = \lim_n d(\mathbf{y}, \mathbf{y}_n) = 0$$

Según la desigualdad triangular  $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{x}_n) + d(\mathbf{x}_n, \mathbf{y}_n) + d(\mathbf{y}_n, \mathbf{y})$  luego  $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{x}_n) + \text{diam}(A) + d(\mathbf{y}_n, \mathbf{y})$ . Pasando al límite se obtiene la desigualdad  $d(\mathbf{x}, \mathbf{y}) \leq \text{diam}(A)$ . Como esta desigualdad es válida para todo  $\mathbf{x} \in \overline{A}$  y todo  $\mathbf{y} \in \overline{A}$  se obtiene que  $\text{diam}(\overline{A}) \leq \text{diam}(A)$ . ■

La propiedad b) que interviene en el siguiente teorema, *principio de encaje métrico*, similar a la que se ha visto en los espacios compactos, sirve para caracterizar a los espacios métricos completos:

**Teorema 2.14** *Para un espacio métrico  $(E, d)$  son equivalentes:*

a)  $(E, d)$  es completo.

b) Toda sucesión decreciente de conjuntos cerrados no vacíos  $C_n$  que cumple  $\lim_n \text{diam}(C_n) = 0$  tiene intersección no vacía (que se reduce a un punto).

DEM: a)  $\Rightarrow$  b): Sea  $C_n$  una sucesión decreciente de conjuntos cerrados no vacíos tal que  $\lim_n \text{diam}(C_n) = 0$ . Para cada  $n \in \mathbb{N}$  podemos elegir  $\mathbf{x}_n \in C_n$ . Así obtenemos una sucesión de Cauchy: Efectivamente, dado  $\epsilon > 0$  existe  $n \in \mathbb{N}$  tal que  $\text{diam}(C_n) < \epsilon$ , y si  $p > q \geq n$  se cumple  $\mathbf{x}_p \in C_p \subset C_n$ ,  $\mathbf{x}_q \in C_q \subset C_n$ , luego  $d(\mathbf{x}_p, \mathbf{x}_q) \leq \text{diam}(C_n) < \epsilon$ . Por hipótesis  $(E, d)$  es completo, luego la sucesión  $\mathbf{x}_n$

converge hacia un punto  $\mathbf{x} \in E$ . Puesto que cada  $C_n$  es cerrado y  $\mathbf{x}_k \in C_n$  para todo  $k \geq n$ , podemos asegurar que  $\mathbf{x} = \lim_n \mathbf{x}_n \in C_n$  luego  $\mathbf{x} \in \bigcap_n C_n$ , y queda demostrado que la intersección no es vacía. Por otra parte, la condición  $\lim_n \text{diam}(C_n) = 0$ , implica que la intersección se reduce a un punto.

b)  $\Rightarrow$  a): Si  $(\mathbf{x}_n)$  es una sucesión de Cauchy en  $(E, d)$  la sucesión de conjuntos  $A_n = \{\mathbf{x}_k : k \geq n\}$  cumple  $\lim_n \text{diam}(A_n) = 0$ . Según el lema 2.13 la sucesión decreciente de conjuntos cerrados no vacíos  $C_n = \overline{A_n}$  también verifica  $\text{diam}(C_n) = \text{diam}(A_n) \rightarrow 0$ . En virtud de la hipótesis b) podemos asegurar que

$$\bigcap_n \overline{\{\mathbf{x}_k : k \geq n\}} \neq \emptyset$$

y con el lema 2.5 se obtiene que la sucesión de Cauchy  $(\mathbf{x}_n)$  posee una subsucesión convergente y por lo tanto es convergente. ■

NOTA. La condición  $\lim_n \text{diam}(C_n) = 0$  es esencial para la validez del principio de encaje: En  $\mathbb{R}$ , con la distancia usual, la sucesión decreciente de cerrados  $C_n = [n, +\infty)$  tiene intersección vacía. También es esencial que el espacio métrico sea completo, pues en  $E = (0, 1)$ , con la distancia usual,  $C_n = (0, 1/n]$  es una sucesión decreciente de conjuntos cerrados (relativos a  $E$ ), que cumple  $\lim_n \text{diam}(C_n) = 0$  y tiene intersección vacía.

## 2.4. Normas en $C[a, b]$

Con las operaciones habituales de suma de funciones y de producto de un número por una función el conjunto de las funciones continuas  $f : [a, b] \rightarrow \mathbb{R}$ , denotado  $C[a, b]$  en lo que sigue, es un espacio vectorial real infinito dimensional pues las funciones  $1, x, x^2, \dots, x^n, \dots$  son linealmente independientes. Sobre este espacio vectorial vamos a considerar tres normas

$$\| \cdot \|_\infty, \quad \| \cdot \|_2, \quad \| \cdot \|_1$$

análogas a las que se han considerado en  $\mathbb{R}^n$ . Se comprueba fácilmente que

$$\|f\|_\infty = \text{máx}\{|f(t)| : t \in [a, b]\}$$

define una norma sobre el espacio vectorial real  $C[a, b]$ , llamada *norma de la convergencia uniforme*. La razón de este nombre es la siguiente:

Una sucesión de funciones  $f_n \in C[a, b]$  converge uniformemente hacia  $f \in C[a, b]$  si y sólo si  $\lim_n \|f_n - f\|_\infty = 0$ . El teorema A.6 es la clave para obtener:

**Corolario 2.15** *El espacio normado  $(C[a, b], \| \cdot \|_\infty)$  es completo.*

DEM: Sea  $f_n \in C[a, b]$  una sucesión de Cauchy para la norma  $\| \cdot \|_\infty$ . Fijado  $t \in [a, b]$ , la desigualdad  $|f_p(t) - f_q(t)| \leq \|f_p - f_q\|_\infty$  implica que la sucesión de números reales  $f_n(t)$  es de Cauchy y por lo tanto existe el límite  $\lim_n f_n(t) = f(t)$ .

Demostrando que  $f$  es continua y que  $\lim_n \|f_n - f\|_\infty = 0$  quedará establecido que el espacio  $(C[a, b], \|\cdot\|_\infty)$  es completo.

Dado  $\epsilon > 0$  existe  $m \in \mathbb{N}$  tal que para  $p > q \geq m$ , y todo  $t \in [a, b]$  se cumple  $|f_p(t) - f_q(t)| \leq \|f_p - f_q\|_\infty < \epsilon$ . Pasando al límite cuando  $p \rightarrow \infty$  se deduce que para todo  $q \geq m$  y todo  $t \in [a, b]$  se verifica  $|f(t) - f_q(t)| \leq \epsilon$ . Esto significa que la sucesión de funciones continuas  $(f_q)$  converge uniformemente hacia  $f$ . Según el teorema A.6 la función  $f$  es continua y es claro que  $q \geq m \Rightarrow \|f - f_q\|_\infty \leq \epsilon$ , es decir  $\lim_n \|f_n - f\|_\infty = 0$ . ■

Además de la norma de la convergencia uniforme, en el espacio vectorial  $C[a, b]$  también tienen especial interés las siguientes normas,

$$\|f\|_1 = \int_a^b |f(t)| dt, \quad \|f\|_2 = \sqrt{\int_a^b |f(t)|^2 dt}$$

llamadas, respectivamente, *norma de la convergencia en media* y *norma de la convergencia en media cuadrática*.

Para establecer que  $\|\cdot\|_1$  es una norma basta tener en cuenta las propiedades de linealidad y monotonía de la integral y el siguiente resultado bien conocido:

**Lema 2.16** *Si una función continua  $\varphi : [a, b] \rightarrow [0, +\infty)$  cumple  $\int_a^b \varphi(t) dt = 0$  entonces  $\varphi(t) = 0$  para todo  $t \in [a, b]$ .*

DEM: Se deja como ejercicio. ■

Con la linealidad de la integral se comprueba fácilmente que  $\langle f | g \rangle = \int_a^b f(t)g(t) dt$  define en  $C[a, b]$  una aplicación bilineal simétrica que verifica  $\langle f | f \rangle \geq 0$ . Volviendo a usar el lema 2.16 se obtiene que  $\langle f | f \rangle = 0 \Rightarrow f = 0$ , es decir,  $(f, g) \rightarrow \langle f | g \rangle$  es un producto escalar sobre  $C[a, b]$ . Con la proposición 2.2 se concluye que

$$\|f\|_2 = \sqrt{\langle f | f \rangle}$$

es una norma sobre  $C[a, b]$ .

Como el espacio  $(C[a, b], \|\cdot\|_\infty)$  es completo y  $(C[a, b], \|\cdot\|_1)$  no lo es (véase el ejercicio 2.17), en virtud de la proposición 2.11 podemos afirmar que las normas  $\|\cdot\|_\infty, \|\cdot\|_1$  no son equivalentes. Por la misma razón las normas  $\|\cdot\|_\infty, \|\cdot\|_2$  tampoco son equivalentes. Se puede ver directamente que para las topologías  $\mathcal{G}_i$  asociadas a las normas  $\|\cdot\|_i, (i = 1, 2, \infty)$ , se verifican las inclusiones estrictas  $\mathcal{G}_1 \subset \mathcal{G}_2 \subset \mathcal{G}_\infty$  (Véase el problema 2.6.30).

## 2.5. Ejercicios resueltos

**Ejercicio 2.17** *Sea  $a_n = 1/2 + 1/n$  y  $(f_n)$  la sucesión en  $C[0, 1]$  definida por*

$$f_n(x) = 1 \quad \text{si } 0 \leq x \leq 1/2,$$

$$f_n(x) = n(a_n - x) \quad \text{si } 1/2 \leq x \leq a_n,$$

$$f_n(x) = 0 \quad \text{si } a_n \leq x \leq 1.$$

Utilice esta sucesión para obtener que los espacios  $(C[a, b], \| \cdot \|_1)$ ,  $(C[a, b], \| \cdot \|_2)$ , no son completos.

SOLUCIÓN

(véase [6] pág. 76). Representando gráficamente las funciones  $f_n$  se observa que

$$p > q \Rightarrow 0 \leq f_q(t) - f_p(t) \leq f_q(t) \leq 1$$

y teniendo en cuenta que  $f_p(t) - f_q(t)$  se anula fuera del intervalo  $[1/2, a_q]$  resulta

$$\|f_p - f_q\|_2^2 = \int_{1/2}^{a_q} (f_p(t) - f_q(t))^2 dt \leq a_q - 1/2 = 1/q \leq 1/n$$

es decir,  $p > q \geq n \Rightarrow \|f_p - f_q\|_2 \leq 1/\sqrt{n}$ , y por lo tanto  $(f_n)$  es una sucesión de Cauchy para la norma  $\| \cdot \|_2$ . Quedará establecido que  $(C[0, 1], \| \cdot \|_2)$  no es completo viendo que esta sucesión no converge con esta norma.

Razonaremos por reducción al absurdo suponiendo que hay una función continua  $f \in C[0, 1]$  que verifica  $\lim_n \|f_n - f\|_2 = 0$ . Teniendo en cuenta las desigualdades

$$0 \leq \int_0^{1/2} (1 - f(t))^2 dt = \int_0^{1/2} (f_n(t) - f(t))^2 dt \leq \|f - f_n\|_2^2$$

y que la sucesión de la derecha converge hacia 0 se obtiene que  $\int_0^{1/2} (1 - f(t))^2 dt = 0$  luego, según el lema 2.16, la función continua  $(1 - f(t))^2$  es idénticamente nula en el intervalo cerrado  $[0, 1/2]$ , es decir  $f(t) = 1$  si  $t \in [0, 1/2]$ . Por otra parte, fijado  $\alpha \in (1/2, 1]$  podemos asegurar que las funciones  $f_n$  con  $n > 1/\alpha$  son idénticamente nulas en  $[\alpha, 1]$ , luego para todo  $n > 1/\alpha$  podemos escribir:

$$0 \leq \int_\alpha^1 f(t)^2 dt = \int_\alpha^1 (f_n(t) - f(t))^2 dt \leq \|f - f_n\|_2^2$$

y de manera análoga se obtiene que  $f$  es idénticamente nula en el intervalo  $[\alpha, 1]$ . Como  $\alpha \in (1/2, 1)$  es arbitrario se concluye que  $f(t) = 0$  para todo  $t \in (1/2, 1]$ , lo que contradice la continuidad de  $f$ .

La misma sucesión  $f_n$  sirve para demostrar que el espacio normado  $(C[a, b], \| \cdot \|_1)$  no es completo (los razonamientos, análogos a los realizados con la norma  $\| \cdot \|_2$ , se dejan al cuidado del lector). ■

**Ejercicio 2.18** Sea  $C_0[0, +\infty)$  el espacio vectorial de las funciones continuas  $f : [0, +\infty) \rightarrow \mathbb{R}$  que tienen límite 0 cuando  $x \rightarrow +\infty$ .

Demuestre que  $\|f\| = \sup\{|f(x)| : x \geq 0\}$  define una norma en  $C_0[0, +\infty)$  con la que este espacio resulta completo.

## SOLUCIÓN

En primer lugar hay que justificar brevemente que cada  $f \in C_0[0, +\infty)$  es acotada para tener garantizado que el supremo  $\sup\{|f(x)| : x \geq 0\}$  es finito (Como  $\lim_{x \rightarrow +\infty} f(x) = 0$ , podemos asegurar que existe  $C > 0$  tal que  $|f(x)| \leq 1$  para todo  $x \geq C$ . Usando que  $f$  es continua, se obtiene que  $f$  está acotada en el compacto  $[0, C]$  y combinando las dos cosas se concluye que  $f$  está acotada en  $[0, +\infty)$ .)

Como en el caso del espacio  $(C[0, 1], \|\cdot\|_\infty)$ , razonamientos rutinarios permiten comprobar que se cumplen las propiedades de norma. La parte esencial del problema es la demostración de que el espacio es completo: Si  $f_n$  es una sucesión de Cauchy en  $(C_0[0, +\infty), \|\cdot\|)$ , se cumple la condición de Cauchy para la convergencia uniforme sobre  $[0, +\infty)$ , luego la sucesión converge uniformemente hacia una función  $f : [0, +\infty) \rightarrow \mathbb{R}$  que resulta continua (porque el límite uniforme de una sucesión de funciones continuas es continuo). El hecho de que todas las funciones de la sucesión tienen límite 0 en  $+\infty$  implica que  $f$  tiene la misma propiedad: Efectivamente, dado  $\epsilon > 0$  existe  $n_\epsilon \in \mathbb{N}$  tal que si  $n \geq n_\epsilon$  entonces  $|f_n(x) - f(x)| < \epsilon/2$  para todo  $x \in [0, +\infty)$ . Se considera un  $n \geq n_\epsilon$  fijo y para la función  $f_n$  existe  $c \geq 0$  tal que  $x \geq c \Rightarrow |f_n(x)| \leq \epsilon/2$  luego

$$|f(x)| \leq |f(x) - f_n(x)| + |f_n(x)| \leq \epsilon/2 + \epsilon/2 = \epsilon$$

Esto demuestra que  $f$  está en  $C_0[0, +\infty)$ . La convergencia uniforme nos dice que  $\|f_n - f\| = \sup\{|f_n(x) - f(x)| : x \in [0, +\infty)\}$  tiende hacia 0, es decir, la sucesión  $f_n$  converge hacia  $f$  en el espacio normado  $C_0[0, +\infty)$ , y por lo tanto el espacio es completo.

■



## 2.6. Ejercicios propuestos

◇ **2.6.1** Sea  $(E, \|\cdot\|)$  un espacio normado cuya norma procede de un producto escalar.

a) Demuestre que la igualdad  $\|\mathbf{x} + \mathbf{y}\| = \|\mathbf{x}\| + \|\mathbf{y}\|$  implica que los vectores  $\mathbf{x}, \mathbf{y}$  son linealmente dependientes.

b) Compruebe la identidad del paralelogramo:

$$\|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2 = 2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2 \quad \text{para todo } \mathbf{x}, \mathbf{y} \in E$$

c) Deduzca que en  $\mathbb{R}^n$  las normas  $\|\cdot\|_1, \|\cdot\|_\infty$  no proceden de un producto escalar.

◇ **2.6.2** Se considera el polinomio  $Q(x, y) = ax^2 + 2bxy + cy^2$ , donde  $a, b, c \in \mathbb{R}$ ,  $a > 0$  y  $ac - b^2 > 0$ . Demuestre que la fórmula  $\|(x, y)\| = \sqrt{Q(x, y)}$  define en  $\mathbb{R}^2$  una norma asociada a un producto escalar. Utilice DpGraph para visualizar, con distintos valores de  $a, b, c$  la forma de las bolas definidas con esta norma.

◇ **2.6.3** Sean  $d_1, d_2$  distancias definidas en un conjunto  $E$  tales que los espacios métricos  $(E, d_1)$   $(E, d_2)$  tienen las mismas sucesiones convergentes. Demuestre que los límites de las sucesiones convergentes son los mismos: Si  $x_n \in E$  converge hacia  $x$  con la distancia  $d_1$ , también converge hacia  $x$  con la distancia  $d_2$ . Deduzca de ello que las dos distancias son equivalentes.

◇ **2.6.4** Si  $f: \mathbb{R} \rightarrow \mathbb{R}$  es una función estrictamente creciente compruebe que  $d'(x, y) = |f(x) - f(y)|$  define en  $\mathbb{R}$  una distancia. Demuestre que  $d$  y  $d'$  son equivalentes si y sólo si  $f$  es continua.

◇ **2.6.5** Sean  $d, d'$  dos distancias sobre un conjunto  $M$ . Demuestre que la condición: Existen  $\alpha > 0, \beta > 0$  tales que  $\alpha d(\mathbf{x}, \mathbf{y}) \leq d'(\mathbf{x}, \mathbf{y}) \leq \beta d(\mathbf{x}, \mathbf{y})$  para cada  $\mathbf{x}, \mathbf{y} \in M$ . implica que las dos distancias son equivalentes.

Demuestre que  $d'(x, y) = |f(x) - f(y)|$ , con  $f(x) = x/(1 + |x|)$ , define en  $\mathbb{R}$  una distancia equivalente a la usual  $d(x, y) = |x - y|$ , pero  $d$  y  $d'$  no cumplen la condición anterior. Compruebe que  $(\mathbb{R}, d')$  es una sucesión de Cauchy para la distancia  $d'$  y por lo tanto  $(\mathbb{R}, d')$  no es completo.

◇ **2.6.6** Sea  $(E, d)$  un espacio métrico y  $M \subset E$ . Demuestre que  $M'$  es cerrado y que  $M$  es cerrado si y sólo si  $M' \subset M$ .

◇ **2.6.7** Calcule los puntos de acumulación, en  $\mathbb{R}$ , del conjunto

$$M = \{1/n + 1/m : n \in \mathbb{N}, m \in \mathbb{N}\}$$

◇ **2.6.8** En el espacio normado  $(\mathbb{R}^3, \|\cdot\|_2)$  obtenga el interior, la frontera y el conjunto de puntos de acumulación de cada uno de los conjuntos:

$$\begin{aligned}
 A &:= \{(1/n, y) \in \mathbb{R}^2 : n \in \mathbb{N}, y \in \mathbb{R}\} \\
 B &:= \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0, x^2 + y^2 + z^2 \neq 1, 0 \leq z \leq 3\} \\
 C &:= \{\mathbf{x} \in \mathbb{R}^3 : x^2 + y^2 \leq 1\} \\
 D &:= \{\mathbf{x} \in \mathbb{R}^3 : x^2 + y^2 = z^2\}
 \end{aligned}$$

◇ **2.6.9** Si  $A$  es un subconjunto abierto de un espacio métrico  $(E, d)$  demuestre que para todo  $B \subset E$  se cumple  $A \cap \overline{B} \subset \overline{A \cap B}$ . Muestre con un ejemplo que la inclusión puede ser estricta.

◇ **2.6.10** Dado un conjunto no vacío  $M \subset \mathbb{R}^2$  sea  $d_M$  la distancia que induce en  $M$  la distancia usual de  $\mathbb{R}^2$ . Obtenga un conjunto  $M \subset \mathbb{R}^2$  tal que en el espacio métrico  $(M, d_M)$  exista una bola abierta  $\{\mathbf{x} \in M : d_M(\mathbf{x}, \mathbf{a}) < r\}$  que es un conjunto cerrado pero no es una bola cerrada, y una bola cerrada  $\{\mathbf{x} \in M : d_M(\mathbf{z}, \mathbf{a}) \leq R\}$  que es un conjunto abierto pero no es una bola abierta.

◇ **2.6.11** Demuestre que en un espacio normado  $(E, \|\cdot\|)$  la clausura de la bola abierta  $B(\mathbf{a}, r)$  es la bola cerrada  $\widehat{B}(\mathbf{a}, r) = \{\mathbf{x} \in E : \|\mathbf{x} - \mathbf{a}\| \leq r\}$ , el interior de la bola cerrada  $\widehat{B}(\mathbf{a}, r)$  es la bola abierta  $B(\mathbf{a}, r)$  y ambas bolas tienen la misma frontera:

$$\partial B(\mathbf{a}, r) = \partial \widehat{B}(\mathbf{a}, r) = \{\mathbf{x} \in E : \|\mathbf{x} - \mathbf{a}\| = r\}$$

Indique un espacio métrico donde no se cumplan estas propiedades.

◇ **2.6.12** Demuestre que en un espacio normado  $(E, \|\cdot\|)$  todo subespacio vectorial propio  $M \subsetneq E$  tiene interior vacío.

◇ **2.6.13** Si  $(E, d)$  es un espacio métrico y  $A, B \subset E$ , demuestre las siguientes relaciones de inclusión (donde  $\partial M$  denota la frontera de  $M \subset E$ ).

- $\partial(\overline{A}) \subset \partial A$ ,  $\partial(A^\circ) \subset \partial A$ , pero los conjuntos  $A$ ,  $A^\circ$ ,  $\overline{A}$  pueden ser distintos.
- $\partial(A \cup B) \subset \partial A \cup \partial B$ , y puede ocurrir que la inclusión sea estricta.
- Si  $\overline{A} \cap \overline{B} = \emptyset$  entonces  $\partial(A \cup B) \subset \partial A \cup \partial B$ , pero el recíproco es falso.
- Si  $A$  y  $B$  son abiertos se cumple  $(A \cap \partial B) \cup (B \cap \partial A) \subset \partial(A \cap B)$ , y puede ocurrir que la inclusión sea estricta.
- $\overline{A \cap B} \subset \overline{A} \cap \overline{B}$ , y puede ocurrir que la inclusión sea estricta.
- $\overline{A \cup B} = \overline{A} \cup \overline{B}$ .
- Si  $A$  es abierto entonces  $A \cap \overline{B} \subset \overline{A \cap B}$ . Muestre con un ejemplo que la inclusión es falsa cuando no se supone que  $A$  es abierto.
- Existen conjuntos abiertos  $A, B \subset \mathbb{R}$ , tales que los conjuntos  $A \cap \overline{B}$ ,  $\overline{A} \cap B$ ,  $\overline{A \cap B}$ ,  $\overline{A} \cap \overline{B}$  son distintos dos a dos.

◇ **2.6.14** Si  $(\mathbf{x}_n)$ ,  $(\mathbf{y}_n)$  son sucesiones en un espacio métrico  $(E, d)$ , demuestre

- a) Si ambas sucesiones son de Cauchy entonces existe el límite  $\lim_n d(\mathbf{x}_n, \mathbf{y}_n)$ .
- b) Si la sucesión  $(\mathbf{x}_n)$  es de Cauchy y  $\lim_n d(\mathbf{x}_n, \mathbf{y}_n) = 0$  entonces la sucesión  $(\mathbf{y}_n)$  también es de Cauchy.
- c) Si las subsucesiones  $(x_{2n}), (x_{2n+1}), (x_{3n})$ , son convergentes entonces la sucesión  $(x_n)$  es convergente.

◇ **2.6.15** Sea  $\sum_{n=1}^{\infty} a_n$  una serie convergente de números reales positivos no nulos. Demuestre que en el conjunto  $E$  de las sucesiones de números reales  $\mathbf{x} = (x_n)_{n \in \mathbb{N}}$  la fórmula

$$d(\mathbf{x}, \mathbf{y}) = \sum_{n=1}^{\infty} a_n \frac{|x_n - y_n|}{1 + |x_n - y_n|}$$

define una distancia. Estudie si el espacio  $(E, d)$  es completo. Muestre que en  $(E, d)$  hay una sucesión acotada que no tiene subsucesiones convergentes.

◇ **2.6.16** Sea  $(E, d)$  un espacio métrico y  $(x_n)$  una sucesión en  $E$  que verifica

$$[T]: \quad d(x_n, x_{n+1}) < 2^{-n} \quad \text{para todo } n \in \mathbb{N}$$

Demuestre  $(x_n)$  es de Cauchy y que de cada sucesión de Cauchy en  $E$  se puede extraer una subsucesión que verifica la condición  $[T]$ .

Demuestre que  $(E, d)$  es completo si cada sucesión que verifica  $[T]$  es convergente.

◇ **2.6.17** Demuestre que un espacio métrico  $(E, d)$  es completo si y sólo si toda sucesión  $(\mathbf{x}_n)$  en  $E$  que verifique  $d(\mathbf{x}_n, \mathbf{x}_{n+1}) < 2^{-n}$  para todo  $n \in \mathbb{N}$  es convergente.

Demuestre que un espacio normado  $(E, \|\cdot\|)$  es completo si y sólo si toda serie  $\sum_{n=1}^{\infty} \mathbf{x}_n$  en  $E$  que verifique  $\sum_{n=1}^{\infty} \|\mathbf{x}_n\| < +\infty$ , es convergente.

◇ **2.6.18** Si un espacio métrico tiene la propiedad de que las bolas cerradas son compactas demuestre que el espacio es completo. ¿Es cierto que un espacio métrico localmente compacto es completo?

◇ **2.6.19** Demuestre que son compactos los siguientes subconjuntos de  $\mathbb{R}^3$ :

- a)  $\{\mathbf{x}^n : n \in \mathbb{N}\} \cup \{(0, 0, 0)\}$ , donde  $\mathbf{x}^n = (2^{-n}, 2^{-2n}, 2^{-3n})$ .
- b) Un triángulo incluidos los lados (contenido en un plano  $ax + by + cz = c$ ).
- c) Una esfera  $S(\mathbf{a}, r) := \{\mathbf{x} \in \mathbb{R}^3 : \|\mathbf{x} - \mathbf{a}\| = r\}$ , definida por una norma  $\|\cdot\|$ .

◇ **2.6.20** Si  $\|\cdot\|$  es una norma arbitraria sobre  $\mathbb{R}^3$ , y  $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$ , demuestre que el conjunto  $\{\mathbf{x} \in \mathbb{R}^3 : \|\mathbf{x} - \mathbf{a}\| + \|\mathbf{x} - \mathbf{b}\| + \|\mathbf{x} - \mathbf{c}\| \leq 10\}$  es compacto.

◇ **2.6.21** Si  $A, B$  son subconjuntos de un espacio normado  $(E, \|\cdot\|)$  demuestre las siguientes afirmaciones sobre los conjuntos

$$A + B = \{\mathbf{a} + \mathbf{b} : \mathbf{a} \in A, \mathbf{b} \in B\}; \quad [A, B] = \bigcup \{[\mathbf{a}, \mathbf{b}] : \mathbf{a} \in A, \mathbf{b} \in B\}$$

- i) Si  $A$  es abierto,  $A + B$  también lo es.

ii) Si  $A$  es cerrado y  $B$  es compacto entonces  $A+B$  es cerrado pero la conclusión es falsa cuando sólo se supone que  $A$  y  $B$  son cerrados.

iii) Si  $A$  y  $B$  son compactos entonces  $[A, B]$  también lo es.

◇ **2.6.22** Si  $T \subset \mathbb{R}$  y  $K \subset \mathbb{R}^n$  son compactos demuestre que también lo es

$$TK = \{tx : t \in T, \mathbf{x} \in K\}$$

◇ **2.6.23** Sea  $(\mathbf{x}_n)$  una sucesión convergente hacia  $\mathbf{x}$  en el espacio métrico  $(E, d)$ . Demuestre que el conjunto formado por los términos de la sucesión y su límite,  $\{\mathbf{x}\} \cup \{\mathbf{x}_n : n \in \mathbb{N}\}$ , es compacto.

◇ **2.6.24** Sea  $(E, d)$  un espacio métrico y  $A \subset E$ . Demuestre que  $\overline{A}$  es compacto si y sólo si toda sucesión en  $A$  posee una subsucesión convergente.

◇ **2.6.25** En el espacio métrico  $(\mathbb{Q}, d)$ , donde  $d$  es la distancia usual, compruebe que  $F = \{x \in \mathbb{Q} : 2 < x^2 < 3\}$  es un conjunto cerrado y acotado que no es compacto.

◇ **2.6.26** A una sucesión estrictamente creciente  $a_n \in [0, 1]$  se le asocia la sucesión de funciones  $f_n : [0, 1] \rightarrow \mathbb{R}$ , definida por

$$f_n(x) = \frac{(x - a_n)(x - a_{n+1})}{(a_{n+1} - a_n)^2} \quad \text{si } x \in [a_n, a_{n+1}]$$

$$f_n(x) = 0 \quad \text{si } x \notin [a_n, a_{n+1}].$$

Calcule  $\|f_n - f_m\|_\infty$  y deduzca que  $M = \{f_n : n \in \mathbb{N}\}$  es un subconjunto cerrado y acotado de  $(C[0, 1], \|\cdot\|_\infty)$  que no es compacto.

◇ **2.6.27** En el espacio normado  $(C[0, 1], \|\cdot\|_\infty)$  obtenga una sucesión acotada que no posea ninguna subsucesión convergente.

◇ **2.6.28** En el espacio normado  $(C[0, 1], \|\cdot\|_\infty)$  sea  $A_n$  el conjunto formado por las funciones  $f \in C[0, 1]$  que verifican

$$f(0) = 1; \quad 0 \leq f(t) \leq 1 \quad \text{si } 0 \leq t \leq 1/n; \quad f(t) = 0 \quad \text{si } 1/n \leq t \leq 1.$$

Compruebe que  $A_n$  es una sucesión decreciente de conjuntos cerrados con intersección es vacía y obtenga que la bola cerrada  $\{f \in C[0, 1] : \|f\|_\infty \leq 1\}$  no es compacta.

◇ **2.6.29** Sea  $C^1[0, 1]$  el espacio de las funciones  $f : [0, 1] \rightarrow \mathbb{R}$  que tienen derivada continua. Demuestre que  $\|f\| = |f(0)| + \|f'\|_\infty$  es una norma sobre  $C^1[0, 1]$ , y que el espacio normado  $(C^1[0, 1], \|\cdot\|)$  es completo. Demuestre que toda sucesión convergente en este espacio es uniformemente convergente. ¿Es cierto el recíproco?.

◇ **2.6.30** Obtenga la siguiente relación entre las tres normas de  $f \in C[a, b]$

$$\|f\|_1 \leq \sqrt{b-a} \|f\|_2 \leq (b-a) \|f\|_\infty$$

y deduzca las inclusiones  $\mathcal{G}_1 \subset \mathcal{G}_2 \subset \mathcal{G}_\infty$  entre las respectivas topologías. Compruebe que las inclusiones son estrictas.

◇ **2.6.31** Sea  $C_0(\mathbb{R})$  el espacio vectorial de las funciones continuas  $f : \mathbb{R} \rightarrow \mathbb{R}$  que cumplen  $\lim_{|x| \rightarrow +\infty} f(x) = 0$ , dotado de la norma  $\|f\|_\infty = \sup\{|f(x)| : x \in \mathbb{R}\}$  (véase 2.18). Demuestre que  $\overline{\mathcal{K}(\mathbb{R})} = C_0(\mathbb{R})$ , donde  $\mathcal{K}(\mathbb{R}) \subset C_0(\mathbb{R})$  es el subespacio de las funciones continuas  $f : \mathbb{R} \rightarrow \mathbb{R}$  que se anulan fuera de algún intervalo acotado.

◇ **2.6.32** Sea  $(E, d)$  un espacio métrico y  $C_b(E, \mathbb{R})$  el espacio vectorial de las funciones continuas acotadas  $f : E \rightarrow \mathbb{R}$ .

a) Demuestre que  $\|f\|_\infty = \sup\{|f(\mathbf{t})| : \mathbf{t} \in E\}$  define en  $C_b(E, \mathbb{R})$  una norma con la que  $(C_b(E, \mathbb{R}), \|\cdot\|_\infty)$  es completo.

b) Fijado  $\mathbf{x}_0 \in E$ , a cada  $\mathbf{x} \in E$  se le asocia la función  $f_{\mathbf{x}} : E \rightarrow \mathbb{R}$ , definida por  $f_{\mathbf{x}}(\mathbf{t}) = d(\mathbf{t}, \mathbf{x}) - d(\mathbf{t}, \mathbf{x}_0)$ . Compruebe que  $f_{\mathbf{x}} \in C_b(E, \mathbb{R})$  y que para cada par de puntos  $\mathbf{x}, \mathbf{y} \in E$  se verifica  $\|f_{\mathbf{x}} - f_{\mathbf{y}}\|_\infty = d(\mathbf{x}, \mathbf{y})$ .

Como consecuencia de lo anterior demuestre que existe un espacio métrico completo  $(\widehat{E}, \widehat{d})$  tal que  $E$  es isométrico a un subconjunto denso de  $\widehat{E}$ .

(Considere  $\widehat{E} = \overline{\{f_{\mathbf{a}} : \mathbf{a} \in E\}}$ , clausura en  $(C_b(E, \mathbb{R}), \|\cdot\|_\infty)$ )

◇ **2.6.33** Sea  $(E, \|\cdot\|)$  un espacio normado completo cuya norma procede de un producto escalar. Utilice la identidad del paralelogramo (problema 2.6.1):

$$\|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2 = 2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2 \quad \text{para todo } \mathbf{x}, \mathbf{y} \in E$$

para demostrar que si  $A \subset E$  es cerrado y convexo entonces existe un único  $\mathbf{a} \in A$  que verifica  $\|\mathbf{a}\| = \min\{\|\mathbf{x}\| : \mathbf{x} \in A\}$ .

# Capítulo 3

## Límites y continuidad

*Límite funcional; el problema de su existencia. Continuidad en un punto y continuidad global. Funciones continuas en conjuntos compactos. Existencia de extremos absolutos de funciones reales continuas. Espacios normados de dimensión finita. Norma de una aplicación lineal continua. Continuidad uniforme y convergencia uniforme.*

Los fundamentos del Análisis Matemático se basan esencialmente en las nociones de límite y continuidad, a las que está dedicado este capítulo. La continuidad se estudia con detalle en los cursos de topología y aquí el estudio se centra en los resultados, básicos para la teoría de funciones, referentes a las propiedades de las funciones continuas sobre los conjuntos compactos. Como aplicación se demuestra que en los espacios vectoriales de dimensión finita todas las normas son equivalentes. La caracterización de los espacios normados de dimensión finita como aquellos para los que sus bolas cerradas son compactas es un resultado interesante que se ofrece como material de carácter complementario.

Como la noción de límite interviene de manera esencial en la mayor parte de las definiciones y resultados fundamentales del Análisis Matemático, en este capítulo se consideran algunos aspectos que no se suelen enseñar en los cursos de topología general: Técnicas de cálculo de límites; condición de Cauchy para el límite funcional; utilización del límite para extender aplicaciones uniformemente continuas; papel de la convergencia uniforme en el problema del intercambio de límites. Para hacer alguna propaganda de los resultados que conciernen a este problema basta decir que la demostración de teoremas importantes del Análisis Matemático (tanto de este curso como de los posteriores) se reducen en última instancia a conseguir cambiar el orden en dos procesos iterados de paso al límite.

En el apéndice C.1 se expone el problema de intercambio de límites desde un punto de vista general haciendo intervenir la noción de límite a través de una base de filtro. Esta noción general de límite engloba a las diferentes nociones de límite que intervienen en la teoría de funciones.

### 3.1. Definiciones y resultados básicos

**Definición 3.1** Sean  $(E, d)$ ,  $(F, \rho)$  espacios métricos,  $M \subset E$ , y  $\mathbf{a} \in M'$  un punto de acumulación de  $M$  (que puede pertenecer o no a  $M$ ). Una aplicación  $\mathbf{f} : M \rightarrow F$  tiene límite  $\mathbf{b} \in F$ , cuando  $\mathbf{x}$  tiende hacia  $\mathbf{a}$ , si verifica:

Para cada  $\epsilon > 0$  existe  $\delta > 0$  tal que  $[\mathbf{x} \in M, 0 < d(\mathbf{x}, \mathbf{a}) < \delta] \Rightarrow \rho(\mathbf{f}(\mathbf{x}), \mathbf{b}) < \epsilon$

También se dice que existe el límite de  $\mathbf{f}(\mathbf{x})$ , cuando  $\mathbf{x} \rightarrow \mathbf{a}$ , y se escribe

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

Obsérvese que el límite, si existe, es único y que en la definición de límite, cuando  $\mathbf{a} \in M$ , no interviene el valor  $\mathbf{f}(\mathbf{a})$ .

Si  $\mathbf{a}$  es punto de acumulación de un subconjunto  $A \subset M$  y la restricción  $\mathbf{f}|_A : A \rightarrow F$  tiene límite  $\mathbf{b}_A$ , cuando  $\mathbf{x}$  tiende hacia  $\mathbf{a}$ , se dice que  $\mathbf{b}_A$  es el límite de  $\mathbf{f}(\mathbf{x})$ , cuando  $\mathbf{x}$  tiende hacia  $\mathbf{a}$  a través del conjunto  $A$ , y se escribe

$$\lim_{A \ni \mathbf{x} \rightarrow \mathbf{a}} \mathbf{f}(\mathbf{x}) = \mathbf{b}_A$$

Es obvio que si existe el límite  $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{f}(\mathbf{x}) = \mathbf{b}$  también existe  $\lim_{A \ni \mathbf{x} \rightarrow \mathbf{a}} \mathbf{f}(\mathbf{x}) = \mathbf{b}$ .

Esta observación es muy útil en la práctica para decidir que el límite no existe: Basta encontrar un conjunto  $A \subset M$ , con  $\mathbf{a} \in A'$ , a través del cual no exista el límite, o dos conjuntos  $A_1, A_2 \subset M$ , con  $\mathbf{a} \in A_1'$ ,  $\mathbf{a} \in A_2'$ , a través de los cuales existan y sean distintos los límites.

Para estudiar la existencia de límite la tarea se simplifica cuando disponemos de un candidato a límite (un punto  $\mathbf{b} \in F$  del se puede asegurar que si existe límite vale  $\mathbf{b}$ ) con el fin de someterlo a la definición. Si a través de algún conjunto  $A \subset M$  existe el límite y vale  $\mathbf{b}_A$ , este será el candidato a límite. Para funciones de dos variables reales hay otro método, basado en la siguiente proposición sobre límites iterados, que puede ser útil para decidir la no existencia de límite o para obtener un candidato a límite. En 3.11 y 3.12 se pueden ver ejemplos sencillos que ilustran la aplicación de estos métodos.

**Proposición 3.2** Sea  $f : M \rightarrow \mathbb{R}$  una función de dos variables reales, definida en  $M \supset \{(s, t) \in \mathbb{R}^2 : 0 < |s - a| < r, 0 < |t - b| < r\}$ , que cumple

- i) Existe el límite doble,  $\lim_{(s,t) \rightarrow (a,b)} f(s, t) = L$ ;
- ii) Cuando  $0 < |s - a| < r$ , existen los límite parciales  $\lim_{t \rightarrow b} f(s, t) = \alpha(s)$ ;

Entonces existe el límite  $\lim_{s \rightarrow a} \alpha(s) = L$ , es decir, existe el límite iterado

$$\lim_{s \rightarrow a} \left( \lim_{t \rightarrow b} f(s, t) \right) = \lim_{(s,t) \rightarrow (a,b)} f(s, t)$$

y vale lo mismo que el límite doble.

DEM: En  $\mathbb{R}^2$  utilizamos la norma  $\| \cdot \|_\infty$ . Según i), dado  $\epsilon > 0$ , existe  $0 < \delta < r$  tal que

$$0 < \max\{|s - a|, |t - b|\} < \delta \Rightarrow |f(s, t) - L| < \epsilon/2$$

Para demostrar que  $\lim_{s \rightarrow a} \alpha(s) = L$  basta ver que con este  $\delta$  se cumple

$$0 < |s - a| < \delta \Rightarrow |\alpha(s) - L| < \epsilon$$

En efecto, si  $0 < |s - a| < \delta$ , en virtud de ii), existe un punto  $t_s$ ,  $0 < |t_s - b| < \delta$ , tal que  $|f(s, t_s) - \alpha(s)| < \epsilon/2$ . Como  $0 < \max\{|s - a|, |t_s - b|\} < \delta$ , se cumple  $|f(s, t_s) - L| < \epsilon/2$ , luego  $|\alpha(s) - L| \leq |\alpha(s) - f(s, t_s)| + |f(s, t_s) - L| \leq \epsilon/2 + \epsilon/2 = \epsilon$ .

■

OBSERVACIÓN: En la proposición 3.2, si en vez de suponer ii) se supone

ii') Cuando  $0 < |t - b| < r$ , existen los límite parciales  $\lim_{s \rightarrow a} f(s, t) = \beta(t)$ , se obtiene análogamente que existe el límite  $\lim_{t \rightarrow b} \beta(t) = L$ . Por lo tanto, si  $f$  cumple las condiciones i), ii) y ii'), entonces existen los dos límites iterados y son iguales al límite doble:

$$\lim_{s \rightarrow a} (\lim_{t \rightarrow b} f(s, t)) = \lim_{t \rightarrow b} (\lim_{s \rightarrow a} f(s, t)) = L$$

Esta observación es la base de un método muy popular para decidir que una función de dos variables no tiene límite en un punto: Cuando existen y son distintos los límites iterados

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

se puede asegurar que  $f(x, y)$  no tiene límite cuando  $(x, y) \rightarrow (a, b)$ .

La proposición 3.2 también proporciona un método alternativo para obtener un candidato a límite doble. Si existe alguno de los límites iterados, y vale  $L$ , este valor es candidato a límite doble. Ejemplos sencillos ponen de manifiesto el alcance limitado de esta proposición como herramienta para abordar estudio de límites de funciones de dos variables: La existencia e igualdad de los límites iterados

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

no implica la existencia del límite ordinario  $\lim_{(x,y) \rightarrow (a,b)} f(x, y)$  (véase 3.11 a)). También puede ocurrir que exista el límite ordinario aunque no existan los límites iterados (véase 3.13).

Cuando no se dispone del candidato a límite (como ocurre al estudiar la convergencia de integrales impropias) y el espacio de llegada  $(F, \rho)$  es completo, hay un criterio de Cauchy para el límite funcional muy útil para determinar la existencia de límite, sin conocer su valor.

**Teorema 3.3** Sean  $(E, d)$ ,  $(F, \rho)$  espacios métricos y  $\mathbf{a} \in M'$  un punto de acumulación de  $M \subset E$ . Si  $(F, \rho)$  es completo, la siguiente condición (de Cauchy) es necesaria y suficiente para que  $\mathbf{f} : M \rightarrow F$  tenga límite cuando  $\mathbf{x} \rightarrow \mathbf{a}$ :

Para cada  $\epsilon > 0$  existe  $\delta > 0$  tal que

$$[\mathbf{x}, \mathbf{y} \in M, \quad 0 < d(\mathbf{x}, \mathbf{a}) < \delta, \quad 0 < d(\mathbf{y}, \mathbf{a}) < \delta] \Rightarrow \rho(\mathbf{f}(\mathbf{x}), \mathbf{f}(\mathbf{y})) < \epsilon$$



DEM: Se deja al cuidado del lector la demostración de que la condición de Cauchy es necesaria para la existencia del límite (aunque  $(F, \rho)$  no sea completo). Veamos que la condición de Cauchy es suficiente cuando  $(F, \rho)$  es completo.

Como  $\mathbf{a}$  es punto de acumulación de  $M$ , para cada  $n \in \mathbb{N}$  existe  $\mathbf{x}_n \in M$  con  $0 < d(\mathbf{x}_n, \mathbf{a}) < 1/n$ . La sucesión  $\mathbf{f}(\mathbf{x}_n) \in F$  es de Cauchy: En efecto, dado  $\epsilon > 0$ , sea  $\delta > 0$  el número proporcionado por la condición del enunciado. Eligiendo  $m \in \mathbb{N}$  de modo que  $1/m < \delta$  podemos asegurar que para  $p > q \geq m$  se cumple  $0 < d(\mathbf{x}_p, \mathbf{a}) < \delta$ ,  $0 < d(\mathbf{x}_q, \mathbf{a}) < \delta$ , luego  $\rho(\mathbf{f}(\mathbf{x}_p), \mathbf{f}(\mathbf{x}_q)) < \epsilon$ .

Como el espacio métrico  $(F, \rho)$  es completo, existe el límite  $\lim_n \mathbf{f}(\mathbf{x}_n) = \mathbf{b}$ , y para terminar basta ver que  $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{f}(\mathbf{x}) = \mathbf{b}$ . Efectivamente, si  $\mathbf{x} \in M$  y  $0 < d(\mathbf{x}, \mathbf{a}) < \delta$ , para todo  $n \geq m$  se cumple  $0 < d(\mathbf{x}_n, \mathbf{a}) \leq 1/n \leq 1/m < \delta$  luego  $\rho(\mathbf{f}(\mathbf{x}), \mathbf{f}(\mathbf{x}_n)) < \epsilon$  y aplicando la desigualdad triangular se obtiene:

$$\rho(\mathbf{f}(\mathbf{x}), \mathbf{b}) \leq \rho(\mathbf{f}(\mathbf{x}), \mathbf{f}(\mathbf{x}_n)) + \rho(\mathbf{f}(\mathbf{x}_n), \mathbf{b}) < \epsilon + \rho(\mathbf{f}(\mathbf{x}_n), \mathbf{b})$$

Cuando  $n \rightarrow +\infty$  la desigualdad  $\rho(\mathbf{f}(\mathbf{x}), \mathbf{b}) < \epsilon + \rho(\mathbf{f}(\mathbf{x}_n), \mathbf{b})$ , válida para todo  $n \geq m$ , se convierte en  $\rho(\mathbf{f}(\mathbf{x}), \mathbf{b}) \leq \epsilon$ . ■

**Definición 3.4** Sean  $(E, d)$ ,  $(F, \rho)$  espacios métricos y  $M \subset E$ . Una aplicación  $\mathbf{f} : M \rightarrow F$  es continua en  $\mathbf{a} \in M$  cuando para cada  $\epsilon > 0$  existe  $\delta > 0$  tal que

$$[\mathbf{x} \in M, d(\mathbf{x}, \mathbf{a}) < \delta] \Rightarrow \rho(\mathbf{f}(\mathbf{x}), \mathbf{f}(\mathbf{a})) < \epsilon$$

Se dice que  $\mathbf{f}$  es continua cuando es continua en cada punto de su dominio.

Si  $\mathbf{a} \in M'$  es claro que  $\mathbf{f}$  es continua en  $\mathbf{a}$  si y sólo si  $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{a})$ . Cuando  $\mathbf{a}$  es punto aislado de  $M$ , toda aplicación  $\mathbf{f} : M \rightarrow F$  es continua en  $\mathbf{a}$ .

En las condiciones de la definición 3.4 es fácil obtener la caracterización habitual de la continuidad global:  $\mathbf{f}$  es continua si y sólo si para cada abierto  $V \subset F$  existe un abierto  $U \subset E$  tal que  $\mathbf{f}^{-1}(V) = M \cap U$ .

Por otra parte, en la situación considerada en 3.4, dado  $A \subset M$ , conviene distinguir entre las afirmaciones: “ $\mathbf{f}$  es continua en cada  $\mathbf{a} \in A$ ”, y “ $\mathbf{f}|_A$  es continua”. La primera implica la segunda, pero el recíproco es falso (como se pone de manifiesto considerando  $M = E = \mathbb{R}$ ,  $A = \mathbb{Q}$ , y  $f$  la función que vale 0 sobre los racionales y 1 sobre los irracionales).

En el contexto de los espacios métricos, es útil la siguiente caracterización del límite y la continuidad mediante sucesiones:

**Proposición 3.5** Sean  $(E, d)$ ,  $(F, \rho)$  espacios métricos y  $\mathbf{f} : M \rightarrow F$  definida en  $M \subset E$ . Si  $\mathbf{a} \in M$  (resp.  $\mathbf{a} \in M'$ ), son equivalentes:

- $\mathbf{f}$  es continua en  $\mathbf{a}$  (resp. existe  $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{f}(\mathbf{x})$ ).
- $\lim_n \mathbf{f}(\mathbf{x}_n) = \mathbf{f}(\mathbf{a})$  para cada sucesión  $\mathbf{x}_n \in M$  convergente hacia  $\mathbf{a}$ . (resp.  $\mathbf{f}(\mathbf{x}_n)$  converge para cada sucesión  $\mathbf{x}_n \in M \setminus \{\mathbf{a}\}$ , convergente hacia  $\mathbf{a}$ ).

DEM: Demostramos la equivalencia de las afirmaciones sobre límites (las referentes a continuidad, que son más sencillas, se dejan al cuidado del lector).

a)  $\Rightarrow$  b): Si  $\mathbf{b} = \lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{f}(\mathbf{x})$  es inmediato que  $\mathbf{f}(\mathbf{x}_n)$  converge hacia  $\mathbf{b}$  para toda sucesión  $\mathbf{x}_n \in M \setminus \{\mathbf{a}\}$ , convergente hacia  $\mathbf{a}$ .

b)  $\Rightarrow$  a): Comencemos viendo que todas las sucesiones  $\mathbf{x}_n \in M \setminus \{\mathbf{a}\}$  que convergen hacia  $\mathbf{a}$  proporcionan el mismo límite de  $\mathbf{f}(\mathbf{x}_n)$ : Si  $\mathbf{x}_n, \mathbf{y}_n$  son sucesiones en  $M \setminus \{\mathbf{a}\}$  convergentes hacia  $\mathbf{a}$ , la sucesión mezclada  $(\mathbf{x}_1, \mathbf{y}_1, \mathbf{x}_2, \mathbf{y}_2, \dots, \mathbf{x}_n, \mathbf{y}_n, \dots)$  también converge hacia  $\mathbf{a}$  luego, por hipótesis, su imagen mediante  $\mathbf{f}$  es una sucesión convergente. Como  $\mathbf{f}(\mathbf{x}_n), \mathbf{f}(\mathbf{y}_n)$  son subsucesiones de esta sucesión, sus límites deben ser iguales.

Para terminar demostramos, por reducción al absurdo, que el límite común  $\mathbf{b}$  de las sucesiones  $\mathbf{f}(\mathbf{x}_n)$  consideradas en b) es el límite de  $\mathbf{f}(\mathbf{x})$  cuando  $\mathbf{x} \rightarrow \mathbf{a}$ . En caso contrario existe  $\epsilon > 0$  tal que para todo  $\delta > 0$  hay un  $\mathbf{x}_\delta \in M$ , con  $0 < d(\mathbf{x}_\delta, \mathbf{a}) < \delta$ , que verifica  $\rho(\mathbf{f}(\mathbf{x}_\delta), \mathbf{b}) \geq \epsilon$ . Tomando  $\delta$  de la forma  $1/n$  encontramos una sucesión  $\mathbf{x}_n \in M \setminus \{\mathbf{a}\}$ , convergente hacia  $\mathbf{a}$  tal que  $\mathbf{f}(\mathbf{x}_n)$  no converge hacia  $\mathbf{b}$ , y con esta contradicción termina la demostración. ■

Hay una caracterización similar de la continuidad, mediante la conservación de puntos adherentes, que se deja como ejercicio al cuidado del lector:

**Proposición 3.6** Si  $(E, d), (F, \rho)$  son espacios métricos y  $\mathbf{f} : M \rightarrow F$  está definida en  $M \subset E$ , se verifica:

a)  $f$  es continua en  $\mathbf{a} \in M$  si y sólo si para cada  $A \subset M$  con  $\mathbf{a} \in \overline{A}$  se cumple  $\mathbf{f}(\mathbf{a}) \in \overline{\mathbf{f}(A)}$ .

b)  $\mathbf{f}$  es continua si y sólo si  $\mathbf{f}(\overline{A} \cap M) \subset \overline{\mathbf{f}(A)}$  para cada  $A \subset M$ .

## 3.2. Reglas para obtener el límite y la continuidad

Las proposiciones que siguen, de carácter elemental, son útiles para obtener el límite o la continuidad de funciones concretas sin acudir a la definición. La primera de ellas pone de manifiesto que para funciones con valores en  $\mathbb{R}^n$ , el estudio se reduce al caso  $n = 1$ .

**Proposición 3.7** Sea  $\mathbf{f} : M \rightarrow \mathbb{R}^n$ ,  $\mathbf{f} = (f_1, f_2, \dots, f_n)$ , definida en un subconjunto  $M$  del espacio métrico  $(E, d)$ , y  $\mathbf{a} \in M'$  (resp.  $\mathbf{a} \in M$ ). Son equivalentes:

i) Existe el límite  $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{f}(\mathbf{x}) = (b_1, b_2, \dots, b_n)$ , (resp.  $\mathbf{f}$  es continua en  $\mathbf{a}$ ).

ii) Para  $1 \leq j \leq n$  existe  $\lim_{\mathbf{x} \rightarrow \mathbf{a}} f_j(\mathbf{x}) = b_j$  (resp.  $f_j$  es continua en  $\mathbf{a}$ ).

DEM: Si  $\mathbf{b} = (b_1, b_2, \dots, b_n)$  la afirmación sobre límites es consecuencia de las desigualdades

$$|f_j(\mathbf{x}) - b_j| \leq \|\mathbf{f}(\mathbf{x}) - \mathbf{b}\|_2 \leq \sum_{j=1}^n |f_j(\mathbf{x}) - b_j|$$

válidas para todo  $j \in \{1, \dots, n\}$ . (La afirmación alternativa sobre continuidad se obtiene con  $\mathbf{b} = \mathbf{f}(\mathbf{a})$ ). ■

**Proposición 3.8** Sea  $(E, d)$  un espacio métrico,  $M \subset E$ ,  $\mathbf{a} \in M'$ , y  $(F, \|\cdot\|)$  un espacio normado sobre  $\mathbb{K}$  ( $= \mathbb{R}$  o  $\mathbb{C}$ ). Dadas las funciones  $\mathbf{f}, \mathbf{g} : M \rightarrow F$ ,  $\alpha : M \rightarrow \mathbb{K}$ , se verifica:

a) Si  $\mathbf{f}, \mathbf{g}$  tienen límite cuando  $\mathbf{x} \rightarrow \mathbf{a}$ , entonces también lo tiene  $\mathbf{f} + \mathbf{g}$ , y

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

Si las funciones  $\mathbf{f}, \mathbf{g}$  son continuas en  $\mathbf{a} \in M$ , también lo es  $\mathbf{f} + \mathbf{g}$ .

b) Si  $\alpha, \mathbf{g}$  tienen límite cuando  $\mathbf{x} \rightarrow \mathbf{a}$ , también lo tiene su producto  $\alpha\mathbf{g}$ , y

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} (\alpha(\mathbf{x})\mathbf{g}(\mathbf{x})) = \left( \lim_{\mathbf{x} \rightarrow \mathbf{a}} \alpha(\mathbf{x}) \right) \left( \lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{g}(\mathbf{x}) \right)$$

Si las funciones  $\alpha, \mathbf{g}$  son continuas en  $\mathbf{a} \in M$  también lo es  $\alpha\mathbf{g}$ .

c) Si las funciones  $\alpha, \mathbf{g}$  tienen límite cuando  $\mathbf{x} \rightarrow \mathbf{a}$  y  $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \alpha(\mathbf{x}) \neq 0$ , también tiene límite  $(1/\alpha)\mathbf{g}$  (definida en  $M \cap B(\mathbf{a}, r)$  para algún  $r > 0$ ) y

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} (\alpha(\mathbf{x})^{-1}\mathbf{g}(\mathbf{x})) = \left( \lim_{\mathbf{x} \rightarrow \mathbf{a}} \alpha(\mathbf{x}) \right)^{-1} \left( \lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{g}(\mathbf{x}) \right)$$

Si las funciones  $\alpha, \mathbf{g}$  son continuas en  $\mathbf{a} \in M$  y  $\alpha(\mathbf{a}) \neq 0$ , también lo es  $\alpha^{-1}\mathbf{g}$ .

d) Si la norma de  $F$  procede de un producto escalar  $\langle \cdot | \cdot \rangle$ , y las funciones  $\mathbf{f}, \mathbf{g}$  tienen límite cuando  $\mathbf{x} \rightarrow \mathbf{a}$ , también lo tiene el producto  $\langle \mathbf{f}(\mathbf{x}) | \mathbf{g}(\mathbf{x}) \rangle$ , y

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} \langle \mathbf{f}(\mathbf{x}) | \mathbf{g}(\mathbf{x}) \rangle = \left\langle \lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{f}(\mathbf{x}) \mid \lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{g}(\mathbf{x}) \right\rangle$$

En particular, si las funciones  $\mathbf{f}, \mathbf{g}$  son continuas en  $\mathbf{a} \in M$ , también lo es el producto  $\langle \mathbf{f}(\mathbf{x}) | \mathbf{g}(\mathbf{x}) \rangle$ .

DEM: a), b) y c) se demuestran como en el caso de las funciones reales de variable real. La demostración de d) es análoga a la de b), pero utilizando la desigualdad de Cauchy Schwarz: Sea  $\mathbf{b} = \lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{f}(\mathbf{x})$ , y  $\mathbf{b}' = \lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{g}(\mathbf{x})$ . Existe  $r > 0$  tal que si  $\mathbf{x} \in M$  y  $0 < d(\mathbf{x}, \mathbf{a}) < r$  entonces  $\|\mathbf{f}(\mathbf{x}) - \mathbf{b}\| < 1$ , luego  $\|\mathbf{f}(\mathbf{x})\| \leq 1 + \|\mathbf{b}\|$ . Si escribimos

$$\langle \mathbf{f}(\mathbf{x}) | \mathbf{g}(\mathbf{x}) \rangle = \langle \mathbf{f}(\mathbf{x}) | \mathbf{g}(\mathbf{x}) - \mathbf{b}' \rangle + \langle \mathbf{f}(\mathbf{x}) | \mathbf{b}' \rangle$$

basta observar que, cuando  $\mathbf{x} \rightarrow \mathbf{a}$ , el primer sumando de la derecha tiende hacia 0 y el segundo tiende hacia  $\langle \mathbf{b} | \mathbf{b}' \rangle$ . Esto es consecuencia de las desigualdades

$$\text{i) } |\langle \mathbf{f}(\mathbf{x}) | \mathbf{g}(\mathbf{x}) - \mathbf{b}' \rangle| \leq \|\mathbf{f}(\mathbf{x})\| \|\mathbf{g}(\mathbf{x}) - \mathbf{b}'\| \leq (1 + \|\mathbf{b}\|) \|\mathbf{g}(\mathbf{x}) - \mathbf{b}'\|.$$

$$\text{ii) } |\langle \mathbf{f}(\mathbf{x}) | \mathbf{b}' \rangle - \langle \mathbf{b} | \mathbf{b}' \rangle| = |\langle \mathbf{f}(\mathbf{x}) - \mathbf{b} | \mathbf{b}' \rangle| \leq \|\mathbf{f}(\mathbf{x}) - \mathbf{b}\| \|\mathbf{b}'\|.$$

que se cumplen cuando  $\mathbf{x} \in M$ , y  $0 < d(\mathbf{x}, \mathbf{a}) < r$ . ■

**Proposición 3.9** [Regla de la cadena] Para  $j = 1, 2, 3$  sea  $(E_j, d_j)$ , un espacio métrico y  $M_j$  un subconjunto de  $E_j$ . Sea supone que  $\mathbf{f} : M_1 \rightarrow M_2$  tiene límite  $\mathbf{b} \in M_2$  cuando  $\mathbf{x} \rightarrow \mathbf{a} \in M_1'$ . Si  $\mathbf{g} : M_2 \rightarrow E_3$  es continua en  $\mathbf{b}$  entonces existe el límite de la función compuesta

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

Si  $\mathbf{a} \in M_1$ , y  $\mathbf{f}$  es continua en  $\mathbf{a}$ , entonces  $\mathbf{g} \circ \mathbf{f}$  también es continua en  $\mathbf{a}$ .

DEM: Se deja como ejercicio. ■

Las siguientes observaciones son útiles para establecer la continuidad de funciones de  $n$  variables reales sin acudir a la definición  $\epsilon, \delta$ :

i) Sea  $L = \{n_1, n_2, \dots, n_k\} \subset \{1, 2, \dots, n\}$ , donde  $1 \leq n_1 < n_2 < \dots < n_k \leq n$ . La proyección  $\pi_L : \mathbb{R}^n \rightarrow \mathbb{R}^k$ ,  $\pi_L(x_1, x_2, \dots, x_n) = (x_{n_1}, x_{n_2}, \dots, x_{n_k})$ , es continua en todo punto.

ii) Si una función  $f$  de  $k$  variables reales es continua en  $\mathbf{a} = (a_1, a_2, \dots, a_k)$  entonces para cada  $n > k$ , y cada conjunto finito  $L = \{n_1, n_2, \dots, n_k\} \subset \{1, 2, \dots, n\}$ , la función de  $n$  variables reales  $F(x_1, x_2, \dots, x_n) = f(x_{n_1}, x_{n_2}, \dots, x_{n_k})$ , es continua en cada punto  $\mathbf{p} \in \mathbb{R}^n$  tal que  $\pi_L(\mathbf{p}) = \mathbf{a}$ , es decir, la continuidad se conserva cuando una función de  $k$  variables se considera como función de más variables.

iii) Los polinomios en  $n$  variables reales definen funciones continuas en todo  $\mathbb{R}^n$ : Como las proyecciones  $\pi_k : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $\pi_k(x_1, x_2, \dots, x_n) = x_k$  son continuas, también son continuas las funciones definidas mediante un producto finito de proyecciones

$$(x_1, x_2, \dots, x_n) \rightarrow x_1^{p_1} x_2^{p_2} \dots x_n^{p_n}, \quad (p_k \in \{0, 1, 2, \dots\})$$

y sus combinaciones lineales, que son los polinomios en  $n$  variables reales.

Combinando estas observaciones con los resultados de las proposiciones 3.8 y 3.9 se obtiene fácilmente la continuidad en todo punto de funciones de dos y tres variables,  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $g : \mathbb{R}^3 \rightarrow \mathbb{R}$ , como las siguientes

$$f(x, y) = \sin x + \cos xy^2 + e^y; \quad g(x, y, z) = \frac{xe^y}{1+z^2} + x + z$$

que se expresan combinando polinomios con funciones elementales de una variable mediante operaciones que preservan la continuidad. Para funciones de este tipo no es recomendable intentar establecer la continuidad acudiendo a la definición  $\epsilon, \delta$ . Sólo se debe recurrir a la definición cuando las reglas anteriores no son aplicables. Incluso en este caso conviene examinar atentamente la función pues a veces alguna consideración preliminar permite aplicarlas, como ocurre en el siguiente ejemplo.

**Ejemplo 3.10** La función  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  definida por

$$f(x, y) = \frac{\sin(x^2 + y^2)}{x^2 + y^2}, \quad \text{si } (x, y) \neq (0, 0); \quad f(0, 0) = 1$$

es continua en todo  $\mathbb{R}^2$ .

Las reglas anteriores, que se aplican directamente en cada punto  $(x, y) \neq (0, 0)$ , permiten afirmar que  $f$  es continua en todo punto de  $\mathbb{R}^2 \setminus \{(0, 0)\}$ . Para obtener la continuidad en el punto  $(0, 0)$ , donde parece que no se pueden aplicar estas reglas, basta considerar la función real de variable real

$$\varphi(t) = \frac{\operatorname{sen} t}{t}, \quad \text{si } t \neq 0; \quad \varphi(0) = 1$$

que es continua en  $t = 0$ , ya que  $\lim_{t \rightarrow 0} \varphi(t) = 1 = \varphi(0)$ . Entonces, en virtud de la regla de la cadena, la función compuesta  $f(x, y) = \varphi(x^2 + y^2)$  es continua en  $(0, 0)$

Con un razonamiento similar se establece que si  $g$  es una función de dos variables definida y continua en  $\mathbb{R}^2$ , entonces la función

$$f(x, y) = \frac{\operatorname{sen} g(x, y)}{g(x, y)} \quad \text{si } g(x, y) \neq 0; \quad f(x, y) = 1 \quad \text{si } g(x, y) = 0;$$

es continua en todo punto. ■

**Ejemplos 3.11** *Estudio de la existencia de límite, cuando  $(x, y) \rightarrow (0, 0)$ , de las siguientes funciones de dos variables definidas en  $M = \mathbb{R}^2 \setminus \{(0, 0)\}$ :*

a)  $f(x, y) = \frac{xy}{x^2 + y^2};$

b)  $g(x, y) = 1$  si  $0 < |y| < 2x^2$ ,  $g(x, y) = 0$  en los restantes puntos;

c)  $h(x, y) = \frac{xy^2}{x^2 + y^2}.$

a) Aunque existen y son iguales los límites iterados

$$\lim_{x \rightarrow 0} \left( \lim_{y \rightarrow 0} f(x, y) \right) = \lim_{y \rightarrow 0} \left( \lim_{x \rightarrow 0} f(x, y) \right) = 0$$

sin embargo no existe el límite de  $f(x, y)$  cuando  $(x, y) \rightarrow (0, 0)$  porque a través de cada  $A_m = \{(x, y) : y = mx, x \neq 0\} \subset M$  el límite de  $f(x, y)$  existe y vale  $m/(1 + m^2)$ . (Si existiese el límite de  $f(x, y)$  estos límites deberían ser iguales).

La no existencia del límite de  $f(x, y)$  en  $(0, 0)$  significa que es imposible definir  $f(0, 0)$  para conseguir una extensión continua de  $f$  a todo  $\mathbb{R}^2$ .

Definiendo  $f(0, 0) = 0$ , se comprueba fácilmente que todas las funciones parciales  $x \rightarrow f(x, b)$ ,  $y \rightarrow f(a, y)$ , son continuas en todo punto. Con este ejemplo se pone de manifiesto que la continuidad, en cualquier punto, de las las funciones parciales no implica que la función dada sea continua en todo punto.

Representando gráficamente la función  $t/(1 + t^2)$  se observa que todo  $z \in (-1/2, 1/2)$  es de la forma  $z = m/(1 + m^2)$  para algún  $m \in \mathbb{R}$ , luego, en todo entorno del punto  $(0, 0, z)$  hay puntos de la forma  $(x, mx, f(x, mx)) = (x, mx, z)$ ,  $x \neq 0$ , que pertenecen a la gráfica  $G(f)$ .

Es decir, el segmento  $S = \{(0, 0, z) : |z| < 1/2\}$  está contenido en  $\overline{G(f)}$ , y lo mismo le ocurre al segmento cerrado  $\overline{S} = \{(0, 0, z) : |z| \leq 1/2\}$ . Este comportamiento

de  $G(f)$  cerca de  $(0, 0)$  se puede visualizar con DPgraph.

b) Sean  $A_m$  los conjuntos considerados en a). Dibujando un esquema de la región  $U = \{(x, y) : g(x, y) = 0\}$  se observa que para cada  $m \in \mathbb{R}$  existe  $\delta_m > 0$  tal que si  $(x, y) \in A_m$  y  $\|(x, y)\|_2 < \delta_m$  entonces  $(x, y) \in U$ , luego  $g(x, y) = 0$ . Vemos así que a través de cada  $A_m$  existe el límite de  $g$  y vale 0 y esto permite afirmar que 0 es un candidato a límite. En este caso el límite no existe porque a través de  $A = \{(x, y) : y = x^2, x \neq 0\}$  la función  $g$  tiene límite 1 (ya que vale 1 en los puntos de  $A$ ). Definiendo  $g(0, 0) = 0$  se consigue una extensión de  $g$  a todo  $\mathbb{R}^2$ , que no es continua en  $(0, 0)$ , pero tiene una restricción continua a cada recta que pasa por  $(0, 0)$ .

c)  $h$  tiene límite 0 a través de cada  $A_m = \{(x, y) : y = mx, x \neq 0\} \subset M$  pues  $h(x, mx) = xm^2/(1 + m^2)$ , luego 0 es el candidato a límite. En este caso es fácil ver que  $h(x, y)$  tiene límite 0 pues  $|h(x, y)| \leq |x| \leq \|(x, y)\|_2$ . Por lo tanto, definiendo  $h(0, 0) = 0$  se consigue una extensión continua de  $h$  a todo  $\mathbb{R}^2$ . ■

**Ejemplos 3.12** *Estudio de la existencia de límite, cuando  $(x, y) \rightarrow (0, 0)$ , de las siguientes funciones de dos variables definidas en  $M = \mathbb{R}^2 \setminus \{(0, 0)\}$ :*

a)  $f(x, y) = (x^2 - y^2)/(x^2 + y^2)$ ;

b)  $g(x, y) = (x^2 + y^2)/(x^2 + y)$  si  $x^2 + y \neq 0$ ,  $g(x, -x^2) = 0$ ;

a) La función  $f$  no tiene límite en  $(0, 0)$  porque existen y son distintos los límites iterados:  $\lim_{y \rightarrow 0} f(x, y) = 1$  si  $x \neq 0$ ;  $\lim_{x \rightarrow 0} f(x, y) = -1$  si  $y \neq 0$ , luego

$$\lim_{x \rightarrow 0} \left( \lim_{y \rightarrow 0} f(x, y) \right) = 1, \quad \lim_{y \rightarrow 0} \left( \lim_{x \rightarrow 0} f(x, y) \right) = -1.$$

También se puede razonar considerando los límites a través de las rectas perforadas  $A_m = \{(x, mx) : x \neq 0\}$ , que existen y son distintos.

b) La función  $g$  tampoco tiene límite en  $(0, 0)$ . Efectivamente:

$$\lim_{y \rightarrow 0} g(x, y) = 1, \text{ si } x \neq 0, \text{ luego } \lim_{x \rightarrow 0} \left( \lim_{y \rightarrow 0} g(x, y) \right) = 1.$$

$$\lim_{x \rightarrow 0} g(x, y) = y, \text{ si } y \neq 0, \text{ luego } \lim_{y \rightarrow 0} \left( \lim_{x \rightarrow 0} g(x, y) \right) = 0.$$

Se llega a la misma conclusión observando que los límites a través de las parábolas perforadas  $P_m = \{(x, mx^2) : x \neq 0\}$  existen y son distintos. ■

**Ejemplo 3.13** *Una función de dos variables  $g(x, y)$  que tiene límite ordinario en  $(0, 0)$ , aunque no existen los correspondientes límites iterados.*

Sea  $g(x, y) = x \operatorname{sen}(x/y) + y \operatorname{sen}(y/x)$ , si  $x \neq 0, y \neq 0$ ;  $g(0, y) = g(x, 0) = 0$ .

Esta función, definida en  $\mathbb{R}^2 \setminus \{(0, 0)\}$ , tiene límite 0 cuando  $(x, y) \rightarrow (0, 0)$ , porque  $|g(x, y)| \leq |x| + |y| = \|(x, y)\|_1$ . Sin embargo no existen los límites iterados en  $(0, 0)$ , porque ni siquiera existen los límites parciales: Si  $x \neq 0$  no existe  $\lim_{y \rightarrow 0} g(x, y)$  (porque  $\lim_{y \rightarrow 0} y \operatorname{sen}(y/x) = 0$ , pero no existe  $\lim_{y \rightarrow 0} x \operatorname{sen}(x/y) = 0$ ). Análogamente, si  $y \neq 0$ , tampoco existe  $\lim_{x \rightarrow 0} g(x, y)$ . ■

### 3.3. Funciones continuas en conjuntos compactos

En esta sección se demuestran, usando la técnica de las sucesiones, las propiedades básicas de las funciones continuas sobre los subconjuntos compactos de los espacios métricos.

**Teorema 3.14** Sean  $(E, d)$ ,  $(F, \rho)$  espacios métricos. Si  $K \subset E$  es compacto y  $f : K \rightarrow F$  es continua entonces  $f(K)$  es compacto.

DEM: La demostración que se ofrece está basada en el teorema 2.7 y en la caracterización de la continuidad mediante sucesiones (proposición 3.5). Si  $K \subset E$  es compacto, para demostrar que  $f(K)$  también lo es basta ver que toda sucesión  $y_n \in f(K)$  tiene una subsucesión convergente hacia un punto de  $f(K)$ . Efectivamente,  $y_n = f(x_n)$  donde  $x_n \in K$ . Como  $K$  es compacto, la sucesión  $x_n$  posee una subsucesión  $x_{n_k}$  que converge hacia un punto  $x \in K$ . Como  $f$  es continua, la subsucesión  $y_{n_k} = f(x_{n_k})$  converge hacia  $f(x) \in f(K)$ . ■

Aunque la siguiente proposición se puede obtener como corolario directo del teorema 3.14 ofrecemos una demostración con la técnica de las sucesiones para insistir en ella, dando un ejemplo de cómo se suele utilizar el corolario 2.8.

**Proposición 3.15** Sean  $(E, d)$ ,  $(F, \rho)$  espacios métricos y  $f : X \rightarrow F$  una aplicación continua e inyectiva definida en un compacto  $X \subset E$ . Entonces la inversa,  $f^{-1} : Y \rightarrow X$ , definida en  $Y = f(X)$ , es continua.

DEM: Según la proposición 3.5 basta demostrar que si  $y_n$  es una sucesión en  $Y$  que converge hacia  $y$  entonces  $x_n = f^{-1}(y_n)$  converge hacia  $x = f^{-1}(y)$ .

Como la sucesión  $x_n$  está contenida en el compacto  $X$ , basta demostrar que  $x$  es su único punto de aglomeración (2.8): Si  $x' = \lim_k x_{n_k}$  es un punto de aglomeración de esta sucesión, en virtud de la continuidad de  $f$ ,

$$f(x') = \lim_k f(x_{n_k}) = \lim_k y_{n_k} = y = f(x)$$

y usando que  $f$  es inyectiva se concluye que  $x = x'$ . ■

**Teorema 3.16** Toda función continua  $f : K \rightarrow \mathbb{R}$ , definida en un subconjunto compacto  $K$  de un espacio métrico, alcanza en  $K$  un máximo y un mínimo absoluto: Existen  $\mathbf{a}, \mathbf{b} \in K$  tales que  $f(\mathbf{a}) \leq f(\mathbf{x}) \leq f(\mathbf{b})$ , para todo  $\mathbf{x} \in K$ .

DEM: Es consecuencia inmediata del teorema 3.14 teniendo en cuenta que el extremo inferior y el extremo superior de un conjunto acotado de números reales son valores adherentes a dicho conjunto. Como  $f(K) \subset \mathbb{R}$  es cerrado y acotado los números  $\alpha = \inf f(K)$ ,  $\beta = \sup f(K)$  pertenecen a  $f(K)$  es decir, existen  $\mathbf{a}, \mathbf{b} \in K$  tales que  $\alpha = f(\mathbf{a})$  y  $\beta = f(\mathbf{b})$ . ■

En ciertas situaciones, en ausencia de compacidad, es posible utilizar el resultado anterior para justificar que una función real continua alcanza un máximo o un mínimo absoluto. En la siguiente proposición se describe una situación que se presenta con bastante frecuencia.

**Proposición 3.17** Si  $f : A \rightarrow \mathbb{R}$  es una función continua definida en un conjunto cerrado no acotado  $A \subset \mathbb{R}^n$ , se verifica:

- a) Si  $\lim_{\|\mathbf{x}\|_2 \rightarrow +\infty} f(\mathbf{x}) = +\infty$  entonces  $f$  alcanza en  $A$  un mínimo absoluto.
- b) Si  $\lim_{\|\mathbf{x}\|_2 \rightarrow +\infty} f(\mathbf{x}) = -\infty$  entonces  $f$  alcanza en  $A$  un máximo absoluto.
- c) Si  $\lim_{\|\mathbf{x}\|_2 \rightarrow +\infty} f(\mathbf{x}) = L$  entonces  $f$  alcanza en  $A$  un máximo absoluto o un mínimo absoluto.

DEM: a) Elegimos un punto  $\mathbf{z} \in A$  y una constante  $M > f(\mathbf{z})$ . En virtud de la hipótesis existe  $r > 0$  tal que si  $\mathbf{x} \in A$  y  $\|\mathbf{x}\|_2 > r$  entonces  $f(\mathbf{x}) > M$ . El conjunto  $K = \{\mathbf{x} \in A : \|\mathbf{x}\|_2 \leq r\}$  es compacto, (porque es un subconjunto cerrado y acotado de  $\mathbb{R}^n$ ) y por lo tanto  $f|_K$  alcanza en  $K$  un mínimo absoluto  $f(\mathbf{a})$  en algún punto  $\mathbf{a} \in K \subset A$ . Es claro que  $\mathbf{z} \in K$ , luego  $f(\mathbf{a}) \leq f(\mathbf{z})$  y se sigue que para todo  $\mathbf{x} \in A \setminus K$  también se cumple  $f(\mathbf{x}) \geq f(\mathbf{a})$  porque  $f(\mathbf{x}) > M > f(\mathbf{z}) \geq f(\mathbf{a})$ . Queda demostrado que  $f$  alcanza en  $\mathbf{a}$  un mínimo absoluto.

b) Basta aplicar el resultado anterior a la función  $-f$ .

c) Si  $f$  es constante  $= L$  el resultado es trivial. En otro caso, si existe  $\mathbf{z} \in A$  con  $f(\mathbf{z}) > L$ , con un razonamiento análogo al realizado en el apartado a) se demuestra que  $f$  alcanza en  $A$  un máximo absoluto. Análogamente, si  $f(\mathbf{w}) < L$  para algún  $\mathbf{w} \in A$  entonces  $f$  alcanza en  $A$  un mínimo absoluto. ■

Dado un cerrado  $A \subset \mathbb{R}^n$  y un punto  $\mathbf{p} \in \mathbb{R}^n \setminus A$ , en virtud de la proposición 3.17 podemos asegurar que la función continua  $f(\mathbf{x}) = \|\mathbf{p} - \mathbf{x}\|_2$  alcanza en  $A$  un mínimo absoluto. En particular, si  $\mathbf{g} : \mathbb{R}^n \rightarrow \mathbb{R}^m$  es continua y el conjunto  $A = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{g}(\mathbf{x}) = \mathbf{0}\}$  no es vacío, entre todos los puntos de  $A$  hay uno  $\mathbf{a} \in A$  que es el más cercano a  $\mathbf{p} \in \mathbb{R}^n \setminus A$  con la distancia usual. Este resultado se aplica, en particular, cuando  $A \subset \mathbb{R}^3$  es una curva o superficie definida de forma implícita.

### 3.4. Espacios normados de dimensión finita

Una consecuencia importante del teorema 3.16 es el siguiente resultado que fue anunciado en el capítulo 2

**Teorema 3.18** Sobre un espacio vectorial de dimensión finita (real o complejo) todas las normas son equivalentes.

DEM: Comenzamos demostrando que cualquier norma  $\|\cdot\|$  sobre  $\mathbb{R}^n$  es equivalente a la norma euclídea. Según la proposición 2.4 debemos encontrar  $\alpha > 0$ ,  $\beta > 0$ , tales que para todo  $\mathbf{x} \in \mathbb{R}^n$  se cumpla:

$$\alpha \|\mathbf{x}\|_2 \leq \|\mathbf{x}\| \leq \beta \|\mathbf{x}\|_2$$

Si  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  es la base canónica de  $\mathbb{R}^n$ , aplicando la desigualdad triangular se obtiene que  $\beta = \sum_{k=1}^n \|\mathbf{e}_k\|$  cumple la desigualdad de la derecha:

$$\|\mathbf{x}\| = \left\| \sum_{k=1}^n x_k \mathbf{e}_k \right\| \leq \sum_{k=1}^n |x_k| \|\mathbf{e}_k\| \leq \sum_{k=1}^n \|\mathbf{x}\|_2 \|\mathbf{e}_k\| = \beta \|\mathbf{x}\|_2$$



Esta desigualdad implica que la función  $\mathbf{x} \rightarrow \|\mathbf{x}\|$  es continua sobre  $\mathbb{R}^n$  con su topología usual (la asociada a  $\|\cdot\|_2$ ) y por lo tanto alcanza un mínimo absoluto sobre el compacto  $K = \{\mathbf{y} \in \mathbb{R}^n : \|\mathbf{y}\|_2 = 1\}$ , es decir, existe  $\mathbf{a} \in K$  tal que  $\|\mathbf{a}\| \leq \|\mathbf{y}\|$  para todo  $\mathbf{y} \in K$ . El número  $\alpha = \|\mathbf{a}\| > 0$  cumple la desigualdad de la izquierda: Basta considerar  $\mathbf{x} \neq 0$ ; como  $\mathbf{y} = \mathbf{x}/\|\mathbf{x}\|_2 \in K$ , se obtiene  $\alpha \leq \|\mathbf{y}\| = \|\mathbf{x}\|/\|\mathbf{x}\|_2$ .

Sea ahora  $E$  un espacio vectorial de dimensión  $n$  sobre el cuerpo  $\mathbb{R}$ . Fijada una base  $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$  de  $E$  podemos identificar  $E$  con  $\mathbb{R}^n$  mediante el isomorfismo algebraico  $S: \mathbb{R}^n \rightarrow E$ , que asocia a  $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$  el vector  $S(\mathbf{x}) = \sum_{k=1}^n x_k \mathbf{u}_k$ .

Si  $\|\cdot\|, \|\cdot\|'$  son normas sobre  $E$  entonces  $\|\mathbf{x}\|_S = \|S(\mathbf{x})\|$ , y  $\|\mathbf{x}\|'_S = \|S(\mathbf{x})\|'$  son normas equivalentes sobre  $\mathbb{R}^n$ . Como  $S$  es sobreyectiva, se sigue que las normas  $\|\cdot\|, \|\cdot\|'$  son equivalentes. ■

**Proposición 3.19** *Todo espacio normado de dimensión finita  $(E, \|\cdot\|)$  es completo y su bola unidad  $B_E = \{\mathbf{x} \in E : \|\mathbf{x}\| \leq 1\}$  es compacta.*

DEM: Sea  $n$  la dimensión de  $E$  como espacio vectorial sobre  $\mathbb{R}$ ,  $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$  una base de  $E$  y  $S: \mathbb{R}^n \rightarrow E$  el isomorfismo asociado  $S(\mathbf{x}) = \sum_{k=1}^n x_k \mathbf{u}_k$ .

Es fácil ver que  $\|\mathbf{x}\|_S = \|S(\mathbf{x})\|_E$  es una norma sobre  $\mathbb{R}^n$  cuya bola unidad

$$K = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_S \leq 1\} = S^{-1}(B_E)$$

es un conjunto cerrado y acotado. Como esta norma es equivalente a la norma euclídea,  $K$  también es cerrado y acotado para ella, y por lo tanto es compacto. Para obtener que  $B_E = S(K)$  es compacto basta tener en cuenta que  $S$  es continua (es una isometría entre  $(\mathbb{R}^n, \|\cdot\|_S)$  y  $(E, \|\cdot\|)$ ), y por lo tanto transforma compactos en compactos.

El hecho de que la bola  $B_E$  sea compacta implica que las bolas cerradas de  $E$  son compactas. Toda sucesión de Cauchy  $(\mathbf{y}_n)$  en  $E$  es acotada luego está contenida en alguna bola compacta  $\{\mathbf{y} \in E : \|\mathbf{y}\| \leq R\}$  y por lo tanto posee una subsucesión convergente. Basta recordar que toda sucesión de Cauchy con una subsucesión convergente es convergente para obtener que  $(E, \|\cdot\|)$  es completo. ■

**Teorema 3.20** *Un espacio normado  $(E, \|\cdot\|)$  (real o complejo) es de dimensión finita si y sólo si su bola unidad  $B_E = \{\mathbf{x} \in E : \|\mathbf{x}\| \leq 1\}$  es compacta.*

DEM: Después de la proposición 3.19 queda por demostrar que si la bola  $B_E$  es compacta entonces  $E$  finito dimensional. Efectivamente, fijado  $0 < \epsilon < 1$ , podemos recubrir el compacto  $B_E$  con un número finito de bolas abiertas de radio  $\epsilon > 0$ ,

$$B_E \subset \bigcup_{j=1}^m B(\mathbf{a}_j, \epsilon)$$

Si  $F \subset E$  es el subespacio finito dimensional engendrado por  $\{\mathbf{a}_j : 1 \leq j \leq m\}$  demostraremos que  $E = F$ . Lo haremos por reducción al absurdo suponiendo que existe  $\mathbf{x} \in E \setminus F$ . En virtud de la proposición 3.19  $F$  es un subespacio completo y por lo tanto cerrado. Entonces podemos asegurar que

$$\alpha = d(\mathbf{x}, F) := \inf\{\|\mathbf{x} - \mathbf{y}\| : \mathbf{y} \in F\}$$

no es nulo. Como  $\alpha < \alpha/\epsilon$ , la definición de extremo inferior asegura que existe  $\mathbf{y} \in F$  tal que  $\alpha \leq \|\mathbf{x} - \mathbf{y}\| < \alpha/\epsilon$ . El vector  $\mathbf{z} = (\mathbf{x} - \mathbf{y})/\|\mathbf{x} - \mathbf{y}\|$  pertenece a  $B_E$ , luego  $\|\mathbf{z} - \mathbf{a}_j\| < \epsilon$  para algún  $j \in \{1, 2, \dots, m\}$ . Entonces el vector  $\mathbf{x}$  se puede escribir en la forma

$$\mathbf{x} = \mathbf{y} + \|\mathbf{x} - \mathbf{y}\| \mathbf{z} = (\mathbf{y} + \|\mathbf{x} - \mathbf{y}\| \mathbf{a}_j) + \|\mathbf{x} - \mathbf{y}\| (\mathbf{z} - \mathbf{a}_j) = \mathbf{u} + \mathbf{v}$$

donde  $\mathbf{u} = \mathbf{y} + \|\mathbf{x} - \mathbf{y}\| \mathbf{a}_j$  pertenece a  $F$ . Se sigue que

$$\alpha \leq \|\mathbf{x} - \mathbf{u}\| = \|\mathbf{v}\| = \|\mathbf{x} - \mathbf{y}\| \|\mathbf{z} - \mathbf{a}_j\| \leq \|\mathbf{x} - \mathbf{y}\| \epsilon < \alpha/\epsilon = \alpha$$

y con esta contradicción concluye la prueba. ■

### Aplicaciones lineales continuas

**Proposición 3.21** Sean  $(E, \|\cdot\|)$ ,  $(F, \|\cdot\|')$  espacios normados sobre el mismo cuerpo ( $\mathbb{R}$  o  $\mathbb{C}$ ). Para una aplicación lineal  $T : E \rightarrow F$ , son equivalentes

- a)  $T$  es continua;
- b)  $T$  es continua en  $0$ ;
- c) Existe  $C > 0$  tal que  $\|T(\mathbf{x})\|' \leq C \|\mathbf{x}\|$  para todo  $\mathbf{x} \in E$ .

Si  $E$  es de dimensión finita, toda aplicación lineal  $T : E \rightarrow F$  es continua.

DEM: Es inmediato que c)  $\Rightarrow$  a)  $\Rightarrow$  b) y basta demostrar b)  $\Rightarrow$  c). Si  $T$  es continua en  $0$  existe  $\delta > 0$  tal que

$$\|\mathbf{y}\| < \delta \Rightarrow \|T(\mathbf{y})\|' = \|T(\mathbf{y}) - T(\mathbf{0})\|' < 1$$

La constante  $C = 2/\delta$  cumple c): Es evidente si  $\mathbf{x} = 0$ , y si  $\mathbf{x} \neq 0$  basta considerar  $\mathbf{y} = (C \|\mathbf{x}\|)^{-1} \mathbf{x}$ , que cumple  $\|\mathbf{y}\| = 1/C < \delta$ , para obtener

$$1 > \|T(\mathbf{y})\|' = (C \|\mathbf{x}\|)^{-1} \|T(\mathbf{x})\|'$$

es decir,  $\|T(\mathbf{x})\|' \leq C \|\mathbf{x}\|$ .

Si  $E$  es de dimensión finita  $n$  sobre el cuerpo  $\mathbb{R}$ , dada una base de  $E$ ,  $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ , se considera el isomorfismo algebraico  $S : \mathbb{R}^n \rightarrow E$ :

$$S(x_1, x_2, \dots, x_n) = \sum_{k=1}^n x_k \mathbf{u}_k$$

Para cada  $\mathbf{x} = \sum_{k=1}^n x_k \mathbf{u}_k \in E$ , es claro que  $\|\mathbf{x}\|^* = \sum_{k=1}^n |x_k| = \|S^{-1}(\mathbf{x})\|_1$  define una norma sobre  $E$ , equivalente a la norma inicial  $\|\cdot\|$  (véase 3.18). La constante  $C = \max\{\|T(\mathbf{u}_k)\| : 1 \leq k \leq n\}$  verifica:

$$\|T(\mathbf{x})\|' = \left\| \sum_{k=1}^n x_k T(\mathbf{u}_k) \right\|' \leq \sum_{k=1}^n |x_k| \|T(\mathbf{u}_k)\| \leq C \|\mathbf{x}\|^*$$

En virtud de b)  $T$  es continua para la norma  $\|\cdot\|^*$  y también para la norma equivalente  $\|\cdot\|$ . ■

Si  $(E, \|\cdot\|)$ ,  $(F, \|\cdot\|')$  son espacios normados sobre  $\mathbb{K}$  el conjunto de las aplicaciones lineales continuas  $T : E \rightarrow F$ , denotado  $\mathcal{L}(E, F)$ , también es un espacio vectorial sobre  $\mathbb{K}$ . En virtud de la proposición 3.21, para cada  $T \in \mathcal{L}(E, F)$ , el número

$$\|T\| = \sup\{\|T(\mathbf{x})\|' : \|\mathbf{x}\| \leq 1\}$$

es finito y se comprueba fácilmente que define una norma sobre  $\mathcal{L}(E, F)$ . Es fácil ver que  $\|T(\mathbf{x})\| \leq \|T\| \|\mathbf{x}\|$  para cada  $\mathbf{x} \in E$ . También es claro que  $\|T\| \leq C$ , donde  $C$  es cualquier constante que verifique la condición c) de la proposición 3.21, luego  $\|T\|$  es la mejor constante (la mínima) que cumple esta condición.

Conviene advertir que el valor numérico de la norma de una aplicación lineal continua  $T \in \mathcal{L}(E, F)$  depende de las normas de los espacios  $E$  y  $F$ . Si en estos espacios se cambian las normas por otras equivalentes, en  $\mathcal{L}(E, F)$  resulta otra norma equivalente a la antigua (la comprobación de esta afirmación se deja como ejercicio). Cuando  $E = \mathbb{R}^n$  y  $F = \mathbb{R}^m$ , cada  $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$  se puede identificar con su matriz  $A = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$ , respecto a las bases canónicas de  $\mathbb{R}^n$  y  $\mathbb{R}^m$ :

$$T(\mathbf{x}) = \mathbf{y} \quad \text{donde} \quad y_i = \sum_{j=1}^n a_{ij} x_j, \quad 1 \leq i \leq m$$

En este caso, en vez de hablar de la norma de la aplicación lineal  $T$  se suele hablar de la norma de la matriz asociada  $A$ . La norma de la matriz  $A$  depende de las normas consideradas en  $\mathbb{R}^n$  y en  $\mathbb{R}^m$  (véanse los ejercicios 3.8.25, 3.8.26).

### 3.5. Continuidad uniforme

**Definición 3.22** Sean  $(E, d)$ ,  $(F, \rho)$  espacios métricos. Una aplicación  $\mathbf{f} : M \rightarrow F$  definida en  $M \subset E$  se dice que es uniformemente continua cuando para cada  $\epsilon > 0$  existe  $\delta > 0$  tal que  $[\mathbf{x}, \mathbf{y} \in M, d(\mathbf{x}, \mathbf{y}) < \delta] \Rightarrow \rho(\mathbf{f}(\mathbf{x}), \mathbf{f}(\mathbf{y})) < \epsilon$ .

Si  $K \subset M$  y  $\mathbf{f}|_K$  es uniformemente continua se dice que  $\mathbf{f}$  es uniformemente continua sobre  $K$ .

La continuidad uniforme es una noción que se refiere al comportamiento global de la aplicación  $\mathbf{f}$  en todo su dominio. Se puede formular de modo equivalente en los siguientes términos: Para cada  $\epsilon > 0$  existe  $\delta > 0$  tal que para todo  $A \subset M$  con  $d\text{-diam}(A) < \delta$  se verifica  $\rho\text{-diam}(\mathbf{f}(A)) < \epsilon$ .

La diferencia entre la continuidad y la continuidad uniforme es la siguiente:  $\mathbf{f}$  es continua si para cada  $\epsilon > 0$  y cada  $\mathbf{a} \in M$  existe  $\delta(\mathbf{a}, \epsilon) > 0$  (que depende de  $\mathbf{a}$  y de  $\epsilon$ ) tal que todo  $\mathbf{x} \in M$  con  $d(\mathbf{x}, \mathbf{a}) < \delta(\mathbf{a}, \epsilon)$  verifica  $\rho(\mathbf{f}(\mathbf{x}), \mathbf{f}(\mathbf{a})) < \epsilon$ . Puede ocurrir que  $\mathbf{f}$  sea continua en cada punto de  $M$  pero al cambiar de un punto  $\mathbf{a} \in M$  a otro  $\mathbf{a}' \in M$  el número  $\delta(\mathbf{a}, \epsilon)$  que servía para el primer punto no sirva para el otro. La continuidad uniforme exige que para cada  $\epsilon > 0$  exista  $\delta(\epsilon) > 0$ , (dependiendo sólo de  $\epsilon$ ) tal que la condición de continuidad se cumple de modo simultáneo en todos los puntos:

*Para todo  $\mathbf{a} \in M$  y todo  $\mathbf{x} \in M$  con  $d(\mathbf{x}, \mathbf{a}) < \delta(\epsilon)$  se cumple  $\rho(\mathbf{f}(\mathbf{x}), \mathbf{f}(\mathbf{a})) < \epsilon$ .* Toda aplicación uniformemente continua es continua pero el recíproco no es cierto, como se pone de manifiesto con los ejemplos b) y c) que siguen.

### Ejemplos 3.23

a) Se dice que  $\mathbf{f} : M \rightarrow F$  es *Lipschitziana* con constante de Lipschitz  $C > 0$  si para cada par de puntos  $\mathbf{x}, \mathbf{y} \in M$  se cumple  $\rho(\mathbf{f}(\mathbf{x}), \mathbf{f}(\mathbf{y})) \leq Cd(\mathbf{x}, \mathbf{y})$ .

Es inmediato que toda aplicación Lipschitziana es uniformemente continua. En particular, toda función derivable  $f : (a, b) \rightarrow \mathbb{R}$  con derivada acotada es uniformemente continua ya que, en virtud del teorema del incremento finito es Lipschitziana con constante  $C = \sup\{|f'(x)| : a < x < b\}$ .

b) La función real de variable real  $f(x) = x^2$  no es uniformemente continua en  $\mathbb{R}$  porque transforma los intervalos  $A_\epsilon = (1/\epsilon, \epsilon + 1/\epsilon)$ , de diámetro  $\epsilon$  en los intervalos  $f(A_\epsilon) = (1/\epsilon^2, (\epsilon + 1/\epsilon)^2)$  de diámetro  $2 + \epsilon^2 > 2$ .

c) La función real de variable real  $g(x) = \text{sen}(1/x)$  no es uniformemente continua en  $(0, 1)$  ya que transforma los intervalos  $(0, \epsilon)$ , en conjuntos de diámetro 2.

**Teorema 3.24** Sean  $(E, d)$ ,  $(F, \rho)$  espacios métricos. Toda aplicación continua  $\mathbf{f} : M \rightarrow F$  definida en  $M \subset E$ , es uniformemente continua sobre cada compacto  $K \subset M$ .

DEM: Lo demostramos por reducción al absurdo suponiendo que hay una función continua  $\mathbf{f} : M \rightarrow F$ , que no es uniformemente continua sobre un compacto  $K \subset M$ . Entonces existe  $\epsilon > 0$  tal que para cada  $\delta > 0$  hay puntos  $\mathbf{x}_\delta, \mathbf{y}_\delta \in K$  tales que  $d(\mathbf{x}_\delta, \mathbf{y}_\delta) < \delta$  pero  $\rho(\mathbf{f}(\mathbf{x}_\delta), \mathbf{f}(\mathbf{y}_\delta)) \geq \epsilon$ . Con  $\delta = 1/n$ ,  $n \in \mathbb{N}$ , se obtienen sucesiones  $\mathbf{x}_n, \mathbf{y}_n \in K$  tales que  $d(\mathbf{x}_n, \mathbf{y}_n) < 1/n$ , pero  $\rho(\mathbf{f}(\mathbf{x}_n), \mathbf{f}(\mathbf{y}_n)) \geq \epsilon$ .

Como  $K$  es compacto la sucesión  $(\mathbf{x}_n)$  posee una subsucesión  $(\mathbf{x}_{n_k})$  convergente hacia un punto  $\mathbf{x} \in K$  (teorema 2.7). La subsucesión  $(\mathbf{y}_{n_k})$  también converge hacia  $\mathbf{x}$  porque  $0 \leq d(\mathbf{y}_{n_k}, \mathbf{x}) \leq d(\mathbf{y}_{n_k}, \mathbf{x}_{n_k}) + d(\mathbf{x}_{n_k}, \mathbf{x}) \leq 1/n_k + d(\mathbf{x}_{n_k}, \mathbf{x}) = \alpha_k$ , donde  $\alpha_k$  tiende hacia 0. En virtud de la continuidad de  $\mathbf{f}$  las sucesiones  $\mathbf{f}(\mathbf{x}_{n_k}), \mathbf{f}(\mathbf{y}_{n_k})$  convergen hacia  $\mathbf{f}(\mathbf{x})$ , luego  $\lim_k \rho(\mathbf{f}(\mathbf{x}_{n_k}), \mathbf{f}(\mathbf{y}_{n_k})) = 0$ . Se llega así a una contradicción porque las sucesiones  $\mathbf{x}_n, \mathbf{y}_n$  fueron elegidas verificando  $\rho(\mathbf{f}(\mathbf{x}_n), \mathbf{f}(\mathbf{y}_n)) \geq \epsilon$  para todo  $n \in \mathbb{N}$ . ■

OBSERVACIÓN: En la demostración del teorema anterior no se ha usado que la sucesión  $\mathbf{y}_n$  esté en el compacto  $K$ . La misma demostración permite mejorar la conclusión del teorema en la forma siguiente: Para cada  $\epsilon > 0$  existe  $\delta > 0$  tal que  $[\mathbf{x} \in K, \mathbf{y} \in M, d(\mathbf{x}, \mathbf{y}) < \delta] \Rightarrow \rho(\mathbf{f}(\mathbf{x}), \mathbf{f}(\mathbf{y})) < \epsilon$ . Esta mejora del teorema 3.24 se suele llamar *teorema de la continuidad uniforme mejorado*.

Con hipótesis adicionales, en ausencia de compacidad, es posible utilizar el teorema 3.24 para justificar que una función real continua es uniformemente continua. Así ocurre en la situación considerada en la siguiente proposición, cuya demostración sirve además para mostrar la simplificación que se consigue en algunos razonamientos cuando se tiene presente el teorema de la continuidad uniforme mejorado

**Proposición 3.25** *Sea  $f : A \rightarrow \mathbb{R}$  una función continua definida en un conjunto cerrado no acotado  $A \subset \mathbb{R}^n$ . Si existe y es finito el límite  $\lim_{\|\mathbf{x}\|_2 \rightarrow +\infty} f(\mathbf{x}) = L$  entonces  $f$  es uniformemente continua.*

DEM: Según la definición de límite, dado  $\epsilon > 0$  existe  $r > 0$  tal que

$$[\mathbf{x} \in A, \|\mathbf{x}\|_2 > r] \Rightarrow |f(\mathbf{x}) - L| < \epsilon/2$$

El conjunto  $K = \{\mathbf{x} \in A : \|\mathbf{x}\|_2 \leq r\}$  es compacto, (porque es un subconjunto cerrado y acotado de  $\mathbb{R}^n$ ) y utilizando el teorema de la continuidad uniforme mejorado podemos asegurar la existencia de un número  $\delta > 0$  tal que si uno de los puntos  $\mathbf{x}, \mathbf{y} \in A$  está en  $K$  y  $\|\mathbf{x} - \mathbf{y}\|_2 < \delta$  entonces  $|f(\mathbf{x}) - f(\mathbf{y})| < \epsilon$ . En caso contrario, si ninguno de los puntos está en  $K$ , también se cumple

$$|f(\mathbf{x}) - f(\mathbf{y})| \leq |f(\mathbf{x}) - L| + |L - f(\mathbf{y})| \leq \epsilon/2 + \epsilon/2 = \epsilon$$

En definitiva, cualquier pareja de puntos  $\mathbf{x}, \mathbf{y} \in A$  con  $\|\mathbf{x} - \mathbf{y}\|_2 < \delta$  hace que se cumpla la desigualdad  $|f(\mathbf{x}) - f(\mathbf{y})| < \epsilon$  y así queda demostrado que  $f$  es uniformemente continua. ■

En el ejercicio resuelto 3.37 se puede ver un resultado similar al expuesto en la proposición 3.25. La solución propuesta muestra como habría que proceder sin tener en cuenta el teorema de la continuidad uniforme mejorado. Se invita al lector a que lo utilice para dar una solución más breve.

**Teorema 3.26** *Sean  $(E, d)$ ,  $(F, \rho)$  espacios métricos y  $M \subset E$ . Si  $(F, \rho)$  es completo, cada aplicación uniformemente continua  $\mathbf{f} : M \rightarrow F$  se puede extender a una única aplicación uniformemente continua  $\bar{\mathbf{f}} : \bar{M} \rightarrow F$ .*

DEM: Si  $(F, \rho)$  es completo la extensión  $\bar{\mathbf{f}}$  se define por

$$\bar{\mathbf{f}}(\mathbf{a}) = \mathbf{f}(\mathbf{a}) \text{ si } \mathbf{a} \in M; \quad \bar{\mathbf{f}}(\mathbf{a}) = \lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{f}(\mathbf{x}) \text{ si } \mathbf{a} \in \bar{M} \setminus M \subset M'.$$

una vez que se haya demostrado que el límite existe. Para ello, según el teorema 3.3 basta ver que se cumple la condición de Cauchy:

Dado  $\epsilon > 0$  sea  $\delta > 0$  es el suministrado por la continuidad uniforme de  $\mathbf{f}$  que hace que se cumpla  $[\mathbf{x}, \mathbf{y} \in M, d(\mathbf{x}, \mathbf{y}) < \delta] \Rightarrow \rho(\mathbf{f}(\mathbf{x}), \mathbf{f}(\mathbf{y})) < \epsilon$ . Entonces, para toda pareja  $\mathbf{x}, \mathbf{y} \in M$ , con  $0 < d(\mathbf{x}, \mathbf{a}) < \delta/2$ ,  $0 < d(\mathbf{y}, \mathbf{a}) < \delta/2$ , se verifica  $d(\mathbf{x}, \mathbf{y}) < \delta$ , luego  $\rho(\mathbf{f}(\mathbf{x}), \mathbf{f}(\mathbf{y})) < \epsilon$ . Queda visto que  $\bar{\mathbf{f}}$  está bien definida y sólo queda demostrar que es uniformemente continua.

Veremos que, dado  $\epsilon > 0$ , el número  $\delta > 0$  proporcionado por la continuidad uniforme de  $\mathbf{f}$  sirve para establecer la continuidad uniforme de  $\bar{\mathbf{f}}$ . Efectivamente, si  $\mathbf{x}, \mathbf{y} \in \bar{M}$ , y  $d(\mathbf{x}, \mathbf{y}) < \delta$ , existen sendas sucesiones  $\mathbf{x}_n, \mathbf{y}_n \in M$ , convergentes hacia  $\mathbf{x}, \mathbf{y}$  respectivamente (si  $\mathbf{x} \in M$  tomamos  $\mathbf{x}_n = \mathbf{x}$  para cada  $n \in \mathbb{N}$ , y si  $\mathbf{y} \in M$  tomamos  $\mathbf{y}_n = \mathbf{y}$  para cada  $n \in \mathbb{N}$ ). En virtud de la desigualdad triangular

$$d(\mathbf{x}_n, \mathbf{y}_n) \leq d(\mathbf{x}_n, \mathbf{x}) + d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{y}_n) = \alpha_n$$

donde la sucesión  $\alpha_n$  converge hacia  $d(\mathbf{x}, \mathbf{y}) < \delta$ . Se sigue que existe  $m \in \mathbb{N}$  tal que todo  $n \geq m$  cumple  $d(\mathbf{x}_n, \mathbf{y}_n) \leq \alpha_n < \delta$  y por lo tanto  $\rho(\mathbf{f}(\mathbf{x}_n), \mathbf{f}(\mathbf{y}_n)) \leq \epsilon$ . Utilizando otra vez la desigualdad triangular obtenemos

$$\rho(\bar{\mathbf{f}}(\mathbf{x}), \bar{\mathbf{f}}(\mathbf{y})) \leq \rho(\bar{\mathbf{f}}(\mathbf{x}), \mathbf{f}(\mathbf{x}_n)) + \rho(\mathbf{f}(\mathbf{x}_n), \mathbf{f}(\mathbf{y}_n)) + \rho(\mathbf{f}(\mathbf{y}_n), \bar{\mathbf{f}}(\mathbf{y}))$$

y podemos asegurar que para todo  $n \geq m$  se cumple

$$\rho(\bar{\mathbf{f}}(\mathbf{x}), \bar{\mathbf{f}}(\mathbf{y})) \leq \rho(\bar{\mathbf{f}}(\mathbf{x}), \mathbf{f}(\mathbf{x}_n)) + \epsilon + \rho(\mathbf{f}(\mathbf{y}_n), \bar{\mathbf{f}}(\mathbf{y}))$$

Pasando al límite cuando  $n \rightarrow +\infty$ , y usando, cuando  $\mathbf{x}$  ó  $\mathbf{y}$  no pertenecen a  $M$ , la caracterización del límite mediante sucesiones (3.5) se obtiene que

$$\mathbf{x}, \mathbf{y} \in \bar{M}, d(\mathbf{x}, \mathbf{y}) < \delta \Rightarrow \rho(\bar{\mathbf{f}}(\mathbf{x}), \bar{\mathbf{f}}(\mathbf{y})) \leq \epsilon$$

La unicidad de la extensión  $\bar{\mathbf{f}}$  es inmediata: Si  $\mathbf{g} : \bar{M} \rightarrow F$  es una extensión continua de  $\mathbf{f}$ , para todo  $\mathbf{a} \in \bar{M} \setminus M$  se cumple  $\mathbf{g}(\mathbf{a}) = \lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{g}(\mathbf{x})$ , y considerando el límite a través de  $M$  resulta  $\mathbf{g}(\mathbf{a}) = \lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{g}|_M(\mathbf{x}) = \lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{f}(\mathbf{x}) = \bar{\mathbf{f}}(\mathbf{a})$ . ■

OBSERVACIÓN: Si  $(E, \|\cdot\|)$ ,  $(F, \|\cdot\|)$  son espacios normados según la proposición 3.21, cada aplicación lineal continua  $T : E \rightarrow F$  es Lipschitziana, con constante de Lipschitz  $\|T\|$ , luego es uniformemente continua. Por lo tanto, cuando  $F$  es completo, se puede aplicar el teorema de extensión 3.26 y es fácil comprobar que la extensión sigue siendo lineal y tiene la misma norma que  $T$ .

**Corolario 3.27** *Sea  $\mathbf{f} : M \rightarrow F$  una función continua, definida en un conjunto acotado  $M \subset \mathbb{R}^n$ , con valores en un espacio métrico completo  $(F, \rho)$ .*

*Son equivalentes:*

- i)  $\mathbf{f}$  se puede extender a una (única) aplicación continua  $\bar{\mathbf{f}} : \bar{M} \rightarrow F$ ;*
- ii)  $\mathbf{f}$  es uniformemente continua.*

DEM: i)  $\Rightarrow$  ii): Como  $\overline{M}$  es un subconjunto compacto de  $\mathbb{R}^n$  (por ser cerrado y acotado), la extensión continua  $\overline{\mathbf{f}} : \overline{M} \rightarrow F$  es uniformemente continua (en virtud del teorema 3.24) y por lo tanto  $\mathbf{f} = \overline{\mathbf{f}}|_M$  también lo es. El recíproco ii)  $\Rightarrow$  i) es el teorema 3.26. ■

### Propiedades topológicas y propiedades uniformes.

**Definición 3.28** *Un isomorfismo uniforme entre espacios métricos  $(E, d)$ ,  $(F, \rho)$  es una aplicación biyectiva uniformemente continua  $\mathbf{f} : (E, d) \rightarrow (F, \rho)$  con inversa uniformemente continua. Dos distancias  $d, d'$  sobre un conjunto  $E$  se dice que son uniformemente equivalentes cuando la identidad  $i : (E, d) \rightarrow (E, d')$  es un isomorfismo uniforme.*

Una noción referente a un espacio métrico  $(E, d)$  se dice que es *topológica* cuando sólo depende de la topología del espacio y por lo tanto permanece cuando se cambia la distancia  $d$  por otra equivalente. Una noción referente a un espacio métrico  $(E, d)$  se dice que es *uniforme* cuando permanece al cambiar la distancia  $d$  por otra uniformemente equivalente.

Las nociones de conjunto abierto, cerrado, compacto, adherencia, interior, sucesión convergente y función continua son topológicas. Todas las nociones topológicas son uniformes pero el recíproco es falso: Con el ejercicio 2.6.5 se pone de manifiesto que las nociones de sucesión de Cauchy y de espacio completo no son topológicas, pero son uniformes como consecuencia directa del siguiente ejercicio:

**Ejercicio 3.29** *Si  $\mathbf{f} : (E, d) \rightarrow (F, \rho)$  es uniformemente continua demuestre que  $\mathbf{f}$  transforma sucesiones de Cauchy en sucesiones de Cauchy. Si  $\mathbf{f} : (E, d) \rightarrow (F, \rho)$  es una biyección uniformemente continua con inversa continua y  $(F, \rho)$  es completo demuestre que  $(E, d)$  también es completo.*

La noción de conjunto totalmente acotado (o precompacto) que se considera en el apéndice B.2 también es una noción uniforme que no es topológica.

La noción de conjunto acotado no es uniforme pues en  $\mathbb{R}$  la distancia usual  $d$  y la distancia  $d^*(x, y) = \min\{1, d(x, y)\}$  son uniformemente equivalentes pero no producen los mismos conjuntos acotados.

Todo isomorfismo uniforme entre dos espacios métricos es un homeomorfismo, y dos distancias  $d, d'$  uniformemente equivalentes son equivalentes, pero los recíprocos son falsos: En  $\mathbb{R}$  las distancias  $d, d'$  consideradas en el problema 2.6.5 son equivalentes pero no son uniformemente equivalentes porque si lo fuesen producirían las mismas sucesiones de Cauchy en virtud del ejercicio 3.29.

Si dos espacios métricos son homeomorfos y uno de ellos es compacto el otro también lo es (en virtud de 3.14) y el homeomorfismo es un isomorfismo uniforme en virtud del teorema 3.24. En particular, si dos distancias en  $E$  son equivalentes y  $E$  es compacto para la topología asociada a las distancias entonces las distancias son uniformemente equivalentes. Un resultado análogo se verifica cuando  $E$  es un espacio vectorial (real o complejo): Si  $d, d'$  son distancias equivalentes sobre  $E$ , asociadas a normas, entonces son uniformemente equivalentes en virtud de 2.4.

Como la composición de aplicaciones uniformemente continuas es uniformemente continua se sigue que si tenemos una aplicación  $\mathbf{f} : (E, d) \rightarrow (F, \rho)$  uniformemente continua y cambiamos las distancias  $d$  y  $\rho$  por otras uniformemente equivalentes entonces  $\mathbf{f}$  sigue siendo uniformemente continua respecto a las nuevas distancias. Esto significa que la continuidad uniforme es una noción uniforme respecto a ambos espacios  $E$  y  $F$ . Otra noción uniforme de especial interés es la convergencia uniforme que se estudia a continuación.

### 3.6. Convergencia uniforme

En esta sección se extienden al contexto de los espacios métricos los resultados básicos sobre convergencia uniforme. En particular se muestra que la convergencia uniforme es una noción uniforme.

Sea  $(F, \rho)$  un espacio métrico y  $\mathbf{f}_n : T \rightarrow F$  una sucesión de funciones definidas en un conjunto no vacío  $T$ . Se dice que la sucesión  $\mathbf{f}_n$  *converge puntualmente* cuando para cada  $t \in T$  la sucesión  $\mathbf{f}_n(t)$  converge en el espacio métrico  $(F, \rho)$ . En este caso el límite puntual es la función  $\mathbf{f} : T \rightarrow F$ , definida por  $\mathbf{f}(t) = \lim_n \mathbf{f}_n(t)$ .

En el contexto de las funciones reales de variable real ejemplos sencillos, como  $f_n(x) = 1/(1 + x^{2n})$ , muestran que el límite puntual de una sucesión de funciones continuas puede no ser continuo.

La convergencia uniforme es una noción más fuerte que la convergencia puntual pues se exige que el valor de  $n$  a partir del cual se logra la aproximación  $\rho(\mathbf{f}_n(t), \mathbf{f}(t)) \leq \epsilon$ , no depende del punto  $t \in T$ , es decir se exige una aproximación al límite uniforme en todos los puntos de  $T$ . La convergencia uniforme, con la que se consigue la continuidad del límite de una sucesión de funciones continuas (3.31), en el contexto de los espacios métricos se formula así:

**Definición 3.30** *Sea  $(F, \rho)$  un espacio métrico y  $T$  un conjunto. La sucesión  $\mathbf{f}_n : T \rightarrow F$  converge uniformemente hacia  $\mathbf{f} : T \rightarrow F$  cuando para cada  $\epsilon > 0$  existe  $n(\epsilon) \in \mathbb{N}$  (que depende sólo de  $\epsilon$ ) tal que para todo  $n \geq n(\epsilon)$  y todo  $t \in T$  se cumple  $\rho(\mathbf{f}_n(t), \mathbf{f}(t)) \leq \epsilon$ .*

Si  $K \subset T$  y la sucesión  $\mathbf{f}_n|_K$  converge puntualmente (resp. uniformemente) se dice, más brevemente, que la sucesión  $\mathbf{f}_n$  converge puntualmente (resp. uniformemente) sobre  $K$ . A veces ocurre que una sucesión de funciones  $\mathbf{f}_n : T \rightarrow F$ , no converge uniformemente sobre todo  $T$ , pero la convergencia es uniforme sobre los conjuntos de una familia  $\mathcal{C}$  de subconjuntos de  $T$ . Un caso particular, cuando  $\mathcal{C}$  es la familia de los subconjuntos compactos de  $T$ , es el de la convergencia uniforme sobre compactos. La convergencia uniforme sobre  $T$  implica la convergencia uniforme sobre cada  $C \in \mathcal{C}$ , lo que implica la convergencia puntual sobre  $\bigcup \mathcal{C}$ , pero las afirmaciones recíprocas son falsas.

Con el fin de formular la condición de convergencia uniforme de modo más conciso conviene introducir la siguiente notación: Si  $K \subset T$ , dadas  $\mathbf{f}, \mathbf{g} : T \rightarrow F$ , se define

$$\rho_K(\mathbf{f}, \mathbf{g}) = \sup\{\rho(\mathbf{f}(t), \mathbf{g}(t)) : t \in K\} \quad (\leq +\infty)$$



El hecho de que la sucesión  $\mathbf{f}_n : T \rightarrow F$  sea uniformemente convergente hacia  $\mathbf{f} : T \rightarrow F$  se escribe en la forma  $\lim_n \rho_T(\mathbf{f}_n, \mathbf{f}) = 0$ . Análogamente, la convergencia uniforme sobre  $K \subset T$  se expresa mediante la condición  $\lim_n \rho_K(\mathbf{f}_n, \mathbf{f}) = 0$ .

**Teorema 3.31** Sean  $(T, d), (F, \rho)$  espacios métricos. Si la sucesión  $\mathbf{f}_n : T \rightarrow F$  converge uniformemente hacia  $\mathbf{f} : T \rightarrow F$  y cada  $\mathbf{f}_n$  es continua en  $a \in T$  entonces el límite  $\mathbf{f}$  también lo es. En particular, si las funciones  $\mathbf{f}_n$  son continuas en todo punto, el límite uniforme  $\mathbf{f}$  también lo es.

DEM: Dado  $\epsilon > 0$ , en virtud de la convergencia uniforme, existe  $m \in \mathbb{N}$  tal que para todo  $t \in T$  se cumple  $\rho(\mathbf{f}_m(t), \mathbf{f}(t)) \leq \epsilon/3$ .

Por la continuidad de  $\mathbf{f}_m$  en  $a$ , existe una bola abierta  $B(a, r) \subset T$ , tal que todo  $t \in B(a, r)$  cumple  $\rho(\mathbf{f}_m(t), \mathbf{f}_m(a)) \leq \epsilon/3$ , luego

$$\begin{aligned} \rho(\mathbf{f}(t), \mathbf{f}(a)) &\leq \rho(\mathbf{f}(t), \mathbf{f}_m(t)) + \rho(\mathbf{f}_m(t), \mathbf{f}_m(a)) + \rho(\mathbf{f}_m(a), \mathbf{f}(a)) \leq \\ &\leq \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon \end{aligned}$$

■

En las condiciones del teorema anterior, para conseguir la continuidad de  $\mathbf{f}$  en un punto  $a \in T$  basta suponer que, desde un valor de  $n$  en adelante, las funciones  $\mathbf{f}_n$  son continuas en  $a$ , y que la convergencia es uniforme en algún entorno de  $a$ . En particular, para conseguir la continuidad global del límite basta la convergencia uniforme local, lo que significa que cada  $t \in T$  tiene un entorno abierto  $V_t$  donde la convergencia es uniforme. La convergencia uniforme local implica la convergencia uniforme sobre compactos (porque todo compacto  $K \subset T$  se puede recubrir con una cantidad finita de entornos abiertos sobre los que hay convergencia uniforme).

Aunque el teorema 3.31 sigue valiendo para un espacio topológico general  $T$ , el siguiente resultado utiliza que  $T$  es un espacio métrico.

**Corolario 3.32** Sean  $(T, d), (F, \rho)$  espacios métricos. Si una sucesión de funciones continuas  $\mathbf{f}_n : T \rightarrow F$ , converge uniformemente sobre compactos hacia  $\mathbf{f} : T \rightarrow F$ , entonces  $\mathbf{f}$  es continua.

DEM: Demostraremos que  $\mathbf{f}$  es continua en cada  $t \in T$  viendo que si  $t_n \in T$  es una sucesión convergente hacia  $t$  entonces la sucesión  $\mathbf{f}(t_n)$ , converge hacia  $\mathbf{f}(t)$ : Como el conjunto  $K = \{t\} \cup \{t_n : n \in \mathbb{N}\}$  es compacto (véase el problema 2.6.23) la sucesión  $\mathbf{f}_n$  converge uniformemente sobre  $K$  hacia  $\mathbf{f}$ , y aplicando el teorema 3.31 a la sucesión  $\mathbf{f}_n|_K$  se obtiene que  $\mathbf{f}|_K$  es continua. Como  $t_n$  es una sucesión en  $K$  que converge hacia  $t \in K$  se concluye que  $\mathbf{f}(t) = \mathbf{f}|_K(t) = \lim_n \mathbf{f}|_K(t_n) = \lim_n \mathbf{f}(t_n)$ .

■

Las proposiciones siguientes relacionan la continuidad uniforme con la convergencia uniforme.

**Proposición 3.33** Sean  $(T, d), (F, \rho)$  espacios métricos y  $\mathbf{f}_n : T \rightarrow F$  una sucesión uniformemente convergente de funciones uniformemente continuas. Entonces la función límite  $\mathbf{f} : T \rightarrow F$  es uniformemente continua.

DEM: La demostración, análoga a la del teorema 3.31, se deja como ejercicio. ■

**Proposición 3.34** Sean  $(E, d), (F, \rho)$  espacios métricos. Si  $\mathbf{g} : E \rightarrow F$  es uniformemente continua y la sucesión  $\mathbf{f}_n : T \rightarrow E$  converge uniformemente hacia  $\mathbf{f} : T \rightarrow E$  entonces  $\mathbf{g} \circ \mathbf{f}_n : T \rightarrow F$  converge uniformemente hacia  $\mathbf{g} \circ \mathbf{f} : T \rightarrow F$ .

DEM: Es consecuencia directa de las definiciones y se deja como ejercicio. ■

Obsérvese que la proposición 3.34 nos dice que la convergencia uniforme, como indica su nombre, es una noción uniforme. Es decir, si  $\rho, \rho'$  son dos distancias uniformemente equivalentes en  $F$  y la sucesión  $\mathbf{f}_n : T \rightarrow F$  converge uniformemente hacia  $\mathbf{f} : T \rightarrow F$  con la distancia  $\rho$ , también converge uniformemente con la distancia  $\rho'$ .

**Proposición 3.35** [Condición de Cauchy] Si el espacio métrico  $(F, \rho)$  es completo, una condición necesaria y suficiente para que la sucesión  $\mathbf{f}_n : T \rightarrow F$  sea uniformemente convergente es que cumpla la condición de Cauchy: Para cada  $\epsilon > 0$  existe  $n(\epsilon) \in \mathbb{N}$  tal que  $[q > p \geq n(\epsilon), t \in T] \Rightarrow \rho(\mathbf{f}_p(t), \mathbf{f}_q(t)) \leq \epsilon$ .

DEM: La demostración de que la condición es necesaria es inmediata y se deja al cuidado del lector.

La condición es suficiente: Para cada  $t \in T$  es  $\rho(\mathbf{f}_p(t), \mathbf{f}_q(t)) \leq \rho_T(\mathbf{f}_p, \mathbf{f}_q)$ , luego  $\mathbf{f}_n(t)$  es una sucesión de Cauchy en el espacio métrico completo  $(F, \rho)$  y por lo tanto converge hacia un punto  $\mathbf{f}(t) \in F$ . Así queda establecido que la sucesión es puntualmente convergente y debemos verificar que la convergencia es uniforme.

Dado  $\epsilon > 0$  si  $q > p \geq n(\epsilon)$ , para todo  $t \in T$  se cumple  $\rho(\mathbf{f}_p(t), \mathbf{f}_q(t)) \leq \epsilon$ . Fijando  $t \in T$  y pasando al límite cuando  $q \rightarrow +\infty$  la última desigualdad se convierte en  $\rho(\mathbf{f}_p(t), \mathbf{f}(t)) \leq \epsilon$ , que resulta válida para todo  $t \in T$  y todo  $p > n(\epsilon)$ . ■

**Métrica de la convergencia uniforme.** Sea  $F^T$  el conjunto de las aplicaciones  $\mathbf{f} : T \rightarrow F$ . Cuando la distancia  $\rho$  es acotada, para cada  $\mathbf{f}, \mathbf{g} \in F^T$  el supremo

$$\rho_T(\mathbf{f}, \mathbf{g}) = \sup\{\rho(\mathbf{f}(t), \mathbf{g}(t)) : t \in T\} < +\infty$$

es finito y define una distancia en  $F^T$  tal que las sucesiones convergentes en el espacio métrico  $(F^T, \rho_T)$  son precisamente las uniformemente convergentes, es decir la convergencia uniforme se metriza con la distancia  $\rho_T$ .

Cuando la distancia  $\rho$  no es acotada, podemos sustituir  $\rho$  por una distancia acotada uniformemente equivalente (p.e.  $\rho'(\mathbf{x}, \mathbf{y}) = \min\{1, \rho(\mathbf{x}, \mathbf{y})\}$ ). Como la convergencia uniforme con una distancia equivale a la convergencia uniforme con otra distancia uniformemente equivalente también es posible, en este caso, metrizar la convergencia uniforme. Cuando  $T$  es un subconjunto abierto de  $\mathbb{R}^n$  también se puede demostrar que la convergencia uniforme sobre compactos es metrizable.

Cuando  $T$  es un espacio métrico (o más generalmente, un espacio topológico) el teorema 3.31 asegura que el conjunto  $C(T, F)$  formado por las funciones continuas  $\mathbf{f} : T \rightarrow F$ , es un subconjunto cerrado del espacio métrico  $(F^T, \rho_T)$ .

Si la distancia  $\rho$  es acotada la proposición 3.35 se puede reformular diciendo que si el espacio métrico  $(F, \rho)$  es completo entonces  $(F^T, \rho_T)$  también lo es.

Cuando  $(F, \|\cdot\|)$  es un espacio normado el subconjunto  $l^\infty(T, E)$  de  $F^T$  formado por las aplicaciones acotadas de  $T$  en  $F$ , es un espacio vectorial en el que se puede definir la norma de la convergencia uniforme

$$\|\mathbf{f}\|_T = \sup\{\|\mathbf{f}(t)\| : t \in T\}$$

llamada así porque con ella se metriza la convergencia uniforme en  $l^\infty(T, F)$ . En virtud de 3.35 el espacio normado  $(l^\infty(T, F), \|\cdot\|_T)$  es completo si  $(F, \|\cdot\|)$  lo es. Cuando  $T$  es un espacio métrico compacto  $C(T, F)$  es un subespacio vectorial cerrado de  $(l^\infty(T, F), \|\cdot\|_T)$  y por lo tanto será completo cuando  $(F, \|\cdot\|)$  lo sea.

**Convergencia uniforme de series.** Para funciones con valores en un espacio normado  $(F, \|\cdot\|)$  se pueden considerar series  $\sum_{n=1}^{\infty} \mathbf{f}_n$  de funciones  $\mathbf{f}_n : T \rightarrow F$ . Se dice que la serie converge puntualmente (resp. uniformemente) si la sucesión de sumas parciales  $\mathbf{s}_m(t) = \sum_{n=1}^m \mathbf{f}_n(t)$  tiene la correspondiente propiedad. En ese caso queda definida en  $T$  la función suma

$$\mathbf{f}(t) = \sum_{n=1}^{\infty} \mathbf{f}_n(t)$$

Con demostraciones análogas a las del caso de funciones numéricas (véase el apéndice A) se establecen los criterios de convergencia uniforme de series que llevan los nombres de Weierstrass, Abel y Dirichlet, que se pueden ver en el apéndice C.2.

### 3.7. Ejercicios resueltos

**Ejercicio 3.36** Sean  $(E, d)$ ,  $(F, d')$  espacios métricos y  $\mathbf{f} : E \rightarrow F$  una biyección continua tal que  $\mathbf{f}^{-1}(K)$  es compacto en  $E$  para cada compacto  $K \subset F$ . Demuestre que  $\mathbf{f}^{-1}$  es continua.

SOLUCIÓN

Demostraremos que  $\mathbf{f}^{-1}$  es continua usando la caracterización de la continuidad por sucesiones 3.5. Basta demostrar que si la sucesión  $\mathbf{y}_n \in F$  converge hacia  $\mathbf{b} \in F$  entonces  $\mathbf{x}_n = \mathbf{f}^{-1}(\mathbf{y}_n)$  converge hacia  $\mathbf{a} = \mathbf{f}^{-1}(\mathbf{b})$ . Según 2.6.23 el conjunto  $K = \{\mathbf{y}_n : n \in \mathbb{N}\} \cup \{\mathbf{b}\}$  es compacto en  $F$  y en virtud de la hipótesis  $\mathbf{f}^{-1}(K)$  es compacto en  $E$ . Como la sucesión  $(\mathbf{x}_n)$  está en este compacto, para ver que converge hacia  $\mathbf{a} = \mathbf{f}^{-1}(\mathbf{b})$  basta comprobar que  $\mathbf{a}$  es su único punto de aglomeración (véase 2.8).

En efecto, si  $\mathbf{x} = \lim_k \mathbf{x}_{n_k}$  es un punto de aglomeración de la sucesión  $(\mathbf{x}_n)$ , en virtud de la continuidad de  $\mathbf{f}$  en  $\mathbf{x}$  se tiene  $\mathbf{f}(\mathbf{x}) = \lim_k \mathbf{f}(\mathbf{x}_{n_k}) = \lim_k \mathbf{y}_{n_k}$ . Como la sucesión  $\mathbf{y}_n$  era convergente hacia  $\mathbf{b}$  lo mismo le ocurre a la subsucesión  $\mathbf{y}_{n_k}$ , luego  $\mathbf{f}(\mathbf{x}) = \mathbf{b}$  y por lo tanto  $\mathbf{x} = \mathbf{a}$  (ya que  $\mathbf{f}$  es inyectiva). ■

**Ejercicio 3.37** Demuestre que toda función monótona continua y acotada  $f : I \rightarrow \mathbb{R}$  definida en un intervalo cerrado no acotado  $I$  (de la forma  $[a, +\infty)$ ,  $(-\infty, a]$ , ó  $(-\infty, +\infty)$ ) es uniformemente continua.

SOLUCIÓN

Consideremos primero el caso  $I = [a, +\infty)$ .

Al ser  $f$  monótona y acotada, existe el límite  $\lim_{x \rightarrow +\infty} f(x) = l$  luego dado  $\epsilon > 0$  existe  $b > a$  tal que  $x \geq b \Rightarrow |f(x) - l| < \epsilon/2$ . Por lo tanto, en virtud de la desigualdad triangular

$$x, y \in [b, +\infty) \Rightarrow |f(x) - f(y)| \leq |f(x) - l| + |f(y) - l| < \epsilon/2 + \epsilon/2 = \epsilon$$

Por otra parte, como la función continua  $f$  es uniformemente continua sobre el intervalo compacto  $[a, b + 1]$  existe  $\delta \in (0, 1)$  tal que

$$x, y \in [a, b + 1] \Rightarrow |f(x) - f(y)| < \epsilon$$

Entonces se cumple

$$x, y \in [a, +\infty), \quad |x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon$$

Para ver esto, basta tener en cuenta que al ser  $|x - y| < \delta < 1$ , o bien ambos puntos  $x, y$  están en  $[a, b + 1]$  ó ambos puntos están en  $[b, +\infty)$  (si un punto  $x$  cumple  $x \geq b + 1 > b$  entonces el otro punto  $y$  cumple  $y \geq b$ ). ■

**Ejercicio 3.38** Sean  $(E, d)$ ,  $(F, \rho)$  espacios métricos,  $M \subset E$ , y  $a \in M'$ . Se supone que la sucesión  $\mathbf{f}_n : M \rightarrow F$  converge uniformemente hacia  $\mathbf{f} : M \rightarrow F$  y que cada  $\mathbf{f}_n$  tiene límite  $\mathbf{b}_n$  cuando  $t \rightarrow a$ . Si el espacio métrico  $(F, \rho)$  es completo, demuestre que existen y son iguales los límites  $\lim_{t \rightarrow a} \mathbf{f}(t) = \lim_n \mathbf{b}_n$ , es decir

$$\lim_{t \rightarrow a} (\lim_n \mathbf{f}_n(t)) = \lim_n (\lim_{t \rightarrow a} \mathbf{f}_n(t))$$

SOLUCIÓN

Como  $(F, \rho)$  se supone completo, para ver que la sucesión  $(\mathbf{b}_n)$  converge basta demostrar que es de Cauchy. Dado  $\epsilon > 0$ , según la condición de Cauchy para la convergencia uniforme (proposición 3.35) existe  $n(\epsilon) \in \mathbb{N}$  tal que si  $p > q \geq n(\epsilon)$  entonces

$$\rho(\mathbf{f}_p(t), \mathbf{f}_q(t)) \leq \epsilon/3 \quad \text{para todo } t \in M$$

Por otra parte, una vez que hemos fijado  $p > q \geq n(\epsilon)$ , usando la definición de  $\mathbf{b}_p$ ,  $\mathbf{b}_q$  como límites, podemos encontrar  $z \in M$ , próximo al punto  $a$ , verificando las dos desigualdades

$$\rho(\mathbf{b}_p, \mathbf{f}_p(z)) < \epsilon/3, \quad \rho(\mathbf{b}_q, \mathbf{f}_q(z)) < \epsilon/3.$$

Combinando las desigualdades anteriores, se obtiene

$$\rho(\mathbf{b}_p, \mathbf{b}_q) \leq \rho(\mathbf{b}_p, \mathbf{f}_p(z)) + \rho(\mathbf{f}_p(z), \mathbf{f}_q(z)) + \rho(\mathbf{f}_q(z), \mathbf{b}_q) \leq \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon$$

Queda demostrado que  $(\mathbf{b}_n)$  es una sucesión de Cauchy, luego existe el límite  $\mathbf{b} = \lim_n \mathbf{b}_n$ , y para terminar tenemos que demostrar que  $\lim_{t \rightarrow a} \mathbf{f}(t) = \mathbf{b}$ .

Fijado  $\epsilon > 0$ , en virtud de la convergencia de  $(\mathbf{b}_n)$  y de la convergencia uniforme de  $(\mathbf{f}_n)$ , existe  $m \in \mathbb{N}$  que verifica simultáneamente

$$\rho(\mathbf{b}_m, \mathbf{b}) < \epsilon/3, \quad \rho(\mathbf{f}_m(t), \mathbf{f}(t)) \leq \epsilon/3 \quad \text{para todo } t \in M.$$

Como  $\mathbf{b}_m = \lim_{t \rightarrow a} \mathbf{f}_m(t)$ , existe  $\delta > 0$  tal que si  $t \in M$  y  $0 < d(t, a) < \delta$  se verifica  $\rho(\mathbf{f}_m(t), \mathbf{b}_m) < \epsilon/3$ . Por lo tanto, cuando  $t \in M$ , y  $0 < d(t, a) < \delta$ , se cumple

$$\rho(\mathbf{f}(t), \mathbf{b}) \leq \rho(\mathbf{f}(t), \mathbf{f}_m(t)) + \rho(\mathbf{f}_m(t), \mathbf{b}_m) + \rho(\mathbf{b}_m, \mathbf{b}) \leq \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon.$$

■

**Ejercicio 3.39** Sean  $(E, d)$ ,  $(F, \rho)$  espacios métricos y  $M \subset E$ . Una sucesión  $\mathbf{f}_n : M \rightarrow F$  se dice que es equicontinua en  $a \in M$  si para cada  $\epsilon > 0$  existe  $\delta(\epsilon) > 0$  tal que para todo  $t \in M$  con  $d(t, a) < \delta$  y todo  $n \in \mathbb{N}$  se cumple  $\rho(\mathbf{f}_n(t), \mathbf{f}_n(a)) < \epsilon$ . (Es decir, todas las funciones  $\mathbf{f}_n$  son continuas en  $a$  y la condición de continuidad se cumple con un número  $\delta(\epsilon)$  que sirve para todas las funciones a la vez) Demuestre las siguientes afirmaciones:

- Si  $\mathbf{f}_n : M \rightarrow F$  es una sucesión equicontinua en  $a \in M$  que converge puntualmente hacia  $\mathbf{f} : M \rightarrow F$ , entonces  $\mathbf{f}$  es continua en  $a$ .
- Si  $\mathbf{f}_n : M \rightarrow F$  es una sucesión equicontinua en cada  $t \in M$  que converge puntualmente hacia  $\mathbf{f} : M \rightarrow F$ , entonces  $\mathbf{f}_n$  converge uniformemente sobre compactos hacia  $\mathbf{f}$ .

SOLUCIÓN

- a) Por la equicontinuidad en  $a$ , dado  $\epsilon > 0$ , existe  $\delta(\epsilon) > 0$  tal que

$$[t \in M, d(t, a) < \delta(\epsilon)] \Rightarrow [\rho(\mathbf{f}_n(t), \mathbf{f}_n(a)) < \epsilon \quad \text{para todo } n \in \mathbb{N}]$$

Fijado  $t \in M$  con  $d(t, a) < \delta(\epsilon)$ , la desigualdad  $\rho(\mathbf{f}_n(t), \mathbf{f}_n(a)) \leq \epsilon$  se cumple para todo  $n \in \mathbb{N}$ , y pasando al límite la desigualdad se conserva. Se obtiene así que  $[t \in M, d(t, a) < \delta(\epsilon)] \Rightarrow \rho(\mathbf{f}(t), \mathbf{f}(a)) \leq \epsilon$ , luego  $\mathbf{f}$  es continua en  $a$ .

- b) Por la equicontinuidad, para cada  $\epsilon > 0$ , y cada  $t \in M$  existe  $\delta_t > 0$  tal que

$$[s \in M, d(s, t) < \delta_t, n \in \mathbb{N}] \Rightarrow \rho(\mathbf{f}_n(s), \mathbf{f}_n(t)) < \epsilon$$

Según el apartado a)  $\mathbf{f}$  es continua en cada  $t \in M$ , luego podemos suponer que cada  $\delta_t$  ha sido elegido de modo que también cumpla la condición

$$s \in M, d(s, t) < \delta_t \Rightarrow \rho(\mathbf{f}(s), \mathbf{f}(t)) < \epsilon$$

Dado un compacto  $K \subset M$ , la familia de las bolas abiertas  $\{B(t, \delta_t) : t \in K\}$  recubre  $K$  luego existe un conjunto finito  $H \subset K$  tal que  $K \subset \bigcup_{a \in H} B(a, \delta_a)$ . Como  $H$  es finito, por la convergencia puntual de la sucesión, existe  $m \in \mathbb{N}$  tal

que para todo  $n \geq m$  y todo  $t \in H$  se cumple  $\rho(\mathbf{f}_n(t), \mathbf{f}(t)) < \epsilon$ .

Entonces, para todo  $n \geq m$  y todo  $s \in K$  se cumple  $\rho(\mathbf{f}_n(s), \mathbf{f}(s)) \leq 3\epsilon$ , pues dado  $s \in K$  existe  $t \in H$  tal que  $s \in B(t, \delta_t)$ , y por lo tanto

$$\rho(\mathbf{f}_n(s), \mathbf{f}(s)) \leq \rho(\mathbf{f}_n(s), \mathbf{f}_n(t)) + \rho(\mathbf{f}_n(t), \mathbf{f}(t)) + \rho(\mathbf{f}(t), \mathbf{f}(s)) \leq \epsilon + \epsilon + \epsilon = 3\epsilon$$

■

**Ejercicio 3.40** Sea  $\mathcal{P}_m$  el conjunto de los polinomios de grado  $\leq m$ . Demuestre que para cada  $k \in [1, 2, \dots, m]$  existe  $C_k > 0$  tal que para todo  $p \in \mathcal{P}_m$ , y todo  $a \in [0, 1]$  se cumple

$$|p^{(k)}(a)| \leq C_k \max\{|p(x)| : 0 \leq x \leq 1\}$$

SOLUCIÓN

$\mathcal{P}_m$ , con la norma de la convergencia uniforme  $\|p\|_\infty = \max\{|p(x)| : 0 \leq x \leq 1\}$  es un espacio normado de dimensión finita. Por lo tanto, la aplicación lineal

$$T_k : (\mathcal{P}_m, \|\cdot\|_\infty) \rightarrow (C[0, 1], \|\cdot\|_\infty), \quad T_k(p) = p^{(k)}$$

es continua y esto significa que existe  $C_k > 0$  tal que para todo  $p \in \mathcal{P}_m$  se cumple  $\|T_k(p)\|_\infty \leq C_k \|p\|_\infty$ , luego, para cada  $p \in \mathcal{P}_m$  y cada  $a \in [0, 1]$ , se cumple la desigualdad del enunciado. ■

### 3.8. Ejercicios propuestos

◇ **3.8.1** Calcule, si existen, los límites en  $(0,0)$  de las siguientes funciones

$$i) f(x, y) = \frac{y}{x^2 + y} \quad \text{si } x^2 + y \neq 0, \quad f(x, -x^2) = 0.$$

$$ii) f(x, y) = \frac{x^2 + y^2}{x^2 + y} \quad \text{si } x^2 + y \neq 0, \quad f(x, -x^2) = 0.$$

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

$$iv) f(x, y) = \frac{x - y}{x + y} \quad \text{si } x + y \neq 0, \quad f(x, -x) = 0.$$

◇ **3.8.2** Sea  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ , donde  $f(x, y) = \frac{xy}{x^2 + y^2}$  si  $(x, y) \neq (0, 0)$ ,  $f(0, 0) = 0$ .

Compruebe que  $f$  no es continua pero todas las funciones parciales  $x \rightarrow f(x, b)$ ,  $y \rightarrow f(a, y)$ , son continuas. Obtenga la clausura de la gráfica de  $f$ .

◇ **3.8.3** Estudie la continuidad de las siguientes funciones

$$i) f(x, y) = \frac{x \operatorname{sen} y}{x^2 + y^2} \quad \text{si } (x, y) \neq (0, 0), \quad f(0, 0) = 0.$$

$$ii) f(x, y) = \frac{\operatorname{sen}(x^2 + y^2)}{x^2 + y^2} \quad \text{si } (x, y) \neq (0, 0), \quad f(0, 0) = 1.$$

$$iii) f(x, y) = x \quad \text{si } |x| \leq |y|, \quad f(x, y) = y \quad \text{si } |x| > |y|.$$

$$iv) f(x, y) = \frac{x}{4x^2 + y^2 - 1} \quad \text{si } 4x^2 + y^2 \neq 1, \quad f(x, y) = 1 \quad \text{si } 4x^2 + y^2 = 1.$$

$$v) f(x, y) = \frac{x^2 y^2}{x^2 y^2 + (x - y)^2} \quad \text{si } (x, y) \neq (0, 0), \quad f(0, 0) = 1.$$

$$vi) f(x, y) = (x + y) \operatorname{sen}(1/x) \operatorname{sen}(1/y) \quad \text{si } x \neq 0, y \neq 0, \quad f(0, y) = f(x, 0) = 0.$$

$$vii) f(x, y) = \frac{x + \operatorname{sen}(x + y)}{x + y} \quad \text{si } x + y \neq 0, \quad f(x, -x) = 0.$$

$$viii) f(x, y) = \frac{x}{y} \operatorname{sen}(x^2 + y^2) \quad \text{si } y \neq 0, \quad f(x, 0) = 0.$$

$$ix) f(x, y) = \frac{x^3}{x^2 - y^2} \quad \text{si } x^2 - y^2 \neq 0, \quad f(x, y) = 0 \quad \text{si } x^2 - y^2 = 0.$$

◇ **3.8.4** Si  $f : \mathbb{R} \rightarrow \mathbb{R}$  es derivable se le asocia la función de dos variables

$$F(x, y) = \frac{f(x) - f(y)}{x - y} \quad \text{si } x \neq y; \quad F(x, x) = f'(x).$$

Si  $f'$  es continua demuestre que  $F$  es continua en  $\mathbb{R}^2$ . Considerando la función  $f(t) = t^2 \operatorname{sen}(1/t)$  si  $t \neq 0$ ,  $f(0) = 0$ , muestre que el resultado es falso cuando  $f'$  no es continua.

◇ **3.8.5** Demuestre que  $K \subset \mathbb{R}^n$  es compacto si y sólo si toda función continua de  $K$  en  $\mathbb{R}$  es acotada.

◇ **3.8.6** Si  $(E, d)$  es un espacio métrico,  $\mathbf{x} \in E$  y  $A, B$  son subconjuntos no vacíos de  $E$ , se define

$$d(\mathbf{x}, B) = \inf\{d(\mathbf{x}, \mathbf{b}) : \mathbf{b} \in B\}; \quad d(A, B) = \inf\{d(\mathbf{a}, \mathbf{b}) : \mathbf{a} \in A, \mathbf{b} \in B\}.$$

Demuestre las siguientes afirmaciones:

- La aplicación  $\mathbf{x} \rightarrow d(\mathbf{x}, A)$  es uniformemente continua y  $\mathbf{x} \in \overline{A}$  si y sólo si  $d(\mathbf{x}, A) = 0$ .
- Si  $A, B \subset E$  son cerrados disjuntos, existen abiertos disjuntos  $U, V$  tales que  $A \subset U$  y  $B \subset V$ .
- Si  $A$  es compacto,  $d(A, B) = d(\mathbf{a}, B)$  para algún  $\mathbf{a} \in A$ . Si además  $B$  es cerrado, entonces  $d(A, B) > 0$  si y sólo si  $A \cap B = \emptyset$ . Muestre que esta afirmación es falsa cuando sólo se supone que  $A$  y  $B$  son cerrados: Obtenga dos cerrados disjuntos  $A, B \subset \mathbb{R}^2$  tales que no existen  $\mathbf{a} \in A, \mathbf{b} \in B$  con  $d(\mathbf{a}, \mathbf{b}) = d(A, B) = 0$ .
- Si  $A \subset \mathbb{R}^n$  es cerrado y  $d$  es la distancia usual en  $\mathbb{R}^n$ , entonces para cada compacto  $B \subset \mathbb{R}^n$  existen  $\mathbf{a} \in A$  y  $\mathbf{b} \in B$  tales que  $d(\mathbf{a}, \mathbf{b}) = d(A, B)$ . (en particular, para cada  $\mathbf{b} \in \mathbb{R}^n$  existe  $\mathbf{a} \in A$  tal que  $d(\mathbf{b}, A) = d(\mathbf{b}, \mathbf{a})$ ).

◇ **3.8.7** Un subespacio  $M$  de un espacio normado  $(E, \|\cdot\|)$  se dice que tiene la propiedad de aproximación óptima si para cada  $\mathbf{x} \in E$  existe  $\mathbf{p}(\mathbf{x}) \in M$  tal que

$$\|\mathbf{x} - \mathbf{p}(\mathbf{x})\| = \min\{\|\mathbf{x} - \mathbf{y}\| : \mathbf{y} \in M\}$$

Demuestre que todo subespacio finito dimensional tiene esta propiedad.

Si  $\mathcal{P}_m$  es el conjunto de los polinomios de grado  $\leq m$  y  $f : [0, 1] \rightarrow \mathbb{R}$  es continua, demuestre que existe  $q \in \mathcal{P}_m$  tal que para todo  $p \in \mathcal{P}_m$  se cumple

$$\max\{|f(t) - q(t)| : t \in [0, 1]\} \leq \max\{|f(t) - p(t)| : t \in [0, 1]\}$$

◇ **3.8.8** Sea  $(E, d)$  un espacio métrico completo. A un conjunto abierto  $V \subset E$  se le asocia la función  $\rho_V(\mathbf{x}) = 1/d(\mathbf{x}, V^c)$ . Demuestre que

$$d_V(\mathbf{x}, \mathbf{y}) = d(\mathbf{x}, \mathbf{y}) + |\rho_V(\mathbf{x}) - \rho_V(\mathbf{y})|$$

define en  $V$  una distancia, equivalente a la inducida por  $d$  en  $V$ , con la cual  $(V, d_V)$  es un espacio métrico completo.

◇ **3.8.9** Demuestre que  $f(\mathbf{x}) = \frac{\|\mathbf{x}\|}{1+\|\mathbf{x}\|}$  es uniformemente continua en todo  $\mathbb{R}^n$ .



◇ **3.8.10** Sea  $M = A \cup B \subset \mathbb{R}^n$  donde  $A$  y  $B$  son conjuntos cerrados. Se supone que  $f : M \rightarrow \mathbb{R}$  es continua y que  $f|_A, f|_B$  son uniformemente continuas. Demuestre que  $f$  es uniformemente continua cuando alguno de los conjuntos  $A, B$  es acotado. Utilice el siguiente ejemplo para mostrar que este resultado es falso cuando  $A$  y  $B$  no son acotados.

EJEMPLO:  $A = \{(x, y) \in \mathbb{R}^2 : xy \geq 1\}$ ,  $B = \{(x, y) \in \mathbb{R}^2 : xy \leq -1\}$  y  $f : A \cup B \rightarrow \mathbb{R}$  definida por  $f(x, y) = x$  si  $(x, y) \in A$ ,  $f(x, y) = y$  si  $(x, y) \in B$ .

◇ **3.8.11** Demuestre que la función  $\frac{\operatorname{sen} x + \operatorname{sen} y}{1 + x^2 + y^2}$  es uniformemente continua en  $\mathbb{R}^2$  y alcanza un máximo y un mínimo absoluto en  $\mathbb{R}^2$ .

◇ **3.8.12** Sea  $f : [0, 1] \times [0, +\infty) \rightarrow \mathbb{R}$  una función continua tal que  $|f(x, y)| \leq \varphi(y)$ , para todo  $x \in [0, 1]$ , y todo  $y \geq 0$ , donde  $\varphi : [0, +\infty) \rightarrow [0, +\infty)$ , verifica  $\lim_{y \rightarrow +\infty} \varphi(y) = 0$ . Demuestre que  $f$  es uniformemente continua. y que  $|f|$  alcanza un máximo absoluto.

◇ **3.8.13** Estudie la continuidad uniforme de las siguientes funciones en el conjunto  $A$  que se indica en cada caso:

i)  $f(x, y) = \frac{x}{1 + x^2} + \frac{y}{1 + y^2}$ ;  $A = \mathbb{R}^2$ .

ii)  $f(x, y) = \frac{\operatorname{sen}^2(x^2 + y^2 - 1)}{\cos(x^2 + y^2 - 1)}$ ;  $A = \{(x, y) : x^2 + y^2 < 1 + \pi/2\}$ .

iii)  $f(x, y) = (x + y) \operatorname{sen}(1/x) \operatorname{sen}(1/y)$  si  $x \neq 0$  e  $y \neq 0$ ,  $f(0, y) = f(x, 0) = 0$ ;  
 $A = \{(x, y) : 0 < x \leq 1, 0 < y \leq 1\}$ .

iv)  $f(x, y) = x + y/x$ ;  $A = \{(x, y) : x \neq 0\}$ .

v)  $f(x, y) = \frac{xy}{x - 1}$ ;  $A = \{(x, y) : x^2 + y^2 < 1\}$ .

vi)  $f(x, y) = \frac{x^3 - y^3}{x^2 + y^2}$  si  $(x, y) \neq (0, 0)$ ,  $f(0, 0) = 0$ ;  $A = \mathbb{R}^2$ .

vii)  $f(x, y) = \cos^3\left(\frac{1}{x^2 + y^2}\right)$ ;  $A = \{(x, y) : x^2 + y^2 > 1\}$ .

viii)  $f(x, y) = (x^2 + y^2) \operatorname{sen}^3\left(\frac{1}{x^2 + y^2}\right)$  si  $(x, y) \neq (0, 0)$ ,  $f(0, 0) = 0$ ;  $A = \mathbb{R}^2$

ix)  $f(x, y) = \frac{\operatorname{sen}(x + y)}{|x| + |y|}$ ,  $A = \{(x, y) : x \geq 0, y \geq 0, x + y > 0\}$ .

◇ **3.8.14** Sea  $f : M \rightarrow \mathbb{R}$  una función continua en  $M = \{\mathbf{x} \in \mathbb{R}^n : 0 < \|\mathbf{x}\|_2 \leq 1\}$ . Demuestre que  $f$  es uniformemente continua si y sólo si existe el límite  $\lim_{\mathbf{x} \rightarrow 0} f(\mathbf{x})$ .

◇ **3.8.15** Si  $f : [a, b] \times [c, d] \rightarrow \mathbb{R}$  es continua y la sucesión  $x_n \in [a, b]$  converge demuestre que la sucesión  $f_n(t) = f(x_n, t)$  converge uniformemente sobre  $[c, d]$ . Demuestre que la sucesión  $f_n(t) = (n/t) \log(1 + t/n)$  converge uniformemente sobre  $[0, d]$ , pero no converge uniformemente sobre  $[0, +\infty)$ .

◇ **3.8.16** a) Si  $f : [a, b] \times [0, +\infty) \rightarrow \mathbb{R}$  es uniformemente continua y la sucesión  $x_n \in [a, b]$  converge demuestre que la sucesión  $f_n(t) = f(x_n, t)$  converge uniformemente sobre  $[0, +\infty)$ .

b) Sea  $g : [0, 1] \times [0, +\infty) \rightarrow \mathbb{R}$  definida por  $g(x, y) = \frac{\log(1+xy)}{xy}$  si  $xy \neq 0$ ;  $g(x, y) = 1$  si  $xy = 0$ . Utilice a) para demostrar que la sucesión  $f_n(t) = (n/t) \log(1 + t/n)$  converge uniformemente sobre cada intervalo  $[0, r]$ . Compruebe que la sucesión  $f_n$  no converge uniformemente sobre  $[0, +\infty)$  y obtenga que  $g$  no es uniformemente continua.

◇ **3.8.17** Sean  $T, E$  espacios métricos y  $f_n : T \rightarrow E$  una sucesión de funciones continuas que converge uniformemente sobre  $A \subset T$ . Si  $E$  es completo, demuestre que la sucesión también converge uniformemente sobre  $\overline{A}$ .

◇ **3.8.18** Sean  $(E, d), (F, d')$  espacios métricos y  $f_n : E \rightarrow F$  una sucesión que converge uniformemente sobre compactos hacia  $f : E \rightarrow F$ . Si  $g : F \rightarrow \mathbb{R}$  es uniformemente continua compruebe que  $g \circ f_n$  converge uniformemente sobre compactos. Demuestre el mismo resultado cuando sólo se supone que  $g$  es continua, pero todas las funciones  $f_n$  son continuas.

◇ **3.8.19** Sea  $K$  un subconjunto compacto de un espacio métrico  $(E, d)$  y  $f_n : K \rightarrow \mathbb{R}$  una sucesión de funciones continuas que converge puntualmente hacia una función continua  $f : K \rightarrow \mathbb{R}$ . Si la sucesión  $f_n(\mathbf{x})$  es decreciente para todo  $\mathbf{x} \in K$ , demuestre que la sucesión  $f_n$  es uniformemente convergente sobre  $K$ .

◇ **3.8.20** Sean  $(E, d), (F, d')$  espacios métricos y  $f_n : E \rightarrow F$  una sucesión que converge uniformemente hacia  $f : E \rightarrow F$ . Demuestre las siguientes afirmaciones:

a) Si cada  $f_n$  es uniformemente continua entonces  $f$  también lo es.

b) Si  $g : F \rightarrow \mathbb{R}$  es uniformemente continua entonces  $g \circ f_n$  es uniformemente convergente.

◇ **3.8.21** Se dice que  $\mathbf{g} : [a, b] \rightarrow E$  es escalonada cuando existe una subdivisión  $p = (t_0 < t_1 < t_2 \cdots t_m)$  de  $[a, b]$  tal que  $\mathbf{g}$  es constante en cada intervalo  $(t_{i-1}, t_i)$ , ( $1 \leq i \leq m$ ). Una función  $\mathbf{f} : [a, b] \rightarrow E$  con valores en un espacio normado  $E$  se dice que es reglada cuando es límite uniforme de una sucesión de funciones escalonadas. Demuestre que a)  $\Rightarrow$  b)  $\Rightarrow$  c), y que si  $E$  es completo entonces a)  $\Leftrightarrow$  b).

a) Todas las discontinuidades de  $\mathbf{f} : [a, b] \rightarrow E$  son de primera especie (e.d. en cada  $t \in (a, b)$  existen los límites laterales  $\mathbf{f}(t-), \mathbf{f}(t+)$  y cuando  $x = t$  ó  $t = b$  existe el correspondiente límite lateral).

b)  $\mathbf{f}$  es reglada.

c) El conjunto de las discontinuidades de  $\mathbf{f}$  es numerable.

◇ **3.8.22** Demuestre que la gráfica  $G(\mathbf{f}) = \{(\mathbf{x}, \mathbf{y}) \in E \times F : \mathbf{f}(\mathbf{x}) = \mathbf{y}\}$  de una función continua  $\mathbf{f} : (E, d) \rightarrow (F, d')$  es un subconjunto cerrado de  $E \times F$ . Muestre que el recíproco es falso en general pero es cierto cuando  $F$  es compacto. Demuestre que  $G(\mathbf{f})$  es compacto en  $E \times F$  si y sólo si  $E$  es compacto.

◇ **3.8.23** Si el espacio métrico  $(E, d)$  es compacto y  $\mathbf{f} : E \rightarrow E$  verifica

$$d(\mathbf{f}(\mathbf{x}), \mathbf{f}(\mathbf{y})) < d(\mathbf{x}, \mathbf{y}) \quad \text{para todo } (\mathbf{x}, \mathbf{y}) \in E \times E, \mathbf{x} \neq \mathbf{y}$$

demuestre que  $\mathbf{f}$  tiene un punto fijo. (Indicación: Considere  $g(\mathbf{x}) = d(\mathbf{x}, \mathbf{f}(\mathbf{x}))$ )

◇ **3.8.24** Sea  $(E, d)$  un espacio métrico compacto. Si  $\mathbf{g} : E \rightarrow E$  cumple

$$d(\mathbf{g}(\mathbf{x}), \mathbf{g}(\mathbf{y})) \geq d(\mathbf{x}, \mathbf{y}) \quad \text{para todo } (\mathbf{x}, \mathbf{y}) \in E \times E$$

demuestre que  $\mathbf{g}$  es una isometría.

◇ **3.8.25** Sea  $\mathbf{a} = (a_1, a_2, \dots, a_n) \in \mathbb{R}^n$  y  $f_{\mathbf{a}} : \mathbb{R}^n \rightarrow \mathbb{R}$  la forma lineal definida por  $f_{\mathbf{a}}(\mathbf{x}) = \sum_{i=1}^n a_i x_i$ . Demuestre que  $\|f_{\mathbf{a}}\|_1 = \|\mathbf{a}\|_{\infty}$ ,  $\|f_{\mathbf{a}}\|_2 = \|\mathbf{a}\|_2$ , y  $\|f_{\mathbf{a}}\|_{\infty} = \|\mathbf{a}\|_1$ .

◇ **3.8.26** Sea  $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$  la aplicación lineal definida por  $A(\mathbf{x}) = \mathbf{y}$ , donde  $y_i = \sum_{j=1}^n a_{ij} x_j$ ,  $1 \leq i \leq m$ . Compruebe, para  $p = 1, \infty$ , que la norma  $\|A\|_p = \sup\{\|A(\mathbf{x})\|_p : \|\mathbf{x}\|_p \leq 1\}$ , viene dada por

$$\|A\|_1 = \max\left\{\sum_{i=1}^m |a_{ij}| : 1 \leq j \leq n\right\}; \quad \|A\|_{\infty} = \max\left\{\sum_{j=1}^n |a_{ij}| : 1 \leq i \leq m\right\}.$$

◇ **3.8.27** Sea  $(E, \|\cdot\|)$  un espacio normado, que no se supone completo, cuya norma procede de un producto escalar. Utilice el resultado del problema 3.8.27 para demostrar que si  $A \subset E$  es un conjunto cerrado convexo contenido en un subespacio finito dimensional de  $E$  entonces existe un único  $\mathbf{a} \in A$  que verifica

$$\|\mathbf{a}\| = \min\{\|\mathbf{x}\| : \mathbf{x} \in A\}$$

◇ **3.8.28** Si  $g : [a, b] \rightarrow \mathbb{R}$  es continua, para cada  $f \in C[a, b]$  se define  $T_g(f) = \int_a^b f(t)g(t)dt$ . Demuestre que la forma lineal  $T_g : C[a, b] \rightarrow \mathbb{R}$  es continua para las tres normas  $\|\cdot\|_1, \|\cdot\|_2, \|\cdot\|_{\infty}$ .

◇ **3.8.29** Sea  $\mathcal{P}[0, 1]$  el subespacio de  $C[0, 1]$  formado por las funciones polinomiales, dotado con la norma de la convergencia uniforme  $\|\cdot\|_{\infty}$ . Compruebe que la aplicación lineal  $T : \mathcal{P}[0, 1] \rightarrow \mathbb{R}$  definida por  $T(p) = p(2)$  no es continua.

◇ **3.8.30** Demuestre que una sucesión de polinomios de grado  $\leq m$ ,  $p_n = \sum_{k=0}^m a_k(n)x^k$  es uniformemente convergente en un intervalo  $[a, b]$  si y sólo si para cada  $k \in \{0, 1, 2, \dots, m\}$  la sucesión  $(a_k(n))_{n \in \mathbb{N}}$  converge, y en ese caso el límite uniforme es el polinomio  $p(x) = \sum_{k=0}^m a_k x^k$  de coeficientes  $a_k = \lim_n a_k(n)$ ,  $1 \leq k \leq m$ .

# Capítulo 4

## Funciones vectoriales de una variable

*Derivación de funciones vectoriales de una variable. Teorema del incremento finito y desarrollo de Taylor. Longitud de un arco de curva. Integral respecto al arco. Aplicaciones*

Este capítulo está dedicado al cálculo diferencial e integral para funciones vectoriales de variable real. En este contexto la noción de derivada está motivada por el problema de trazar tangentes a curvas dadas en forma paramétrica. También sirve para formular la noción física de velocidad instantánea de una partícula que se mueve en el espacio.

Para una función de variable real con valores vectoriales la derivada se define como en el caso de funciones con valores reales. Para que la definición tenga sentido basta que la función tome valores en un espacio vectorial dotado de una norma, lo que permite formar el cociente incremental y considerar la existencia de su límite.

El cálculo diferencial de funciones vectoriales de variable real se desarrolla de modo paralelo al de las funciones con valores reales, con pequeñas diferencias que surgen en relación con el teorema del incremento finito. Naturalmente que en el caso de funciones vectoriales hay una serie de cuestiones que no se plantean, como son las relativas al signo de la derivada, crecimiento y decrecimiento, convexidad, etc. La primera novedad respecto al caso de las funciones con valores reales surge en que ahora hay interpretaciones geométricas y físicas muy interesantes: La derivada primera y la derivada segunda proporcionan, respectivamente, la velocidad y la aceleración de una partícula que se mueve en el espacio.

Uno de los resultados centrales del capítulo es la versión del teorema del incremento finito para funciones vectoriales de variable real (4.7, 4.8). En este contexto no es válida la formulación habitual en términos del valor de la derivada en un punto intermedio, y la dificultad se resuelve con una formulación en términos de desigualdad en la que interviene una cota de la derivada en el intervalo donde se aplica. La versión con desigualdad del teorema del incremento finito es suficiente para proseguir con el cálculo diferencial y obtener el desarrollo de Taylor de una función vectorial de una variable en un punto donde la función es derivable  $m$  veces, con el resto (o

término complementario) en forma infinitesimal. Igual que ocurre con el teorema del incremento finito, para las funciones vectoriales de variable real de clase  $C^{m+1}$ , no es válida la forma habitual del resto en la forma de Lagrange, donde interviene un punto intermedio del intervalo. Sin embargo sigue valiendo la fórmula integral del resto y con el fin de obtenerla se hace una breve incursión en la integración vectorial en el ámbito de las funciones continuas con valores en el espacio euclídeo  $\mathbb{R}^n$ .

En esta situación es fácil comprobar, razonando componente a componente, que siguen valiendo los resultados básicos del cálculo integral: Integrabilidad de las funciones continuas, desigualdad triangular para las integrales, teorema fundamental del cálculo, regla de Barrow, cambio de variable e integración por partes.

Como material complementario relacionado con este tema se puede consultar en el apéndice D la definición de integral para el caso general de funciones de variable real con valores en un espacio normado completo.

Con el teorema del incremento finito se demuestra que si el movimiento de una partícula lo describe un camino de clase  $C^1$  entonces la norma del vector velocidad coincide, en cada instante, con la celeridad de la partícula (la magnitud escalar que mide el cuenta kilómetros de un automóvil). Una consecuencia inmediata de este hecho es la fórmula integral para el cálculo de la longitud de la trayectoria recorrida por un punto que se mueve siguiendo una trayectoria regular a trozos (véase 4.24).

La noción de camino rectificable se expone como caso particular de la noción más general de función de variación acotada. Estas funciones cuando toman valores en  $\mathbb{R}$  o  $\mathbb{R}^n$  se caracterizan fácilmente en términos de funciones crecientes (véase 4.30 y 4.7.14). Uno de los objetivos de este capítulo es el de ir progresando en la noción de camino o arco de curva "razonable". La clase de los caminos rectificables es razonable desde varios puntos de vista, ya que incluye a los caminos regulares a trozos y no contiene los ejemplos patológicos considerados en el apéndice A, como la curva de Peano A.18 y la gráfica de la función de Weierstrass A.17. Un resultado profundo de la teoría de funciones afirma que los caminos rectificables con valores en  $\mathbb{R}^n$  son derivables en casi todo punto, hecho que contrasta con el camino continuo que describe la gráfica de la función de Weierstrass, que no es derivable en ningún punto.

Para caminos rectificables, usando la función abscisa curvilínea  $s = v(t)$  que da la longitud  $s$  del camino recorrido en el instante  $t$  se formula la definición de integral de una función respecto al arco del camino. Para curvas planas, esta integral se puede interpretar como el área de un trozo de superficie cilíndrica cuyas generatrices son ortogonales a la curva dada. Como otra aplicación de esta noción de integral se puede citar el cálculo de la masa total y del centro de masa de un alambre que sigue una curva regular, cuando se conoce la función de densidad que describe la distribución de la masa.

## 4.1. Derivada de una función vectorial

En lo que sigue las funciones se suponen definidas en un abierto  $\Omega \subset \mathbb{R}$ , o en un intervalo  $I \subset \mathbb{R}$ , con valores en un espacio normado  $(F, \|\cdot\|)$ , aunque el lector que lo desee puede suponer la situación habitual donde  $F$  es  $\mathbb{R}^n$  dotado de cualquier norma (recuérdese que en  $\mathbb{R}^n$  todas las normas son equivalentes).

**Definición 4.1** Sea  $\mathbf{f} : \Omega \rightarrow F$  una función de variable real, definida en un abierto  $\Omega \subset \mathbb{R}$ , con valores en un espacio normado  $(F, \|\cdot\|)$ . La derivada de  $\mathbf{f}$  en  $a \in \Omega$  es el vector

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

en el supuesto de que el límite exista. Si  $\mathbf{f}$  es derivable en cada  $a \in \Omega$ , se dice que  $\mathbf{f}$  es derivable en  $\Omega$ .

Si una función es derivable en  $a$  y se cambia la norma de  $F$  por otra equivalente, la función sigue siendo derivable en  $a$ , con la misma derivada, ya que el límite es una noción topológica que no cambia al reemplazar una norma por otra equivalente.

Como en el caso de funciones con valores reales, se definen las derivadas laterales, por la derecha y por la izquierda:

$$\mathbf{f}'_d(a) = \lim_{t \rightarrow a^+} \frac{\mathbf{f}(t) - \mathbf{f}(a)}{t-a}; \quad \mathbf{f}'_i(a) = \lim_{t \rightarrow a^-} \frac{\mathbf{f}(t) - \mathbf{f}(a)}{t-a}$$

y es inmediato comprobar que  $\mathbf{f}$  es derivable en  $a$  si y sólo si existen y son iguales las derivadas laterales  $\mathbf{f}'_i(a) = \mathbf{f}'_d(a)$ .

Una función  $\mathbf{f} : I \rightarrow F$  que está definida en un intervalo  $I \subset \mathbb{R}$  de extremos  $\alpha = \inf I \geq -\infty$ ,  $\beta = \sup I \leq +\infty$ , se dice que es derivable cuando es derivable en cada  $x \in (\alpha, \beta)$ , derivable por la derecha en  $\alpha$  cuando  $\alpha \in I$ , y derivable por la izquierda en  $\beta$  cuando  $\beta \in I$ .

Si  $\mathbf{f}$  es derivable en cada punto de un abierto  $\Omega \subset \mathbb{R}$  queda definida la función derivada  $\mathbf{f}' : \Omega \rightarrow F$ . Si esta función es continua se dice que  $\mathbf{f}$  es de clase  $C^1$  en  $\Omega$ , y se escribe  $\mathbf{f} \in C^1(\Omega, F)$ . Si  $\mathbf{f}'$  es derivable en  $a \in \Omega$  su derivada se llama derivada segunda  $\mathbf{f}''(a) = (\mathbf{f}')'(a)$ . Si existe  $\mathbf{f}''(t)$  en cada  $t \in \Omega$  queda definida la función derivada segunda  $\mathbf{f}'' : \Omega \rightarrow F$ , y si esta función es continua se dice que  $\mathbf{f}$  es de clase  $C^2$  y se escribe  $\mathbf{f} \in C^2(\Omega, F)$ . De manera recurrente se define la derivada  $m$ -ésima, denotada  $\mathbf{f}^{(m)}(a)$ . Obsérvese que la existencia de  $\mathbf{f}^{(m)}(a)$  lleva implícita la existencia de  $\mathbf{f}^{(m-1)}(t)$  en todos los puntos de algún entorno de  $a$ .

El espacio de las funciones  $\mathbf{f} : \Omega \rightarrow F$  con derivada  $m$ -ésima continua se denota  $C^m(\Omega, F)$ . Análogamente, cuando el dominio de  $\mathbf{f}$  es un intervalo  $I \subset \mathbb{R}$ , haciendo intervenir las derivadas laterales en los extremos del intervalo que le pertenecen se definen las derivadas sucesivas y el espacio  $C^m(I, F)$ .

**Proposición 4.2** Sea  $A \subset \mathbb{R}$  un intervalo, o un conjunto abierto. Toda función  $\mathbf{f} : A \rightarrow F$  derivable en  $a \in A$  es continua en  $a$ .

Toda función  $\mathbf{f} : A \rightarrow F$  derivable por la derecha (resp. izquierda) en  $a \in A$  es continua por la derecha (resp. izquierda) en  $a$ .

DEM: Cuando  $t \rightarrow a$  el cociente  $\Delta \mathbf{f}(t) = [\mathbf{f}(t) - \mathbf{f}(a)]/(t - a)$  tiene límite luego, en virtud de 3.8,  $(t - a)\Delta \mathbf{f}(t)$  tiende hacia cero, es decir  $\lim_{t \rightarrow a} \mathbf{f}(t) = \mathbf{f}(a)$ .

El mismo razonamiento se aplica para las derivadas laterales con las que se obtiene la correspondiente continuidad lateral. ■

NOTA: El recíproco de la proposición 4.2 es falso: La función real de variable real  $f(x) = |x|$  es continua pero no es derivable en  $a = 0$ .

Cuando  $F = \mathbb{R}^n$ , dotado de cualquier norma, (recuérdese que en  $\mathbb{R}^n$  todas las normas son equivalentes) se tiene:

**Proposición 4.3** *Sea  $A \subset \mathbb{R}$  un intervalo, o un conjunto abierto y  $\mathbf{f} : A \rightarrow \mathbb{R}^n$ , de componentes  $\mathbf{f}(t) = (f_1(t), f_2(t), \dots, f_n(t))$ . Entonces  $\mathbf{f}$  es derivable en  $a \in A$  si y sólo si todas sus componentes lo son, y en este caso  $\mathbf{f}'(a) = (f'_1(a), f'_2(a), \dots, f'_n(a))$ .*

DEM: Las componentes del cociente incremental  $\Delta \mathbf{f}(t) = [\mathbf{f}(t) - \mathbf{f}(a)]/(t - a)$  son los cocientes incrementales de las componentes de  $\mathbf{f}$ :

$$\Delta \mathbf{f}(t) = (\Delta f_1(t), \Delta f_2(t), \dots, \Delta f_n(t))$$

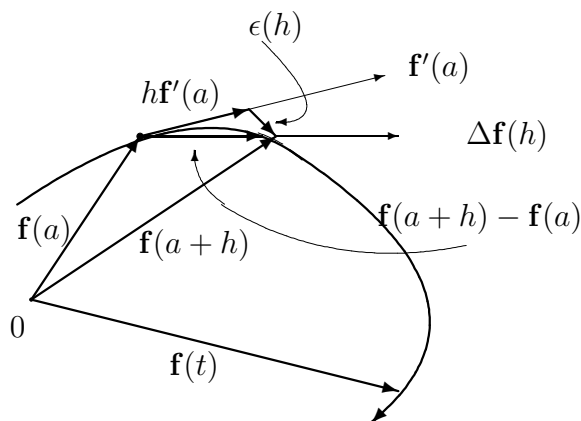
Según 3.7, una condición necesaria y suficiente para que  $\Delta \mathbf{f}(t)$  tenga límite, cuando  $t \rightarrow a$ , es que todas sus componentes  $\Delta f_i(t)$ ,  $1 \leq i \leq n$ , tengan límite, y en este caso, el límite es el vector cuyas componentes son los límites de las componentes, es decir  $(f'_1(a), f'_2(a), \dots, f'_n(a))$ . ■

**Interpretaciones física y geométrica de la derivada.** Cuando  $F$  es el espacio euclídeo  $\mathbb{R}^3$  (resp.  $\mathbb{R}^2$ ) para una función de variable real  $\mathbf{f} : (\alpha, \beta) \rightarrow F$  se puede interpretar que  $t$  es el tiempo y que  $\mathbf{f}(t)$  es la posición, en el instante  $t$ , de una partícula que se mueve en el espacio (resp. en el plano).

El cambio de posición de la partícula, desde el instante  $t = a$  hasta el instante  $t = a + h$ , viene dado por el vector cuerda  $\mathbf{f}(a + h) - \mathbf{f}(a)$  que se representa mediante una flecha con origen en  $\mathbf{f}(a)$  y extremo en  $\mathbf{f}(a + h)$ . La velocidad media de la partícula, durante el intervalo de tiempo  $(a, a + h)$  viene dada por el cociente incremental  $\Delta \mathbf{f}(h) = [\mathbf{f}(a + h) - \mathbf{f}(a)]/h$ . Si este cociente tiene límite cuando  $h$  tiende hacia 0 el límite  $\mathbf{f}'(a)$  es un vector que representa la velocidad instantánea de la partícula en el instante  $t = a$ . En esta interpretación física la derivada segunda  $\mathbf{f}''(a)$  proporciona la aceleración de la partícula en ese instante.

Como el cociente incremental  $[\mathbf{f}(a + h) - \mathbf{f}(a)]/h$  tiene la dirección del vector cuerda,  $\mathbf{f}(a + h) - \mathbf{f}(a)$ , cuando no es nulo el límite  $\mathbf{f}'(a)$ , la dirección de este vector será el límite de las direcciones de las cuerdas determinadas por  $\mathbf{f}(a)$  y los puntos cada vez más próximos  $\mathbf{f}(a + h)$ . Por esta razón se dice que  $\mathbf{f}'(a)$  es un vector tangente a la trayectoria  $\mathbf{f}(t)$  en el instante  $t = a$ , que se suele representar mediante una flecha con origen en el punto  $\mathbf{f}(a)$ .

Más adelante se demostrará que la longitud del vector velocidad  $\|\mathbf{f}'(t)\|$  es la *celeridad* o *rapidez* de la partícula (derivada respecto al tiempo, del espacio recorrido sobre la trayectoria).



**Definición 4.4** Si  $\mathbf{f} : \Omega \rightarrow F$  es derivable en  $a \in \Omega$ , la recta tangente a la trayectoria  $\mathbf{f}(t)$  en el instante  $t = a$  es la que pasa por  $\mathbf{f}(a)$  en la dirección del vector  $\mathbf{f}'(a)$ , de ecuación paramétrica

$$\mathbf{r}(t) = \mathbf{f}(a) + (t - a)\mathbf{f}'(a)$$

Si se utiliza la recta tangente  $\mathbf{r}(t)$  como aproximación local de la trayectoria  $\mathbf{f}(t)$  en un entorno de  $t = a$ , el error cometido al pasar del valor  $t = a$  al valor  $t = a + h$  viene dado por

$$\boldsymbol{\epsilon}(h) = \mathbf{f}(a + h) - \mathbf{r}(a + h) = \mathbf{f}(a + h) - \mathbf{f}(a) - h\mathbf{f}'(a)$$

Cuando  $h$  es muy pequeño, el tamaño de error  $\|\boldsymbol{\epsilon}(h)\|$  es despreciable frente a  $h$  ya que, en virtud de la definición de derivada, se cumple

$$\lim_{h \rightarrow 0} \frac{\|\boldsymbol{\epsilon}(h)\|}{h} = \mathbf{0}$$

condición que se suele denotar escribiendo  $\boldsymbol{\epsilon}(h) = o(h)$ .

Cuando dos funciones de variable real  $\mathbf{f}(t)$ ,  $\mathbf{g}(t)$ , definidas en un entorno de  $t = a$  con  $\mathbf{f}(a) = \mathbf{g}(a)$  verifican la condición  $\mathbf{f}(a + h) - \mathbf{g}(a + h) = o(h)$  se suele decir que  $\mathbf{f}$  y  $\mathbf{g}$  presentan en  $a$  una tangencia de primer orden, y también que  $\mathbf{g}$  es una aproximación local de primer orden de  $\mathbf{f}$  en el punto  $a$ . Es fácil comprobar que una condición suficiente para que esto ocurra es que  $\mathbf{f}$  y  $\mathbf{g}$  sean derivables en  $a$ , con la misma derivada,  $\mathbf{f}'(a) = \mathbf{g}'(a)$ . Cuando  $\mathbf{f}$  es derivable en el punto  $a$ , la recta  $\mathbf{r}$  tangente a  $\mathbf{f}$  en  $a$  es una aproximación local de primer orden de  $\mathbf{f}$  en ese punto.

### Operaciones con funciones derivables

**Proposición 4.5** Sea  $(F, \|\cdot\|)$  un espacio vectorial normado, y  $\Omega \subset \mathbb{R}$  abierto. Si  $\mathbf{f}, \mathbf{g} : \Omega \rightarrow F$  y  $\alpha : \Omega \rightarrow \mathbb{R}$  son derivables en  $a \in \Omega$ , entonces la suma  $\mathbf{f} + \mathbf{g}$ , el producto  $\alpha\mathbf{f}$  y el cociente  $\mathbf{f}/\alpha := \alpha^{-1}\mathbf{f}$  (cuando  $\alpha(a) \neq 0$ ) también son derivables en  $a$ , y se verifica:



$$i) (\mathbf{f} + \mathbf{g})'(a) = \mathbf{f}'(a) + \mathbf{g}'(a);$$

$$ii) (\alpha \mathbf{f})'(a) = \alpha'(a)\mathbf{f}(a) + \alpha(a)\mathbf{f}'(a);$$

$$iii) (\mathbf{f}/\alpha)'(a) = [\alpha(a)\mathbf{f}'(a) - \alpha'(a)\mathbf{f}(a)]/\alpha(a)^2$$

Si la norma de  $F$  procede de un producto escalar  $\langle \cdot | \cdot \rangle$ , también se cumple

iv) La función producto escalar  $h(t) = \langle \mathbf{f}(t) | \mathbf{g}(t) \rangle$  es derivable en  $a$ , y

$$h'(a) = \langle \mathbf{f}(a) | \mathbf{g}'(a) \rangle + \langle \mathbf{f}'(a) | \mathbf{g}(a) \rangle$$

DEM: Las demostraciones i), ii) y iii) son análogas a las del caso de funciones con valores reales y se dejan al cuidado del lector. En relación con iii) se debe señalar que al ser  $\alpha$  continua en  $a$ , con  $\alpha(a) \neq 0$ , debe existir un entorno de  $a$  donde  $\alpha(t)$  no se anula y en él está definido el cociente  $\mathbf{f}(t)/\alpha(t) = \alpha(t)^{-1}\mathbf{f}(t)$ .

Para demostrar iv), en virtud de la bilinealidad del producto escalar podemos escribir

$$h(t) - h(a) = \langle \mathbf{f}(t) | \mathbf{g}(t) - \mathbf{g}(a) \rangle + \langle \mathbf{f}(t) - \mathbf{f}(a) | \mathbf{g}(a) \rangle$$

y obtenemos la siguiente expresión para  $\Delta h(t) = [h(t) - h(a)]/(t - a)$ :

$$\Delta h(t) = \langle \mathbf{f}(t) | \Delta \mathbf{g}(t) \rangle + \langle \Delta \mathbf{f}(t) | \mathbf{g}(a) \rangle$$

Pasando al límite cuando  $t \rightarrow a$  se obtiene el resultado, ya que el producto escalar de dos funciones vectoriales con límite tiene límite y su valor es el producto escalar de los límites (véase 3.8) ■

**Proposición 4.6** [Regla de la cadena] Sean  $\Omega, V \subset \mathbb{R}$  abiertos y  $(F, \|\cdot\|)$  un espacio normado. Si  $\varphi : \Omega \rightarrow V$  es derivable en  $a \in \Omega$ , y  $\mathbf{f} : V \rightarrow F$  es derivable en  $b = \varphi(a)$ , entonces la función compuesta,  $\mathbf{g} : \Omega \rightarrow F$ ,  $\mathbf{g}(t) = \mathbf{f}(\varphi(t))$ , es derivable en  $a$  y  $\mathbf{g}'(a) = \varphi'(a)\mathbf{f}'(\varphi(a))$ .

DEM: En virtud de la definición de derivada las funciones  $\delta : \Omega \rightarrow \mathbb{R}$ ,  $\Delta : V \rightarrow F$ , definidas por

$$\delta(t) = \frac{\varphi(t) - \varphi(a)}{t - a} \quad \text{si } t \in \Omega \setminus \{a\}, \quad \delta(a) = \varphi'(a)$$

$$\Delta(x) = \frac{\mathbf{f}(x) - \mathbf{f}(b)}{x - b} \quad \text{si } x \in V \setminus \{b\}, \quad \Delta(b) = \mathbf{f}'(b)$$

son continuas en  $a$  y en  $b$  respectivamente. Para todo  $t \in \Omega$  y todo  $x \in V$  se cumple:

$$\mathbf{f}(x) - \mathbf{f}(b) = (x - b)\Delta(x); \quad \varphi(t) - \varphi(a) = (t - a)\delta(t);$$

y sustituyendo en la primera igualdad  $x = \varphi(t)$ ,  $b = \varphi(a)$ , resulta

$$\mathbf{g}(t) - \mathbf{g}(a) = \mathbf{f}(\varphi(t)) - \mathbf{f}(\varphi(a)) = [\varphi(t) - \varphi(a)]\Delta(\varphi(t)) = (t - a)\delta(t)\Delta(\varphi(t))$$

Como  $\varphi(t)$  es continua en  $a$  (por ser derivable), y  $\Delta(x)$  es continua en  $b = \varphi(a)$ , su composición  $\Delta(\varphi(t))$  es continua en  $a$ , luego  $\lim_{t \rightarrow a} \Delta(\varphi(t)) = \Delta(b) = \mathbf{f}'(b)$ , y por lo tanto existe el límite

$$\mathbf{g}'(a) = \lim_{t \rightarrow a} \frac{\mathbf{g}(t) - \mathbf{g}(a)}{t - a} = \lim_{t \rightarrow a} \delta(t) \Delta(\varphi(t)) = \varphi'(a) \mathbf{f}'(b)$$

■

El teorema del valor medio para funciones reales de variable real asegura que si  $f : [a, b] \rightarrow \mathbb{R}$  es continua y derivable en  $(a, b)$  entonces existe  $\eta \in (a, b)$  tal que

$$f(b) - f(a) = f'(\eta)(b - a) \quad (*)$$

Cuando  $f'$  está acotada,  $|f'(x)| \leq M$  para todo  $x \in (a, b)$ , se obtiene que

$$|f(b) - f(a)| \leq M(b - a) \quad (**)$$

Para funciones vectoriales de variable real no se cumple el teorema del valor medio: Para  $\mathbf{f} : [0, 2\pi] \rightarrow \mathbb{R}^2$  definida por  $\mathbf{f}(t) = (\cos t, \sin t)$  se cumple  $\mathbf{f}(2\pi) - \mathbf{f}(0) = (0, 0)$ , pero  $\mathbf{f}'(t) = (-\sin t, \cos t) \neq (0, 0)$  para todo  $t \in [0, 2\pi]$ , es decir, no hay un punto intermedio  $\eta \in [0, 2\pi]$  donde se satisfaga la igualdad (\*). Un ejemplo similar lo suministra la aplicación  $\mathbf{g} : \mathbb{R} \rightarrow \mathbb{R}^3$ ,  $\mathbf{g}(t) = (\cos t, \sin t, t)$ , cuya imagen es la curva llamada hélice. El vector tangente a la hélice  $\mathbf{g}'(t) = (-\sin t, \cos t, 1)$  nunca es vertical, mientras que el vector  $\mathbf{g}(2\pi) - \mathbf{g}(0) = (0, 0, 2\pi)$  sí lo es.

Aunque el teorema del valor medio no subsiste para funciones con valores en un espacio normado  $F$  de dimensión  $\geq 2$ , sin embargo la acotación (\*\*), cuando la derivada es acotada, sigue valiendo para el caso de funciones con valores en cualquier espacio normado. Esto se obtendrá como corolario del siguiente teorema:

**Teorema 4.7** [Incremento finito] *Sea  $(F, \|\cdot\|)$  un espacio normado. Se supone que  $\mathbf{f} : [a, b] \rightarrow F$ , y  $g : [a, b] \rightarrow \mathbb{R}$  son funciones continuas en  $[a, b]$ , derivables por la derecha en  $(a, b)$  que verifican  $\|\mathbf{f}'_d(t)\| \leq g'_d(t)$  para todo  $t \in (a, b)$ . Entonces*

$$\|\mathbf{f}(b) - \mathbf{f}(a)\| \leq g(b) - g(a)$$

DEM: Para cada  $\epsilon > 0$  sea  $S_\epsilon \subset [a, b]$  el subconjunto formado por los  $x \in [a, b]$  que cumplen  $\|\mathbf{f}(x) - \mathbf{f}(a)\| \leq g(x) - g(a) + \epsilon(x - a) + \epsilon$ . Si se demuestra que  $b \in S_\epsilon$ , pasando al límite, cuando  $\epsilon \rightarrow 0$ , en la desigualdad

$$\|\mathbf{f}(b) - \mathbf{f}(a)\| \leq g(b) - g(a) + \epsilon(b - a) + \epsilon$$

se obtiene el resultado deseado.

La función  $h(x) = \|\mathbf{f}(x) - \mathbf{f}(a)\| - [g(x) - g(a) + \epsilon(x - a)]$  es continua en  $[a, b]$ , y  $h(a) = 0$  luego existe  $\delta > 0$  tal que  $h(t) < \epsilon$  para todo  $t \in [a, a + \delta)$ , es decir  $S_\epsilon$  contiene al intervalo  $[a, a + \delta)$  y por lo tanto  $S_\epsilon \neq \emptyset$ . Entonces podemos considerar el supremo  $\mu = \sup S_\epsilon$  que es adherente a  $S_\epsilon$ . La continuidad de  $h$

implica que  $S_\epsilon = \{x \in [a, b] : h(x) \leq \epsilon\}$  es un subconjunto cerrado de  $[a, b]$  y por lo tanto  $\mu \in S_\epsilon$ . Para terminar basta que ver  $\mu = b$ , y esto lo demostraremos por reducción al absurdo suponiendo  $\mu < b$ .

Como  $\mu \geq a + \delta > a$ , las funciones  $\mathbf{f}$ ,  $g$  son derivables por la derecha en  $\mu$ , y por lo tanto existe  $\mu < r < b$  tal que si  $\mu < t < r$  se cumple

$$\left\| \frac{\mathbf{f}(t) - \mathbf{f}(\mu)}{t - \mu} - \mathbf{f}'_d(\mu) \right\| < \epsilon/2, \quad \left| \frac{g(t) - g(\mu)}{t - \mu} - g'_d(\mu) \right| < \epsilon/2$$

y en virtud de la desigualdad triangular

$$\left\| \frac{\mathbf{f}(t) - \mathbf{f}(\mu)}{t - \mu} \right\| \leq \|\mathbf{f}'_d(\mu)\| + \epsilon/2 \leq g'_d(\mu) + \epsilon/2 \leq \frac{g(t) - g(\mu)}{t - \mu} + \epsilon/2 + \epsilon/2$$

luego

$$\|\mathbf{f}(t) - \mathbf{f}(\mu)\| \leq g(t) - g(\mu) + \epsilon(t - \mu)$$

Por otra parte, como  $\mu \in S_\epsilon$  se cumple

$$\|\mathbf{f}(\mu) - \mathbf{f}(a)\| \leq g(\mu) - g(a) + \epsilon(\mu - a) + \epsilon$$

Usando la desigualdad triangular y sumando miembro a miembro las últimas desigualdades se obtiene

$$\begin{aligned} \|\mathbf{f}(t) - \mathbf{f}(a)\| &\leq \|\mathbf{f}(t) - \mathbf{f}(\mu)\| + \|\mathbf{f}(\mu) - \mathbf{f}(a)\| \leq \\ &\leq g(t) - g(a) + \epsilon(t - a) + \epsilon \end{aligned}$$

es decir,  $t \in S_\epsilon$ , lo que es absurdo porque  $t > \mu = \sup S_\epsilon$ . ■

NOTA: Con las modificaciones obvias, que se dejan al cuidado del lector, se puede obtener otra versión del teorema anterior reemplazando las derivadas por la derecha por las derivadas por la izquierda.

**Corolario 4.8** Sea  $(F, \|\cdot\|)$  un espacio normado y  $\mathbf{f} : [a, b] \rightarrow F$ , una función continua en  $[a, b]$  y derivable por la derecha en  $(a, b)$ , tal que  $\|\mathbf{f}'_d(t)\| \leq M$  para todo  $t \in (a, b)$ . Entonces  $\|\mathbf{f}(b) - \mathbf{f}(a)\| \leq M(b - a)$ .

DEM: Basta aplicar el teorema 4.7 con  $g(t) = Mt$ . ■

**Corolario 4.9** Sea  $(F, \|\cdot\|)$  un espacio normado y  $\mathbf{f} : [a, b] \rightarrow F$ , una función continua en  $[a, b]$  y derivable en  $(a, b)$ . Si  $\mathbf{f}'(t) = 0$  para todo  $t \in (a, b)$  entonces  $\mathbf{f}$  es constante.

DEM: En cada  $[a, x] \subset [a, b]$  se aplica el corolario 4.8 con  $M = 0$  y se obtiene  $\mathbf{f}(a) = \mathbf{f}(x)$ . ■

**Corolario 4.10** Si  $g : [a, b] \rightarrow \mathbb{R}$  es continua y derivable por la derecha en cada  $t \in (a, b)$ , con  $g'_d(t) \geq 0$ , entonces  $g$  es creciente.

DEM: Basta aplicar el teorema 4.7, con  $\mathbf{f} = 0$ , en cada  $[x, y] \subset [a, b]$ . ■

## 4.2. Desarrollo de Taylor

Sea  $(F, \|\cdot\|)$  un espacio normado y  $\mathbf{f} : \Omega \rightarrow F$  una función definida en un abierto  $\Omega \subset \mathbb{R}$ , derivable  $m$  veces en  $a \in \Omega$ . Llamaremos "polinomio" de Taylor de  $\mathbf{f}$  en  $a$ , de grado  $m$ , a la función vectorial de variable real

$$\mathbf{P}_m(x-a) = \mathbf{f}(a) + \mathbf{f}'(a)(x-a) + \frac{1}{2!}\mathbf{f}''(a)(x-a)^2 + \frac{1}{3!}\mathbf{f}'''(a)(x-a)^3 + \dots + \frac{1}{m!}\mathbf{f}^{(m)}(a)(x-a)^m$$

(donde se ha utilizado el convenio de escribir  $\frac{1}{k!}\mathbf{f}^{(k)}(a)(x-a)^k$  en lugar de  $\frac{(x-a)^k}{k!}\mathbf{f}^{(k)}(a)$ ). Obsérvese que se trata de un polinomio en un sentido generalizado, pues los coeficientes  $\frac{1}{k!}\mathbf{f}^{(k)}(a)$  son ahora vectores del espacio normado  $F$ .

**Teorema 4.11** *Sea  $(F, \|\cdot\|)$  un espacio normado y  $\mathbf{f} : \Omega \rightarrow F$  una función definida en un abierto  $\Omega \subset \mathbb{R}$ . Si  $\mathbf{f}$  es derivable  $m$  veces en  $a \in \Omega$  y  $\mathbf{P}_m(x-a)$  es su polinomio de Taylor en  $a$  de grado  $m$ , entonces  $\mathbf{f}(x) = \mathbf{P}_m(x-a) + \mathbf{R}_m(x-a)$  donde  $\mathbf{R}_m(x-a) = o(|x-a|^m)$ , es decir*

$$\lim_{x \rightarrow a} \frac{\mathbf{R}_m(x-a)}{|x-a|^m} = 0$$

DEM: En lo que sigue escribimos  $h = x - a$ . La demostración se hace por inducción sobre  $m$ . El resultado es inmediato cuando  $m = 1$  pues  $\mathbf{P}_1(h) = \mathbf{f}(a) + h\mathbf{f}'(a)$  y en virtud de la definición de derivada,  $\mathbf{R}_1(h) = \mathbf{f}(a+h) - \mathbf{f}(a) - h\mathbf{f}'(a)$  cumple la condición requerida.

Se comprueba fácilmente que el polinomio de Taylor de grado  $m-1$  de la función  $\mathbf{f}'$  en el punto  $a$  es  $\mathbf{P}'_m(x-a)$ . Suponiendo el teorema cierto para funciones derivables  $m-1$  veces en  $a$ , (con  $m \geq 2$ ) y aplicándolo a  $\mathbf{f}'$ , definida en un cierto entorno de  $a$ , resulta  $\mathbf{f}'(a+h) = \mathbf{P}'_m(h) + \mathbf{r}_{m-1}(h)$ , donde  $\mathbf{r}_{m-1}(h) = o(|h|^{m-1})$ .

Derivando respecto a la variable  $h$  en la igualdad  $\mathbf{f}(a+h) = \mathbf{P}_m(h) + \mathbf{R}_m(h)$  se obtiene que  $\mathbf{R}'_m(h) = \mathbf{r}_{m-1}(h)$ , luego

$$\lim_{h \rightarrow 0} \frac{\mathbf{R}'_m(h)}{|h|^{m-1}} = 0$$

y para terminar basta ver que esta condición implica que  $\mathbf{R}_m(h) = o(|h|^m)$ .

Efectivamente, como  $\mathbf{R}_m(h) = \mathbf{f}(a+h) - \mathbf{P}_m(h)$  es derivable dos veces en 0 (porque  $m \geq 2$ ) podemos asegurar que existe  $\eta > 0$  tal que  $\mathbf{R}_m(h)$  es derivable en  $(-\eta, \eta)$ . Ahora, por la definición de límite, dado  $\epsilon > 0$  podemos encontrar  $0 < \delta < \eta$  tal que para todo  $t \in (0, \delta)$  se cumple

$$\left\| \frac{\mathbf{R}'_m(t)}{t^{m-1}} \right\| < \epsilon$$

Si  $0 < s < \delta$ , para todo  $t \in (a, s)$  es cierta la desigualdad  $\|\mathbf{R}'_m(t)\| \leq \epsilon t^{m-1} = g'(t)$ , con  $g(t) = (\epsilon/m)t^m$ . Aplicando el teorema 4.7 a las funciones  $\mathbf{R}_m, g$  en el intervalo  $[0, s]$  resulta

$$\|\mathbf{R}_m(s) - \mathbf{R}_m(0)\| \leq \frac{\epsilon}{m}s^m \leq \epsilon s^m$$

Como  $\mathbf{R}_m(0) = \mathbf{0}$ , se puede asegurar que

$$0 < s < \delta \Rightarrow \left\| \frac{\mathbf{R}_m(s)}{s^m} \right\| < \epsilon.$$

Análogamente, podemos obtener  $\delta' < \eta$  tal que

$$-\delta' < s < 0 \Rightarrow \left\| \frac{\mathbf{R}_m(s)}{s^m} \right\| < \epsilon$$

Queda demostrado que  $\lim_{s \rightarrow 0} \frac{\mathbf{R}_m(s)}{|s|^m} = 0$ , luego  $\mathbf{R}_m(h) = o(|h|^m)$ . ■

**Proposición 4.12** *Sea  $(F, \|\cdot\|)$  un espacio normado,  $\mathbf{f} : \Omega \rightarrow F$  una función de clase  $C^m$  y  $\mathbf{P}_m(x-a)$  su polinomio de Taylor en  $a$  de grado  $m$ . Sea  $[a, x] \subset \Omega$  un intervalo tal que en cada  $t \in (a, x)$  existe la derivada  $\mathbf{f}^{(m+1)}(t)$  y verifica  $\|\mathbf{f}^{(m+1)}(t)\| \leq M$ . Entonces para el término complementario o resto  $\mathbf{R}_m(x-a) = \mathbf{f}(x) - \mathbf{P}_m(x-a)$  vale la acotación*

$$\|\mathbf{R}_m(x-a)\| \leq \frac{M}{(m+1)!} (x-a)^{m+1}$$

DEM: Ofrecemos una demostración por inducción sobre  $m$ . El resultado es cierto para  $m = 0$  en virtud de 4.8. Suponemos que el teorema es cierto hasta el orden  $m$  y demostraremos que también lo es hasta el orden  $m+1$ . Para ello se considera la función auxiliar

$$\mathbf{v}(t) = \mathbf{f}(a+th) - \left[ \mathbf{f}(a) + th\mathbf{f}'(a) + \cdots + t^m h^m \frac{1}{m!} \mathbf{f}^{(m)}(a) \right]$$

donde  $h = x - a > 0$ , cuya derivada es

$$\mathbf{v}'(t) = h\mathbf{f}'(a+th) - h \left[ \mathbf{f}'(a) + th\mathbf{f}''(a) + \cdots + t^{m-1} h^{m-1} \frac{1}{(m-1)!} \mathbf{f}^{(m)}(a) \right]$$

Si aplicamos la hipótesis de inducción a la función  $\mathbf{f}'$  en el intervalo  $[a, a+th]$  se obtiene

$$\mathbf{f}'(a+th) - \left[ \mathbf{f}'(a) + \sum_{k=1}^{m-1} \frac{1}{k!} \mathbf{f}^{(k+1)}(a) t^k h^k \right] = \mathbf{r}_{m-1}(th)$$

donde  $\|\mathbf{r}_{m-1}(th)\| \leq \frac{M}{m!} t^m h^m$ . Como  $\mathbf{v}'(t) = h\mathbf{r}_{m-1}(th)$  resulta

$$\|\mathbf{v}'(t)\| \leq h \|\mathbf{r}_{m-1}(th)\| \leq \frac{M}{m!} t^m h^{m+1}$$

y aplicando 4.7 con  $g(t) = \frac{M}{(m+1)!} t^{m+1} h^{m+1}$  se obtiene

$$\|\mathbf{v}(1) - \mathbf{v}(0)\| \leq g(1) - g(0) = \frac{M}{(m+1)!} h^{m+1}$$

es decir

$$\|\mathbf{R}_m(x-a)\| \leq \frac{M}{(m+1)!} (x-a)^{m+1}$$
■

### 4.3. Integral de una función vectorial

La integral de Riemann de una función acotada  $\mathbf{f} : [a, b] \rightarrow \mathbb{R}^n$  se puede definir en términos de sus componentes. Se dice que  $\mathbf{f} = (f_1, f_2, \dots, f_n)$  es integrable Riemann cuando todas sus componentes lo son y en ese caso se define

$$\int_a^b \mathbf{f}(t) dt = \left( \int_a^b f_1(t) dt, \int_a^b f_2(t) dt, \dots, \int_a^b f_n(t) dt \right)$$

es decir, la integral de  $\mathbf{f}$  es el vector cuyas componentes son las integrales de sus componentes. La teoría de la integral de Riemann para funciones de una variable real con valores en  $\mathbb{R}^n$  se desarrolla directamente, a partir de la definición, tomando como base los resultados conocidos para el caso  $n = 1$ . Es inmediata la integrabilidad de las funciones continuas y también lo es la linealidad de la integral y la aditividad de la integral respecto al intervalo:

- i) Si las funciones  $\mathbf{f}, \mathbf{g} : [a, b] \rightarrow \mathbb{R}^n$  son integrables y  $\alpha, \beta \in \mathbb{R}$ , entonces  $\alpha\mathbf{f} + \beta\mathbf{g}$  es integrable y  $\int_a^b (\alpha\mathbf{f}(t) + \beta\mathbf{g}(t)) dt = \alpha \int_a^b \mathbf{f}(t) dt + \beta \int_a^b \mathbf{g}(t) dt$ .
- ii) Si  $a \leq c \leq b$ , la función  $\mathbf{f} : [a, b] \rightarrow \mathbb{R}^n$ , es integrable si y sólo si  $\mathbf{f}|_{[a,c]}$  y  $\mathbf{f}|_{[c,b]}$  lo son, y en ese caso  $\int_a^b \mathbf{f}(t) dt = \int_a^c \mathbf{f}(t) dt + \int_c^b \mathbf{f}(t) dt$ .

Sea  $\mathcal{P}([a, b])$  el conjunto de las subdivisiones de  $[a, b]$ . Para una subdivisión  $p \in \mathcal{P}([a, b])$ ,  $p = (a = t_0 < t_1 < \dots < t_m = b)$ , diremos que  $\eta = (\eta_1 \leq \eta_2 \leq \dots, \leq \eta_m)$  es un sistema de puntos asociado a  $p$  si  $\eta_j \in [t_{j-1}, t_j]$  para cada  $j \in \{1, \dots, m\}$ . Una *suma de Riemann* asociada a  $p$  es una suma de la forma

$$\Sigma(\mathbf{f}, p, \eta) = \sum_{k=1}^m (t_k - t_{k-1}) \mathbf{f}(\eta_k)$$

donde  $\eta = (\eta_1 \leq \eta_2 \leq \dots, \leq \eta_m)$  es un sistema de puntos asociado a  $p$ . Obsérvese que las componentes del vector  $\Sigma(\mathbf{f}, p, \eta)$  son las sumas de Riemann de las funciones componentes,  $\Sigma(\mathbf{f}, p, \eta) = (\Sigma(f_1, p, \eta), \Sigma(f_2, p, \eta), \dots, \Sigma(f_n, p, \eta))$ . En lo que sigue la suma de Riemann formada con los puntos  $\eta_j = t_{j-1}$ ,  $1 \leq j \leq m$ , la denotaremos más brevemente  $\Sigma(\mathbf{f}, p)$ .

**Lema 4.13** Si  $\mathbf{f} : [a, b] \rightarrow \mathbb{R}^n$  es integrable Riemann, para cada  $\epsilon > 0$  existe  $p_\epsilon \in \mathcal{P}([a, b])$  tal que si  $p \in \mathcal{P}([a, b])$  es más fina que  $p_\epsilon$  se cumple

$$\left\| \Sigma(\mathbf{f}, p) - \int_a^b \mathbf{f}(t) dt \right\|_\infty < \epsilon$$

DEM: El resultado, que es bien conocido para funciones con valores reales, se extiende fácilmente al caso de una función  $\mathbf{f} = (f_1, f_2, \dots, f_n)$  con valores en  $\mathbb{R}^n$ : Para cada  $j \in \{1, 2, \dots, n\}$  existe  $p_\epsilon^j \in \mathcal{P}([a, b])$  tal que si  $p \in \mathcal{P}([a, b])$  es más fina que  $p_\epsilon^j$  se cumple

$$[*] \quad \left| \Sigma(f_j, p) - \int_a^b f_j(t) dt \right| < \epsilon$$

Sea  $p_\epsilon \in \mathcal{P}([a, b])$  más fina que todas las  $p_\epsilon^j$ ,  $1 \leq j \leq n$ . Entonces, si  $p \in \mathcal{P}([a, b])$  es más fina que  $p_\epsilon$  podemos asegurar que para todo  $j \in \{1, 2, \dots, n\}$  se cumple  $[*]$ , y teniendo en cuenta que  $\Sigma(f_j, p)$  son las componentes del vector  $\Sigma(\mathbf{f}, p)$ , resulta

$$\left\| \int_a^b \mathbf{f}(t) dt - \Sigma(\mathbf{f}, p) \right\|_\infty < \epsilon$$

■

Puesto que en  $\mathbb{R}^n$  todas las normas son equivalentes, el resultado del lema anterior se sigue verificando para cualquier norma en  $\mathbb{R}^n$ .

NOTA: Dada una subdivisión  $p \in \mathcal{P}([a, b])$ ,  $p = (a = t_0 < t_1 < \dots < t_m = b)$  se define  $\Delta(p) = \max\{t_j - t_{j-1} : 1 \leq j \leq m\}$ . Es bien conocido que si la función  $f : [a, b] \rightarrow \mathbb{R}$  es integrable Riemann entonces  $\lim_{\Delta(p) \rightarrow 0} \Sigma(f, p) = \int_a^b f(t) dt$  (véase [12]) III, 1.4.2). Lo mismo se sigue cumpliendo en el caso de funciones con valores en  $\mathbb{R}^n$ , pero este resultado, cuya demostración es sencilla para el caso de funciones continuas, no será utilizado aquí.

En el caso de funciones vectoriales, la propiedad de monotonía de la integral no tiene sentido, pero se sigue verificando:

**Proposición 4.14** *Si  $\mathbf{f} : [a, b] \rightarrow \mathbb{R}^n$  es continua y  $\|\cdot\|$  es una norma sobre  $\mathbb{R}^n$  entonces la función  $\|\mathbf{f}\|$  también es continua y se verifica*

- a)  $\left\| \int_a^b \mathbf{f}(t) dt \right\| \leq \int_a^b \|\mathbf{f}(t)\| dt$ ;
- b)  $\int_a^b \mathbf{f}(t) dt = \int_a^c \mathbf{f}(t) dt + \int_c^b \mathbf{f}(t) dt$  si  $a \leq c \leq b$ .
- c) La función  $\mathbf{g} : [a, b] \rightarrow \mathbb{R}^n$  definida por  $\mathbf{g}(x) = \int_a^x \mathbf{f}(t) dt$  es derivable en  $[a, b]$  y  $\mathbf{g}'(x) = \mathbf{f}(x)$  para todo  $x \in [a, b]$  (en  $a$  y  $b$  se consideran las derivadas laterales correspondientes).

DEM: a) Vemos en primer lugar una demostración elemental para el caso particular de la norma euclídea  $\|\cdot\|_2$ . La desigualdad del enunciado basta establecerla cuando el vector  $\mathbf{v} = \int_a^b \mathbf{f}(t) dt$  /no es nulo. En este caso, considerando el vector unitario  $\mathbf{u} = \mathbf{v} / \|\mathbf{v}\|_2$  podemos escribir

$$\begin{aligned} \|\mathbf{v}\|_2 &= \langle \mathbf{u} | \mathbf{v} \rangle = \sum_{i=1}^n u_i \int_a^b f_i(t) dt = \int_a^b \left( \sum_{i=1}^n u_i f_i(t) \right) dt \leq \\ &\leq \int_a^b \left| \sum_{i=1}^n u_i f_i(t) \right| dt \leq \int_a^b \|\mathbf{u}\|_2 \|\mathbf{f}(t)\|_2 dt = \int_a^b \|\mathbf{f}(t)\|_2 dt \end{aligned}$$

donde la última desigualdad es consecuencia de la desigualdad de Cauchy-Schwarz.

En el caso de una norma arbitraria en  $\mathbb{R}^n$ , con el lema 4.13 podemos conseguir una sucesión  $p_k \in \mathcal{P}([a, b])$ , donde cada  $p_k$  es más fina que  $p_{k-1}$ , tal que

$$\int_a^b \mathbf{f}(t) dt = \lim_k \Sigma(\mathbf{f}, p_k) \quad \text{y} \quad \int_a^b \|\mathbf{f}(t)\| dt = \lim_k \Sigma(\|\mathbf{f}\|, p_k)$$

En virtud de la desigualdad triangular  $\|\Sigma(\mathbf{f}, p_k)\| \leq \Sigma(\|\mathbf{f}\|, p_k)$  y usando la continuidad de la norma se obtiene

$$\left\| \int_a^b \mathbf{f}(t) dt \right\| = \lim_k \|\Sigma(\mathbf{f}, p_k)\| \leq \lim_k \Sigma(\|\mathbf{f}\|, p_k) = \int_a^b \|\mathbf{f}(t)\| dt.$$

La demostración de b) y c) es inmediata razonando componente a componente. ■

NOTA: En el capítulo 10 se verá el teorema de Lebesgue que caracteriza las funciones integrables Riemann. Usando este teorema es fácil demostrar que si  $\mathbf{f} : [a, b] \rightarrow \mathbb{R}^n$  es integrable Riemann entonces la función escalar  $t \rightarrow \|\mathbf{f}(t)\|$  también es integrable y se cumple a) (véase D.14). Para la norma euclídea se puede dar una demostración elemental de este hecho basada en la siguiente propiedad: Si  $\varphi : [a, b] \rightarrow [0, +\infty)$  es integrable Riemann, también lo es su raíz cuadrada  $\sqrt{\varphi}$  (que se deja como ejercicio).

#### Corolario 4.15

-[Regla de Barrow] Si  $\mathbf{g} : [a, b] \rightarrow \mathbb{R}^n$  es derivable con derivada continua se verifica  $\mathbf{g}(b) - \mathbf{g}(a) = \int_a^b \mathbf{g}'(t) dt$ .

-[Integración por partes] Si las funciones  $\mathbf{f} : [a, b] \rightarrow \mathbb{R}^n$ ,  $\varphi : [a, b] \rightarrow \mathbb{R}$  son derivables con derivada continua, se verifica

$$\varphi(b)\mathbf{f}(b) - \varphi(a)\mathbf{f}(a) = \int_a^b \varphi'(t)\mathbf{f}(t) dt + \int_a^b \varphi(t)\mathbf{f}'(t) dt$$

**Teorema 4.16** Sea  $f : \Omega \rightarrow \mathbb{R}^n$  de clase  $C^{m+1}$ , donde  $\Omega \subset \mathbb{R}$  es abierto y  $\mathbf{P}_m(x-a)$  su polinomio de Taylor de grado  $m$  en  $a \in \Omega$ . Si  $[a, x] \subset \Omega$  entonces vale la siguiente fórmula integral para el resto  $\mathbf{R}_m(x-a) = \mathbf{f}(x) - \mathbf{P}_m(x-a)$

$$\mathbf{R}_m(x-a) = \frac{(x-a)^{m+1}}{m!} \int_0^1 (1-t)^m \mathbf{f}^{(m+1)}(a+t(x-a)) dt$$

DEM: Si  $h = (x-a)$ , la función  $\mathbf{v}(t) = \mathbf{f}(a+th)$ , está definida en el abierto  $U = \{t : a+th \in \Omega\} \supset [0, 1]$ , donde admite derivadas sucesivas continuas, hasta la de orden  $m+1$ , que vienen dadas por:

$$\mathbf{v}'(t) = h\mathbf{f}'(a+th), \quad \mathbf{v}''(t) = h^2\mathbf{f}''(a+th), \dots, \mathbf{v}^{(m+1)}(t) = h^{m+1}\mathbf{f}^{(m+1)}(a+th)$$

La función  $\mathbf{g}(t) = \mathbf{v}(t) + (1-t)\mathbf{v}'(t) + \frac{1}{2!}(1-t)^2\mathbf{v}''(t) + \dots + \frac{1}{m!}(1-t)^m\mathbf{v}^{(m)}(t)$  es derivable con derivada continua en  $[0, 1]$ :

$$\mathbf{g}'(t) = \mathbf{v}'(t) + [(1-t)\mathbf{v}''(t) - \mathbf{v}'(t)] + \dots + \left[ \frac{(1-t)^m}{m!} \mathbf{v}^{(m+1)}(t) - \frac{(1-t)^{m-1}}{(m-1)!} \mathbf{v}^{(m)}(t) \right]$$

y cancelando términos en esta suma telescópica resulta

$$\mathbf{g}'(t) = \frac{(1-t)^m}{m!} \mathbf{v}^{(m+1)}(t) = \frac{(1-t)^m}{m!} h^{m+1} \mathbf{f}^{(m+1)}(a+th)$$



y aplicando D.12 se obtiene

$$\mathbf{g}(1) - \mathbf{g}(0) = \int_0^1 \mathbf{g}'(t) dt = \frac{h^{m+1}}{m!} \int_0^1 (1-t)^m \mathbf{f}^{(m+1)}(a+th) dt$$

El resultado se obtiene observando que

$$\begin{aligned} \mathbf{g}(1) - \mathbf{g}(0) &= \mathbf{v}(1) - \left[ \mathbf{v}(0) + \mathbf{v}'(0) + \cdots + \frac{1}{m!} \mathbf{v}^{(m)}(0) \right] = \\ &= \mathbf{f}(a+h) - \left[ \mathbf{f}(a) + h\mathbf{f}'(a) + \cdots + \frac{h^m}{m!} \mathbf{f}^{(m)}(a) \right] = \mathbf{R}_m(h) \end{aligned}$$

■

- En el apéndice D se exponen dos alternativas para definir la integral en el caso de funciones con valores en un espacio normado completo. El lector interesado podrá apreciar que el teorema anterior sigue valiendo en este contexto más general.

- Para funciones con valores en  $\mathbb{R}^n$  (o más generalmente en un espacio normado completo) el resultado 4.12, con la hipótesis algo más restrictiva de que la función sea de clase  $C^{m+1}$ , se puede obtener como consecuencia inmediata de 4.16 usando la propiedad 4.14 a). ■

## 4.4. Caminos rectificables

Comenzamos con la terminología asociada a una aplicación  $\gamma : [a, b] \rightarrow E$ , con valores en un espacio normado  $(E, \|\cdot\|)$ . Si  $\gamma : [a, b] \rightarrow E$  es continua y  $\gamma([a, b]) \subset \Omega$  se dice que  $\gamma$  es un *camino* o *trayectoria* en  $\Omega \subset E$ . El origen (resp. extremo) del camino es el punto  $\gamma(a)$  (resp.  $\gamma(b)$ ) y si  $\gamma(a) = \gamma(b)$  se dice que el camino es cerrado. La *multiplicidad* de un punto  $\mathbf{x} \in \gamma([a, b])$  es el número  $\text{Card}\{t \in [a, b] : \gamma(t) = \mathbf{x}\}$ . Los puntos de multiplicidad 1 se llaman *simples*. Un camino es *simple* cuando todos los puntos de su imagen son simples, es decir, cuando es inyectivo. Un camino cerrado  $\gamma : [a, b] \rightarrow E$  se dice que es simple cuando cada  $\mathbf{x} \in \gamma((a, b))$  es simple y  $\gamma(a) = \gamma(b)$  tiene multiplicidad 2.

Cuando  $E = \mathbb{R}^3$  podemos interpretar que el parámetro  $t$  es el tiempo y que  $\gamma(t) \in \mathbb{R}^3$  es la posición, en el instante  $t$ , de un punto que se mueve en el espacio recorriendo una trayectoria que puede pasar por un mismo punto varias veces.

Dos caminos  $\gamma_i : [a_i, b_i] \rightarrow E$  se dice que son *topológicamente equivalentes* cuando existe una biyección continua  $h : [a_1, b_1] \rightarrow [a_2, b_2]$  tal que  $\gamma_1 = \gamma_2 \circ h$ . Esta biyección necesariamente es estrictamente monótona. Cuando es creciente se dice que los caminos son *topológicamente equivalentes con la misma orientación* y cuando es decreciente se dice que los caminos son *topológicamente equivalentes con orientaciones opuestas*. En el primer caso los dos caminos tienen el mismo origen y el mismo extremo. En el segundo caso el origen de un camino coincide con el extremo del otro y los dos caminos recorren el mismo trayecto, pero en sentidos opuestos.

En el conjunto de los caminos en el espacio normado  $E$  queda definida así una relación de equivalencia. Cada clase de equivalencia se dice que es una curva o un

*arco de curva.* Cada representante de la clase se dice que es una *representación paramétrica* de la curva. La clase de equivalencia  $[\gamma]$  del camino  $\gamma$  está formada por los caminos que se deducen de  $\gamma$  efectuando un cambio de parámetro continuo estrictamente monótono. Si  $\gamma : [a, b] \rightarrow E$  es derivable y la derivada  $t \rightarrow \gamma'(t)$  es continua en  $[a, b]$  se dice que  $\gamma$  es un camino de *clase*  $C^1$ . Entre caminos de clase  $C^1$ , procediendo en forma similar, pero considerando cambios de parámetro de clase  $C^1$  con derivada no nula  $h : [a_1, b_1] \rightarrow [a_2, b_2]$ , se pueden definir caminos  $C^1$ -*equivalentes*, relación de equivalencia que da lugar a la noción de arco de curva de clase  $C^1$ .

**Proposición 4.17** *Si dos caminos simples  $\gamma_i : [a_i, b_i] \rightarrow E$ ,  $i = 1, 2$ , tienen la misma imagen son topológicamente equivalentes y definen el mismo arco de curva.*

DEM:  $K = \gamma_1([a_1, b_1]) = \gamma_2([a_2, b_2])$  es compacto y en virtud de 3.15 las biyecciones continuas  $\gamma_i : [a_i, b_i] \rightarrow K$  ( $i=1,2$ ) tienen inversa continua. La biyección continua  $h = \gamma_2^{-1} \circ \gamma_1 : [a_1, b_1] \rightarrow [a_2, b_2]$  cumple  $\gamma_1 = \gamma_2 \circ h$ . ■

Una partición o subdivisión del intervalo  $[a, b]$  es una familia finita de puntos del intervalo que contiene a los extremos,  $p = (a = t_0 < t_1 < t_2 \cdots t_m = b)$ . En lo que sigue denotaremos por  $\mathcal{P}(I)$  al conjunto de las subdivisiones del intervalo  $I = [a, b]$ . Si  $p', p \in \mathcal{P}(I)$ , se dice que  $p'$  es más fina que  $p$  cuando  $p' \supset p$ .

Dada una aplicación  $\mathbf{f} : [a, b] \rightarrow (E, \|\cdot\|)$ , para cada  $p \in \mathcal{P}(I)$ , la variación de  $\mathbf{f}$  relativa a  $p$  es la suma

$$V(\mathbf{f}, p) = \sum_{j=1}^m \|\mathbf{f}(t_j) - \mathbf{f}(t_{j-1})\|$$

Usando la desigualdad triangular se comprueba fácilmente que si  $p' \in \mathcal{P}(I)$  es más fina que  $p \in \mathcal{P}(I)$  entonces  $V(\mathbf{f}, p) \leq V(\mathbf{f}, p')$  (basta comprobarlo cuando  $p'$  se obtiene añadiendo un punto a  $p$ ). La *variación total* de  $\mathbf{f}$  sobre  $[a, b]$  es la cantidad

$$V(\mathbf{f}, [a, b]) = \sup\{V(\mathbf{f}, p) : p \in \mathcal{P}(I)\} \leq +\infty$$

Cuando  $V(\mathbf{f}, [a, b]) < +\infty$  se dice que  $\mathbf{f} : [a, b] \rightarrow E$  es de *variación acotada* o de *variación total finita*. El conjunto de las funciones  $\mathbf{f} : [a, b] \rightarrow E$  que son de variación acotada se suele denotar  $BV([a, b], E)$ . Si  $\mathbf{f} \in BV([a, b], E)$ , para cada par de puntos  $a \leq x < y \leq b$  se cumple  $\|\mathbf{f}(x) - \mathbf{f}(y)\| \leq V(\mathbf{f}, [a, b])$ . De esta desigualdad se deduce que  $\mathbf{f}$  es acotada ya que

$$\|\mathbf{f}(x)\| \leq \|\mathbf{f}(a)\| + V(\mathbf{f}, [a, b]) \quad \text{para todo } x \in [a, b]$$

También se deduce que  $\mathbf{f}$  es constante si y sólo si  $V(\mathbf{f}, [a, b]) = 0$ .

Si  $\mathbf{f} : [a, b] \rightarrow (E, \|\cdot\|)$  es de variación acotada es fácil ver que también lo es respecto a cualquier norma  $\|\cdot\|'$  equivalente a la norma  $\|\cdot\|$  dada en  $E$ . La noción de aplicación de variación acotada también se puede definir de modo análogo para funciones con valores en un espacio métrico  $(E, d)$ , utilizando las sumas  $V(\mathbf{f}, p) = \sum_{j=1}^m d(\mathbf{f}(t_j), \mathbf{f}(t_{j-1}))$ , pero puede ocurrir que  $\mathbf{f} : [a, b] \rightarrow (E, d)$  sea de variación

acotada para la distancia  $d$  pero no lo sea para otra distancia equivalente (véase el ejercicio 4.31).

NOTA: (Para el lector familiarizado con la teoría de redes):  $\mathcal{P}(I)$  es un conjunto dirigido por refinamiento y  $(V(\mathbf{f}, p))_{p \in \mathcal{P}(I)}$  es una red creciente de números reales que converge, en la recta real ampliada, hacia

$$\lim_{p \in \mathcal{P}(I)} V(\mathbf{f}, p) = V(\mathbf{f}, [a, b])$$

**Definición 4.18** *Un camino  $\gamma : [a, b] \rightarrow E$ , en un espacio normado  $(E, \|\cdot\|)$  se dice que es rectificable cuando es de variación acotada. En ese caso a la variación total se le llama longitud del camino:  $\text{Long}(\gamma) = V(\gamma, [a, b])$ .*

Sabemos que si un camino  $\gamma$  en un espacio normado  $(E, \|\cdot\|)$ , es rectificable también lo es para cualquier norma equivalente  $\|\cdot\|'$  que se considere en  $E$ . Sin embargo el valor numérico de su longitud depende esencialmente de la norma que se esté usando. Obsérvese que un segmento  $\sigma(t) = t\mathbf{y} + (1-t)\mathbf{x}$ ,  $0 \leq t \leq 1$ , de origen  $\mathbf{x}$  y extremo  $\mathbf{y}$ , es un camino rectificable de longitud  $\text{Long}(\sigma) = \|\mathbf{y} - \mathbf{x}\|$ .

No obstante, cuando se hable de la longitud de un camino en el espacio  $E = \mathbb{R}^n$  se entenderá, salvo mención expresa de otra cosa, que se trata de su longitud euclídea, es decir, la longitud calculada usando la norma euclídea de  $\mathbb{R}^n$ .

**Proposición 4.19** *Si  $\gamma_i : [a_i, b_i] \rightarrow E$ ,  $i = 1, 2$ , son caminos topológicamente equivalentes, entonces  $\gamma_1$  es rectificable si y sólo si lo es  $\gamma_2$  y en este caso*

$$V(\gamma_1, [a_1, b_1]) = V(\gamma_2, [a_2, b_2]) \quad \text{es decir} \quad \text{Long}(\gamma_1) = \text{Long}(\gamma_2)$$

DEM: Por hipótesis existe una biyección continua  $h : [a_1, b_1] \rightarrow [a_2, b_2]$  tal que  $\gamma_1 = \gamma_2 \circ h$ . Como  $h$  es estrictamente monótona, queda establecida una biyección natural  $p \rightarrow \bar{p}$ , entre las subdivisiones de  $[a_1, b_1]$  y las de  $[a_2, b_2]$  (si  $h$  es decreciente viene dada por  $p = (t_0 < t_1 < \dots < t_m) \rightarrow \bar{p} = (h(t_m) < h(t_{m-1}) < \dots < h(t_0))$ ).

La conclusión se obtiene observando que  $V(\gamma_1, p) = V(\gamma_2, \bar{p})$ . ■

NOTA: El resultado de la proposición anterior también se cumple cuando  $\gamma_1 = \gamma_2 \circ h$  donde  $h : [a_1, b_1] \rightarrow [a_2, b_2]$  es monótona continua y sobreyectiva, pero no se supone inyectiva. Basta tener en cuenta que en este caso la aplicación  $p \rightarrow \bar{p}$  aunque no es biyectiva, es sobreyectiva y se sigue verificando  $V(\gamma_1, p) = V(\gamma_2, \bar{p})$ .

En virtud de la proposición 4.19 se puede decir que un arco de curva es rectificable cuando una (y por consiguiente todas) sus representaciones paramétricas lo son. En ese caso, todas las representaciones paramétricas tienen la misma longitud que, por definición, es la longitud del arco de curva.

### Proposición 4.20

a) *Toda función monótona  $f : [a, b] \rightarrow \mathbb{R}$  es de variación acotada y*

$$V(f, [a, b]) = |f(b) - f(a)|$$

b) Si  $\mathbf{f} : [a, b] \rightarrow E$  cumple la condición de Lipschitz,

$$\|\mathbf{f}(x) - \mathbf{f}(y)\| \leq M|x - y| \text{ para cada } x, y \in [a, b]$$

entonces  $\mathbf{f}$  es un camino rectificable y  $V(\mathbf{f}, [a, b]) \leq M|b - a|$ .

c) Todo camino derivable  $\mathbf{f} : [a, b] \rightarrow E$  con derivada acotada (en particular, todo camino de clase  $C^1$ ) es rectificable.

DEM: Las demostraciones de a) y b) son rutinarias y se dejan al cuidado del lector. c) es consecuencia de b): Si  $\mathbf{f}$  es derivable con derivada acotada, en virtud de 4.8 se cumple la condición de Lipschitz con  $M = \sup\{\|\mathbf{f}'(x)\| : x \in [a, b]\}$ . ■

En la siguiente proposición, si  $\mathbf{f}$  no es de variación acotada, la igualdad se cumple con el convenio:  $+\infty + (+\infty) = +\infty$ ,  $+\infty + u = u + (+\infty) = +\infty$  si  $u \in \mathbb{R}$ .

**Proposición 4.21** Si  $a \leq x \leq b$ , entonces  $V(\mathbf{f}, [a, b]) = V(\mathbf{f}, [a, x]) + V(\mathbf{f}, [x, b])$ . Por lo tanto, si  $\mathbf{f}$  es de variación acotada (en particular, un camino rectificable) también lo es su restricción a cualquier intervalo  $[x, y] \subset [a, b]$ .

DEM: Si  $x = a$  ó  $x = b$  el resultado es evidente. Supongamos  $a < x < b$ . Para cada  $p \in \mathcal{P}([a, b])$  sea  $p_x$  la subdivisión obtenida añadiendo a  $p$  el punto  $x$ , y  $p'$ ,  $p''$  las subdivisiones que  $p_x$  induce en  $[a, x]$  y en  $[x, b]$  respectivamente.

$$V(\mathbf{f}, p) \leq V(\mathbf{f}, p_x) = V(\mathbf{f}, p') + V(\mathbf{f}, p'') \leq V(\mathbf{f}, [a, x]) + V(\mathbf{f}, [x, b])$$

y considerando el supremo de las cantidades  $V(\mathbf{f}, p)$  resulta

$$V(\mathbf{f}, [a, b]) \leq V(\mathbf{f}, [a, x]) + V(\mathbf{f}, [x, b])$$

Recíprocamente, para cada  $p' \in \mathcal{P}([a, x])$  y cada  $p'' \in \mathcal{P}([x, b])$  sea  $p \in \mathcal{P}([a, b])$  la subdivisión formada con los puntos de  $p'$  y  $p''$ .

$$V(\mathbf{f}, p') + V(\mathbf{f}, p'') = V(\mathbf{f}, p) \leq V(\mathbf{f}, [a, b])$$

Considerando primero el supremo de las cantidades  $V(\mathbf{f}, p')$  y luego el supremo de las cantidades  $V(\mathbf{f}, p'')$  resulta

$$V(\mathbf{f}, [a, x]) + V(\mathbf{f}, [x, b]) \leq V(\mathbf{f}, [a, b])$$

■

**Definición 4.22** Si  $\mathbf{f} : [a, b] \rightarrow E$  es de variación acotada se llama *variación indefinida de  $\mathbf{f}$*  a la función  $v : [a, b] \rightarrow [0, +\infty)$  definida por

$$v(x) = V(\mathbf{f}, [a, x]) \text{ si } a < x \leq b, \quad v(a) = 0.$$

En virtud de la proposición 4.21, la variación indefinida  $v$  es una función creciente que cumple  $\|\mathbf{f}(x) - \mathbf{f}(y)\| \leq V(\mathbf{f}, [x, y]) = v(y) - v(x)$  para todo  $[x, y] \subset [a, b]$ .

**Teorema 4.23** *En las condiciones de la definición 4.21,  $\mathbf{f}$  es continua por la izquierda (resp. derecha) en  $x \in (a, b]$  (resp.  $x \in [a, b)$ ), si y sólo si  $v$  es continua por la izquierda (resp. derecha) en  $x$ . En particular, si  $\mathbf{f}$  es un camino rectificable de longitud  $L$ , su variación indefinida  $v : [a, b] \rightarrow L$  es una función creciente continua y sobreyectiva.*

DEM: Supongamos que  $\mathbf{f}$  es continua por la izquierda en  $x \in (a, b]$ . Según la definición de  $V(\mathbf{f}, [a, x])$ , dado  $\epsilon > 0$  existe  $p \in \mathcal{P}([a, x])$  tal que

$$V(\mathbf{f}, p) \geq V(\mathbf{f}, [a, x]) - \epsilon/2$$

Al refinar  $p$  se conserva la desigualdad anterior y añadiendo un punto si es necesario podemos suponer que  $p = (t_0 < t_1 < \dots < t_{n-1} < t_n = x)$ , donde  $t_{n-1}$  ha sido elegido suficientemente próximo a  $x$  para que, en virtud de la continuidad por la izquierda de  $\mathbf{f}$  en  $x$ , se pueda asegurar que

$$t_{n-1} \leq t \leq x \Rightarrow \|\mathbf{f}(t) - \mathbf{f}(x)\| < \epsilon/2$$

En estas condiciones, si  $t_{n-1} < t < x$  se cumple

$$v(t) = V(\mathbf{f}, [a, t]) \geq V(\mathbf{f}, [a, t_{n-1}]) \geq \sum_{i=1}^{n-1} \|\mathbf{f}(t_i) - \mathbf{f}(t_{i-1})\|$$

$$= V(\mathbf{f}, p) - \|\mathbf{f}(x) - \mathbf{f}(t_{n-1})\| \geq V(\mathbf{f}, [a, x]) - \epsilon/2 - \epsilon/2 = v(x) - \epsilon$$

y queda demostrado que  $v$  es continua por la izquierda en  $x$ . Análogamente demuestra la continuidad por la derecha, cuando  $x \in [a, b)$ .

Recíprocamente, teniendo en cuenta que para cada par de puntos  $a \leq x < y \leq b$  se verifica  $\|\mathbf{f}(x) - \mathbf{f}(y)\| \leq V(\mathbf{f}, [x, y]) = v(y) - v(x)$ , es inmediato que  $\mathbf{f}$  es continua por la izquierda (resp. derecha) en cada punto donde  $v$  sea continua por la izquierda (resp. derecha). ■

**Teorema 4.24** *Si  $\mathbf{f} : [a, b] \rightarrow E$  es un camino de clase  $C^1$  su variación indefinida  $v(t) = V(\mathbf{f}, [a, t])$  también lo es y  $v'(x) = \|\mathbf{f}'(x)\|$  para cada  $x \in [a, b]$ . Por lo tanto*

$$V(\mathbf{f}, [a, b]) = \int_a^b \|\mathbf{f}'(x)\| dx$$

DEM: Si  $\mathbf{f}$  es derivable con derivada continua ya hemos indicado en 4.20 que  $\mathbf{f}$  es de variación acotada y por lo tanto podemos considerar su variación indefinida. Probaremos en primer lugar que si  $x < b$  entonces  $v$  es derivable por la derecha en  $x$  con  $v'_d(x) = \|\mathbf{f}'(x)\|$ .

Si  $x < s \leq b$  sea  $M(x, s) = \max\{\|\mathbf{f}'(t)\| : x \leq t \leq s\}$ . En virtud del teorema del incremento finito, para todo  $t, t' \in [x, s]$  se verifica  $\|\mathbf{f}(t) - \mathbf{f}(t')\| \leq M(x, s)|t - t'|$  luego, en virtud de 4.20 b)

$$V(\mathbf{f}, [x, s]) \leq M(x, s)|s - x|$$

Como  $\mathbf{f}'$  es continua en  $x$ , dado  $\epsilon$  existe  $\delta > 0$  tal que

$$t \in [a, b], \quad |t - x| < \delta \Rightarrow \|\mathbf{f}'(t) - \mathbf{f}'(x)\| < \epsilon$$

Sea  $s \in (a, b]$  tal que  $|s - x| < \delta$ . Entonces para todo  $t \in [x, s]$  se cumple  $\|\mathbf{f}'(t)\| \leq \|\mathbf{f}'(x)\| + \epsilon$  luego  $M(x, s) \leq \|\mathbf{f}'(x)\| + \epsilon$  y se obtiene

$$v(s) - v(x) = V(\mathbf{f}, [x, s]) \leq (\|\mathbf{f}'(x)\| + \epsilon)|s - x| \quad (4.1)$$

Por otra parte, utilizando la definición de derivada y la continuidad de la norma

$$\|\mathbf{f}'(x)\| = \lim_{s \rightarrow x} \left\| \frac{\mathbf{f}(s) - \mathbf{f}(x)}{s - x} \right\|$$

y podemos suponer que el número  $\delta > 0$  ha sido elegido de modo que para todo  $s \in (x, b]$  con  $|s - x| < \delta$  se cumple la desigualdad

$$\|\mathbf{f}'(x)\| - \epsilon \leq \left\| \frac{\mathbf{f}(s) - \mathbf{f}(x)}{s - x} \right\| \quad (4.2)$$

Combinando 4.1 y 4.2 con la desigualdad  $\|\mathbf{f}(s) - \mathbf{f}(x)\| \leq v(s) - v(x)$  se concluye que para todo  $s \in (x, b]$  con  $|s - x| < \delta$  se verifica

$$\|\mathbf{f}'(x)\| - \epsilon \leq \frac{v(s) - v(x)}{s - x} \leq \|\mathbf{f}'(x)\| + \epsilon$$

Esto prueba que  $v$  es derivable por la derecha en  $x$  y que  $v'_d(x) = \|\mathbf{f}'(x)\|$ . Análogamente se prueba, cuando  $a < x$ , que  $v$  es derivable por la izquierda en  $x$  con  $v'_i(x) = \|\mathbf{f}'(x)\|$ . Finalmente, en virtud del teorema fundamental del cálculo

$$V(\mathbf{f}, [a, b]) = v(b) - v(a) = \int_a^b v'(x) dx = \int_a^b \|\mathbf{f}'(x)\| dx$$

■

Dados dos caminos  $\gamma_1 : [a_1, b_1] \rightarrow E$ , y  $\gamma_2 : [a_2, b_2] \rightarrow E$  tales que  $b_1 = a_2$ , si el extremo del primero coincide con el origen del segundo, la *yuxtaposición*  $\gamma = \gamma_1 \vee \gamma_2$ , es el camino  $\gamma : [a_1, b_2] \rightarrow E$  definido por

$$\gamma(t) = \gamma_1(t) \quad \text{si } t \in [a_1, b_1]; \quad \gamma(t) = \gamma_2(t) \quad \text{si } t \in [a_2, b_2]$$

Análogamente se define la yuxtaposición  $\gamma = \gamma_1 \vee \gamma_2 \vee \cdots \vee \gamma_n$  de un número finito de caminos. que, en virtud de 4.21, será rectificable, si y sólo si cada  $\gamma_k$  lo es, y en ese caso  $\text{Long}(\gamma) = \sum_{k=1}^m \text{Long}(\gamma_k)$ . Se dice que un camino  $\gamma : [a, b] \rightarrow E$  es de *clase  $C^1$  a trozos* ó *regular a trozos* cuando se puede expresar como yuxtaposición de un número finito de caminos de clase  $C^1$ , es decir, cuando existe una subdivisión  $a = x_0 < x_1 < x_2 < \cdots < x_n = b$  tal que cada  $\gamma_k = \gamma|_{[x_{k-1}, x_k]}$  es de clase  $C^1$ .

En los puntos  $x_k$ ,  $1 < k < n$ , existen las derivadas laterales  $\gamma'_i(x_k)$ ,  $\gamma'_d(x_k)$  pero no está asegurada la derivabilidad. Es decir, los caminos regulares a trozos son derivables excepto en un conjunto finito de puntos, donde existen las derivadas laterales y son distintas. Por ello, a los correspondientes puntos de la imagen  $\gamma([a, b])$  se les llama *vértices del camino*.

El teorema 4.24 se extiende fácilmente al caso de los caminos regulares a trozos.

**Teorema 4.25** *Todo camino  $\gamma : [a, b] \rightarrow E$  regular a trozos es rectificable y su longitud viene dada por la integral*

$$\text{Long}(\gamma) = \int_a^b \|\gamma'(t)\| dt$$

(La función  $\|\gamma'(t)\|$  se supone definida de modo arbitrario en los puntos donde la derivada no existe).

DEM: Por hipótesis, hay una subdivisión  $a = x_0 < x_1 < x_2 < \dots < x_m = b$  tal que cada  $\gamma_k = \gamma|_{[x_{k-1}, x_k]}$  es de clase  $C^1$ . Según 4.24 cada  $\gamma_k$  es rectificable con

$$\text{Long}(\gamma_k) = \int_{x_{k-1}}^{x_k} \|\gamma'_k(t)\| dt$$

(en los extremos de  $[x_{k-1}, x_k]$  el valor de  $\gamma'_k$  es el de la correspondiente derivada lateral). Entonces  $\gamma = \gamma_1 \vee \gamma_2 \vee \dots \vee \gamma_n$  también es rectificable y

$$\text{Long}(\gamma) = \sum_{k=1}^n \text{Long}(\gamma_k) = \sum_{k=1}^n \int_{x_{k-1}}^{x_k} \|\gamma'_k(t)\| dt$$

La función  $\|\gamma'(t)\|$  no está definida en los puntos  $x_k$ ,  $1 < k < n$ . Si en estos puntos se supone definida de modo arbitrario se obtiene una función integrable en cada intervalo  $[x_{k-1}, x_k]$  con

$$\int_{x_{k-1}}^{x_k} \|\gamma'(t)\| dt = \int_{x_{k-1}}^{x_k} \|\gamma'_k(t)\| dt$$

(obsérvese que  $\|\gamma'_k\|$  es continua en  $[x_{k-1}, x_k]$  y coincide con  $\|\gamma'\|$  en  $(x_{k-1}, x_k)$ ).

En virtud del teorema de adición de la integral respecto al intervalo se concluye que  $\|\gamma'\|$  es integrable en  $[a, b]$  y que

$$\text{Long}(\gamma) = \sum_{k=1}^n \int_{x_{k-1}}^{x_k} \|\gamma'_k(t)\| dt = \sum_{k=1}^n \int_{x_{k-1}}^{x_k} \|\gamma'(t)\| dt = \int_a^b \|\gamma'(t)\| dt$$

■

**Caminos referidos al arco.** Cuando  $\gamma : [a, b] \rightarrow E$  es un camino rectificable de longitud  $L$ , su variación indefinida  $v(t) = V(\gamma, [a, t])$  se suele llamar *abscisa curvilínea del camino*. En virtud de 4.21 y 4.23 la abscisa curvilínea es una función creciente continua y sobreyectiva  $v : [a, b] \rightarrow [0, L]$  cuya interpretación física es obvia: Si se piensa que  $t$  es el tiempo, entonces  $s_0 = v(t_0)$  es la longitud del camino que ha recorrido el punto  $\gamma(t)$  desde el instante inicial  $t = a$  hasta el instante  $t = t_0$ . Cuando el camino es de clase  $C^1$ , según esta interpretación física,  $\Delta s \approx v'(t)\Delta t$  es un valor aproximado de la longitud que recorre la partícula en un intervalo de tiempo pequeño  $(t, t + \Delta t)$ , luego  $v'(t)$  es la razón de cambio instantánea de la longitud

recorrida, dentro de la curva, frente al tiempo. A esta razón de cambio se le suele llamar *celeridad* o *rapidez*. Según el teorema 4.24 la rapidez  $v'(t)$  es precisamente la longitud  $\|\gamma'(t)\|$  del vector velocidad.

Si  $\gamma$  no es inyectiva puede ocurrir que  $\gamma(t)$  pase por un mismo punto  $\mathbf{x} = \gamma(t_1) = \gamma(t_2)$  en dos instantes distintos  $t_1 < t_2$ , de modo  $\mathbf{x}$  puede tener dos abscisas curvilíneas distintas y en ese caso convendrá precisar diciendo que  $s_i = v(t_i)$  es la abscisa curvilínea de  $\mathbf{x}$  que corresponde al valor  $t_i$  del parámetro.

Si  $v$  es estrictamente creciente con el cambio de variable  $t = v^{-1}(s)$  se obtiene la representación paramétrica  $\tilde{\gamma}(s) = \gamma(v^{-1}(s))$ , definida en  $[0, L]$ . El camino  $\tilde{\gamma}$  se dice que está referido al arco como parámetro ya que el punto  $\tilde{\gamma}(s)$  es el punto de la curva al que se llega después de recorrer sobre la misma un trayecto de longitud  $s$ . También se dice que el parámetro de  $\tilde{\gamma}$  es el arco. Se puede formalizar esta definición diciendo que un camino rectificable  $\gamma$ , de longitud  $L$ , está referido al arco como parámetro cuando su dominio es  $[0, L]$  y para cada  $s \in [0, L]$  la longitud de  $\gamma|_{[0,s]}$  es  $s$ .

En particular, cuando  $\gamma$  es de clase  $C^1$  y  $\gamma'(t) \neq 0$  para todo  $t \in [a, b]$  se cumple que la abscisa curvilínea  $v(t)$  es estrictamente creciente, ya que  $v'(t) = \|\gamma'(t)\| > 0$  para todo  $t \in [a, b]$ . Si  $s = v(t)$ , en virtud de la regla de la cadena,  $v'(t) = \|\gamma'(t)\| = \|\tilde{\gamma}'(s)v'(t)\| = \|\tilde{\gamma}'(s)\|v'(t)$ , luego  $\|\tilde{\gamma}'(s)\| = 1$  para todo  $s \in [0, L]$ . Se puede obtener lo mismo usando el teorema 4.24 ya que la variación indefinida de  $\tilde{\gamma}$  es  $v(s) = s$ .

## 4.5. Integral respecto al arco

Cuando la abscisa curvilínea del camino rectificable  $\gamma$  no es estrictamente creciente, aunque no se puede definir el camino equivalente  $\tilde{\gamma}$ , sin embargo es posible definir un camino, que seguimos denotando  $\tilde{\gamma}$ , que está referido al arco como parámetro, tiene su misma longitud y verifica  $\tilde{\gamma} \circ v = \gamma$ . Este camino se obtiene de modo informal olvidando los intervalos de tiempo durante los que  $\gamma(t)$  está parado. Así se consigue un camino sin paradas cuya abscisa curvilínea es estrictamente creciente y por lo tanto admite una parametrización equivalente  $\tilde{\gamma}$  cuyo parámetro es el arco. Formalmente  $\tilde{\gamma}$  queda definido mediante la siguiente proposición:

**Proposición 4.26** *Sea  $\gamma : [a, b] \rightarrow E$  un camino rectificable de longitud  $L$  y  $v(t) = V(\gamma, [a, t])$  su abscisa curvilínea. Entonces existe un único camino  $\tilde{\gamma} : [0, L] \rightarrow E$  tal que  $\tilde{\gamma}(v(t)) = \gamma(t)$  para todo  $t \in [a, b]$ . El camino  $\tilde{\gamma}$  es rectificable, de longitud  $L$  y está referido al arco como parámetro.*

DEM: Como  $v : [a, b] \rightarrow [0, L]$  es continua y sobreyectiva, para cada  $s \in [0, L]$  existe  $t \in [a, b]$  tal que  $s = v(t)$ . Si  $s = v(t) = v(t')$  con  $t < t'$  entonces  $V(\gamma, [t, t']) = v(t') - v(t) = 0$  luego  $\gamma$  es constante en  $[t, t']$ . Esto prueba que  $\gamma$  es constante en el intervalo  $\{t \in [a, b] : v(t) = s\}$ , con lo cual se puede definir  $\tilde{\gamma}(s)$  como ese valor constante. Evidentemente  $\tilde{\gamma}(s) = \gamma(t)$  siempre que  $s = v(t)$ . En virtud de la nota que sigue a la proposición 4.19 el camino  $\tilde{\gamma}$  es rectificable de longitud  $L$  y está referido al arco como parámetro ya que si  $s \in [0, L]$  y  $v(t) = s$  se cumple

$$V(\tilde{\gamma}, [0, s]) = V(\tilde{\gamma} \circ v, [a, t]) = V(\gamma, [a, t]) = v(t) = s$$





**Definición 4.27** Si  $\gamma([a, b]) \rightarrow E$  es rectificable y  $g : \gamma([a, b]) \rightarrow \mathbb{R}$  es acotada, la integral de  $g$  respecto al arco del camino  $\gamma$  se define mediante la integral de Riemann  $\int_0^L g(\tilde{\gamma}(s))ds$  siempre que esta integral exista. Para ella se usa la notación

$$\int g \|d\gamma\| = \int_0^L g(\tilde{\gamma}(s))ds$$

Toda función continua sobre un arco de curva rectificable es integrable respecto al arco. Cuando  $g$  es la función 1 el valor de la integral es la longitud del camino  $\gamma$ .

Si  $\gamma$  es de clase  $C^1$ , en virtud del teorema 4.24 la abscisa curvilínea  $v(t)$  también lo es, y efectuando el cambio de variable  $s = v(t)$  resulta

$$\int g \|d\gamma\| = \int_0^L g(\tilde{\gamma}(s))ds = \int_a^b g(\gamma(t))v'(t)dt = \int_a^b g(\gamma(t)) \|\gamma'(t)\| dt$$

y esta última fórmula es la que ha motivado la notación utilizada para la integral respecto al arco. Más generalmente, si  $\gamma$  es regular a trozos se obtiene una expresión análoga pero teniendo en cuenta que ahora  $\|\gamma'(t)\|$ , no está definida en un conjunto finito de puntos y se pueden hacer las mismas observaciones que se hicieron en la demostración de 4.25. Para relacionar el concepto que se acaba de definir con algo concreto consideremos varios ejemplos:

a) A lo largo de un arco de curva plana simple  $C$ , situada en el suelo y de ecuaciones paramétricas  $x = x(t)$ ,  $y = y(t)$ ,  $t \in [0, 1]$  se levanta una valla  $V$  de altura variable. Si  $h(x, y) \geq 0$  es la altura de la valla en el punto  $(x, y) \in C$ , podemos tomar como área de la valla el valor de la integral de  $h$  sobre  $C$ . Este es un caso particular de la fórmula general, que veremos más adelante, para hallar el área de un trozo de superficie:

$$\text{Área}(V) = \int_0^1 h(x(t), y(t))\sqrt{x'(t)^2 + y'(t)^2}dt$$

b) Un alambre en el espacio tridimensional tiene la forma del arco de curva  $C$  cuya representación paramétrica es  $\gamma(t) = (x(t), y(t), z(t))$ ,  $t \in [0, 1]$ . Si  $p(\mathbf{x})$  es la densidad de masa en el punto  $\mathbf{x} \in C$  entonces la integral de  $p$  sobre  $C$  proporciona la masa total del alambre:

$$M = \int p \|d\gamma\| = \int_0^1 p(\gamma(t))\sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2}dt$$

c) Análogamente, si  $T = g(\mathbf{x})$  es la temperatura del punto  $\mathbf{x}$  del alambre y  $L = \text{Long}(\gamma)$  es su longitud, podemos usar la integral de  $g$  sobre  $C$  para obtener la temperatura media del alambre

$$T_m = \frac{1}{L} \int g \|d\gamma\|$$

Antes de demostrar la siguiente proposición conviene hacer una observación preliminar recogida en el siguiente lema

**Lema 4.28** Sean  $\gamma_i : [a_i, b_i] \rightarrow E$ ,  $i = 1, 2$ , caminos rectificables topológicamente equivalentes y sea  $h : [a_1, b_1] \rightarrow [a_2, b_2]$  una biyección continua tal que  $\gamma_1 = \gamma_2 \circ h$ . Si  $h$  es creciente se verifica  $\tilde{\gamma}_1 = \tilde{\gamma}_2$ , y si  $h$  es decreciente entonces  $\tilde{\gamma}_1(s) = \tilde{\gamma}_2(L-s)$ , donde  $L = \text{Long}(\gamma_1) = \text{Long}(\gamma_2)$ .

DEM: Si  $v_i$  es la abscisa curvilínea de  $\gamma_i$ , ( $i = 1, 2$ ), para cada  $0 \leq s \leq L$  existe  $t \in [a_1, b_1]$  tal que  $s = v_1(t)$ . Si  $h$  es creciente y  $t' = h(t)$ , en virtud de la nota que sigue a 4.19,

$$v_1(t) = V(\gamma_1, [a_1, t]) = V(\gamma_2, [a_2, t']) = v_2(t')$$

luego  $\tilde{\gamma}_1(s) = \gamma_1(t) = \gamma_2(t') = \tilde{\gamma}_2(s)$ .

Cuando  $h$  es decreciente, la imagen de  $[a_1, t]$  es  $[t', b_2]$ , luego

$$v_1(t) = V(\gamma_1, [a_1, t]) = V(\gamma_2, [t', b_2]) = v_2(b_2) - v_2(t') = L - v_2(t')$$

es decir,  $v_2(t') = L - s$ . Por consiguiente  $\tilde{\gamma}_1(s) = \gamma_1(t) = \gamma_2(t') = \tilde{\gamma}_2(L - s)$ . ■

**Proposición 4.29** Sea  $\gamma : [a, b] \rightarrow E$  un camino rectificable y  $g : \gamma([a, b]) \rightarrow \mathbb{R}$  una función integrable respecto al arco.

- a) Si  $h : [c, d] \rightarrow [a, b]$  es un homeomorfismo y  $\gamma_1 = \gamma \circ h$  entonces las integrales  $\int g \|d\gamma\|$ ,  $\int g \|d\gamma_1\|$ , existen simultáneamente y tienen el mismo valor.
- b) Si  $\gamma = \gamma_1 \vee \gamma_2$  entonces  $g$  es integrable sobre  $\gamma_1$  y sobre  $\gamma_2$  y

$$\int g \|d\gamma\| = \int g \|d\gamma_1\| + \int g \|d\gamma_2\|$$

- c) Si  $M = \sup\{|g(\mathbf{x})| : \mathbf{x} \in \gamma([a, b])\}$  entonces  $\left| \int g \|d\gamma\| \right| \leq M \text{Long}(\gamma)$

DEM: a) Según el lema 4.28  $\tilde{\gamma}_1(s) = \tilde{\gamma}(s)$  si  $h$  es creciente y  $\tilde{\gamma}_1(s) = \tilde{\gamma}(L-s)$  si  $h$  es decreciente. En el primer caso a) es consecuencia directa de la definición. En el segundo caso también lo es después de hacer el cambio de variable  $x = L - s$  en la integral  $\int_0^L g(\tilde{\gamma}_1(x)) ds$ .

b) Si  $L_i = \text{Long}(\gamma_i)$  y  $L = \text{Long}(\gamma)$  se tiene  $L = L_1 + L_2$  luego

$$\int_0^L g(\tilde{\gamma}(s)) ds = \int_0^{L_1} g(\tilde{\gamma}(s)) ds + \int_{L_1}^L g(\tilde{\gamma}(s)) ds$$

Es fácil ver que  $\tilde{\gamma}_1(s) = \tilde{\gamma}(s)$  si  $s \in [0, L_1]$  y  $\tilde{\gamma}_2(x) = \tilde{\gamma}(x + L_1)$  si  $x \in [0, L_2]$ . Haciendo el cambio de variable  $s = x + L_1$  en la segunda integral se obtiene el resultado.

c) Es inmediato. ■

La propiedad a) en la proposición anterior permite definir la integral de una función sobre un arco de curva rectificable a través de cualquier representación paramétrica ya que todas sus representaciones paramétricas proporcionan la misma integral. Es decir, la integral de una función sobre un camino rectificable realmente es una noción asociada al arco de curva definido por el camino.

## 4.6. Ejercicios resueltos

**Ejercicio 4.30** *Demuestre que una función real  $f : [a, b] \rightarrow \mathbb{R}$  es de variación acotada si y sólo si existe una pareja de funciones crecientes  $g, h : [a, b] \rightarrow \mathbb{R}$  tal que  $f = g - h$ . Las funciones  $g, h$  se pueden elegir continuas si  $f$  es continua.*

SOLUCIÓN

Basta tomar  $g = v$  y  $h = v - g$  con  $v(x) = V(f, [a, x])$ . Obsérvese que  $h = v - f$  es creciente pues si  $a \leq x < y \leq b$  se verifica

$$v(y) - v(x) = V(f, [x, y]) \geq |f(y) - f(x)| \geq f(y) - f(x)$$

es decir  $h(x) \leq h(y)$ . Después del teorema 4.23 es claro que  $g$  y  $h$  son continuas en los puntos donde  $f$  es continua. ■

**Ejercicio 4.31** *Es bien sabido que  $d'(x, y) = |x^3 - y^3|$  es una distancia equivalente a la distancia usual de  $\mathbb{R}$ ,  $d(x, y) = |x - y|$ . Compruebe que la función  $f : [0, 1] \rightarrow \mathbb{R}$ ,  $f(t) = t \cos(\pi/t)$ , si  $t \in (0, 1]$ ,  $f(0) = 0$ , no es de variación acotada para la distancia  $d$  pero es de variación acotada para la distancia  $d'$ .*

SOLUCIÓN

Si  $p_n = (0 < 1/n < 1/(n-1) < \dots < 1/2 < 1)$ , como  $\cos \pi k = -\cos \pi(k-1) = \pm 1$ , resulta  $|f(1/k) - f(1/(k-1))| = 1/k + 1/(k-1) \geq 2/k$  luego la sucesión  $V(f, p_n) \geq \sum_{k=1}^n (2/k)$  no está acotada, y por lo tanto  $f$  no es de variación acotada para la distancia  $d$ .

Es claro que  $f$  es de variación acotada para la distancia  $d'$  si y sólo si  $f^3$  es de variación acotada para la distancia  $d$ . Es fácil ver que  $f^3$  es derivable con derivada acotada, y utilizando 4.20 b) se obtiene que  $f^3$  es de variación acotada para la distancia  $d$ , lo que significa que  $f$  es de variación acotada para la distancia  $d'$ . ■

**Ejercicio 4.32** *Sea  $(E, \| \cdot \|)$  un espacio normado y  $\mathbf{f} : [a, b] \rightarrow E$  es de variación acotada demuestre que el conjunto de sus puntos de discontinuidad es numerable. Si  $E$  es completo demuestre también que todas las discontinuidades de  $\mathbf{f}$  son de primera especie (e.d. en los puntos de discontinuidad existen los límites laterales)*

SOLUCIÓN

La primera afirmación es consecuencia directa del teorema 4.23 pues el conjunto de los puntos de discontinuidad de la función creciente  $v$  es numerable. La función creciente  $v$  tiene límites laterales en todos los puntos, y por lo tanto cumple en todos ellos la condición de Cauchy para la existencia de los límites laterales. Entonces, utilizando la desigualdad  $\|\mathbf{f}(x) - \mathbf{f}(y)\| \leq v(y) - v(x)$  se obtiene que  $\mathbf{f}$  también cumple, en todos los puntos, la condición de Cauchy para la existencia de los límites laterales. Por lo tanto, cuando  $E$  sea completo, se puede asegurar que  $\mathbf{f}$  tiene límites

laterales en todo punto.

NOTA: Las funciones con límites laterales en todo punto son regladas (límites uniformes de funciones escalonadas) y las discontinuidades de estas funciones también forman un conjunto numerable (véase el ejercicio propuesto 3.8.21).

Por otra parte, cuando  $E = \mathbb{R}^n$ , si  $\mathbf{f} : [a, b] \rightarrow \mathbb{R}^n$  es de variación acotada, en virtud de 4.30 y del ejercicio propuesto 4.7.14, cada componente de  $\mathbf{f}$  es diferencia de dos funciones crecientes de modo que, en este caso, se puede obtener directamente la existencia de los límites laterales sin acudir a la condición de Cauchy. ■

## 4.7. Ejercicios propuestos

◇ 4.7.1 Sea  $(E, \| \cdot \|)$  un espacio normado real cuya norma procede de un producto escalar y  $\mathbf{f} : (a, b) \rightarrow E$  una función derivable: Demuestre las siguientes afirmaciones

- La función  $t \rightarrow \|\mathbf{f}(t)\|$  es constante si y sólo si los vectores  $\mathbf{f}(t)$  y  $\mathbf{f}'(t)$  son ortogonales para todo  $t \in (a, b)$ .
- Si para cada  $t \in (a, b)$  es  $\mathbf{f}(t) \neq 0$  y los vectores  $\mathbf{f}(t)$ ,  $\mathbf{f}'(t)$  son linealmente dependientes entonces  $\mathbf{f}$  es de la forma  $\mathbf{f}(t) = \alpha(t)\mathbf{v}$  donde  $\alpha : (a, b) \rightarrow \mathbb{R}$  es derivable y  $\mathbf{v} \in E$ .

◇ 4.7.2 Sea  $\mathbf{f} : (a, b) \rightarrow \mathbb{R}^n$  derivable tal que  $\mathbf{f}'(t) \neq 0$  para todo  $t \in (a, b)$  y  $\mathbf{p} \in \mathbb{R}^n \setminus \mathbf{f}((a, b))$ . Se supone que existe  $t_0 \in (a, b)$  tal que  $\mathbf{q} = \mathbf{f}(t_0)$  es el punto de  $\mathbf{f}((a, b))$  más cercano a  $\mathbf{p}$ , es decir,

$$\|\mathbf{p} - \mathbf{q}\|_2 \leq \|\mathbf{p} - \mathbf{f}(t)\|_2 \text{ para todo } t \in (a, b)$$

Demuestre que el vector  $\mathbf{p} - \mathbf{q}$  es ortogonal a la curva  $\mathbf{f}(t)$  en el punto  $\mathbf{q} = \mathbf{f}(t_0)$ .

◇ 4.7.3 Demuestre las siguientes propiedades de reflexión de las cónicas:

- En un reflector parabólico, los rayos paralelos al eje se reflejan pasando por el foco.
- Los rayos luminosos que parten de uno de los focos de un reflector elíptico se reflejan pasando por el otro foco.
- Los rayos luminosos que dirigidos a uno de los focos de un reflector hiperbólico se reflejan pasando por el otro foco.

◇ 4.7.4 Demuestre que en el movimiento de una partícula el producto escalar del vector velocidad por el vector aceleración es igual a la mitad de la derivada del cuadrado de la celeridad.

◇ 4.7.5 Una partícula se mueve recorriendo con velocidad escalar uniforme  $v$  la circunferencia de centro  $(0, 0)$  y radio  $r$ . Si  $\mathbf{r}(t)$  es la posición de la partícula en el instante  $t$  demuestre que los vectores  $\mathbf{r}(t)$  y  $\mathbf{r}''(t)$  son ortogonales a  $\mathbf{r}(t)$ , luego  $\mathbf{r}''(t) = k(t)\mathbf{r}(t)$  donde  $k(t) \in \mathbb{R}$ . Deduzca que  $k(t)$  es constante y que el vector aceleración  $\mathbf{r}''(t)$  apunta hacia el origen y que su longitud es  $v^2/r$ . (Indicación: Derivar  $\langle \mathbf{r}(t) | \mathbf{r}'(t) \rangle = 0$ ).

◇ 4.7.6 En un movimiento plano, una partícula  $\mathbf{r}(t)$  se mueve de modo que el vector aceleración siempre es radial (e.d.  $\mathbf{r}''(t) = \alpha(t)\mathbf{f}(t)$  con  $\alpha(t) > 0$ ). Demuestre que esto ocurre si y sólo si el área barrida por el vector de posición  $\mathbf{r}(t)$  es proporcional al tiempo empleado.

◇ 4.7.7 Sea  $(E, \| \cdot \|)$  un espacio normado completo. Se supone que  $\mathbf{f} : (a, b) \rightarrow E$  es derivable por la derecha en cada  $t \in (a, b)$  con  $\|\mathbf{f}'_d(t)\| \leq M$  para todo  $t \in (a, b)$ . Demuestre las siguientes afirmaciones:

- a)  $\mathbf{f}$  se puede extender a una función continua  $\bar{\mathbf{f}}: [a, b] \rightarrow E$ .
- b) Si existe el límite  $\lim_{t \rightarrow b} \mathbf{f}'_d(t) = \mathbf{w}$  entonces  $\bar{\mathbf{f}}$  es derivable por la izquierda en  $b$  con  $\bar{\mathbf{f}}'_i(b) = \mathbf{w}$ .

Deduzca de lo anterior que toda función continua  $\mathbf{f}: (a, b) \rightarrow E$  con derivada por la derecha continua es derivable.

◇ 4.7.8 Obtenga la longitud de los siguientes arcos de curva:

- a) Gráfica de la función  $f(x) = |x|^{3/2}$ , sobre el intervalo  $-a \leq x \leq a$ .
- b) Arco de cicloide  $\mathbf{f}(t) = (t - \operatorname{sen} t, 1 - \operatorname{cos} t)$ ,  $t \in [0, 2\pi]$ .

◇ 4.7.9 Obtenga una parametrización equivalente del arco de hélice

$$\gamma(t) = (\operatorname{cos} t, \operatorname{sen} t, t), 0 \leq t \leq 2\pi$$

en la que el parámetro sea el arco.

◇ 4.7.10 Un alambre tiene la forma de la curva  $C \subset \mathbb{R}^3$  y su densidad de masa en  $(x, y, z) \in C$  viene dada por la función continua  $p(x, y, z)$ . La masa total del alambre es  $M = \int p(x, y, z) ds$ , su centro de gravedad es el punto  $(x_0, y_0, z_0)$  de coordenadas

$$x_0 = \frac{1}{M} \int_C xp(x, y, z) ds; \quad y_0 = \frac{1}{M} \int_C yp(x, y, z) ds; \quad z_0 = \frac{1}{M} \int_C zp(x, y, z) ds$$

y su momento de inercia respecto a un eje  $E$  es

$$I_E = \int_C \delta^2(x, y, z) p(x, y, z) ds$$

donde  $\delta(x, y, z)$  es la distancia del punto  $(x, y, z)$  al eje  $E$ .

- 1) Calcule el centro de gravedad de un alambre, con distribución uniforme de masa, que tiene forma de semicircunferencia.
- 2) Calcule la masa de un alambre que sigue la intersección del plano  $x + y + z = 1$  con la esfera  $x^2 + y^2 + z^2 = 1$ , cuando su densidad de masa viene dada por  $p(x, y, z) = x^2 + y^2 + z^2$  gramos por unidad de longitud.
- 3) Un alambre tiene la forma de la hélice  $\mathbf{f}(t) = (\operatorname{cos} t, \operatorname{sen} t, t)$ ,  $0 \leq t \leq 2\pi$  y su función de densidad es  $p(x, y, z) = x^2 + y^2 + z^2$ . Obtenga:
  - i) La masa total del muelle en forma de hélice.
  - ii) La coordenada  $z_0$  de su centro de gravedad.
  - iii) Su momento de inercia respecto al eje  $OZ$ .

◇ 4.7.11 Calcule el centro de gravedad de un arco de cicloide

$$\mathbf{f}(t) = (t - \operatorname{sen} t, 1 - \operatorname{cos} t), \quad 0 \leq t \leq \pi.$$

(se supone una distribución uniforme de la masa)

◇ **4.7.12** Calcule el área de la valla construida sobre la curva plana  $\mathbf{f}(t) = (\cos^3 t, \sin^3 t)$   $t \in [0, \pi/2]$ , con altura variable  $h(x, y) = 1 + y/3$ .

◇ **4.7.13** Sea  $\mathbf{f} : [a, b] \rightarrow \mathbb{R}^n$  un camino de clase  $C^2$  y longitud  $L$  tal que  $\mathbf{f}'(t) \neq 0$  para cada  $t \in (a, b)$ . Si  $\mathbf{F} : [0, L] \rightarrow \mathbb{R}^n$  es la representación paramétrica canónica que utiliza el arco como parámetro y  $s(t)$  es la función la abscisa curvilínea demuestre:

a)  $s(t)$  es de clase  $C^2$  y  $s''(t) = \frac{1}{s'(t)} \langle \mathbf{f}'(t) | \mathbf{f}''(t) \rangle$

b)  $\|\mathbf{F}''(s(t))\| = \|\mathbf{f}'(t)\|^{-3} \sqrt{\|\mathbf{f}'(t)\|^2 \|\mathbf{f}''(t)\|^2 - \langle \mathbf{f}'(t) | \mathbf{f}''(t) \rangle^2}$

◇ **4.7.14**  $\mathbf{f} : [a, b] \rightarrow \mathbb{R}^n$  es de variación acotada si y sólo si lo son todas sus componentes  $f_j : [a, b] \rightarrow \mathbb{R}$ ,  $1 \leq j \leq n$ . (En  $\mathbb{R}^n$  se puede considerar la norma usual, o cualquier otra, dado que todas las normas son equivalentes).

◇ **4.7.15** Si  $(E, \|\cdot\|)$  es un espacio normado y  $\mathbf{f}, \mathbf{g} : [a, b] \rightarrow E$ , obtenga la desigualdad

$$V(\mathbf{f} + \mathbf{g}, [a, b]) \leq V(\mathbf{f}, [a, b]) + V(\mathbf{g}, [a, b])$$

Como consecuencia la suma de funciones de variación acotada es de variación acotada y  $BV(I, E)$  es un espacio vectorial (sobre el mismo cuerpo que  $E$ ).

◇ **4.7.16** Demuestre que el producto de dos funciones de variación acotada  $f, g : [a, b] \rightarrow \mathbb{R}$  es de variación acotada. Enuncie y demuestre los resultados análogos referentes al producto de funciones con valores complejos, al producto de una función escalar con una función vectorial, y al producto escalar de dos funciones vectoriales (cuando la norma procede del producto escalar).

◇ **4.7.17** Si  $f : [a, b] \rightarrow \mathbb{R}$  (o  $\mathbb{C}$ ) es de variación acotada y  $|f(t)| \geq m > 0$  para todo  $t \in [a, b]$  demuestre que  $1/f$  también es de variación acotada.

◇ **4.7.18** Si  $f : \mathbb{R} \rightarrow \mathbb{R}$  es un polinomio y  $x_1 < x_2 \cdots x_m$  son los ceros de  $f'$  que pertenecen al intervalo  $[a, b]$ , demuestre que

$$V(f, [a, b]) = |f(a) - f(x_1)| + |f(x_2) - f(x_3)| + \cdots + |f(b) - f(x_m)|$$

◇ **4.7.19** Si  $f : [a, b] \rightarrow \mathbb{R}$  es de variación acotada y  $\mathbf{g} : \mathbb{R} \rightarrow \mathbb{R}^n$  es de clase  $C^1$  demuestre que  $\mathbf{g} \circ f : [a, b] \rightarrow \mathbb{R}^n$  es de variación acotada.

◇ **4.7.20** Sea  $f : [a, b] \rightarrow \mathbb{R}$  integrable Riemann. Demuestre que  $g(x) = \int_a^x f(t) dt$  es de variación acotada en  $[a, b]$  y  $V(g, [a, b]) = \int_a^b |f(t)| dt$ .

◇ **4.7.21** Demuestre que para todo  $a > 0$  la función  $f(x) = \sum_{n=1}^{\infty} 2^{-n} \sin(10^n x)$  no es de variación acotada sobre  $[0, a]$ .

◇ **4.7.22** *Dado un espacio normado  $(E, \| \cdot \|)$  completo demuestre que*

$$\|\mathbf{f}\| = \|\mathbf{f}(a)\| + V(\mathbf{f}, [a, b])$$

*defina una norma en el espacio vectorial  $BV([a, b], E)$  de las funciones de variación acotada y que con esta norma el espacio es completo.*

*Si  $NBV([a, b], E)$  es el subespacio vectorial de  $BV([a, b], E)$  formado por las funciones que se anulan en  $a$  y son continuas por la izquierda en cada punto de  $(a, b)$  compruebe que  $NBV([a, b], E)$  es un subespacio cerrado de  $BV([a, b], E)$ .*



# Capítulo 5

## Funciones diferenciables

*Funciones de varias variables: Derivada según un vector y derivadas parciales. Aplicaciones diferenciables. Condición suficiente de diferenciability. Reglas del cálculo diferencial. Regla de la cadena. Espacio tangente. Gradiente.*

En este capítulo comienza el cálculo diferencial para funciones de varias variables reales con valores escalares o vectoriales. Aunque la noción elemental de derivada parcial es insuficiente para poder desarrollar con ella un cálculo diferencial satisfactorio, similar al de funciones de una variable, sin embargo la sola consideración de las derivadas parciales tiene aplicaciones interesantes que se exponen en la primera sección: Las funciones con derivadas parciales acotadas son continuas y las funciones con derivadas parciales nulas son constantes si su dominio es un abierto conexo.

Una de las primeras aplicaciones de la noción de derivada parcial la proporciona la condición necesaria de extremo relativo. Con este modesto recurso ya se pueden abordar y resolver algunos problemas de optimización.

Sin embargo la sola consideración de las derivadas parciales no proporciona un marco satisfactorio para desarrollar un cálculo diferencial con resultados similares al caso de las funciones de una variable. El marco adecuado lo proporciona la noción de aplicación diferenciable. En este caso la diferencial en un punto proporciona una aproximación local de la función mediante un polinomio de primer grado (lo que equivale a una aproximación local, mediante una aplicación lineal, del incremento de la función). Esta noción, para funciones de dos variables, se motiva con el problema de definir planos tangentes a 'superficies' dadas en forma explícita  $z = f(x, y)$ .

Después de introducir la noción de aplicación diferenciable y de haber estudiado las condiciones necesarias para la diferenciability se demuestra la condición suficiente de diferenciability en términos de continuidad de las derivadas parciales. Este resultado proporciona una herramienta eficaz para obtener la diferenciability de funciones en las situaciones habituales.

A continuación se establecen las reglas usuales del cálculo diferencial (diferencial de la suma, del producto, etc) haciendo especial énfasis en la regla de la cadena, en la que se apoyan la mayor parte de los resultados útiles del cálculo diferencial.

Después de introducir la matriz Jacobiana y las notaciones habituales del cálculo diferencial, se consideran las funciones con valores reales y para ellas, aprovechando

la estructura euclídea de  $\mathbb{R}^n$ , se expresa la diferencial en términos del vector gradiente. La noción de gradiente, que se formula de modo intrínseco, permite nuevas interpretaciones geométricas y físicas, y con las que se aprecia su interés desde el punto de vista de las aplicaciones.

Finaliza el capítulo con la existencia de espacio tangente a la gráfica de una aplicación diferenciable en un punto. También se considera el problema de la existencia de espacio tangente a conjuntos de nivel de aplicaciones diferenciables (en particular, para superficies dadas en forma implícita) y para conjuntos que se pueden parametrizar mediante una aplicación diferenciable (en particular, para superficies dadas en forma paramétrica). Para ello merece la pena anticipar algunos resultados posteriores sobre existencia de funciones implícitas, con el fin de que al finalizar el capítulo quede clara la noción precisa de espacio tangente y se conozcan los métodos para calcularlo en las situaciones habituales.

En este capítulo se formulan las primeras definiciones en el contexto general de los espacios normados sobre el cuerpo  $\mathbb{R}$ , es decir, para funciones  $\mathbf{f} : \Omega \rightarrow F$ , definidas en un abierto  $\Omega$  de un espacio normado  $(E, \|\cdot\|)$  con valores en otro espacio normado  $(F, \|\cdot\|)$ . (Para simplificar la notación designamos igual las normas de  $E$  y  $F$ ). Este contexto, además de proporcionar mayor generalidad, obliga a usar una notación más compacta que, incluso en el caso  $E = \mathbb{R}^n$ ,  $F = \mathbb{R}^m$ , evita el uso engorroso de coordenadas en aquellas cuestiones en las que estas no desempeñan un papel especial. Como los principales resultados, aplicaciones e interpretaciones geométricas, se refieren casi siempre al caso  $E = \mathbb{R}^n$ ,  $F = \mathbb{R}^m$ , el lector que lo desee puede considerar, desde ahora en adelante, esta situación particular.

## 5.1. Derivada según un vector

Es natural que el estudio local, en el entorno de un punto, de una función de varias variables se apoye en la consideración de las funciones de una variable que se obtienen restringiendo la función a las rectas que pasan por el punto. Esta idea es la que inspira la siguiente definición.

**Definición 5.1** *Sea  $\mathbf{f} : \Omega \rightarrow F$  una aplicación definida en un abierto  $\Omega$  del espacio normado  $(E, \|\cdot\|)$ , con valores en el espacio normado  $(F, \|\cdot\|)$ . Dados  $\mathbf{a} \in \Omega$  y  $\mathbf{v} \in E$ , la derivada de  $\mathbf{f}$  en  $\mathbf{a}$  según el vector  $\mathbf{v}$  es el vector de  $F$  dado por*

$$D_{\mathbf{v}}\mathbf{f}(\mathbf{a}) = \lim_{t \rightarrow 0} \frac{\mathbf{f}(\mathbf{a} + t\mathbf{v}) - \mathbf{f}(\mathbf{a})}{t}$$

en el supuesto de que el límite exista.

La derivada según un vector unitario se llama derivada direccional: Si  $\|\mathbf{v}\| = 1$ , la derivada direccional de  $\mathbf{f}$  en  $\mathbf{a}$  según la dirección  $\mathbf{v}$  es  $D_{\mathbf{v}}\mathbf{f}(\mathbf{a})$ .

Cuando  $E = \mathbb{R}^n$ , las derivadas según los vectores de la base canónica

$$\mathbf{e}_1 = (1, 0, \dots, 0), \mathbf{e}_2 = (0, 1, 0, \dots, 0), \mathbf{e}_n = (0, 0, \dots, 1)$$

tienen un interés especial: La derivada  $D_{\mathbf{e}_1}\mathbf{f}(\mathbf{a})$ , si existe, viene dada por

$$\lim_{t \rightarrow 0} = \frac{\mathbf{f}(\mathbf{a} + t\mathbf{e}_1) - \mathbf{f}(\mathbf{a})}{t} = \lim_{t \rightarrow 0} \frac{\mathbf{f}(a_1 + t, a_2, \dots, a_n) - \mathbf{f}(a_1, a_2, \dots, a_n)}{t}$$

Es decir,  $D_{\mathbf{e}_1}\mathbf{f}(\mathbf{a})$  es la derivada en  $a_1$  de la función parcial  $x_1 \rightarrow \mathbf{f}(x_1, a_2, a_3, \dots, a_n)$ , definida en un entorno abierto de  $a_1$ . Análogamente,  $D_{\mathbf{e}_j}\mathbf{f}(\mathbf{a})$  es la derivada, en  $a_j$ , de la función parcial  $x_j \rightarrow \mathbf{f}(a_1, \dots, a_{j-1}, x_j, a_{j+1}, \dots, a_n)$ .

Por esta razón a la derivada  $D_{\mathbf{e}_j}\mathbf{f}(\mathbf{a})$ , cuando existe, se le llama derivada parcial de  $\mathbf{f}(\mathbf{x}) = \mathbf{f}(x_1, x_2, \dots, x_n)$ , respecto a la variable  $x_j$ , en el punto  $\mathbf{a}$ , para la que se usan las notaciones habituales

$$D_j\mathbf{f}(\mathbf{a}); \quad \frac{\partial \mathbf{f}}{\partial x_j}(\mathbf{a}).$$

Para funciones de dos o tres variables, es habitual llamar  $x, y, z$  a las variables  $x_1, x_2, x_3$  y suele resultar cómodo emplear las notaciones

$$\mathbf{f}_x(\mathbf{a}) = \frac{\partial \mathbf{f}}{\partial x}(\mathbf{a}); \quad \mathbf{f}_y(\mathbf{a}) = \frac{\partial \mathbf{f}}{\partial y}(\mathbf{a}); \quad \mathbf{f}_z(\mathbf{a}) = \frac{\partial \mathbf{f}}{\partial z}(\mathbf{a}).$$

La ventaja de las derivadas parciales respecto a las derivadas según vectores arbitrarios reside en que su cálculo suele ser una tarea mecánica basada en las reglas usuales del cálculo de derivadas de funciones reales de una variable real.

Veremos en la siguiente sección que para las funciones diferenciables siempre existe la derivada según un vector  $\mathbf{v} = (v_1, v_2, \dots, v_n)$ , que se puede calcular en términos de las derivadas parciales usando la fórmula

$$D_{\mathbf{v}}\mathbf{f}(\mathbf{a}) = v_1 D_1\mathbf{f}(\mathbf{a}) + v_2 D_2\mathbf{f}(\mathbf{a}) + \dots + v_n D_n\mathbf{f}(\mathbf{a})$$

La siguiente proposición nos dice que para funciones con valores en  $\mathbb{R}^m$  el cálculo de la derivada según un vector, se puede realizar componente a componente:

**Proposición 5.2** *Si  $F = \mathbb{R}^m$  y  $\mathbf{f} = (f_1, f_2, \dots, f_m)$ , la derivada  $D_{\mathbf{v}}\mathbf{f}(\mathbf{a})$  existe si y sólo si existen las derivadas de las componentes,  $D_{\mathbf{v}}f_1(\mathbf{a}), D_{\mathbf{v}}f_2(\mathbf{a}), \dots, D_{\mathbf{v}}f_m(\mathbf{a})$  y en ese caso  $D_{\mathbf{v}}\mathbf{f}(\mathbf{a}) = (D_{\mathbf{v}}f_1(\mathbf{a}), D_{\mathbf{v}}f_2(\mathbf{a}), \dots, D_{\mathbf{v}}f_m(\mathbf{a}))$ .*

DEM: Basta aplicar la proposición 4.3 a la función  $\gamma_{\mathbf{v}}(t) = \mathbf{f}(\mathbf{a} + t\mathbf{v})$  cuyas componentes son las funciones  $f_j(\mathbf{a} + t\mathbf{v}), 1 \leq j \leq m$ . ■

*Interpretación física de la derivada según un vector:*

Cuando  $E = \mathbb{R}^3$  y  $F = \mathbb{R}$ , es posible dar diversas interpretaciones físicas de la derivada según un vector: Supongamos que en un recinto  $\Omega$  del espacio la temperatura  $T$  (o cualquier otra magnitud física, como la presión) depende del punto  $(x, y, z) \in \Omega$  y viene dada mediante la función  $T = f(x, y, z)$ . Si una partícula recorre una línea recta que pasa por  $\mathbf{p} \in \Omega$  con velocidad constante  $\mathbf{v}$  entonces  $D_{\mathbf{v}}f(\mathbf{p})$  representa la razón de cambio de la temperatura de la partícula respecto al tiempo, cuando la partícula pasa por  $\mathbf{p}$ : En efecto, si suponemos que la partícula pasa por  $\mathbf{p}$  en el instante  $t = 0$ , un poco después, en el instante  $t > 0$  la partícula

se encuentra en la posición  $\mathbf{p} + t\mathbf{v}$ , luego  $[f(\mathbf{p} + t\mathbf{v}) - f(\mathbf{p})]/t$  es la razón de cambio de la temperatura durante el intervalo de tiempo  $(0, t)$ . Pasando al límite cuando  $t \rightarrow 0$  se obtiene que  $D_{\mathbf{v}}f(\mathbf{p})$  es la razón de cambio instantánea en  $t = 0$ , es decir, cuando la partícula pasa por  $\mathbf{p}$ .

*Interpretación geométrica de la derivada según un vector:*

La interpretación geométrica de la derivada según un vector se basa en la siguiente

**Definición 5.3** Sea  $M$  un subconjunto no vacío del espacio normado  $F$  y  $\mathbf{p} \in M$ . Si existe una aplicación  $\gamma : (-\epsilon, \epsilon) \rightarrow F$ , derivable en  $t = 0$ , tal que

$$\gamma(t) \in M \text{ si } |t| < \epsilon, \quad \gamma(0) = \mathbf{p}, \quad \text{y} \quad \gamma'(0) = \mathbf{v}$$

diremos que  $\mathbf{v} \in F$  es un vector tangente a  $M$  en  $\mathbf{p}$ . En lo que sigue  $T_{\mathbf{p}}(M)$  designará el conjunto de los vectores tangentes a  $M$  en  $\mathbf{p}$ .

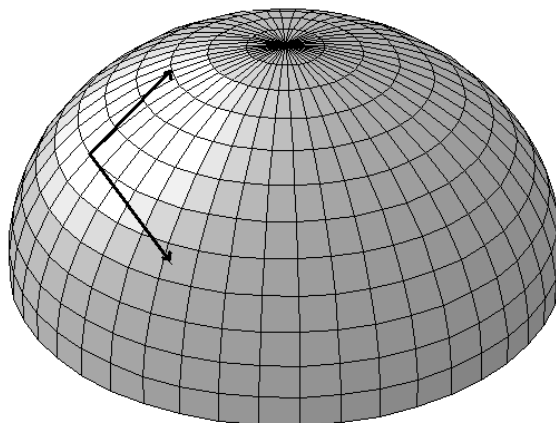
En las condiciones de la definición 5.1,  $D_{\mathbf{v}}f(\mathbf{a}) = \gamma'_{\mathbf{v}}(0)$  donde  $\gamma_{\mathbf{v}}(t) = \mathbf{f}(\mathbf{a} + t\mathbf{v})$  está definida en un entorno de 0, luego  $D_{\mathbf{v}}f(\mathbf{a})$  es un vector tangente a  $\mathbf{f}(\Omega)$  en  $\mathbf{p} = \mathbf{f}(\mathbf{a})$ . Si existe  $D_{\mathbf{v}}f(\mathbf{a})$  es fácil ver que también existe  $D_{\mathbf{u}}f(\mathbf{a})$  donde  $\mathbf{u} = t\mathbf{v}$ , con  $t \in \mathbb{R}$ , y que su valor es  $D_{\mathbf{u}}f(\mathbf{a}) = tD_{\mathbf{v}}f(\mathbf{a})$ . Es decir, cuando  $\mathbf{u} = t\mathbf{v}$  recorre una recta que pasa por 0, los vectores tangentes  $D_{\mathbf{u}}f(\mathbf{a})$  forman un subespacio vectorial unidimensional. Si estos vectores se dibujan con origen en  $\mathbf{p}$ , sus extremos recorren una recta afín, y diremos que esta recta es tangente a  $\mathbf{f}(\Omega)$  en  $\mathbf{p}$ .

a) Comencemos considerando el caso de una 'superficie' paramétrica,  $M = \varphi(\Omega)$  descrita como imagen de una aplicación  $\varphi : \Omega \rightarrow \mathbb{R}^3$ , definida en un abierto  $\Omega \subset \mathbb{R}^2$ . Las coordenadas  $(x, y, z)$  de un punto genérico de la 'superficie' son variables que dependen de dos parámetros independientes  $(s, t) \in \Omega$ . Si  $\varphi_1, \varphi_2, \varphi_3$  son las componentes de  $\varphi$ , se suele decir que  $M = \varphi(\Omega)$  es una 'superficie' de ecuaciones paramétricas  $x = \varphi_1(s, t)$ ,  $y = \varphi_2(s, t)$ ,  $z = \varphi_3(s, t)$ .

Dado un punto  $\mathbf{a} \in \Omega$ , uno de los objetivos del cálculo diferencial es el de definir de manera razonable, cuando sea posible, el plano tangente a la 'superficie' paramétrica  $M = \varphi(\Omega)$  en un punto  $\mathbf{p} = \varphi(\mathbf{a})$  de la misma. La idea básica es que este plano debe estar formado por rectas que pasan por  $\mathbf{p}$  y son tangentes a 'curvas' contenidas en la superficie. Dado un vector  $\mathbf{v} = (v_1, v_2) \in \mathbb{R}^2$ , una 'curva' de este tipo es la que tiene la ecuación paramétrica  $\gamma_{\mathbf{v}}(t) = \varphi(\mathbf{a} + t\mathbf{v})$ , donde  $t$  es un parámetro real que varía en el entorno de 0,  $A_{\mathbf{v}} = \{t \in \mathbb{R} : \mathbf{a} + t\mathbf{v} \in \Omega\}$ .

La derivada  $\gamma'_{\mathbf{v}}(0) = D_{\mathbf{v}}\varphi(\mathbf{a})$ , si existe, es un vector tangente a la 'superficie' paramétrica  $M = \varphi(\Omega)$  en el punto  $\mathbf{p} = \varphi(\mathbf{a})$ , que se suele representar gráficamente mediante una flecha con origen en  $\mathbf{p}$  y extremo en  $\mathbf{p} + D_{\mathbf{v}}\varphi(\mathbf{a})$ . En particular, las derivadas parciales  $D_1\varphi(\mathbf{a})$ ,  $D_2\varphi(\mathbf{a})$ , si existen, son vectores tangentes a  $M$  en  $\mathbf{p}$ .

Más adelante estudiaremos las condiciones que garantizan que  $T_{\mathbf{p}}(M)$  es un espacio vectorial de dimensión 2, con base  $\{D_1\varphi(\mathbf{a}), D_2\varphi(\mathbf{a})\}$ .



Para ver un ejemplo muy concreto, consideremos la parametrización habitual de la esfera unidad,  $\mathbf{r}(\varphi, \theta) = (\cos \varphi \cos \theta, \cos \varphi \sin \theta, \sin \varphi)$ , donde  $\varphi \in (-\pi/2, \pi/2)$ ,  $\theta \in (0, 2\pi)$ , son respectivamente, la latitud y la longitud geográficas del punto  $\mathbf{r}(\varphi, \theta)$ . Los vectores

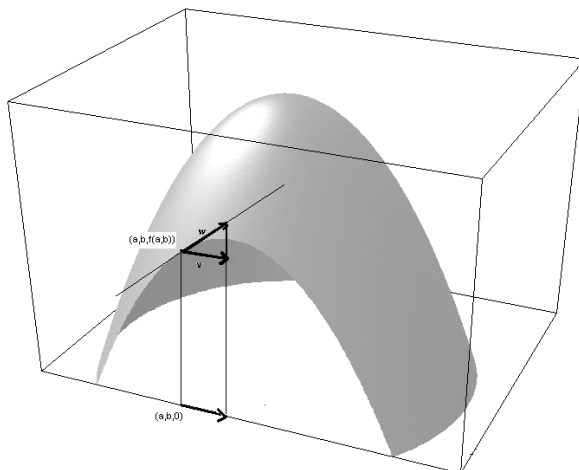
$$\frac{\partial \mathbf{r}}{\partial \varphi}(\varphi, \theta) = (-\sin \varphi \cos \theta, -\sin \varphi \sin \theta, \cos \varphi)$$

$$\frac{\partial \mathbf{r}}{\partial \theta}(\varphi, \theta) = (-\cos \varphi \sin \theta, \cos \varphi \cos \theta, 0)$$

son tangentes, respectivamente, al meridiano y al paralelo de la esfera que pasa por el punto  $\mathbf{r}(\varphi, \theta)$ . Pronto podremos justificar, en este caso particular, que el conjunto de vectores tangentes a la esfera en el punto  $\mathbf{r}(\varphi, \theta)$  forma un espacio vectorial de dimensión 2, generado por los vectores  $\{D_1\mathbf{r}(\varphi, \theta), D_2\mathbf{r}(\varphi, \theta)\}$ .

b) Consideremos ahora el caso de una función real de dos variables reales  $f : \Omega \rightarrow \mathbb{R}$ , definida en un abierto  $\Omega \subset \mathbb{R}^2$  (caso  $E = \mathbb{R}^2$ ,  $F = \mathbb{R}$ ). Su gráfica  $G(f) = \{(x, y, f(x, y)) : (x, y) \in \Omega\}$  es un subconjunto de  $\mathbb{R}^3$  que admite una representación tridimensional como una 'superficie', sobre el plano  $xy$ . Se trata de una 'superficie' de un tipo muy especial, pues cada recta paralela al eje  $z$ , trazada por un punto  $(x, y, 0)$ , con  $(x, y) \in \Omega$  la corta solamente en el punto  $(x, y, f(x, y))$ . Para un punto genérico  $\mathbf{p} = (x, y, z) \in G(f)$ , las dos primeras coordenadas,  $x, y$ , son variables independientes que recorren  $\Omega$ , mientras que la tercera  $z$ , depende de ellas a través de la función  $z = f(x, y)$ . Estas 'superficies' son un caso particular de las consideradas en a) pues admiten la parametrización  $\varphi(x, y) = (x, y, f(x, y))$ .

No obstante merece la pena ver, en este caso, la interpretación geométrica de la derivada direccional: Si  $\mathbf{u} \in \mathbb{R}^2$  es unitario,  $\|\mathbf{u}\| = 1$ , ahora la derivada  $D_{\mathbf{u}}f(a, b)$  es un número que representa la pendiente de la tangente en el punto  $\mathbf{p} = (a, b, f(a, b))$  a la curva que se obtiene como intersección de la gráfica  $G(f)$  con el plano que pasa por  $(a, b, f(a, b))$  y  $(a, b, 0)$  y es paralelo al vector  $\mathbf{v} = (u_1, u_2, 0)$ . (Obsérvese que esta recta tangente tiene la dirección del vector  $\mathbf{w} = (u_1, u_2, D_{\mathbf{u}}f(a, b))$  pues en el triángulo rectángulo que tiene como cateto horizontal el vector unitario  $\mathbf{v}$ , y la hipotenusa en la recta tangente, la longitud del cateto vertical es la pendiente de la tangente, es decir  $D_{\mathbf{u}}f(a, b)$ ).



En particular el número  $D_1f(a, b)$  (resp.  $D_2f(a, b)$ ) da la pendiente de la tangente a la curva que se obtiene cortando la gráfica de  $f$  con el plano que pasa por  $\mathbf{p}$  y es paralelo al plano  $xz$  (resp. al plano  $yz$ ).

Si se parametriza la 'superficie'  $G(f)$  en la forma estandar,  $G(f) = \varphi(\Omega)$ , con  $\varphi(x, y) = (x, y, f(x, y))$ , se tiene  $\varphi_1(x, y) = x$ ,  $\varphi_2(x, y) = y$ ,  $\varphi_3(x, y) = f(x, y)$ , luego  $D_{\mathbf{u}}\varphi_1(a, b) = u_1$ ,  $D_{\mathbf{u}}\varphi_2(a, b) = u_2$ , y según 5.2 volvemos a obtener que  $D_{\mathbf{u}}\varphi(a, b) = (u_1, u_2, D_{\mathbf{u}}f(a, b))$  es un vector tangente a  $G(f)$  en el punto  $\mathbf{p} = (a, b, f(a, b))$ .

El siguiente ejemplo muestra una función  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  que no es continua en  $(0, 0)$  pero existen las derivadas  $D_{\mathbf{v}}f(0, 0)$  según todos los vectores  $\mathbf{v} \in \mathbb{R}^2$ . Así se pone de manifiesto que la noción elemental de derivada según un vector es insuficiente para desarrollar con ella un cálculo diferencial satisfactorio similar al ya conocido para funciones reales de una sola variable.

**Ejemplo 5.4** La función  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  definida por

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

no es continua en  $(0, 0)$ , y para todo  $\mathbf{v} \in \mathbb{R}^2$  existe  $D_{\mathbf{v}}f(0, 0)$ . Además, las derivadas parciales  $D_1f(x, y)$ ,  $D_2f(x, y)$  no están acotadas en ningún entorno de  $(0, 0)$ .

DEM: En  $(0, 0)$  existe la derivada según cualquier vector  $\mathbf{v} = (v_1, v_2)$  pues en el caso no trivial  $(v_1, v_2) \neq (0, 0)$ , es evidente que

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

tiene límite cuando  $t \rightarrow 0$  (si  $v_1 = 0$  el límite es 0, y si  $v_1 \neq 0$  el límite es  $v_2^2/v_1$ ). Por otra parte, como  $f(0, 0) = 0$  y en todo entorno de  $(0, 0)$  hay puntos (de la parábola  $x = y^2$ ) donde el valor de  $f$  es  $1/2$  se sigue que  $f$  no es continua en  $(0, 0)$ .

Para  $y \in \mathbb{R}$ , fijo la función de variable real  $x \rightarrow f(x, y)$  es derivable en todo  $x \in \mathbb{R}$ , y la derivada vale

$$D_1 f(x, y) = \frac{y^6 - x^2 y^2}{(x^2 + y^4)^2}$$

Para  $x \in \mathbb{R}$ , fijo la función de variable real  $y \rightarrow f(x, y)$  es derivable en todo  $y \in \mathbb{R}$ , y la derivada vale

$$D_2 f(x, y) = \frac{2x^3 y - 2xy^5}{(x^2 + y^4)^2}$$

En el punto  $(0, 0)$ , usando la definición es claro que existen las derivadas parciales

$$D_1 f(0, 0) = \lim_{h \rightarrow 0} [f(h, 0) - f(0, 0)]/h = 0; D_2 f(0, 0) = \lim_{k \rightarrow 0} [f(0, k) - f(0, 0)]/k = 0$$

$D_1 f(x, y)$  no está acotada en ningún entorno de  $(0, 0)$  porque  $D_1 f(0, y) = 1/y^2$ .  $D_2 f(x, y)$  no está acotada en ningún entorno de  $(0, 0)$  porque en los puntos de parábola  $x = my^2$ , con  $(m \neq 1)$ , toma los valores  $D_1 f(my^2, y) = c(m)/y$ , donde  $c(m) = 2(m^3 - m)/(m^2 + 1)^2 \neq 0$ . ■

Aunque la sola consideración de las derivadas parciales es insuficiente para desarrollar un cálculo diferencial satisfactorio hay algunas aplicaciones interesantes de esta noción que recogemos a continuación.

**Algunas aplicaciones de la noción de derivada parcial.** La función considerada en el ejercicio 5.4 también sirve para mostrar que la acotación de las derivadas parciales es una hipótesis esencial para la validez del siguiente resultado, que proporciona una condición suficiente para la continuidad.

**Proposición 5.5** *Sea  $\Omega \subset \mathbb{R}^n$  abierto y  $\mathbf{f} : \Omega \rightarrow F$  una aplicación tal que en cada punto  $\mathbf{x} \in \Omega$  existen las derivadas parciales  $D_1 \mathbf{f}(\mathbf{x}), D_2 \mathbf{f}(\mathbf{x}), \dots, D_n \mathbf{f}(\mathbf{x})$ . Si las derivadas parciales  $D_k \mathbf{f}(\mathbf{x}), 1 \leq k \leq n$  están acotadas entonces  $\mathbf{f}$  es continua. Si además  $\Omega = \mathbb{R}^n$  entonces  $\mathbf{f}$  es uniformemente continua.*

DEM: a) Por hipótesis existe  $M > 0$  tal que  $\|D_k \mathbf{f}(\mathbf{x})\| \leq M$  para todo  $\mathbf{x} \in \Omega$  y todo  $k \in \{1, 2, \dots, n\}$ . Dado un punto  $\mathbf{a} \in \Omega$ , sea  $r > 0$  tal que  $B(\mathbf{a}, r) \subset \Omega$ . Para cada  $\mathbf{x} \in B(\mathbf{a}, r)$  el segmento de extremos  $(x_1, x_2, \dots, x_n), (a_1, x_2, \dots, x_n)$  está contenido en  $B(\mathbf{a}, r)$  y aplicando el corolario 4.8 a la función de una variable real  $\varphi_1(t) = \mathbf{f}(t, x_2, x_3, \dots, x_n)$  en el segmento  $[a_1, x_1]$  se obtiene

$$(1) \quad \|\mathbf{f}(x_1, x_2, \dots, x_n) - \mathbf{f}(a_1, x_2, \dots, x_n)\| \leq M|x_1 - a_1|$$

El segmento de extremos  $(a_1, x_2, \dots, x_n), (a_1, a_2, x_3, \dots, x_n)$  también está contenido en  $B(\mathbf{a}, r)$ , y aplicando el teorema del incremento finito a la función de una variable real  $\varphi_2(t) = \mathbf{f}(a_1, t, x_3, \dots, x_n)$  en el segmento  $[a_2, x_2]$  se obtiene

$$(2) \quad \|\mathbf{f}(a_1, x_2, \dots, x_n) - \mathbf{f}(a_1, a_2, \dots, x_n)\| \leq M|x_2 - a_2|$$

Análogamente se acota cada sumando de la suma telescópica

$$\begin{aligned} \mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a}) &= \mathbf{f}(x_1, x_2, x_3, \dots, x_n) - \mathbf{f}(a_1, x_2, x_3, \dots, x_n) + \\ &\quad \mathbf{f}(a_1, x_2, x_3, \dots, x_n) - \mathbf{f}(a_1, a_2, x_3, \dots, x_n) + \\ &\quad \dots \dots \dots + \\ &\quad \mathbf{f}(a_1, a_2, \dots, a_{n-1}, x_n) - \mathbf{f}(a_1, a_2, a_3, \dots, a_n) \end{aligned}$$

y utilizando la desigualdad triangular se obtiene

$$\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a})\| \leq M \sum_{j=1}^n |x_j - a_j| = M \|\mathbf{x} - \mathbf{a}\|_1$$

de donde se sigue que  $\mathbf{f}$  es continua en  $\mathbf{a}$ .

b) Cuando  $\Omega = \mathbb{R}^n$ , con el mismo razonamiento se obtiene que para todo par de puntos  $\mathbf{a}, \mathbf{x} \in \mathbb{R}^n$  se verifica

$$\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a})\| \leq M \|\mathbf{x} - \mathbf{a}\|_1$$

y de aquí se sigue que  $f$  es uniformemente continua. ■

NOTA: El resultado de la proposición 5.5 se puede localizar en un punto  $\mathbf{a}$ : Si sólo se supone que en una bola  $B(\mathbf{a}, r) \subset \Omega$  existen todas las derivadas parciales y están acotadas, entonces  $\mathbf{f}$  es continua en  $B(\mathbf{a}, r)$ , y en particular es continua en  $\mathbf{a}$ .

Por otra parte, la última afirmación de la proposición 5.5 no es cierta cuando  $\Omega \neq \mathbb{R}^n$ : En el ejercicio 5.35 veremos un ejemplo de una función  $g : \Omega \rightarrow \mathbb{R}$ , que no es uniformemente continua en  $\Omega = \mathbb{R}^2 \setminus \{(x, 0) : x \leq 0\}$ , pero tiene derivadas parciales acotadas en  $\Omega$ .

Para el siguiente resultado se usa la noción de conjunto conexo que se estudia en los cursos de Topología General.

**Proposición 5.6** *Sea  $\Omega \subset \mathbb{R}^n$  un abierto conexo y  $\mathbf{f} : \Omega \rightarrow F$  tal que en cada  $\mathbf{x} \in \Omega$  existen y son nulas las derivadas parciales  $D_1\mathbf{f}(\mathbf{x}), D_2\mathbf{f}(\mathbf{x}), \dots, D_n\mathbf{f}(\mathbf{x})$ . Entonces  $\mathbf{f}$  es constante.*

DEM: En virtud de la proposición 5.5  $\mathbf{f}$  es continua. Más aún, la demostración de esta proposición, con  $M = 0$ , conduce a que  $\mathbf{f}$  es constante en cada bola  $B(\mathbf{a}, r) \subset \Omega$ .

El conjunto no vacío  $A = \{\mathbf{x} \in \Omega : \mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{a})\}$  es cerrado para la topología relativa de  $\Omega$  porque  $\mathbf{f}$  es continua. Basta ver que  $A$  también es un subconjunto abierto del abierto conexo  $\Omega$  para concluir que  $A = \Omega$ , y con ello que  $\mathbf{f}$  es constante. Efectivamente, si  $\mathbf{b} \in A$  existe  $\rho > 0$  tal que  $B(\mathbf{b}, \rho) \subset \Omega$ , y en esta bola  $\mathbf{f}$  es constante, luego  $\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{b})$  para todo  $\mathbf{x} \in B(\mathbf{b}, \rho)$ . Como  $\mathbf{b} \in A$  se cumple que  $\mathbf{f}(\mathbf{b}) = \mathbf{f}(\mathbf{a})$ , luego  $B(\mathbf{b}, \rho) \subset A$  y con esto queda establecido que  $A$  es abierto. ■

Como aplicación de la proposición 5.6 mostramos en el siguiente ejemplo una alternativa para definir en  $\Omega = \mathbb{R}^2 \setminus \{(x, 0) : x \leq 0\}$  la función argumento principal que asigna a cada  $(x, y) \in \Omega$  el único argumento  $\theta = f(x, y)$  del número complejo  $x + iy$  que cumple  $-\pi < \theta < \pi$ . La definición utiliza la determinación principal  $\text{Arc tg} : \mathbb{R} \rightarrow (-\pi/2, \pi/2)$  de la función multivaluada  $\arctan$ .



**Ejemplo 5.7** En  $\Omega_1 = \{(x, y) : y > 0\}$ ,  $\Omega_2 = \{(x, y) : x > 0\}$ ,  $\Omega_3 = \{(x, y) : y < 0\}$  se definen, respectivamente, las funciones

$$f_1(x, y) = \text{Arc tg}(-x/y) + \pi/2$$

$$f_2(x, y) = \text{Arc tg}(y/x),$$

$$f_3(x, y) = \text{Arc tg}(-x/y) - \pi/2$$

Entonces en el abierto  $\Omega = \mathbb{R}^2 \setminus \{(x, 0) : x \leq 0\} = \Omega_1 \cup \Omega_2 \cup \Omega_3$ , hay definida una función  $f : \Omega \rightarrow \mathbb{R}$ , con derivadas parciales continuas, tal que  $f|_{\Omega_i} = f_i$ .

DEM: Para  $i = 1, 2, 3$  y todo  $(x, y) \in \Omega_i$ , con un cálculo rutinario se obtiene

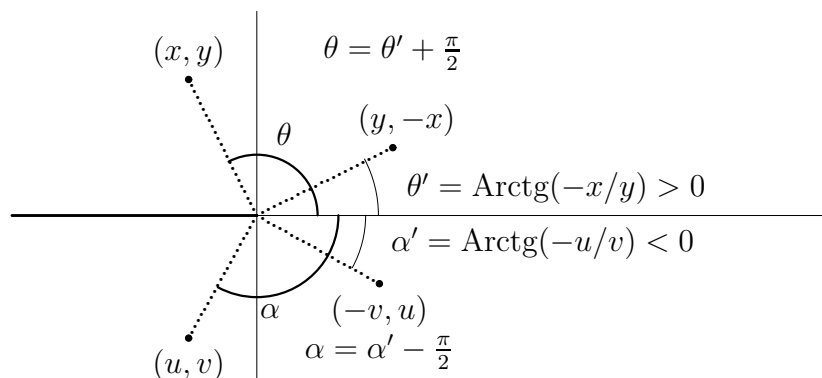
$$D_1 f_i(x, y) = \frac{-y}{x^2 + y^2}; \quad D_2 f_i(x, y) = \frac{x}{x^2 + y^2}$$

Según 5.6,  $f_1 - f_2$  es constante en el abierto conexo  $\Omega_1 \cap \Omega_2 = \{(x, y); x > 0, y > 0\}$  y su valor constante es  $f_1(1, 1) - f_2(1, 1) = (-\pi/4 + \pi/2) - \pi/4 = 0$ .

Análogamente,  $f_2 - f_3$  es constante en  $\Omega_2 \cap \Omega_3 = \{(x, y); x > 0, y < 0\}$  y su valor constante es  $f_2(1, -1) - f_3(1, -1) = -\pi/4 - (\pi/4 - \pi/2) = 0$ .

En definitiva,  $f_1$  y  $f_2$  coinciden en la intersección de sus dominios, y lo mismo ocurre con  $f_2$  y  $f_3$  (los dominios de  $f_1$  y  $f_3$  son disjuntos). Esto permite definir en  $\Omega = \cup_{j=1}^3 \Omega_j = \mathbb{R}^2 \setminus \{(x, 0) : x \leq 0\}$  una función  $f$  cuya restricción a cada  $\Omega_j$  sea  $f_j$ . Las derivadas parciales de esta función siguen siendo

$$D_1 f(x, y) = \frac{-y}{x^2 + y^2}; \quad D_2 f(x, y) = \frac{x}{x^2 + y^2}$$



Según la figura, se tiene

$$0 < \theta = f(x, y) = \text{Arc tg}(-x/y) + \pi/2 = \theta' + \pi/2 < \pi$$

$$-\pi < \alpha = f(u, v) = \text{Arc tg}(-u/v) - \pi/2 = \alpha' - \pi/2 < 0$$

La función  $f : \Omega \rightarrow \mathbb{R}$  tiene un significado geométrico muy claro:  $\theta = f(x, y)$  es el único argumento del número complejo  $(x + iy)$  que cumple  $-\pi < \theta < \pi$ . ■

Una de las aplicaciones más notables de la noción de derivada parcial la proporciona la condición necesaria de extremo relativo para funciones reales de varias variables reales:

**Definición 5.8** Una función  $f : \Omega \rightarrow \mathbb{R}$ , definida en un abierto  $\Omega \subset \mathbb{R}^n$ , presenta en  $\mathbf{a} \in \Omega$  un mínimo relativo (resp. mínimo relativo estricto) si existe una bola  $B(\mathbf{a}, r) \subset \Omega$ , tal que cuando  $\|\mathbf{x} - \mathbf{a}\| < r$  (resp.  $0 < \|\mathbf{x} - \mathbf{a}\| < r$ ) se verifica  $f(\mathbf{a}) \leq f(\mathbf{x})$  (resp.  $f(\mathbf{a}) < f(\mathbf{x})$ ).

La definición de máximo relativo (resp. máximo relativo estricto) es análoga pero usando la desigualdad  $f(\mathbf{a}) \geq f(\mathbf{x})$  (resp.  $f(\mathbf{a}) > f(\mathbf{x})$ ).

Si  $f$  presenta en  $\mathbf{a}$  un máximo o un mínimo relativo se dice que  $f$  presenta en  $\mathbf{a}$  un extremo relativo.

Es obvio que  $f$  presenta en  $\mathbf{a} \in \Omega$  un máximo relativo (resp. máximo relativo estricto) si y sólo si  $-f$  presenta en  $\mathbf{a} \in \Omega$  un mínimo relativo (resp. mínimo relativo estricto). A veces, en estas definiciones, en vez del calificativo "relativo" se emplea el de "local", y se habla de máximos locales, mínimos locales, extremos locales, etc.

**Proposición 5.9** Si  $f : \Omega \rightarrow \mathbb{R}$ , definida en un abierto  $\Omega \subset \mathbb{R}^n$ , presenta en  $\mathbf{a} \in \Omega$  un extremo local y existe la derivada  $D_{\mathbf{u}}f(\mathbf{a})$  entonces  $D_{\mathbf{u}}f(\mathbf{a}) = 0$ . En particular, si existen las derivadas parciales  $D_1f(\mathbf{a}), D_2f(\mathbf{a}), \dots, D_nf(\mathbf{a})$ , todas son nulas.

DEM: La función real de variable real  $\varphi(t) = f(\mathbf{a} + t\mathbf{u})$  está definida en un entorno de  $t = 0$  donde es derivable:

$$\varphi'(0) = \lim_{t \rightarrow 0} \frac{\varphi(t) - \varphi(0)}{t} = \lim_{t \rightarrow 0} \frac{f(\mathbf{a} + t\mathbf{u}) - f(\mathbf{a})}{t} = D_{\mathbf{u}}f(\mathbf{a})$$

Si  $f$  presenta en  $\mathbf{a}$  un máximo (resp. mínimo) relativo entonces  $\varphi$  presenta en 0 un máximo (resp. mínimo) relativo y por lo tanto  $D_{\mathbf{u}}f(\mathbf{a}) = \varphi'(0) = 0$ . ■

Si  $f : \Omega \rightarrow \mathbb{R}$  posee derivadas parciales en todo  $\mathbf{x} \in \Omega$ , se llaman *puntos estacionarios* de  $f$ , a las soluciones (si las hay) del sistema de  $n$  ecuaciones

$$D_1f(x_1, x_2, \dots, x_n) = 0, \quad D_2f(x_1, x_2, \dots, x_n) = 0, \quad \dots, \quad D_nf(x_1, x_2, \dots, x_n) = 0$$

Según la proposición 5.9 los puntos estacionarios de las funciones que tienen derivadas parciales en todos los puntos de su dominio son los puntos donde estas funciones pueden presentar extremos relativos. Con el siguiente ejemplo se muestra que puede haber puntos que sean estacionarios pero no sean puntos de extremo local

**Ejemplo 5.10** El origen  $(0, 0)$  es punto estacionario de  $f(x, y) = (y - x^2)(y - 2x^2)$ , pero  $f$  no presenta en  $(0, 0)$  un extremo relativo:

DEM: Obsérvese que  $f(0, 0) = 0$  y que en todo entorno de  $(0, 0)$  hay puntos  $(x_1, y_1)$ ,  $(x_2, y_2)$  tales que  $f(x_1, y_1) < f(0, 0) < f(x_2, y_2)$ . (Basta tomar  $x_1^2 < y_1 < 2x_1^2$ ,  $2x_2^2 < y_2$ ). ■

*Aplicaciones al cálculo de extremos absolutos.*

A veces ocurre que por algún razonamiento de tipo topológico tenemos la certeza de que una función continua  $f : \Omega \rightarrow \mathbb{R}$ , definida en un abierto  $\Omega \subset \mathbb{R}^n$ , alcanza un máximo absoluto en algún punto  $\mathbf{a} \in \Omega$ . Si  $f$ , además de ser continua, tiene derivadas parciales en todo  $\mathbf{x} \in \Omega$  entonces  $\mathbf{a} \in E$ , donde  $E$  es el conjunto de los puntos estacionarios de  $f$ . Si este conjunto es finito y se puede calcular explícitamente, quedará resuelto del problema de obtener el máximo absoluto de  $f$ , que vendrá dado por  $\max\{f(\mathbf{e}) : \mathbf{e} \in E\}$ .

Cuando alguna de las derivadas parciales de  $f$  no existe en algún punto, y también es finito el conjunto  $H \subset \Omega$  formado por los puntos donde esto ocurre, para obtener el máximo absoluto basta calcular  $\max\{f(\mathbf{e}) : \mathbf{e} \in E \cup H\}$ . El mismo método se puede seguir para calcular el mínimo absoluto de una función continua, cuando se sabe de antemano que la función alcanza en algún punto de su dominio un mínimo absoluto. En los ejercicios resueltos 5.37, 5.38, se pueden ver ejemplos típicos del uso de este método.

Los ejercicios resueltos 5.39, 5.40 son ejemplos típicos de problemas de extremos condicionados: Extremos de una función cuyas variables están sometidas a una condición (llamada condición de ligadura). Aunque el tratamiento general de este tipo de problemas se expone en el capítulo 9, aquí se muestra como algunos problemas de este tipo se pueden resolver con los recursos elementales de este capítulo.

Con los recursos disponibles en este momento también se puede abordar el problema de calcular los extremos absolutos de una función continua de dos variables sobre un compacto  $K \subset \mathbb{R}^2$ , siempre que la función tenga derivadas parciales en los puntos del interior de  $K$  y la frontera  $\partial K$  esté formada por una cantidad finita de curvas con una parametrización sencilla.

La estrategia para calcular los extremos absolutos es la siguiente: Un resultado general de topología asegura que  $f$  alcanza en  $K$  sus extremos absolutos. Un punto  $(a, b)$  donde  $f|_K$  alcanza un extremo absoluto puede estar en el interior de  $K$  o en su frontera. En el primer caso  $(a, b)$  será un punto estacionario de  $f$ , es decir, será solución de las ecuaciones  $D_1f(x, y) = 0$ ,  $D_2f(x, y) = 0$ . Si sabemos resolver estas ecuaciones y en el interior de  $K$  sólo hay un conjunto finito  $S$  de soluciones, cada punto de  $S$  será candidato a punto de extremo absoluto.

En el segundo caso, cuando  $(a, b) \in \partial K$ , es claro que  $f|_{\partial K}$  alcanza un extremo absoluto en  $(a, b)$ . Si la frontera  $\partial K$  se compone de varios arcos de curva con parametrizaciones sencillas, usando estas parametrizaciones, el cálculo de los extremos absolutos de  $f|_{\partial K}$  se transforma en varios problemas de optimización unidimensional que se pueden abordar con las técnicas usuales del cálculo de funciones de una variable real. Los tres casos considerados en el ejercicio 5.41 se han resuelto siguiendo esta estrategia.

## 5.2. Aplicaciones diferenciables

Para motivar la noción de diferencial de una aplicación volvemos a considerar el problema de formular una definición razonable de plano tangente a una 'superficie' explícita  $z = f(x, y)$  en un punto  $\mathbf{p} = (a, b, f(a, b))$  de la misma.

Para lograr una definición razonable se debe cumplir que para cada  $\mathbf{v} \in \mathbb{R}^2$  exista la derivada  $D_{\mathbf{v}}f(a, b)$ , pero ya hemos visto que este requisito es demasiado débil pues ni siquiera garantiza la continuidad de la función en el punto  $\mathbf{a}$  (véase el ejercicio 5.4) Incluso cuando se supone que  $f$  continua, no se puede asegurar que el conjunto de vectores tangentes a  $G(f)$  en  $\mathbf{p} = (a, b, f(a, b))$  forme un plano, como ocurre en la función que interviene en el siguiente ejercicio

**Ejemplo 5.11** La función  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  definida por

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

es continua en  $(0, 0)$  y para todo  $\mathbf{v} \in \mathbb{R}^2$  existe  $D_{\mathbf{v}}f(0, 0) = f(\mathbf{v})$ , y en el punto  $\mathbf{p} = (0, 0, 0)$  de la gráfica  $M = G(f)$  se cumple

$$T_{\mathbf{p}}(M) = \{(v_1, v_2, f(v_1, v_2)) : (v_1, v_2) \in \mathbb{R}^2\}$$

DEM: La continuidad de  $f$  en  $(0, 0)$  es consecuencia inmediata de la desigualdad  $|f(x, y)| \leq |y|$  que se cumple en todo punto  $(x, y) \in \mathbb{R}^2$ .

En  $(0, 0)$  existe la derivada según cualquier vector  $\mathbf{v} = (v_1, v_2)$  pues en el caso no trivial  $(v_1, v_2) \neq (0, 0)$ , es evidente que

$$\frac{f((0, 0) + t(v_1, v_2)) - f(0, 0)}{t} = \frac{v_1^2 v_2}{v_1^2 + v_2^2}$$

luego existe  $D_{\mathbf{v}}f(0, 0) = f(\mathbf{v})$ .

Para establecer la igualdad  $T_{\mathbf{p}}(M) = \{(v_1, v_2, f(v_1, v_2)) : (v_1, v_2) \in \mathbb{R}^2\}$  basta demostrar que todo  $\mathbf{u} = (u_1, u_2, u_3) \in T_{\mathbf{p}}(M)$  satisface  $u_3 = f(u_1, u_2)$ , es decir  $u_3(u_1^2 + u_2^2) = u_1^2 u_2$ . Según la definición de vector tangente existe  $\gamma : (-\epsilon, \epsilon) \rightarrow \mathbb{R}^3$ ,  $\gamma = (\gamma_1, \gamma_2, \gamma_3)$ , derivable en  $t = 0$  tal que  $\gamma_3(t)(\gamma_1(t)^2 + \gamma_2(t)^2) = \gamma_1(t)^2 \gamma_2(t)$ ,  $\gamma_j(0) = 0$ , y  $\gamma_j'(0) = u_j$  para  $1 \leq j \leq 3$ . En virtud de la definición de derivada

$$\gamma_j(t) = t[u_j + \epsilon_j(t)], \quad \text{donde } \lim_{t \rightarrow 0} \epsilon_j(t) = 0$$

Se sigue que

$$(u_3 + \epsilon_3(t))[(u_1 + \epsilon_1(t))^2 + (u_2 + \epsilon_2(t))^2] = (u_1 + \epsilon_1(t))^2 (u_2 + \epsilon_2(t))$$

y pasando al límite cuando  $t \rightarrow 0$  se obtiene que  $u_3(u_1^2 + u_2^2) = u_1^2 u_2$ . ■

Dejando aparte las consideraciones geométricas, para la función del ejercicio 5.11 la derivada según vectores  $\mathbf{v} \rightarrow D_{\mathbf{v}}f(\mathbf{a}) = f(\mathbf{v})$  no proporciona nada nuevo. No se consigue así otra función más sencilla (lineal) que aproxime localmente a  $f$ .

Vemos así que para lograr una definición satisfactoria de plano tangente a la gráfica  $G(f)$  en un punto  $\mathbf{p} = (a, b, f(a, b)) \in G(f)$  debemos conseguir que  $\{(\mathbf{v}, D_{\mathbf{v}}f(\mathbf{a})) : \mathbf{v} \in \mathbb{R}^2\}$  sea un subespacio vectorial de dimensión 2 y que se cumpla la igualdad

$$\{(\mathbf{v}, D_{\mathbf{v}}f(\mathbf{a})) : \mathbf{v} \in \mathbb{R}^2\} = T_{\mathbf{p}}(G(f))$$

Es claro que una condición necesaria y suficiente para ello es que la aplicación  $\mathbf{v} \rightarrow D_{\mathbf{v}}f(a, b)$  sea lineal. En este caso,  $T_{\mathbf{p}}(G(f)) \subset \mathbb{R}^3$  es un candidato a plano tangente. Sin embargo, para la definición satisfactoria de aplicación diferenciable en un punto  $\mathbf{a}$  hay que exigir algo más que la continuidad de  $f$  en  $\mathbf{a}$  y la existencia y linealidad de la aplicación  $\mathbf{v} \rightarrow D_{\mathbf{v}}f(\mathbf{a})$  (véanse los ejemplos 5.15 y 5.42).

**Definición 5.12** Sea  $\mathbf{f} : \Omega \rightarrow F$  definida en un abierto  $\Omega$  de un espacio normado  $(E, \|\cdot\|)$ , con valores en otro espacio normado  $(F, \|\cdot\|)$  y  $\mathbf{a} \in \Omega$ . Se dice que  $\mathbf{f}$  es diferenciable en  $\mathbf{a}$  cuando existe una aplicación lineal continua  $L : E \rightarrow F$  que verifica

$$\lim_{\mathbf{h} \rightarrow 0} \frac{\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a}) - L(\mathbf{h})}{\|\mathbf{h}\|} = 0$$

Si  $\mathbf{f}$  es diferenciable en  $\mathbf{a}$  y se cambian las normas de  $E$  y  $F$  por otras equivalentes, es fácil ver que  $\mathbf{f}$  sigue siendo diferenciable en  $\mathbf{a}$ , con la misma diferencial.

Obsérvese que hay una única aplicación lineal  $L$  que cumpla la condición de la definición anterior: Si  $T \in \mathcal{L}(E, F)$  es otra aplicación lineal que cumple

$$\lim_{\mathbf{h} \rightarrow 0} \frac{\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a}) - T(\mathbf{h})}{\|\mathbf{h}\|} = 0 \quad \text{resulta} \quad \lim_{\mathbf{h} \rightarrow 0} \frac{T(\mathbf{h}) - L(\mathbf{h})}{\|\mathbf{h}\|} = 0$$

y considerando el límite a través de  $\{t\mathbf{v} : t > 0\}$  con  $\mathbf{v} \in E$ ,  $\mathbf{v} \neq 0$ , se obtiene

$$0 = \lim_{t \rightarrow 0^+} \frac{T(t\mathbf{v}) - L(t\mathbf{v})}{t\|\mathbf{v}\|} = \frac{T(\mathbf{v}) - L(\mathbf{v})}{\|\mathbf{v}\|} \quad \text{luego} \quad T(\mathbf{v}) = L(\mathbf{v}).$$

**Definición 5.13** En las condiciones de la definición 5.12, la única aplicación lineal continua  $L : E \rightarrow F$  que interviene en ella se llama diferencial de  $\mathbf{f}$  en el punto  $\mathbf{a}$  y se denota  $d\mathbf{f}(\mathbf{a})$ . Si  $\mathbf{f}$  es diferenciable en cada  $\mathbf{a} \in \Omega$  se dice que  $\mathbf{f}$  es diferenciable en  $\Omega$ . En este caso  $d\mathbf{f} : \Omega \rightarrow \mathcal{L}(E, F)$  es la aplicación que asigna a cada  $\mathbf{a} \in \Omega$  la diferencial  $d\mathbf{f}(\mathbf{a})$ .

La condición de que  $\mathbf{f}$  sea diferenciable en  $\mathbf{a}$  se expresa escribiendo

$$\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a}) - L(\mathbf{h}) = \|\mathbf{h}\| \boldsymbol{\rho}(\mathbf{h}) \quad \text{con} \quad \lim_{\mathbf{h} \rightarrow 0} \boldsymbol{\rho}(\mathbf{h}) = 0$$

El significado de la noción de aplicación diferenciable es el siguiente: La aplicación  $\mathbf{f}$  es diferenciable en  $\mathbf{a}$  cuando es posible obtener una aproximación lineal  $L(\mathbf{h})$  del incremento  $\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a})$  con la propiedad de que el error de la aproximación  $\boldsymbol{\alpha}(\mathbf{h}) = [\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a})] - L(\mathbf{h})$  sea pequeño frente a  $\mathbf{h}$ , en el sentido de que el error relativo  $\boldsymbol{\alpha}(\mathbf{h})/\|\mathbf{h}\|$  tiende hacia 0 cuando  $\mathbf{h}$  tiende hacia 0. Si  $\boldsymbol{\alpha}(\mathbf{h})$  es una función

de  $\mathbf{h} \in E$ , definida en un entorno de 0, se suele escribir  $\alpha(\mathbf{h}) = o(\|\mathbf{h}\|)$  para indicar que se cumple

$$\lim_{\mathbf{h} \rightarrow 0} \frac{\alpha(\mathbf{h})}{\|\mathbf{h}\|} = 0$$

Con esta notación la condición de que  $\mathbf{f}$  sea diferenciable en  $\mathbf{a}$  se expresa en la forma más breve

$$\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a}) - L(\mathbf{h}) = o(\|\mathbf{h}\|)$$

Si dos funciones  $\mathbf{f}$ ,  $\mathbf{g}$ , definidas en un entorno de  $\mathbf{a}$ , con  $\mathbf{f}(\mathbf{a}) = \mathbf{g}(\mathbf{a})$ , verifican la condición  $\mathbf{f}(\mathbf{x}) - \mathbf{g}(\mathbf{x}) = o(\|\mathbf{x} - \mathbf{a}\|)$  se suele decir que  $\mathbf{f}$  y  $\mathbf{g}$  presentan en  $\mathbf{a}$  una tangencia de primer orden, y también que  $\mathbf{g}$  es una aproximación local de primer orden de  $\mathbf{f}$  en el punto  $\mathbf{a}$ . Según esta terminología, cuando  $\mathbf{f}$  es diferenciable en  $\mathbf{a}$ , la aplicación lineal  $L(\mathbf{h})$  es una aproximación local de primer orden del incremento  $\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a})$  (en  $\mathbf{h} = 0$ ) y la aplicación afín  $\mathbf{g}(\mathbf{x}) = \mathbf{f}(\mathbf{a}) + L(\mathbf{x} - \mathbf{a})$  es una aproximación local de primer orden de  $\mathbf{f}$  en el punto  $\mathbf{a}$ .

*Comentarios sobre la definición de aplicación diferenciable.* Con la noción de diferencial en un punto se persigue (igual que con la derivada, en el caso de funciones de una sola variable) una aproximación local de la función mediante un polinomio de primer grado. En el caso de las funciones de dos variables la aproximación local será de la forma  $p(x, y) = Ax + By + C$ . Cuando esto se formula en términos de los incrementos  $h = x - a$ ,  $k = y - b$ , la diferencial proporciona una aplicación lineal  $L(h, k) = Ah + Bk$  que, en un entorno de  $(0, 0)$ , aproxima localmente al incremento  $f(a+h, b+k) - f(a, b)$ . La noción de aproximación local que interviene en la definición de función diferenciable 5.12 exige que la aproximación local de  $L(h, k) = Ah + Bk$  al incremento  $f(a+h, b+k) - f(a, b)$  sea uniforme en todas las direcciones. Esta noción es la que permite desarrollar un cálculo diferencial satisfactorio, análogo al de funciones de una variable. Entre otras cosas permitirá demostrar la existencia de plano tangente a la gráfica de una función diferenciable de dos variables reales.

Para pensar en una situación concreta, que termine de aclarar el significado de la diferenciabilidad sigamos considerando el caso de una función de dos variables reales  $z = f(x, y)$  que deseamos evaluar numéricamente en un punto  $(a, b)$  que sólo se puede manejar en forma aproximada (esto ocurrirá cuando  $a, b$  sean irracionales o procedan de medir magnitudes físicas). Supongamos además que sólo aceptamos resultados con error menor que  $\epsilon = 10^{-12}$  (el tamaño que ya no distingue nuestra calculadora) y que  $f$  es diferenciable en  $(a, b)$ . En este caso la aplicación lineal  $L = df(a, b) : \mathbb{R}^2 \rightarrow \mathbb{R}$  es de la forma  $L(h, k) = Ah + Bk$ .

Según la definición de límite existe  $r > 0$  tal que si  $\max\{|h|, |k|\} < r$  se cumple  $[f(a+h, b+k) - f(a, b) - L(h, k)] < \epsilon \max\{|h|, |k|\}$ . Es decir, para incrementos que cumplen  $\max\{|h|, |k|\} < r$ , si utilizamos  $f(a, b) + L(h, k)$  como una aproximación de  $f(a+h, b+k)$  estaremos seguros de que cometemos un error menor que  $10^{-12} \max\{|h|, |k|\}$ , que es prácticamente despreciable frente a  $\max\{|h|, |k|\}$ .

Por otra parte, si utilizamos un ordenador para visualizar la gráfica de  $z = f(x, y)$  y la de su aproximación local  $z = f(a, b) + A(x - a) + B(y - b)$  y hacemos zoom en el punto  $\mathbf{p} = (a, b, f(a, b))$  hasta que el entorno de  $\mathbf{p}$  de radio  $r$  ocupe toda la

pantalla, podremos ver que, dentro de este entorno, no se aprecia diferencia entre el trozo de plano y el trozo de la gráfica de  $f$  ya que el tamaño de la pantalla del ordenador es  $\simeq r$ , mientras que la diferencia de cotas entre un punto del plano y un punto de la gráfica, es menor que  $10^{-12}r$ .

### Ejemplos

- i) Si  $\mathbf{f}$  es constante, es obvio que  $\mathbf{f}$  es diferenciable en todo  $\mathbf{a} \in \Omega$  y su diferencial es la aplicación lineal nula  $\mathbf{0} \in \mathcal{L}(E, F)$ .
- ii) Si  $L : E \rightarrow F$  es lineal continua entonces  $L$  es diferenciable en todo punto  $\mathbf{x} \in E$  y su diferencial es constante:  $dL(\mathbf{x}) = L$  para todo  $\mathbf{x} \in E$ .

En particular, si  $E = \mathbb{R}^2$ ,  $F = \mathbb{R}$  y  $L(x, y) = Ax + By$ , para todo  $(a, b) \in \mathbb{R}^2$

$$dL(a, b)(h, k) = Ah + Bk$$

- iii) Cuando  $E = \mathbb{R}$ , es fácil ver que  $\mathbf{f}$  es derivable en  $a$  si y sólo si es diferenciable en  $a$  y en ese caso la diferencial es a la aplicación lineal  $d\mathbf{f}(a) : \mathbb{R} \rightarrow F$ , definida por  $d\mathbf{f}(a)(h) = \mathbf{f}'(a)h$ .

NOTA: En iii)  $\mathbf{f}'(a)$  es un vector y  $h$  un número, y aparece escrito  $\mathbf{f}'(a)h$  en lugar de  $h\mathbf{f}'(a)$ . Este convenio de notación para la ley externa de los espacios vectoriales es útil para lograr mayor analogía con las fórmulas habituales del cálculo diferencial de las funciones escalares.

**Proposición 5.14** *Si  $\mathbf{f} : \Omega \rightarrow F$  es diferenciable en  $\mathbf{a} \in \Omega$  se verifica*

a)  $\mathbf{f}$  es continua en  $\mathbf{a}$ .

b) Para cada  $\mathbf{v} \in E$  existe la derivada  $D_{\mathbf{v}}\mathbf{f}(\mathbf{a}) = d\mathbf{f}(\mathbf{a})\mathbf{v}$ .

DEM:

a) Como  $L = d\mathbf{f}(\mathbf{a})$  es una aplicación lineal continua la diferencia  $\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a}) = L(\mathbf{h}) + \|\mathbf{h}\| \boldsymbol{\rho}(\mathbf{h})$  tiende hacia  $\mathbf{0}$  cuando  $\mathbf{h} \rightarrow \mathbf{0}$ .

b) Para cada  $\mathbf{v} \in E$  se tiene

$$\frac{\mathbf{f}(\mathbf{a} + t\mathbf{v}) - \mathbf{f}(\mathbf{a})}{t} = \frac{L(t\mathbf{v})}{t} + \frac{\|t\mathbf{v}\|}{t} \boldsymbol{\rho}(t\mathbf{v}) = L(\mathbf{v}) + \frac{|t|}{t} \|\mathbf{v}\| \boldsymbol{\rho}(t\mathbf{v})$$

Como  $\lim_{t \rightarrow 0} \boldsymbol{\rho}(t\mathbf{v}) = 0$  se concluye que existe el límite

$$D_{\mathbf{v}}\mathbf{f}(\mathbf{a}) = \lim_{t \rightarrow 0} \frac{\mathbf{f}(\mathbf{a} + t\mathbf{v}) - \mathbf{f}(\mathbf{a})}{t} = L(\mathbf{v})$$

■

Cuando  $E = \mathbb{R}^n$ , dotado de cualquier norma (recuérdese que en  $\mathbb{R}^n$  todas las normas son equivalentes) si  $\mathbf{f} : \Omega \rightarrow F$  es diferenciable en  $\mathbf{a} \in \Omega$ , para cada vector  $\mathbf{v} = (v_1, v_2, \dots, v_n) \in \mathbb{R}^n$ , en virtud de la linealidad de la diferencial  $d\mathbf{f}(\mathbf{a})$  y de la proposición 5.14 el cálculo de la derivada  $D_{\mathbf{v}}\mathbf{f}(\mathbf{a}) = d\mathbf{f}(\mathbf{a})\mathbf{v}$  se reduce al de las derivadas parciales  $D_j\mathbf{f}(\mathbf{a})$  ya que

$$D_{\mathbf{v}}\mathbf{f}(\mathbf{a}) = d\mathbf{f}(\mathbf{a}) \left( \sum_{j=1}^n v_j \mathbf{e}_j \right) = \sum_{j=1}^n v_j d\mathbf{f}(\mathbf{a})\mathbf{e}_j = \sum_{j=1}^n v_j D_j\mathbf{f}(\mathbf{a})$$

OBSERVACIÓN: Si  $\mathbf{f}$  diferenciable en  $\mathbf{a}$ , según 5.14,  $\mathbf{f}$  es continua en  $\mathbf{a}$ , para cada  $\mathbf{v} \in E$  existe la derivada  $D_{\mathbf{v}}\mathbf{f}(\mathbf{a})$  y la aplicación  $\mathbf{v} \rightarrow D_{\mathbf{v}}\mathbf{f}(\mathbf{a})$ , es lineal continua. Es natural preguntarse si el recíproco es cierto, es decir si toda función  $\mathbf{f} : \Omega \rightarrow F$  continua en  $\mathbf{a}$  que tenga esta propiedad es diferenciable en  $\mathbf{a}$ . La respuesta negativa la proporciona la función  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  del siguiente ejemplo.

**Ejemplo 5.15** La función  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  definida por

$$f(x, y) = \frac{xy^3}{x^2 + y^4} \quad \text{si } (x, y) \neq (0, 0), \quad f(0, 0) = 0$$

es continua en  $(0, 0)$ , para cada  $\mathbf{v} \in \mathbb{R}^2$  existe la derivada  $D_{\mathbf{v}}f(0, 0) = 0$ , y por lo tanto la aplicación  $\mathbf{v} \rightarrow D_{\mathbf{v}}f(0, 0)$  es lineal, pero  $f$  no es diferenciable en  $(0, 0)$ .

SOLUCIÓN

Aplicando la desigualdad  $|ab| \leq a^2 + b^2$  con  $a = x$ ,  $b = y^2$  se obtiene

$$\left| \frac{xy^3}{x^2 + y^4} \right| \leq |y|$$

de donde se sigue que  $f$  es continua en  $(0, 0)$ .

Por otra parte, si  $\mathbf{v} = (v_1, v_2) \in \mathbb{R}^2 \setminus \{(0, 0)\}$  se verifica

$$\frac{f(tv_1, tv_2) - f(0, 0)}{t} = \frac{tv_1v_2^3}{v_1^2 + t^2v_2^4}$$

y pasando al límite cuando  $t \rightarrow 0$  se obtiene que  $D_{\mathbf{v}}f(0, 0) = 0$  para todo  $\mathbf{v} \in \mathbb{R}^2$

Para ver que  $f$  no es diferenciable en  $\mathbf{a} = (0, 0)$  podemos razonar por reducción al absurdo: Si lo fuese, en virtud de la proposición 5.14, su diferencial debería ser la aplicación lineal nula,  $L(\mathbf{h}) = 0$  para todo  $\mathbf{h} \in \mathbb{R}^2$ , pero esta aplicación lineal no cumple la condición de diferenciability: El cociente

$$\frac{\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a}) - L(\mathbf{h})}{\|\mathbf{h}\|} = \frac{\mathbf{f}(\mathbf{h})}{\|\mathbf{h}\|} = \frac{h_1h_2^3}{(h_1^2 + h_2^4)\sqrt{h_1^2 + h_2^2}}$$

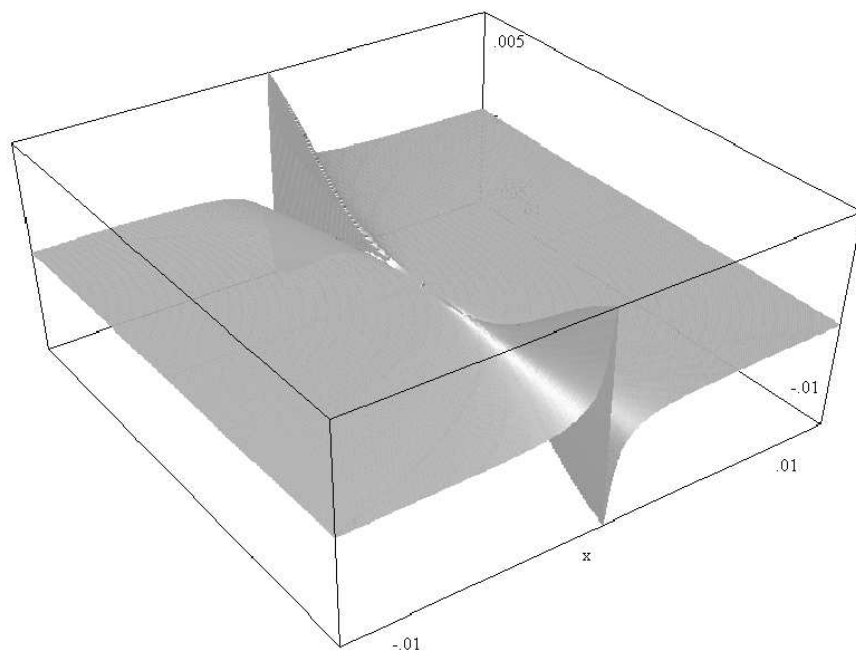
no tiende hacia 0 cuando  $(h_1, h_2) \rightarrow (0, 0)$  ya que a través de la parábola  $h_2 = t$ ,  $h_1 = t^2$  se tiene

$$\frac{\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a})}{\|\mathbf{h}\|} = \frac{t^5}{2t^4\sqrt{t^4 + t^2}} = \frac{t}{2|t|\sqrt{1 + t^2}}$$

que tiende hacia  $1/2$  (resp.  $-1/2$  cuando  $t \rightarrow 0+$  (resp.  $t \rightarrow 0-$ )). ■



La función que interviene en 5.15 proporciona un ejemplo interesante de una función continua tal que los vectores  $\{(\mathbf{v}, D_{\mathbf{v}}f(0,0)) : \mathbf{v} \in \mathbb{R}^2\}$  forman un plano, candidato a ser plano tangente, y sin embargo  $f$  no es diferenciable en  $(0,0)$ . Para ver gráficamente la razón de la no diferenciable de  $f$  en  $(0,0)$  se recomienda utilizar un programa de ordenador como DpGraph que permita visualizar la gráfica de  $f$  en un entorno de  $(0,0)$ .



Podremos apreciar que por muy potente que sea el zoom que apliquemos en un entorno de  $(0,0,0)$  siempre veremos un pliegue a lo largo del eje  $Oy$  y nunca lograremos que la gráfica de  $f$  se confunda con el plano  $z = 0$ . Esto se debe a que en este ejemplo la aplicación  $\mathbf{v} \rightarrow D_{\mathbf{v}}f(0,0)$ , aunque es lineal no proporciona una aproximación local del incremento  $f(v_1, v_2) - f(0,0)$  que sea uniforme en todas las direcciones. El ejercicio resuelto 5.42 tiene como finalidad aclarar lo que acabamos de comentar en relación con la función del ejemplo 5.15

**Condición suficiente de diferenciable.** Si una función de varias variables reales es diferenciable en un punto  $\mathbf{a}$  sabemos que su diferencial tiene que ser la aplicación lineal

$$\mathbf{v} \rightarrow d\mathbf{f}(\mathbf{a})\mathbf{v} = D_{\mathbf{v}}\mathbf{f}(\mathbf{a}) = \sum_{j=1}^n v_j D_j \mathbf{f}(\mathbf{a})$$

Esto nos indica la estrategia natural para estudiar la diferenciable, en un punto  $\mathbf{a}$ , de una función  $\mathbf{f}$  de varias variables reales: Se comienza estudiando la existencia de las derivadas parciales  $D_j \mathbf{f}(\mathbf{a})$ ,  $1 \leq j \leq n$ , y en el caso de que existan todas se considera el candidato natural para la diferencial  $d\mathbf{f}(\mathbf{a})$ , que es la aplicación lineal

$L : \mathbb{R}^n \rightarrow F$  definida por  $L(\mathbf{u}) = \sum_{j=1}^n u_j D_j \mathbf{f}(\mathbf{a})$ . Finalmente se estudia si el cociente  $\rho(\mathbf{h}) = [\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a}) - L(\mathbf{h})] / \|\mathbf{h}\|$  tiende hacia  $\mathbf{0}$  cuando  $\mathbf{h} \rightarrow \mathbf{0}$ .

Esta estrategia es la que se usa para demostrar el siguiente teorema que proporciona una condición suficiente de diferenciabilidad, muy útil en la práctica. En las hipótesis parece que desempeña un papel especial la última variable, pero su papel lo puede desempeñar cualquier otra

**Teorema 5.16** [Condición suficiente de diferenciabilidad] *Sea  $\mathbf{f} : \Omega \rightarrow F$  definida en un abierto  $\Omega \subset \mathbb{R}^n$  con valores en un espacio normado  $(F, \|\cdot\|)$ . Se supone que existe la derivada parcial  $D_n \mathbf{f}(\mathbf{a})$  y que en una bola  $B(\mathbf{a}, r) \subset \Omega$  existen las otras derivadas parciales  $D_1 \mathbf{f}(\mathbf{x}), D_2 \mathbf{f}(\mathbf{x}), \dots, D_{n-1} \mathbf{f}(\mathbf{x})$ , y son continuas en  $\mathbf{a}$ . Entonces  $\mathbf{f}$  es diferenciable en  $\mathbf{a}$ .*

DEM: Para simplificar la exposición usamos en  $\mathbb{R}^n$  la norma  $\|\cdot\|_\infty$ , de modo que

$$B(\mathbf{a}, r) = \{(x_1, x_2, \dots, x_n) : |x_j - a_j| < r, 1 \leq j \leq n\}$$

Como las funciones  $D_j \mathbf{f}(\mathbf{x})$ ,  $1 \leq j < n$  son continuas en  $\mathbf{a}$ , dado  $\epsilon > 0$ , existe  $0 < \delta < r$  tal que

$$(*) \quad \|\mathbf{y} - \mathbf{a}\|_\infty < \delta \Rightarrow \|D_j \mathbf{f}(\mathbf{y}) - D_j \mathbf{f}(\mathbf{a})\| < \epsilon \text{ para } 1 \leq j < n.$$

Para cada  $\mathbf{x} = (x_1, x_2, \dots, x_n) \in B(\mathbf{a}, \delta)$ , la función de variable real

$$\varphi_1(t) = \mathbf{f}(t, x_2, \dots, x_n) - (t - a_1) D_1 \mathbf{f}(a_1, a_2, \dots, a_n)$$

está definida y es derivable para  $|t - a_1| < \delta$ . En virtud de (\*), su derivada

$$\varphi_1'(t) = D_1 \mathbf{f}(t, x_2, x_3, \dots, x_n) - D_1 \mathbf{f}(a_1, a_2, \dots, a_n)$$

verifica

$$\|\varphi_1'(t)\| < \epsilon \quad \text{si} \quad |t - a_1| < \delta.$$

Aplicando el teorema del incremento finito a la función  $\varphi_1$  en el intervalo  $[a_1, x_1]$  se obtiene  $\|\varphi_1(x_1) - \varphi_1(a_1)\| \leq \epsilon |x_1 - a_1|$  es decir

$$(1) \quad \|\mathbf{f}(x_1, x_2, \dots, x_n) - \mathbf{f}(a_1, x_2, \dots, x_n) - (x_1 - a_1) D_1 \mathbf{f}(\mathbf{a})\| \leq \epsilon |x_1 - a_1|$$

Razonando de forma similar con

$$\varphi_2(t) = \mathbf{f}(a_1, t, x_3, \dots, x_n) - (t - a_2) D_2 \mathbf{f}(a_1, a_2, \dots, a_n)$$

se obtiene la desigualdad

$$(2) \quad \|\mathbf{f}(a_1, x_2, x_3, \dots, x_n) - \mathbf{f}(a_1, a_2, x_3, \dots, x_n) - (x_2 - a_2) D_2 \mathbf{f}(\mathbf{a})\| \leq \epsilon |x_2 - a_2|$$

Análogamente se razona con las restantes variables, hasta la variable  $x_{n-1}$

$$(n-1) \quad \|\mathbf{f}(a_1, \dots, a_{n-2}, x_{n-1}, x_n) - \mathbf{f}(a_1, \dots, a_{n-1}, x_n) - (x_{n-1} - a_{n-1}) D_{n-1} \mathbf{f}(\mathbf{a})\| \leq \epsilon |x_{n-1} - a_{n-1}|$$

Por último, usando la definición de  $D_n \mathbf{f}(\mathbf{a})$  como límite, podemos suponer que  $\delta > 0$  se ha elegido de modo que también se cumple

$$\left\| \frac{\mathbf{f}(a_1, \dots, a_{n-1}, x_n) - \mathbf{f}(a_1, \dots, a_n)}{x_n - a_n} - D_n \mathbf{f}(\mathbf{a}) \right\| < \epsilon$$

es decir

$$(n) \quad \|\mathbf{f}(a_1, \dots, a_{n-1}, x_n) - \mathbf{f}(a_1, \dots, a_n) - (x_n - a_n)D_n \mathbf{f}(\mathbf{a})\| < \epsilon|x_n - a_n|$$

Considerando la suma telescópica

$$\begin{aligned} \mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a}) - \sum_{j=1}^n (x_j - a_j)D_j \mathbf{f}(\mathbf{a}) = & \\ \mathbf{f}(x_1, x_2, \dots, x_n) - \mathbf{f}(a_1, x_2, \dots, x_n) - (x_1 - a_1)D_1 \mathbf{f}(\mathbf{a}) & + \\ \mathbf{f}(a_1, x_2, \dots, x_n) - \mathbf{f}(a_1, a_2, \dots, x_n) - (x_2 - a_2)D_2 \mathbf{f}(\mathbf{a}) & + \\ \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots & + \\ \mathbf{f}(a_1, \dots, a_{n-1}, x_n) - \mathbf{f}(a_1, a_2, \dots, a_n) - (x_n - a_n)D_n \mathbf{f}(\mathbf{a}) & \end{aligned}$$

y combinando las desigualdades (1), (2), (n-1), (n), con la desigualdad triangular se obtiene que cuando  $\|\mathbf{x} - \mathbf{a}\|_\infty < \delta$  se cumple

$$\left\| \mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a}) - \sum_{j=1}^n (x_j - a_j)D_j \mathbf{f}(\mathbf{a}) \right\| \leq \epsilon \sum_{j=1}^n |x_j - a_j| \leq n\epsilon \|\mathbf{x} - \mathbf{a}\|_\infty$$

Con esto queda demostrado que  $\mathbf{f}$  es diferenciable en  $\mathbf{a}$  y que su diferencial es la aplicación lineal  $L : \mathbb{R}^n \rightarrow F$  dada por  $L(\mathbf{h}) = \sum_{j=1}^n h_j D_j \mathbf{f}(\mathbf{a})$  ■

**Definición 5.17** Si  $\Omega \subset \mathbb{R}^n$  es abierto y  $\mathbf{f} : \Omega \rightarrow \mathbf{F}$  tiene derivadas parciales continuas en todo  $\mathbf{x} \in \Omega$  se dice que  $\mathbf{f}$  es de clase  $C^1$  en  $\Omega$  y se escribe  $\mathbf{f} \in C^1(\Omega, F)$ . A las aplicaciones de clase  $C^1$  también se les llama continuamente diferenciables.

En virtud del teorema 5.16 cada aplicación de clase  $C^1$  es diferenciable en todos los puntos de su dominio.

### 5.3. Las reglas del cálculo diferencial

En esta sección se exponen los resultados típicos sobre diferenciability de funciones que se obtienen combinando otras funciones diferenciables mediante las operaciones usuales (suma, producto, cociente, composición, etc). Al mismo tiempo obtendremos las reglas para obtener su diferencial en términos de las diferenciales de las funciones que intervienen en su definición. El resultado más notable es la regla de la cadena de la que se desprenden las reglas para el cálculo de derivadas parciales de funciones compuestas. La regla de la cadena también servirá para demostrar que cuando  $\mathbf{f}$  es una función de  $k$  variables reales diferenciable en  $\mathbf{a}$  entonces el conjunto de los vectores tangentes a su gráfica en el punto  $\mathbf{p} = (\mathbf{a}, \mathbf{f}(\mathbf{a}))$  forma un espacio vectorial de dimensión  $k$ . Comenzamos viendo que para funciones con valores en  $\mathbb{R}^m$  el estudio de su diferenciability se puede realizar componente a componente:

**Proposición 5.18** Sea  $\mathbf{f} : \Omega \rightarrow \mathbb{R}^m$  de componentes  $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_m(\mathbf{x}))$ , definida en un abierto  $\Omega \subset E$ . Entonces  $\mathbf{f}$  es diferenciable en  $\mathbf{a} \in \Omega$  si y sólo si todas sus componentes lo son y en este caso  $L = d\mathbf{f}(\mathbf{a}) : E \rightarrow \mathbb{R}^m$  es la aplicación lineal continua cuyas componentes  $L = (L_1, L_2, \dots, L_m)$  son las diferenciales de las componentes de  $\mathbf{f}$ , es decir  $df_j(\mathbf{a}) = L_j$ ,  $j = 1, 2, \dots, m$ .

DEM: Si  $\mathbf{f}$  es diferenciable en  $\mathbf{a}$  su diferencial  $L = d\mathbf{f}(\mathbf{a})$  verifica

$$\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a}) - L(\mathbf{h}) = \|\mathbf{h}\| \boldsymbol{\rho}(\mathbf{h}) \text{ con } \lim_{\mathbf{h} \rightarrow 0} \boldsymbol{\rho}(\mathbf{h}) = 0$$

y pasando a las componentes resulta que para cada  $j \in \{1, 2, \dots, m\}$  se cumple

$$f_j(\mathbf{a} + \mathbf{h}) - f_j(\mathbf{a}) - L_j(\mathbf{h}) = \|\mathbf{h}\| \rho_j(\mathbf{h}) \text{ con } \lim_{\mathbf{h} \rightarrow 0} \rho_j(\mathbf{h}) = 0$$

luego cada  $f_j$  es diferenciable en  $\mathbf{a}$  y  $df_j(\mathbf{a}) = L_j$ .

Recíprocamente, si cada  $f_j$  es diferenciable en  $\mathbf{a}$  y  $L_j = df_j(\mathbf{a})$ , entonces la aplicación lineal continua  $L : E \rightarrow F$  formada con estas componentes,  $L = (L_1, L_2, \dots, L_m)$  cumple

$$\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a}) - L(\mathbf{h}) = \|\mathbf{h}\| \boldsymbol{\rho}(\mathbf{h})$$

donde  $\boldsymbol{\rho}(\mathbf{h}) = (\rho_1(\mathbf{h}), \rho_2(\mathbf{h}), \dots, \rho_m(\mathbf{h}))$  tiende hacia 0 cuando  $\mathbf{h} \rightarrow 0$ . Por lo tanto  $\mathbf{f}$  es diferenciable en  $\mathbf{a}$  y  $d\mathbf{f}(\mathbf{a}) = L$ . ■

**Proposición 5.19** Si  $\mathbf{f}, \mathbf{g} : \Omega \rightarrow F$  son diferenciables en  $\mathbf{a} \in \Omega$ , donde  $\Omega \subset E$  es abierto, se verifica:

i)  $\mathbf{f} + \mathbf{g}$  es diferenciable en  $\mathbf{a}$  y  $d(\mathbf{f} + \mathbf{g})(\mathbf{a}) = d\mathbf{f}(\mathbf{a}) + d\mathbf{g}(\mathbf{a})$ .

ii) Si la norma de  $F$  procede del producto escalar  $\langle \cdot | \cdot \rangle$ , la función

$$\psi(\mathbf{x}) = \langle \mathbf{f}(\mathbf{x}) | \mathbf{g}(\mathbf{x}) \rangle$$

es diferenciable en  $\mathbf{a}$  y su diferencial  $d\psi(\mathbf{a}) : E \rightarrow \mathbb{R}$  es la aplicación lineal

$$d\psi(\mathbf{a})\mathbf{v} = \langle \mathbf{f}(\mathbf{a}) | d\mathbf{g}(\mathbf{a})\mathbf{v} \rangle + \langle d\mathbf{f}(\mathbf{a})\mathbf{v} | \mathbf{g}(\mathbf{a}) \rangle$$

iii) Si  $\alpha : \Omega \rightarrow \mathbb{R}$  es diferenciable en  $\mathbf{a} \in \Omega$  entonces  $\mathbf{g}(\mathbf{x}) = \alpha(\mathbf{x})\mathbf{f}(\mathbf{x})$  es diferenciable en  $\mathbf{a}$  y  $d(\mathbf{g})(\mathbf{a}) : E \rightarrow E$  es la aplicación lineal dada por

$$d\mathbf{g}(\mathbf{a})\mathbf{v} = \alpha(\mathbf{a})d\mathbf{f}(\mathbf{a})\mathbf{v} + [d\alpha(\mathbf{a})\mathbf{v}]\mathbf{f}(\mathbf{a})$$

iv) Si  $\alpha : \Omega \rightarrow \mathbb{R}$  es diferenciable en  $\mathbf{a} \in \Omega$  y  $\alpha(\mathbf{a}) \neq 0$ , entonces la función  $\mathbf{g}(\mathbf{x}) = \mathbf{f}(\mathbf{x})/\alpha(\mathbf{x})$ , (definida en un entorno de  $\mathbf{a}$ ) es diferenciable en  $\mathbf{a}$  y

$$d(\mathbf{g})(\mathbf{a})\mathbf{v} = \frac{\alpha(\mathbf{a})d\mathbf{f}(\mathbf{a})\mathbf{v} - [d\alpha(\mathbf{a})\mathbf{v}]\mathbf{f}(\mathbf{a})}{\alpha(\mathbf{a})^2}$$

DEM: La demostración de i) es inmediata. Para demostrar ii) veamos primero la continuidad de la aplicación lineal  $S : E \rightarrow \mathbb{R}$ , definida por

$$S(\mathbf{v}) = \langle \mathbf{f}(\mathbf{a}) \mid T(\mathbf{v}) \rangle + \langle L(\mathbf{v}) \mid \mathbf{g}(\mathbf{a}) \rangle \quad \text{donde } L = d\mathbf{f}(\mathbf{a}) \text{ y } T = d\mathbf{g}(\mathbf{a})$$

En virtud de la desigualdad de Cauchy-Schwarz

$$|S(\mathbf{v})| \leq \|\mathbf{f}(\mathbf{a})\| \|T(\mathbf{v})\| + \|L(\mathbf{v})\| \|\mathbf{g}(\mathbf{a})\| \leq M \|\mathbf{v}\|$$

con  $M = \|\mathbf{f}(\mathbf{a})\| \|T\| + \|\mathbf{g}(\mathbf{a})\| \|L\|$ , y por lo tanto  $S$  es una aplicación lineal continua. (Cuando  $E$  es finito dimensional no es preciso demostrar esto). Por hipótesis,

$$\mathbf{f}(\mathbf{a} + \mathbf{h}) = \mathbf{f}(\mathbf{a}) + L(\mathbf{h}) + \boldsymbol{\alpha}(\mathbf{h}), \quad \mathbf{g}(\mathbf{a} + \mathbf{h}) = \mathbf{g}(\mathbf{a}) + T(\mathbf{h}) + \boldsymbol{\beta}(\mathbf{h})$$

donde  $\boldsymbol{\alpha}(\mathbf{h}) = o(\|\mathbf{h}\|)$  y  $\boldsymbol{\beta}(\mathbf{h}) = o(\|\mathbf{h}\|)$ . Por la bilinealidad del producto escalar

$$\begin{aligned} \langle \mathbf{f}(\mathbf{a} + \mathbf{h}) \mid \mathbf{g}(\mathbf{a} + \mathbf{h}) \rangle &= \langle \mathbf{f}(\mathbf{a}) + L(\mathbf{h}) \mid \mathbf{g}(\mathbf{a} + \mathbf{h}) \rangle + \langle \boldsymbol{\alpha}(\mathbf{h}) \mid \mathbf{g}(\mathbf{a} + \mathbf{h}) \rangle = \\ &= \langle \mathbf{f}(\mathbf{a}) + L(\mathbf{h}) \mid \mathbf{g}(\mathbf{a}) + T(\mathbf{h}) \rangle + \langle \mathbf{f}(\mathbf{a}) + L(\mathbf{h}) \mid \boldsymbol{\beta}(\mathbf{h}) \rangle + \langle \boldsymbol{\alpha}(\mathbf{h}) \mid \mathbf{g}(\mathbf{a} + \mathbf{h}) \rangle = \\ &= \langle \mathbf{f}(\mathbf{a}) \mid \mathbf{g}(\mathbf{a}) \rangle + S(\mathbf{h}) + \boldsymbol{\rho}(\mathbf{h}) \end{aligned}$$

donde

$$\boldsymbol{\rho}(\mathbf{h}) = \langle L(\mathbf{h}) \mid T(\mathbf{h}) \rangle + \langle \mathbf{f}(\mathbf{a}) + L(\mathbf{h}) \mid \boldsymbol{\beta}(\mathbf{h}) \rangle + \langle \boldsymbol{\alpha}(\mathbf{h}) \mid \mathbf{g}(\mathbf{a} + \mathbf{h}) \rangle$$

y basta ver que  $\boldsymbol{\rho}(\mathbf{h}) = o(\|\mathbf{h}\|)$ : En virtud de desigualdad de Cauchy-Schwarz

$$|\boldsymbol{\rho}(\mathbf{h})| \leq \|L\| \|T\| \|\mathbf{h}\|^2 + \|\mathbf{f}(\mathbf{a}) + L(\mathbf{h})\| \|\boldsymbol{\beta}(\mathbf{h})\| + \|\mathbf{g}(\mathbf{a} + \mathbf{h})\| \|\boldsymbol{\alpha}(\mathbf{h})\|$$

Después de dividir por  $\|\mathbf{h}\|$ , los tres términos de la derecha tienden hacia 0 cuando  $\mathbf{h} \rightarrow \mathbf{0}$ , luego  $|\boldsymbol{\rho}(\mathbf{h})| = o(\|\mathbf{h}\|)$ .

La demostración de iii) que es análoga a la de ii), se deja al cuidado del lector, y iv) es consecuencia directa de iii). ■

**Teorema 5.20** [Regla de la cadena] Sean  $(E, \|\cdot\|)$ ,  $(F, \|\cdot\|)$  y  $(G, \|\cdot\|)$  espacios normados y  $\Omega \subset E$ ,  $V \subset F$  abiertos. Si  $\mathbf{f} : \Omega \rightarrow V$  es diferenciable en  $\mathbf{a} \in \Omega$  y  $\mathbf{g} : V \rightarrow G$  es diferenciable en  $\mathbf{b} = \mathbf{f}(\mathbf{a})$ , entonces  $\mathbf{g} \circ \mathbf{f}$  es diferenciable en  $\mathbf{a}$  y

$$d(\mathbf{g} \circ \mathbf{f})(\mathbf{a}) = d\mathbf{g}(\mathbf{f}(\mathbf{a})) \circ d\mathbf{f}(\mathbf{a})$$

DEM: Si  $L = d\mathbf{f}(\mathbf{a})$  y  $T = d\mathbf{g}(\mathbf{b})$ , en virtud de la hipótesis:

- i)  $\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{a}) + L(\mathbf{x} - \mathbf{a}) + \|\mathbf{x} - \mathbf{a}\| \boldsymbol{\rho}_1(\mathbf{x})$  con  $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \boldsymbol{\rho}_1(\mathbf{x}) = 0$ .
- ii)  $\mathbf{g}(\mathbf{y}) = \mathbf{g}(\mathbf{b}) + T(\mathbf{y} - \mathbf{b}) + \|\mathbf{y} - \mathbf{b}\| \boldsymbol{\rho}_2(\mathbf{y})$ , con  $\lim_{\mathbf{y} \rightarrow \mathbf{b}} \boldsymbol{\rho}_2(\mathbf{y}) = 0$

En lo que sigue conviene suponer que hemos definido  $\rho_1(\mathbf{a}) = \mathbf{0} \in F$  y  $\rho_2(\mathbf{b}) = \mathbf{0} \in G$ . Así la función  $\rho_1(\mathbf{x})$  (resp.  $\rho_2(\mathbf{y})$ ) está definida para todo  $\mathbf{x} \in \Omega$  (resp.  $\mathbf{y} \in V$ ) y es continua en el punto  $\mathbf{x} = \mathbf{a}$  (resp.  $\mathbf{y} = \mathbf{b}$ ). Sustituyendo  $\mathbf{y} = \mathbf{f}(\mathbf{x})$  en ii) resulta

$$\mathbf{g}(\mathbf{f}(\mathbf{x})) = \mathbf{g}(\mathbf{f}(\mathbf{a})) + T(\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a})) + \|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a})\| \rho_2(\mathbf{f}(\mathbf{x}))$$

y reemplazando  $\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a})$  por el valor que proporciona i)

$$(\mathbf{g} \circ \mathbf{f})(\mathbf{x}) = (\mathbf{g} \circ \mathbf{f})(\mathbf{a}) + (T \circ L)(\mathbf{x} - \mathbf{a}) + \|\mathbf{x} - \mathbf{a}\| \mathbf{r}(\mathbf{x})$$

donde

$$\mathbf{r}(\mathbf{x}) = T(\rho_1(\mathbf{x})) + \frac{\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a})\|}{\|\mathbf{x} - \mathbf{a}\|} \rho_2(\mathbf{f}(\mathbf{x}))$$

y para terminar basta demostrar que  $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{r}(\mathbf{x}) = \mathbf{0}$ .

Cuando  $\mathbf{x} \rightarrow \mathbf{a}$ , en virtud de la continuidad de  $T$  y de  $\mathbf{f}$ , podemos asegurar que  $T(\rho_1(\mathbf{x}))$  tiende hacia 0 y que  $\mathbf{f}(\mathbf{x})$  tiende hacia  $\mathbf{b} = \mathbf{f}(\mathbf{a})$ , y teniendo en cuenta que  $\rho_2$  es continua en  $\mathbf{b}$  se obtiene que  $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \rho_2(\mathbf{f}(\mathbf{x})) = \rho_2(\mathbf{b}) = \mathbf{0}$ . Por consiguiente, para demostrar lo que se desea, basta ver que el cociente

$$\frac{\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a})\|}{\|\mathbf{x} - \mathbf{a}\|}$$

se mantiene acotado en un entorno de  $\mathbf{a}$ . Como  $\mathbf{f}$  es diferenciable en  $\mathbf{a}$ , existe  $\delta > 0$  tal que si  $\|\mathbf{x} - \mathbf{a}\| < \delta$  se cumple  $\|\rho_1(\mathbf{x})\| < 1$  luego, en virtud de i)

$$\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a})\| \leq \|L\| \|\mathbf{x} - \mathbf{a}\| + \|\mathbf{x} - \mathbf{a}\| \|\rho_1(\mathbf{x})\| \leq (\|L\| + 1) \|\mathbf{x} - \mathbf{a}\|$$

■

**Corolario 5.21** Sean  $(F, \|\cdot\|)$ ,  $(G, \|\cdot\|)$  espacios normados, y  $\Omega \subset \mathbb{R}$ ,  $V \subset F$  abiertos. Si  $\mathbf{f} : \Omega \rightarrow V$  es derivable en  $t \in \Omega$  y  $\mathbf{g} : V \rightarrow G$  diferenciable en  $\mathbf{f}(t)$ , entonces  $\mathbf{g} \circ \mathbf{f}$  es derivable en  $t$  y

$$(\mathbf{g} \circ \mathbf{f})'(t) = d\mathbf{g}(\mathbf{f}(t))\mathbf{f}'(t)$$

es decir el vector  $(\mathbf{g} \circ \mathbf{f})'(t) \in G$  es la imagen del vector  $\mathbf{f}'(t) \in F$  mediante la aplicación lineal  $d\mathbf{g}(\mathbf{f}(t)) : F \rightarrow G$ .

DEM: Es consecuencia directa de 5.20 pues para cada  $h \in \mathbb{R}$  se tiene

$$h(\mathbf{g} \circ \mathbf{f})'(t) = d(\mathbf{g} \circ \mathbf{f})(t)h = d\mathbf{g}(\mathbf{f}(t))(d\mathbf{f}(t)h) = d\mathbf{g}(\mathbf{f}(t))(h\mathbf{f}'(t)) = hd\mathbf{g}(\mathbf{f}(t))\mathbf{f}'(t)$$

■

**Expresión analítica de la diferencial. Matriz Jacobiana.** Supongamos que  $\Omega \subset \mathbb{R}^n$  es abierto y que  $\mathbf{f} : \Omega \rightarrow \mathbb{R}^m$  es diferenciable en  $\mathbf{a} \in \Omega$ . Su diferencial es una aplicación lineal  $d\mathbf{f}(\mathbf{a}) : \mathbb{R}^n \rightarrow \mathbb{R}^m$  que se suele identificar con una matriz de  $m$  filas y  $n$  columnas. Para obtener la matriz de  $d\mathbf{f}(\mathbf{a}) : \mathbb{R}^n \rightarrow \mathbb{R}^m$  respecto a las bases canónicas de  $\mathbb{R}^n$  y  $\mathbb{R}^m$ , basta considerar las componentes  $(f_1, f_2, \dots, f_m)$  de  $\mathbf{f}$ . Según

5.18 cada componente  $f_j$  es diferenciable en  $\mathbf{a}$  y  $df_j(\mathbf{a}) = L_j$  donde  $(L_1, L_2, \dots, L_m)$  son las componentes de  $L = d\mathbf{f}(\mathbf{a})$ . Entonces, para cada  $\mathbf{v} = (v_1, v_2, \dots, v_n) \in \mathbb{R}^n$ , su imagen  $d\mathbf{f}(\mathbf{a})\mathbf{v} = \mathbf{u} = (u_1, u_2, \dots, u_m)$  viene dada por

$$u_j = L_j(\mathbf{v}) = df_j(\mathbf{a})\mathbf{v} = \sum_{k=1}^m D_k f_j(\mathbf{a})v_k = \sum_{k=1}^m \frac{\partial f_j(\mathbf{a})}{\partial x_k} v_k$$

fórmula que con notación matricial se escribe en la forma

$$\begin{pmatrix} u_1 \\ u_2 \\ \dots \\ u_m \end{pmatrix} = \begin{pmatrix} D_1 f_1(\mathbf{a}) & D_2 f_1(\mathbf{a}) & \dots & D_n f_1(\mathbf{a}) \\ D_1 f_2(\mathbf{a}) & D_2 f_2(\mathbf{a}) & \dots & D_n f_2(\mathbf{a}) \\ \dots & \dots & \dots & \dots \\ D_1 f_m(\mathbf{a}) & D_2 f_m(\mathbf{a}) & \dots & D_n f_m(\mathbf{a}) \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \dots \\ v_n \end{pmatrix}$$

La matriz anterior es la matriz de la diferencial  $d\mathbf{f}(\mathbf{a})$  respecto a las bases canónicas de  $\mathbb{R}^n$  y  $\mathbb{R}^m$ . Se le llama matriz jacobiana y también matriz derivada. Utilizando la notación  $\mathbf{f}'(\mathbf{a})$  para designar esta matriz podemos escribir  $d\mathbf{f}(\mathbf{a})\mathbf{v} = \mathbf{f}'(\mathbf{a})\mathbf{v}$ , con el convenio de que  $\mathbf{f}'(\mathbf{a})\mathbf{v}$  es el producto de la matriz  $\mathbf{f}'(\mathbf{a})$  por el vector  $\mathbf{v}$  escrito en forma de vector columna. Notaciones más habituales para la matriz jacobiana son

$$\frac{D(f_1, f_2, \dots, f_m)}{D(x_1, x_2, \dots, x_n)}(\mathbf{a}), \quad \frac{\partial(f_1, f_2, \dots, f_m)}{\partial(x_1, x_2, \dots, x_n)}(\mathbf{a})$$

En particular, cuando  $n = 1$ , sabemos que  $\mathbf{f}$  es derivable en  $a$  si y sólo si es diferenciable en  $a$  y en ese caso su diferencial es la aplicación lineal  $d\mathbf{f}(a) : \mathbb{R} \rightarrow \mathbb{R}^m$ , definida por  $d\mathbf{f}(a)(h) = \mathbf{f}'(a)h$  y la matriz de la diferencial  $d\mathbf{f}(a)$  se identifica con  $\mathbf{f}'(a)$  escrito como vector columna.

Consideremos ahora el caso de una función compuesta  $\mathbf{g} \circ \mathbf{f}$  donde  $\mathbf{f} : \Omega \rightarrow V$  está definida en un abierto  $\Omega \subset \mathbb{R}^n$  y  $\mathbf{g} : V \rightarrow \mathbb{R}^p$  en un abierto  $V \subset \mathbb{R}^m$ . Se supone que  $\mathbf{f}$  es diferenciable en  $\mathbf{x} \in \Omega$  y que  $\mathbf{g}$  diferenciable en  $\mathbf{y} = \mathbf{f}(\mathbf{x})$ . Entonces, de acuerdo con 5.20, la matriz jacobiana de la función compuesta  $\varphi = \mathbf{g} \circ \mathbf{f}$ , en el punto  $\mathbf{x}$ , es el producto de la matrices jacobianas,  $\varphi'(\mathbf{x}) = \mathbf{g}'(\mathbf{y})\mathbf{f}'(\mathbf{x})$ , es decir

$$\begin{aligned} & \begin{pmatrix} D_1 \varphi_1(\mathbf{x}) & \dots & D_n \varphi_1(\mathbf{x}) \\ D_1 \varphi_2(\mathbf{x}) & \dots & D_n \varphi_2(\mathbf{x}) \\ \dots & \dots & \dots \\ D_1 \varphi_p(\mathbf{x}) & \dots & D_n \varphi_p(\mathbf{x}) \end{pmatrix} = \\ & = \begin{pmatrix} D_1 g_1(\mathbf{y}) & \dots & D_m g_1(\mathbf{y}) \\ D_1 g_2(\mathbf{y}) & \dots & D_m g_2(\mathbf{y}) \\ \dots & \dots & \dots \\ D_1 g_p(\mathbf{y}) & \dots & D_m g_p(\mathbf{y}) \end{pmatrix} \begin{pmatrix} D_1 f_1(\mathbf{x}) & \dots & D_n f_1(\mathbf{x}) \\ D_1 f_2(\mathbf{x}) & \dots & D_n f_2(\mathbf{x}) \\ \dots & \dots & \dots \\ D_1 f_m(\mathbf{x}) & \dots & D_n f_m(\mathbf{x}) \end{pmatrix} \end{aligned}$$

y de acuerdo con la fórmula para el producto de matrices:

$$D_i \varphi_j(\mathbf{x}) = \sum_{k=1}^m D_k \mathbf{g}_j(\mathbf{f}(\mathbf{x})) D_i \mathbf{f}_k(\mathbf{a}), \quad i = 1, \dots, n, \quad j = 1, 2, \dots, p \quad (5.1)$$

Con el fin de recordar fácilmente esta regla para el cálculo de las derivadas parciales de una función compuesta es conveniente adoptar el siguiente convenio: Denotemos por  $(x_1, x_2, \dots, x_n)$  las variables independientes de  $\mathbf{f}$ , y por  $(y_1, y_2, \dots, y_m)$  las variables independientes de  $\mathbf{g}$ , de modo que para obtener la función compuesta

$$(z_1, z_2, \dots, z_p) = (\varphi_1(x_1, x_2, \dots, x_n), \varphi_2(x_1, x_2, \dots, x_n), \dots, \varphi_p(x_1, x_2, \dots, x_n))$$

basta sustituir  $y_k = f_k(x_1, x_2, \dots, x_n)$ ,  $1 \leq k \leq m$ , en la expresión

$$(z_1, z_2, \dots, z_p) = \mathbf{g}(y_1, y_2, \dots, y_m)$$

Si hacemos el convenio de designar del mismo modo a las variables  $y_k$  y a las funciones  $f_k$  (y a las variables  $z_j$  y a las funciones  $g_j$ ) la fórmula para las derivadas parciales de la función compuesta se escribe en una forma fácil de recordar.

$$\frac{\partial z_j}{\partial x_i} = \sum_{k=1}^m \frac{\partial z_j}{\partial y_k} \frac{\partial y_k}{\partial x_i}$$

donde, con el fin de simplificar la expresión, hemos omitido los puntos donde se evalúan las derivadas parciales. Esto es lo que se hace habitualmente cuando los puntos quedan claros por el contexto.

**Las notaciones del cálculo diferencial** Consideremos primero el caso de una función real de  $n$  variables reales que es diferenciable en  $\mathbf{a} \in \Omega \subset \mathbb{R}^n$ . Su diferencial  $df(\mathbf{a})$  pertenece al espacio vectorial  $n$ -dimensional  $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$  y una base de este espacio vectorial son las proyecciones

$$\pi_k : \mathbb{R}^n \rightarrow \mathbb{R}; \quad \pi_k(x_1, x_2, \dots, x_n) = x_k$$

En cálculo diferencial la proyección  $\pi_k$  se acostumbra a denotar  $dx_k$ . La razón de esto es la siguiente: Resulta cómodo denotar una función por la fórmula que se usa para definirla. Así, por ejemplo, la función  $f : \mathbb{R}^3 \rightarrow \mathbb{R}$  que asigna al vector  $\mathbf{x} = (x, y, z)$  el número  $x \cos(y + e^z)$  se suele designar mediante la fórmula que la define. Con este convenio, si estamos trabajando con funciones de tres variables, la función  $x$  será la proyección  $\pi_1 : \mathbb{R}^3 \rightarrow \mathbb{R}$  que asocia al vector  $(x, y, z) \in \mathbb{R}^3$  su primera coordenada  $x$ . Esta proyección es lineal y por lo tanto su diferencial  $dx$  es la aplicación constante que a cada punto  $\mathbf{a} \in \mathbb{R}^3$  le hace corresponder la proyección  $\pi_1$  (véase el ejemplo ii) de la sección 2). Siguiendo el convenio habitual de identificar una aplicación constante con su valor constante, es natural utilizar la notación  $dx$  para designar la proyección  $\pi_1$ . Del mismo modo,  $dy$  y  $dz$  designan las proyecciones  $\pi_2$  y  $\pi_3$ , respectivamente. Análogamente, en  $\mathbb{R}^n$ , la proyección  $\pi_k$  se denota  $dx_k$  ( $k = 1 \dots n$ )

Sea  $\{\mathbf{e}_k : 1 \leq k \leq n\}$  la base canónica de  $\mathbb{R}^n$ . Dado  $\mathbf{v} = (v_1, v_2, \dots, v_n) \in \mathbb{R}^n$ , en virtud de la linealidad de la diferencial  $df(\mathbf{a})$  y de la proposición 5.14 sabemos que

$$df(\mathbf{a})\mathbf{v} = \sum_{k=1}^n D_k f(\mathbf{a})v_k$$



Como  $v_k$  es la imagen de  $\mathbf{v}$  mediante la proyección  $dx_k$  resulta

$$df(\mathbf{a})\mathbf{v} = \left( \sum_{k=1}^n D_k f(\mathbf{a}) dx_k \right) \mathbf{v}$$

luego las derivadas parciales  $(D_1 f(\mathbf{a}), D_2 f(\mathbf{a}), \dots, D_n f(\mathbf{a}))$  son las coordenadas de  $df(\mathbf{a})$  respecto a la base canónica  $(dx_1, dx_2, \dots, \dots, dx_n)$  de  $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$ :

$$df(\mathbf{a}) = \sum_{k=1}^n D_k f(\mathbf{a}) dx_k = \sum_{k=1}^n \frac{\partial f(\mathbf{a})}{\partial x_k} dx_k$$

En particular, cuando  $n = 2$  se acostumbra a escribir

$$df(\mathbf{a}) = \frac{\partial f(\mathbf{a})}{\partial x} dx + \frac{\partial f(\mathbf{a})}{\partial y} dy = f_x(\mathbf{a}) dx + f_y(\mathbf{a}) dy$$

Análogamente, cuando  $n = 3$

$$df(\mathbf{a}) = \frac{\partial f(\mathbf{a})}{\partial x} dx + \frac{\partial f(\mathbf{a})}{\partial y} dy + \frac{\partial f(\mathbf{a})}{\partial z} dz = f_x(\mathbf{a}) dx + f_y(\mathbf{a}) dy + f_z(\mathbf{a}) dz$$

En el caso de funciones de una variable se tiene  $df(a) = f'(a)dx$  lo que está de acuerdo con la notación de Leibniz para la derivada  $\mathbf{f}'(a) = \frac{d\mathbf{f}}{dx}(a)$ , donde la expresión  $\frac{d\mathbf{f}}{dx}$  se debe entender como un símbolo unitario que designa a la derivada  $\mathbf{f}'$  y no como una fracción. Este simbolismo permite escribir, omitiendo el punto  $a$ , la fórmula clásica de Leibniz

$$d\mathbf{f} = \frac{d\mathbf{f}}{dx} dx$$

que actualmente se interpreta como una igualdad entre las aplicaciones lineales  $d\mathbf{f}(a)$  y  $\mathbf{f}'(a)dx$ . La flexibilidad de la notación de Leibniz se pone de manifiesto con la regla de la cadena 4.6: En términos de diferenciales, la regla de la cadena para la composición  $g(t) = f(\varphi(t))$  adopta la forma  $dg(a)h = g'(a)h = f'(\varphi(a))\varphi'(a)h = f'(\varphi(a))[d\varphi(a)h] = [d\mathbf{f}(\varphi(a)) \circ d\varphi(a)]h$  es decir  $dg(a) = d\mathbf{f}(\varphi(a)) \circ d\varphi(a)$ .

**El teorema del incremento finito.** En lo que sigue, si  $E$  es un espacio vectorial normado el segmento cerrado (resp. abierto) de extremos  $\mathbf{a}, \mathbf{b} \in E$ , es

$$[\mathbf{a}, \mathbf{b}] = \{\mathbf{a} + t(\mathbf{b} - \mathbf{a}) : 0 \leq t \leq 1\}, \quad (\text{resp. } (\mathbf{a}, \mathbf{b}) = \{\mathbf{a} + t(\mathbf{b} - \mathbf{a}) : 0 < t < 1\})$$

**Teorema 5.22** [Incremento finito] *Sea  $\mathbf{f} : \Omega \rightarrow F$ , definida en un abierto  $\Omega$  del espacio normado  $(E, \|\cdot\|)$ , con valores en el espacio normado  $(F, \|\cdot\|)$ , continua en cada punto del segmento cerrado  $[\mathbf{a}, \mathbf{b}]$  y diferenciable en cada punto del segmento abierto  $(\mathbf{a}, \mathbf{b})$  con  $\|d\mathbf{f}(\mathbf{x})\| \leq M$  para todo  $\mathbf{x} \in (\mathbf{a}, \mathbf{b})$ . Entonces  $\|\mathbf{f}(\mathbf{b}) - \mathbf{f}(\mathbf{a})\| \leq M \|\mathbf{b} - \mathbf{a}\|$ .*

DEM: En virtud de 5.21 la función de variable real  $\varphi(t) = \mathbf{f}(\mathbf{a} + t(\mathbf{b} - \mathbf{a}))$  es continua en  $[0, 1]$  y derivable en  $(0, 1)$  con derivada  $\varphi'(t) = d\mathbf{f}(\mathbf{a} + t(\mathbf{b} - \mathbf{a}))(\mathbf{b} - \mathbf{a})$ . Para todo  $t \in (0, 1)$  se verifica

$$\|\varphi'(t)\| \leq \|d\mathbf{f}(\mathbf{a} + t(\mathbf{b} - \mathbf{a}))\| \|\mathbf{b} - \mathbf{a}\| \leq M \|\mathbf{b} - \mathbf{a}\|$$

y según 4.8,  $\|\varphi(1) - \varphi(0)\| \leq M \|\mathbf{b} - \mathbf{a}\|$  es decir  $\|\mathbf{f}(\mathbf{b}) - \mathbf{f}(\mathbf{a})\| \leq M \|\mathbf{b} - \mathbf{a}\|$ . ■

Un subconjunto  $A$  del espacio normado  $(E, \|\cdot\|)$  se dice que es *convexo* cuando para cada par de puntos  $\mathbf{x}, \mathbf{y} \in A$ , el segmento  $[\mathbf{x}, \mathbf{y}]$  está contenido en  $A$ .

**Corolario 5.23** *Sea  $\Omega$  un abierto convexo del espacio normado  $(E, \|\cdot\|)$  y  $\mathbf{f} : \Omega \rightarrow F$  una aplicación diferenciable tal que  $\|d\mathbf{f}(\mathbf{x})\| \leq M$  para todo  $\mathbf{x} \in \Omega$ . Entonces  $\|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})\| \leq M \|\mathbf{y} - \mathbf{x}\|$  para cada  $\mathbf{x}, \mathbf{y} \in \Omega$ . En particular, si  $d\mathbf{f}(\mathbf{x}) = 0$  para todo  $\mathbf{x} \in \Omega$ , entonces  $\mathbf{f}$  es constante.*

DEM: Basta aplicar la proposición 5.22 a cada segmento  $[\mathbf{x}, \mathbf{y}] \subset \Omega$ . Si  $d\mathbf{f}(\mathbf{x}) = 0$  para todo  $\mathbf{x} \in \Omega$ , podemos tomar  $M = 0$  y se obtiene que  $\mathbf{f}$  es constante. ■

**Corolario 5.24** *Sea  $\Omega$  un abierto convexo del espacio normado  $(E, \|\cdot\|)$  y  $\mathbf{f} : \Omega \rightarrow F$  una aplicación diferenciable con valores en el espacio normado  $(F, \|\cdot\|)$ . Si  $d\mathbf{f}(\mathbf{x}) = 0$  para todo  $\mathbf{x} \in \Omega$  entonces  $\mathbf{f}$  es constante.*

DEM: Fijado  $\mathbf{a} \in \Omega$ , como  $\mathbf{f}$  es continua  $A = \{\mathbf{x} \in \Omega : \mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{a})\}$  es un subconjunto cerrado para la topología relativa de  $\Omega$ .  $A$  no es vacío, pues  $\mathbf{a} \in A$ , y bastará ver que  $A$  es un subconjunto abierto del espacio convexo  $\Omega$  para concluir que  $A = \Omega$ , y con ello que  $\mathbf{f}$  es constante. Efectivamente, si  $\mathbf{b} \in A$  existe  $r > 0$  tal que  $B(\mathbf{b}, r) \subset \Omega$ . La bola  $B(\mathbf{b}, r)$  es convexa y aplicando el corolario 5.23 se obtiene que  $\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{b})$  para todo  $\mathbf{x} \in B(\mathbf{b}, r)$ . Como  $\mathbf{b} \in A$  se cumple que  $\mathbf{f}(\mathbf{b}) = \mathbf{f}(\mathbf{a})$ , luego  $B(\mathbf{b}, r) \subset A$  y con esto queda probado que  $A$  es abierto. ■

## 5.4. Gradiente

Cada vector  $\mathbf{z} = (z_1, z_2, \dots, z_n) \in \mathbb{R}^n$  define una aplicación lineal  $L_{\mathbf{z}} : \mathbb{R}^n \rightarrow \mathbb{R}$  mediante la fórmula  $L_{\mathbf{z}}(\mathbf{h}) = \sum_{k=1}^n z_k h_k = \langle \mathbf{z} \mid \mathbf{h} \rangle$ .

Recíprocamente, si  $L : \mathbb{R}^n \rightarrow \mathbb{R}$  es lineal el vector  $\mathbf{z} = (z_1, z_2, \dots, z_n) \in \mathbb{R}^n$  definido por  $z_k = L(\mathbf{e}_k)$ , cumple  $L = L_{\mathbf{z}}$  pues

$$L(\mathbf{h}) = L\left(\sum_{k=1}^n h_k \mathbf{e}_k\right) = \sum_{k=1}^n h_k L(\mathbf{e}_k) = \sum_{k=1}^n h_k z_k$$

Es decir, cada aplicación lineal  $L : \mathbb{R}^n \rightarrow \mathbb{R}$  se puede identificar con el único vector  $\mathbf{z} \in \mathbb{R}^n$  que verifica  $L(\mathbf{h}) = \langle \mathbf{z} \mid \mathbf{h} \rangle$  para todo  $\mathbf{h} \in \mathbb{R}^n$ . Cuando  $L$  es la diferencial de una función se define

**Definición 5.25** Sea  $f : \Omega \rightarrow \mathbb{R}$  definida en un abierto  $\Omega \subset \mathbb{R}^n$ . Si  $f$  es diferenciable en  $\mathbf{a} \in \Omega$  el gradiente de  $f$  en  $\mathbf{a}$ , denotado  $\nabla f(\mathbf{a})$ , es el vector de  $\mathbb{R}^n$  asociado a su diferencial  $df(\mathbf{a}) : \mathbb{R}^n \rightarrow \mathbb{R}$ . Es decir,  $\nabla f(\mathbf{a})$  es el único vector de  $\mathbb{R}^n$  que verifica  $df(\mathbf{a})\mathbf{h} = \langle \nabla f(\mathbf{a}) | \mathbf{h} \rangle$  para todo  $\mathbf{h} \in \mathbb{R}^n$ . Sus componentes vienen dadas por  $df(\mathbf{a})\mathbf{e}_k = D_k f(\mathbf{a})$ ,  $1 \leq k \leq n$ , luego

$$\nabla f(\mathbf{a}) = (D_1 f(\mathbf{a}), D_2 f(\mathbf{a}), \dots, D_n f(\mathbf{a})) \in \mathbb{R}^n$$

El interés de la noción de gradiente lo muestran las dos proposiciones siguientes, que son la base de sus interpretaciones físicas y geométricas.

**Proposición 5.26** Sea  $f : \Omega \rightarrow \mathbb{R}$  diferenciable en  $\mathbf{a} \in \Omega$  con  $\nabla f(\mathbf{a}) \neq (0, \dots, 0)$ . El máximo valor de la derivada direccional  $D_{\mathbf{u}}f(\mathbf{a})$  con  $\|\mathbf{u}\|_2 = 1$ , se alcanza en la dirección del gradiente,  $\mathbf{w} = \nabla f(\mathbf{a}) / \|\nabla f(\mathbf{a})\|_2$ , y su valor es  $\|\nabla f(\mathbf{a})\|_2$

DEM: Es consecuencia inmediata de la definición de gradiente y de la desigualdad de Cauchy-Schwarz, pues si  $\mathbf{u} \in \mathbb{R}^n$  y  $\|\mathbf{u}\|_2 = 1$  se verifica

$$D_{\mathbf{u}}f(\mathbf{a}) = df(\mathbf{a})\mathbf{u} = \langle \nabla f(\mathbf{a}) | \mathbf{u} \rangle \leq |\langle \nabla f(\mathbf{a}) | \mathbf{u} \rangle| \leq \|\nabla f(\mathbf{a})\|_2 \|\mathbf{u}\|_2 = \|\nabla f(\mathbf{a})\|_2$$

Cuando  $\mathbf{u} = \mathbf{w} = \nabla f(\mathbf{a}) / \|\nabla f(\mathbf{a})\|_2$  resulta

$$D_{\mathbf{w}}f(\mathbf{a}) = \langle \nabla f(\mathbf{a}) | \mathbf{w} \rangle = \frac{1}{\|\nabla f(\mathbf{a})\|_2} \langle \nabla f(\mathbf{a}) | \nabla f(\mathbf{a}) \rangle = \|\nabla f(\mathbf{a})\|_2$$

■

**Proposición 5.27** Si  $f : \Omega \rightarrow \mathbb{R}$ , definida en un abierto  $\Omega \subset \mathbb{R}^n$ , es diferenciable en  $\mathbf{a} \in \Omega$  y  $f(\mathbf{a}) = c$  entonces el vector gradiente  $\nabla f(\mathbf{a})$  es ortogonal al conjunto de nivel  $N_c = \{\mathbf{x} \in \Omega : f(\mathbf{x}) = c\}$  en el punto  $\mathbf{a}$ . (esto significa que es ortogonal a cada vector tangente a  $N_c$  en  $\mathbf{a}$ ).

DEM: Si  $\mathbf{u} \in \mathbb{R}^n$  es tangente a  $N_c$  en el punto  $\mathbf{a}$  existe un camino  $\gamma : (-r, r) \rightarrow N_c$ , con  $\gamma(0) = \mathbf{a}$  y  $\gamma'(0) = \mathbf{u}$ . La función  $\varphi(t) = f(\gamma(t))$  es constante ( $=c$ ) en el intervalo  $(-r, r)$  luego, en virtud de la regla de la cadena 5.21, se obtiene

$$0 = \varphi'(0) = df(\gamma(0))\gamma'(0) = df(\mathbf{a})\mathbf{u} = \langle \nabla f(\mathbf{a}) | \mathbf{u} \rangle$$

■

*Interpretación física.* Si pensamos que  $T = f(x, y, z)$  es la temperatura del punto  $(x, y, z) \in \Omega \subset \mathbb{R}^3$ , según la proposición 5.26, la dirección según la cual varía más rápidamente la temperatura es la del gradiente  $\nabla f(\mathbf{a})$ , y la proposición 5.27, nos dice que esta dirección es ortogonal a la 'superficie' isoterma que pasa por  $\mathbf{a}$ .

Para una interpretación geométrica podemos pensar que la función de dos variables  $z = f(x, y)$  proporciona la altura del terreno sobre el punto  $(x, y)$  de una región plana  $\Omega \subset \mathbb{R}^2$ . En este caso, si  $\mathbf{a} \in \Omega$  es un punto donde  $h$  es diferenciable, la pendiente del terreno es máxima según la dirección del gradiente  $\nabla f(\mathbf{a})$ , que es

ortogonal a la 'curva' de nivel que pasa por  $\mathbf{a}$ .

NOTA: Utilizando el teorema B.8 se puede dar una definición general e intrínseca de gradiente (donde no intervienen las derivadas parciales) que sirve para el caso infinito dimensional: Sea  $f : \Omega \rightarrow \mathbb{R}$  definida en un abierto  $\Omega$  de un espacio normado completo  $(E, \|\cdot\|)$  cuya norma procede de un producto escalar. Si  $f$  es diferenciable en  $\mathbf{a}$ , en virtud del teorema B.8 se define  $\nabla f(\mathbf{a})$  como el único vector de  $E$  que cumple  $\langle \nabla f(\mathbf{a}) | \mathbf{u} \rangle = df(\mathbf{a})\mathbf{u}$  para todo  $\mathbf{u} \in E$ . Es claro que, en este contexto más general, siguen valiendo las proposiciones 5.26 y 5.27.

## 5.5. Espacio tangente

En esta sección se establece que para una aplicación  $\mathbf{f}$  diferenciable en un punto  $\mathbf{a}$  hay una definición natural de espacio tangente a su gráfica en el punto  $\mathbf{p} = (\mathbf{a}, \mathbf{f}(\mathbf{a}))$  ya que, en este caso, el conjunto de los vectores tangentes a la gráfica en ese punto es un espacio vectorial de dimensión igual al número de variables de la función. También se estudia la existencia de espacio tangente para conjuntos de nivel de aplicaciones diferenciables, y para conjuntos que se pueden parametrizar mediante una aplicación diferenciable.

**Proposición 5.28** *Sea  $M = \{(\mathbf{x}, \mathbf{f}(\mathbf{x})) \in \mathbb{R}^k \times \mathbb{R}^{n-k} : \mathbf{x} \in \Omega\}$  la gráfica de  $\mathbf{f} : \Omega \rightarrow \mathbb{R}^{n-k}$ , definida en un abierto  $\Omega \subset \mathbb{R}^k$ . Si  $\mathbf{f}$  es diferenciable en  $\mathbf{a}$  y  $\mathbf{p} = (\mathbf{a}, \mathbf{f}(\mathbf{a}))$  entonces*

$$T_{\mathbf{p}}(M) = \{(\mathbf{u}, d\mathbf{f}(\mathbf{a})\mathbf{u}) \in \mathbb{R}^k \times \mathbb{R}^{n-k} : \mathbf{u} \in \mathbb{R}^k\}$$

luego  $T_{\mathbf{p}}(M)$  es un subespacio vectorial  $k$ -dimensional de  $\mathbb{R}^n$ .

DEM: Si  $\mathbf{w} \in T_{\mathbf{p}}(M) \subset \mathbb{R}^n$ , según la definición de vector tangente existe una función  $\gamma : (-r, r) \rightarrow \mathbb{R}^n$ , derivable en 0, tal que  $\gamma(-r, r) \subset M$ ,  $\gamma(0) = \mathbf{p}$  y  $\gamma'(0) = \mathbf{w}$ .

En términos de las componentes de  $\gamma(t) = (\gamma_1(t), \gamma_2(t))$ , y  $\mathbf{w} = (\mathbf{u}, \mathbf{v})$ , en  $\mathbb{R}^k \times \mathbb{R}^{n-k}$ , las condiciones anteriores se escriben así:

$$\gamma_2(t) = \mathbf{f}(\gamma_1(t)) \text{ si } |t| < r, \quad \gamma_1(0) = \mathbf{a}, \quad \gamma_2(0) = \mathbf{f}(\mathbf{a}), \quad \gamma_1'(0) = \mathbf{u}, \quad \gamma_2'(0) = \mathbf{v}$$

En virtud de la regla de la cadena  $\gamma_2'(0) = d\mathbf{f}(\gamma_1(0))\gamma_1'(0)$ , luego  $\mathbf{v} = d\mathbf{f}(\mathbf{a})\mathbf{u}$ .

Recíprocamente, si  $\mathbf{w} = (\mathbf{u}, \mathbf{v}) \in \mathbb{R}^k \times \mathbb{R}^{n-k}$  y  $\mathbf{v} = d\mathbf{f}(\mathbf{a})\mathbf{u}$ , podemos considerar la trayectoria  $\gamma_{\mathbf{u}}(t) = (\mathbf{a} + t\mathbf{u}, \mathbf{f}(\mathbf{a} + t\mathbf{u}))$ , definida en un entorno de 0, y contenida en  $M$ . Como  $\gamma_{\mathbf{u}}(0) = \mathbf{p}$ ,  $\gamma_{\mathbf{u}}'(0) = (\mathbf{u}, d\mathbf{f}(\mathbf{a})\mathbf{u}) = (\mathbf{u}, \mathbf{v}) = \mathbf{w}$ , resulta  $\mathbf{w} \in T_{\mathbf{p}}(M)$ . ■

En las condiciones de la proposición 5.28 se dice que  $T_{\mathbf{p}}(M)$  es el *espacio vectorial tangente* a  $M$  en  $\mathbf{p}$ . Si los vectores de este espacio vectorial se colocan con origen en  $\mathbf{p}$ , sus extremos forman  $\mathbf{p} + T_{\mathbf{p}}(M)$ . Esto es el *espacio afín tangente a  $M$  en  $\mathbf{p}$* .

Si un elemento genérico de  $\mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^{n-k}$  lo representamos en la forma  $(\mathbf{x}, \mathbf{y})$  con  $\mathbf{x} = (x_1, x_2, \dots, x_k)$ ,  $\mathbf{y} = (y_1, y_2, \dots, y_{n-k})$ , el espacio afín tangente a la gráfica de  $\mathbf{f}$  en  $\mathbf{p} = (\mathbf{a}, \mathbf{b})$ , ( $\mathbf{b} = \mathbf{f}(\mathbf{a})$ ) viene dado por

$$\{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^k \times \mathbb{R}^{n-k} : \mathbf{y} - \mathbf{b} = d\mathbf{f}(\mathbf{a})(\mathbf{x} - \mathbf{a})\}$$

Si  $m = n - k$ , sus ecuaciones son:

$$\begin{cases} y_1 - b_1 = D_1 f_1(\mathbf{a})(x_1 - a_1) + D_2 f_1(\mathbf{a})(x_2 - a_2) + \cdots + D_k f_1(\mathbf{a})(x_k - a_k) \\ y_2 - b_2 = D_1 f_2(\mathbf{a})(x_1 - a_1) + D_2 f_2(\mathbf{a})(x_2 - a_2) + \cdots + D_k f_2(\mathbf{a})(x_k - a_k) \\ \dots\dots\dots \\ y_m - b_m = D_1 f_m(\mathbf{a})(x_1 - a_1) + D_2 f_m(\mathbf{a})(x_2 - a_2) + \cdots + D_k f_m(\mathbf{a})(x_k - a_k) \end{cases}$$

En particular, cuando  $n = 3$ ,  $k = 2$ ,  $p = 1$ , y  $\mathbf{a} = (x_0, y_0)$ , los vectores tangentes a la gráfica  $G(f)$  en  $\mathbf{p} = (x_0, y_0, f(x_0, y_0))$  forman un subespacio vectorial de  $\mathbb{R}^3$  de dimensión 2, y en este caso el plano afín tangente a  $G(f)$  en  $\mathbf{p}$  es el de ecuación  $z - z_0 = df(x_0, y_0)(x - x_0, y - y_0)$ , es decir

$$z - z_0 = D_1 f(x_0, y_0)(x - x_0) + D_2 f(x_0, y_0)(y - y_0)$$

de modo que  $(D_1 f(x_0, y_0), D_2 f(x_0, y_0), -1)$  es un vector normal al plano tangente.

**Espacio tangente a un conjunto de nivel.** Ahora consideramos el problema de la existencia de plano tangente a un conjunto de nivel  $M = \{\mathbf{x} \in \Omega : g(x, y, z) = c\}$  en un punto  $\mathbf{p} \in M$ , donde  $g : \Omega \rightarrow \mathbb{R}$  es diferenciable. Según la siguiente proposición siempre se verifica  $T_{\mathbf{p}}(M) \subset \{(u_1, u_2, u_3) : D_1 f(\mathbf{p})u_1 + D_2 f(\mathbf{p})u_2 + D_3 f(\mathbf{p})u_3 = 0\}$ .

**Proposición 5.29** *Sea  $M = \{\mathbf{x} \in \Omega : \mathbf{g}(\mathbf{x}) = \mathbf{c}\} \subset \mathbb{R}^n$  un conjunto de nivel de  $\mathbf{g} : \Omega \rightarrow \mathbb{R}^{n-k}$ , definida en un abierto  $\Omega \subset \mathbb{R}^n$ . Si  $\mathbf{g}$  es diferenciable en  $\mathbf{p} \in M$  entonces  $T_{\mathbf{p}}(M) \subset \{\mathbf{u} \in \mathbb{R}^n : d\mathbf{g}(\mathbf{p})\mathbf{u} = \mathbf{0}\}$ .*

DEM: Si  $\mathbf{u} \in T_{\mathbf{p}}(M)$ , según la definición de vector tangente existe  $\gamma : (-r, r) \rightarrow \mathbb{R}^n$  tal que

$$\gamma(t) \in M \text{ si } |t| < r, \quad \gamma(0) = \mathbf{p}, \quad \gamma'(0) = \mathbf{u}$$

Como  $\mathbf{g} \circ \gamma$  es constante resulta  $\mathbf{0} = (\mathbf{g} \circ \gamma)'(0) = d\mathbf{g}(\gamma(0))\gamma'(0) = d\mathbf{g}(\mathbf{p})\mathbf{u}$ . ■

NOTA: En general no es posible garantizar que en la proposición 5.29 se cumpla la igualdad ni tampoco que  $T_{\mathbf{p}}(M)$  sea espacio vectorial. Esto se pone de manifiesto con el siguiente ejemplo.

**Ejemplo 5.30** *Sea  $M = \{(x, y, z) : x^2 + y^2 - z^2 = 0\}$  un cono de revolución de eje  $z$ , con vértice en  $\mathbf{p} = (0, 0, 0)$ . Entonces  $T_{\mathbf{p}}(M) = \{(u_1, u_2, u_3) : u_3^2 = u_1^2 + u_2^2\}$ .*

DEM: Si  $\mathbf{u} = (u_1, u_2, u_3) \in T_{\mathbf{p}}(M)$ , según la definición de vector tangente existe  $\gamma : (-\epsilon, \epsilon) \rightarrow \mathbb{R}^3$ ,  $\gamma = (\gamma_1, \gamma_2, \gamma_3)$ , derivable en  $t = 0$  tal que  $\gamma_3(t)^2 = \gamma_1(t)^2 + \gamma_2(t)^2$ ,  $\gamma_j(0) = 0$ , y  $\gamma_j'(0) = u_j$  para  $1 \leq j \leq 3$ . En virtud de la definición de derivada

$$\gamma_j(t) = t[u_j + \epsilon_j(t)], \quad \text{donde } \lim_{t \rightarrow 0} \epsilon_j(t) = 0$$

Se sigue que  $(u_3 + \epsilon_3(t))^2 = (u_1 + \epsilon_1(t))^2 + (u_2 + \epsilon_2(t))^2$  y pasando al límite cuando  $t \rightarrow 0$  se obtiene que  $u_3^2 = u_1^2 + u_2^2$ .

Recíprocamente, todo vector  $\mathbf{u} = (u_1, u_2, u_3) \in \mathbb{R}^3$  que cumpla  $u_3^2 = u_1^2 + u_2^2$  es tangente a  $M$  en  $(0, 0, 0)$  pues la función  $\gamma(t) = t\mathbf{u}$  cumple los requisitos que intervienen en la definición de vector tangente:  $\gamma(t) \in M$ , para todo  $t \in (-1, 1)$ ,  $\gamma(0) = (0, 0, 0)$  y  $\gamma'(0) = \mathbf{u}$ . ■

En las condiciones de la proposición 5.29, si ocurre que  $T_{\mathbf{p}}(M) \subset \mathbb{R}^n$  es un subespacio vectorial de dimensión  $k$  diremos que existe espacio tangente al conjunto de nivel  $M$  en el punto  $\mathbf{p} \in M$ . En este caso el espacio afín tangente a  $M$  en  $\mathbf{p}$  es

$$\{\mathbf{x} \in \mathbb{R}^n : d\mathbf{g}(\mathbf{p})(\mathbf{x} - \mathbf{p}) = \mathbf{0}\}$$

y si  $m = n - k$ , sus ecuaciones son

$$\begin{cases} D_1g_1(\mathbf{p})(x_1 - p_1) + D_2g_1(\mathbf{p})(x_2 - p_2) + \dots + D_n g_1(\mathbf{p})(x_n - p_n) = 0 \\ D_1g_2(\mathbf{p})(x_1 - p_1) + D_2g_2(\mathbf{p})(x_2 - p_2) + \dots + D_n g_2(\mathbf{p})(x_n - p_n) = 0 \\ \dots\dots\dots \\ D_1g_m(\mathbf{p})(x_1 - p_1) + D_2g_m(\mathbf{p})(x_2 - p_2) + \dots + D_n g_m(\mathbf{p})(x_n - p_n) = 0 \end{cases}$$

Obsérvese que el espacio tangente a una gráfica  $G(\mathbf{f})$  aparece como caso particular de éste considerando la función  $\mathbf{g}(\mathbf{x}, \mathbf{y}) = \mathbf{f}(\mathbf{x}) - \mathbf{y}$ , con  $\mathbf{x} \in \mathbb{R}^k$ ,  $\mathbf{y} \in \mathbb{R}^{n-k}$ .

Cuando sólo hay una ecuación ( $k = n - 1$ ), se dice que  $M = \{\mathbf{x} \in \Omega : g(\mathbf{x}) = c\}$  es una 'hipersuperficie' de nivel. Ahora, el espacio tangente  $T_{\mathbf{p}}(M)$  es el hiperplano ortogonal al vector  $\nabla g(\mathbf{p}) = (D_1g(\mathbf{p}), D_2g(\mathbf{p}), \dots, D_n g(\mathbf{p}))$ . El hiperplano afín tangente a  $M$  en  $\mathbf{p}$ , es el de ecuación

$$D_1g(\mathbf{p})(x_1 - p_1) + D_2g(\mathbf{p})(x_2 - p_2) + \dots + D_n g(\mathbf{p})(x_n - p_n) = 0$$

En el caso particular de  $n = 3$ ,  $M = \{(x, y, z) \in \Omega : g(x, y, z) = c\}$  es la superficie de nivel de una función  $g : \Omega \rightarrow \mathbb{R}$ , definida en un abierto  $\Omega \subset \mathbb{R}^3$ , y diferenciable en  $\mathbf{p} \in M$ . Siempre ocurre que el vector  $\nabla g(\mathbf{p})$  es ortogonal al conjunto  $T_{\mathbf{p}}(M)$ . Cuando  $T_{\mathbf{p}}(M) \subset \mathbb{R}^3$  sea un subespacio vectorial de dimensión 2 se dirá que existe el plano tangente a la 'superficie' de nivel  $M$  en el punto  $\mathbf{p} \in M$ . El correspondiente plano afín tangente es el de ecuación

$$D_1g(\mathbf{p})(x - x_0) + D_2g(\mathbf{p})(y - y_0) + D_3g(\mathbf{p})(z - z_0) = 0$$

Obsérvese que el caso del plano tangente a una 'superficie' explícita  $z = f(x, y)$  es un caso particular cuando  $c = 0$ ,  $g(x, y, z) = f(x, y) - z$  y  $\mathbf{p} = (x_0, y_0, f(x_0, y_0))$ .

Si  $U_{\mathbf{p}}$  es un entorno de  $\mathbf{p} \in M$ , es claro que  $T_{\mathbf{p}}(M) = T_{\mathbf{p}}(M \cap U_{\mathbf{p}})$ . Combinando esta observación con la proposición 5.28 se obtiene que existe plano tangente a una superficie de nivel  $M = \{(x, y, z) \in \Omega : g(x, y, z) = c\}$  en un punto  $\mathbf{p} \in M$  si se cumple la siguiente condición:

(T) Existe un entorno abierto  $U_{\mathbf{p}}$  de  $\mathbf{p}$  tal que  $M \cap U_{\mathbf{p}}$  se puede representar de alguna de las tres formas siguientes

- a)  $\{(x, y, f_3(x, y)) : (x, y) \in A_{12}\};$
- b)  $\{(x, f_2(x, z), z) : (x, z) \in A_{13}\};$
- c)  $\{(f_1(y, z), y, z) : (y, z) \in A_{23}\}.$

donde  $A_{ij}$  es un entorno abierto de  $(p_i, p_j)$  y  $f_k : A_{ij} \rightarrow \mathbb{R}$  es diferenciable en  $(p_i, p_j)$  con  $f_k(p_i, p_j) = p_k$ .  $((i, j, k)$  es una permutación de  $(1, 2, 3)$ ).

En otras palabras, la condición (T) significa que después de realizar una permutación de las variables  $(x, y, z)$ , la superficie de nivel  $M$  se puede representar localmente, en un entorno de  $\mathbf{p}$ , como la gráfica de una función  $z = f(x, y)$  diferenciable en  $(p_1, p_2)$ .

Si se cumple la condición (T), y alguna de las derivadas parciales  $D_j g(\mathbf{p})$  no es nula, entonces los dos conjuntos que intervienen en la inclusión

$$T_{\mathbf{p}}(M) \subset \{(u_1, u_2, u_3) : D_1 g(\mathbf{p})u_1 + D_2 g(\mathbf{p})u_2 + D_3 g(\mathbf{p})u_3 = 0\}$$

son subespacios vectoriales de dimensión 2 y por lo tanto son iguales

Es claro que una esfera  $S = \{(x, y, z) : x^2 + y^2 + z^2 = R^2\}$  cumple la condición (T) en cualquier punto  $\mathbf{p} \in S$  y queda justificado así que el conjunto de los vectores tangentes a la esfera  $S$  en un punto  $\mathbf{p} \in S$  es un espacio vectorial de dimensión 2.

Dejamos al cuidado del lector la formulación de la condición (T) para el caso de una función de  $n$  variables reales. Anticipamos que, en virtud del teorema de la función implícita 8.16 la condición (T) se cumple cuando  $g : \Omega \rightarrow \mathbb{R}$ , tiene derivadas parciales continuas en un entorno de  $\mathbf{p}$  y no es nulo el gradiente  $\nabla g(\mathbf{p})$ .

La siguiente proposición extiende a una situación más general la condición suficiente para la existencia de espacio tangente a un conjunto de nivel. Para simplificar su enunciado hacemos el siguiente convenio: Si  $I = \{i_1, i_2, \dots, i_k\} \subset \{1, 2, \dots, n\}$ , para cada  $\mathbf{x} \in \mathbb{R}^n$  denotaremos por  $\mathbf{x}_I$  el vector de  $\mathbb{R}^k = \mathbb{R}^I$ , definido por  $\mathbf{x}_I = (x_{i_1}, x_{i_2}, \dots, x_{i_k})$ . Si  $J = \{j_1, j_2, \dots, j_{n-k}\} = \{1, 2, \dots, n\} \setminus I$ , podemos identificar  $\mathbb{R}^n$  con  $\mathbb{R}^I \times \mathbb{R}^J$  mediante la biyección natural  $\mathbf{x} \leftrightarrow (\mathbf{x}_I, \mathbf{x}_J)$ .

**Proposición 5.31** Sea  $M = \{\mathbf{x} \in \Omega : \mathbf{g}(\mathbf{x}) = \mathbf{c}\}$  donde  $\mathbf{g} : \Omega \rightarrow \mathbb{R}^{n-k}$ , está definida en un abierto  $\Omega \subset \mathbb{R}^n$  y es diferenciable en  $\mathbf{p} \in M$ . Se supone que existe un entorno  $\Omega_{\mathbf{p}}$  de  $\mathbf{p}$  tal que  $M \cap \Omega_{\mathbf{p}}$  es la gráfica de una aplicación de  $k$  variables reales  $\mathbf{x}_J = \mathbf{f}(\mathbf{x}_I)$  (definida en un abierto de  $\mathbb{R}^I$ , con valores en  $\mathbb{R}^J$ ) diferenciable en  $\mathbf{p}_J$ . Entonces  $T_{\mathbf{p}}(M)$  es un subespacio vectorial de dimensión  $k$ .

DEM: Es fácil ver que  $T_{\mathbf{p}}(M) = T_{\mathbf{p}}(\Omega_{\mathbf{p}} \cap M)$ . En virtud de 5.28 la hipótesis garantiza que  $T_{\mathbf{p}}(\Omega_{\mathbf{p}} \cap M)$  es un subespacio vectorial de dimensión  $k$ . ■

NOTA. El teorema de la función implícita, que se verá más adelante, asegura que si  $\mathbf{g} : \Omega \rightarrow \mathbb{R}^{n-k}$  es de clase  $C^1$  en  $\Omega \subset \mathbb{R}^n$  y  $M = \{\mathbf{x} \in \Omega : \mathbf{g}(\mathbf{x}) = \mathbf{c}\}$  entonces, en cada  $\mathbf{p} \in M$  donde la matriz jacobiana  $\mathbf{g}'(\mathbf{p})$  tenga rango  $n - k$  se cumple la condición requerida en la proposición anterior 5.31.

**Espacio tangente a la imagen de una parametrización.** Consideremos ahora un conjunto de la forma  $M = \varphi(\Omega)$ , donde  $\varphi : \Omega \rightarrow \mathbb{R}^n$ , definida en un abierto  $\Omega \subset \mathbb{R}^k$ , es diferenciable en  $\mathbf{a} \in \Omega$ . Si  $\mathbf{p} = \varphi(\mathbf{a})$ , sabemos que

$$d\varphi(\mathbf{a})(\mathbb{R}^k) = \{D_{\mathbf{u}}\varphi(\mathbf{a}) : \mathbf{u} \in \mathbb{R}^k\} \subset T_{\mathbf{p}}(M)$$

Si los vectores  $D_1\varphi(\mathbf{a}), \dots, D_k\varphi(\mathbf{a})$  son linealmente independientes, forman una base del subespacio  $d\varphi(\mathbf{a})(\mathbb{R}^k)$ , que será de dimensión  $k$ . Cuando se verifique la igualdad  $d\varphi(\mathbf{a})(\mathbb{R}^k) = T_{\mathbf{p}}(M)$ , diremos que existe espacio tangente a  $M = \varphi(\Omega)$  en el punto  $\mathbf{p} \in M$ . En este caso, el espacio afín tangente a  $M$  en  $\mathbf{p}$ , viene dado por

$$\{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} = \mathbf{a} + d\varphi(\mathbf{a})(t_1\mathbf{e}_1 + t_2\mathbf{e}_2 + \dots + t_k\mathbf{e}_k) : (t_1, t_2, \dots, t_k) \in \mathbb{R}^k\}$$

por lo que sus ecuaciones paramétricas son

$$\mathbf{x} = \mathbf{a} + t_1D_1\varphi(\mathbf{a}) + t_2D_2\varphi(\mathbf{a}) + \dots + t_kD_k\varphi(\mathbf{a}), \quad \text{con } (t_1, t_2, \dots, t_k) \in \mathbb{R}^k$$

es decir,

$$\begin{cases} x_1 = a_1 + t_1D_1\varphi_1(\mathbf{a}) + \dots + t_kD_k\varphi_1(\mathbf{a}) \\ x_2 = a_2 + t_1D_1\varphi_2(\mathbf{a}) + \dots + t_kD_k\varphi_2(\mathbf{a}) \\ \dots\dots\dots \\ x_n = a_n + t_1D_1\varphi_n(\mathbf{a}) + \dots + t_kD_k\varphi_n(\mathbf{a}) \end{cases}$$

Obsérvese que el caso del espacio tangente a una gráfica  $G(\mathbf{f})$  aparece como caso particular considerando la parametrización estandar

$$\varphi(\mathbf{t}) = (\mathbf{t}, \mathbf{f}(\mathbf{t})) = (t_1, \dots, t_k, f_1(t_1, \dots, t_k), \dots, f_{n-k}(t_1, \dots, t_k))$$

En general, para un conjunto de la forma  $M = \varphi(\Omega)$ , si  $\mathbf{p} = \varphi(\mathbf{a})$  y  $\varphi$  es diferenciable en  $\mathbf{a}$ , sólo es posible asegurar la inclusión  $d\varphi(\mathbf{a})(\mathbb{R}^k) \subset T_{\mathbf{p}}(M)$  y puede ocurrir que  $T_{\mathbf{p}}(M)$  no sea un espacio vectorial de dimensión  $n$  (véase el ejemplo 5.33). La siguiente proposición da una condición suficiente para la existencia de espacio tangente a un conjunto de esta forma.

**Proposición 5.32** *Sea  $M = \varphi(\Omega)$  donde  $\varphi : \Omega \rightarrow \mathbb{R}^n$  está definida en un abierto  $\Omega \subset \mathbb{R}^k$ , y es diferenciable en  $\mathbf{a} \in \Omega$ . Se supone que los vectores  $D_1\varphi(\mathbf{a}), \dots, D_k\varphi(\mathbf{a})$  son linealmente independientes y que existe un entorno abierto  $U_{\mathbf{p}} \subset \mathbb{R}^n$  de  $\mathbf{p} = \varphi(\mathbf{a})$  tal que  $M \cap U_{\mathbf{p}}$  se puede representar en forma implícita*

$$U_{\mathbf{p}} \cap M = \{\mathbf{x} \in U_{\mathbf{p}} : \mathbf{g}(\mathbf{x}) = 0\}$$

donde  $\mathbf{g} : U_{\mathbf{p}} \rightarrow \mathbb{R}^{n-k}$  es diferenciable en  $\mathbf{p}$  y el rango de la matriz jacobiana  $\mathbf{g}'(\mathbf{p})$  es  $n - k$ . Entonces  $T_{\mathbf{p}}(M)$  es un subespacio vectorial de dimensión  $k$  y por lo tanto existe el plano tangente a  $M$  en el punto  $\mathbf{p}$ .

DEM: Es fácil ver que  $T_{\mathbf{p}}(M) = T_{\mathbf{p}}(U_{\mathbf{p}} \cap M)$ . Sabemos que se verifica

$$d\varphi(\mathbf{a})(\mathbb{R}^k) \subset T_{\mathbf{p}}(M) = T_{\mathbf{p}}(U_{\mathbf{p}} \cap M) \subset \{\mathbf{u} \in \mathbb{R}^n : d\mathbf{g}(\mathbf{p})\mathbf{u} = \mathbf{0}\}$$



Como  $d\varphi(\mathbf{a})(\mathbb{R}^k)$  y  $\{\mathbf{u} \in \mathbb{R}^n : d\mathbf{g}(\mathbf{p})\mathbf{u} = \mathbf{0}\}$  son subespacios vectoriales de dimensión  $k$ , deben ser iguales, y por lo tanto coinciden con  $T_{\mathbf{p}}(M)$ . ■

NOTA En el capítulo 8 se verán condiciones suficientes para que se cumpla la condición requerida en la proposición 5.32 y en consecuencia, para que exista el plano tangente a  $M = \varphi(\Omega)$  en  $\mathbf{p} = \varphi(\mathbf{a})$ .

Consideremos por último la situación más concreta de una 'superficie' paramétrica  $M = \varphi(\Omega)$ , donde  $\varphi : \Omega \rightarrow \mathbb{R}^3$  está definida en un abierto  $\Omega \subset \mathbb{R}^2$  y es diferenciable en  $\mathbf{a} \in \Omega$ . Si  $\mathbf{p} = \varphi(\mathbf{a})$ , sabemos que

$$d\varphi(\mathbf{a})(\mathbb{R}^2) = \{D_{\mathbf{u}}\varphi(\mathbf{a}) : \mathbf{u} \in \mathbb{R}^2\} \subset T_{\mathbf{p}}(M)$$

Si los vectores  $D_1\varphi(\mathbf{a})$ ,  $D_2\varphi(\mathbf{a})$  son linealmente independientes forman una base del subespacio  $d\varphi(\mathbf{a})(\mathbb{R}^2)$ . Cuando se cumpla la igualdad  $d\varphi(\mathbf{a})(\mathbb{R}^2) = T_{\mathbf{p}}(M)$ , y exista el plano tangente, en  $\mathbf{p}$ , a la 'superficie' paramétrica  $M = \varphi(\Omega)$ , el plano afín tangente  $\{\mathbf{x} \in \mathbb{R}^3 : \mathbf{x} = \mathbf{a} + d\varphi(\mathbf{a})(s\mathbf{e}_1 + t\mathbf{e}_2) : (s, t) \in \mathbb{R}^2\}$  tendrá las ecuaciones paramétricas  $\mathbf{x} = \mathbf{a} + sD_1\varphi(\mathbf{a}) + tD_2\varphi(\mathbf{a})$ , con  $(s, t) \in \mathbb{R}^2$ , es decir

$$\begin{cases} x_1 = a_1 + sD_1\varphi_1(\mathbf{a}) + tD_2\varphi_1(\mathbf{a}) \\ x_2 = a_2 + sD_1\varphi_2(\mathbf{a}) + tD_2\varphi_2(\mathbf{a}) \\ x_3 = a_3 + sD_1\varphi_3(\mathbf{a}) + tD_2\varphi_3(\mathbf{a}) \end{cases}$$

El caso del plano tangente a una 'superficie' explícita  $z = f(x, y)$  considerado anteriormente se obtiene como caso particular de este con la parametrización estandar  $\varphi(s, t) = (s, t, f(s, t))$ . Obsérvese que las derivadas parciales de esta parametrización estandar proporcionan dos vectores tangentes linealmente independientes

$$D_1\varphi(x_0, y_0) = (1, 0, D_1f(x_0, y_0)), \quad D_2\varphi(x_0, y_0) = (0, 1, D_2f(x_0, y_0))$$

que, en este caso, forman una base del plano tangente  $T_{\mathbf{p}}(M)$ .

En general, para una superficie paramétrica  $M = \varphi(\Omega)$ , si  $\mathbf{p} = \varphi(\mathbf{a})$  y  $\varphi$  es diferenciable en  $\mathbf{a}$ , sólo es posible asegurar la inclusión  $d\varphi(\mathbf{a})(\mathbb{R}^2) \subset T_{\mathbf{p}}(M)$  y puede ocurrir que  $T_{\mathbf{p}}(M)$  no sea un plano, como se muestra en el siguiente ejemplo.

**Ejemplo 5.33** Sea  $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  definida por  $\varphi(r, t) = (r \cos t, r \sin t, r)$  y  $\mathbf{a} = (0, 0)$ . En este caso  $M$  es un cono de revolución con vértice en  $(0, 0, 0)$ . Ya hemos visto en 5.30 que  $T_{\mathbf{p}}(M)$  no es un espacio vectorial. Los tres vectores linealmente independientes  $(1, 0, 1)$ ,  $(-1, 0, 1)$ ,  $(0, 1, 1)$  pertenecen a  $T_{\mathbf{p}}(M)$  y sin embargo  $d\varphi(\mathbf{a})(\mathbb{R}^2)$  es un subespacio de dimensión 1 porque  $d\varphi(\mathbf{a})\mathbf{e}_1 = D_1\varphi(\mathbf{a}) = (1, 0, 1)$  y  $d\varphi(\mathbf{a})\mathbf{e}_2 = D_2\varphi(\mathbf{a}) = (0, 0, 0)$

## 5.6. Ejercicios resueltos

**Ejercicio 5.34** Demuestre que es uniformemente continua la función  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  definida por

$$f(x, y) = \frac{x^3 - y^3}{x^2 + y^2} \quad \text{si } (x, y) \neq (0, 0); \quad f(0, 0) = 0 \quad \text{si } (x, y) = (0, 0)$$

SOLUCIÓN

Lo demostraremos viendo que en todo punto  $(x, y) \in \mathbb{R}^2$  existen las derivadas parciales  $D_1f(x, y)$ ,  $D_2f(x, y)$ , y las funciones  $D_1f$ ,  $D_2f$  están acotadas. Utilizando lo que se acaba de establecer se demuestra fácilmente que la función

$$g(x, y) = \frac{x^3}{x^2 + y^2} \quad \text{si } (x, y) \neq (0, 0); \quad g(0, 0) = 0 \quad \text{si } (x, y) = (0, 0)$$

es uniformemente continua. Basta observar que en todo punto  $(x, y)$  existen las derivadas parciales

$$\blacksquare \quad D_1g(x, y) = \frac{x^4 + 3x^2y^2}{(x^2 + y^2)^2} \quad \text{si } (x, y) \neq (0, 0); \quad D_1g(0, 0) = 1.$$

$$\blacksquare \quad D_2g(x, y) = \frac{-2x^3y}{(x^2 + y^2)^2} \quad \text{si } (x, y) \neq (0, 0); \quad D_2g(0, 0) = 0.$$

Utilizando coordenadas polares  $(x, y) = (r \cos \varphi, r \sin \varphi)$  se observa que para todo  $(x, y) \in \mathbb{R}^2$  se cumple  $|D_1f(x, y)| \leq 4$  y  $|D_2f(x, y)| \leq 2$  luego  $g$  es uniformemente continua en  $\mathbb{R}^2$ . Se sigue que  $f(x, y) = g(x, y) - g(y, x)$  también es uniformemente continua en  $\mathbb{R}^2$ . ■

**Ejercicio 5.35** Si  $f : \Omega \rightarrow \mathbb{R}$  es la función definida en el ejemplo 5.7 compruebe que  $g(x, y) = xf(x, y)$  no es uniformemente continua en  $\Omega$ , pero tiene derivadas parciales acotadas.

SOLUCIÓN

$$D_1g(x, y) = f(x, y) - xD_1f(x, y); \quad D_2g(x, y) = xD_2f(x, y)$$

y con los valores de  $D_1f(x, y)$ ,  $D_2f(x, y)$  vistos en 5.7 se obtiene

$$D_1g(x, y) = f(x, y) - \frac{xy}{x^2 + y^2}; \quad D_2g(x, y) = \frac{x^2}{x^2 + y^2}$$

luego  $|D_1g(x, y)| \leq \pi + 1$ , y  $|D_2g(x, y)| \leq 1$  para todo  $(x, y) \in \Omega$ .

La función  $g$  no es uniformemente continua en  $\Omega$ , porque no admite una extensión continua a  $\bar{\Omega} = \mathbb{R}^2$ , ya que para  $a < 0$  se cumple

$$\lim_{y \rightarrow 0^+} g(a, y) = a\pi; \quad \lim_{y \rightarrow 0^-} g(a, y) = -a\pi$$

■

**Ejercicio 5.36** *Se considera la función de dos variables*

$$f(x, y) = yx^2 \operatorname{sen}(1/x) \text{ si } x \neq 0; \quad f(0, y) = 0$$

- a) *Demuestre que  $f$  es diferenciable en todo punto  $(x, y) \in \mathbb{R}^2$*   
 b) *Determine un punto  $\mathbf{p}$  de la elipse  $\{(x, y) : x^2 + \pi^2 y^2 = 1\}$  tal que un vector unitario tangente en  $\mathbf{p}$  a la elipse proporcione el máximo valor de la derivada direccional  $D_{\mathbf{u}}f(2/\pi, \pi)$ .*

SOLUCIÓN

a) En todo punto  $(x, y) \in \mathbb{R}^2$  existe la derivada parcial  $D_1 f(x, y)$ :

$$D_1 f(x, y) = 2xy \operatorname{sen}(1/x) - y \cos(1/x) \text{ si } x \neq 0.$$

$$D_1 f(0, y) = \lim_{h \rightarrow 0} [f(h, y) - f(0, y)]/h = 0.$$

En todo punto  $(x, y) \in \mathbb{R}^2$  existe la derivada parcial  $D_2 f(x, y)$ :

$$D_2 f(x, y) = x^2 \operatorname{sen}(1/x) \text{ si } x \neq 0.$$

$$D_2 f(0, y) = \lim_{k \rightarrow 0} [f(0, y+k) - f(0, y)]/k = 0.$$

Se comprueba que  $D_2 f$  es continua en todo punto  $(x, y) \in \mathbb{R}^2$ . Entonces, según la condición suficiente de diferenciabilidad,  $f$  es diferenciable en todo punto. (Observación: Para aplicar la condición suficiente de diferenciabilidad basta que sea continua una de las dos derivadas parciales. En este caso se puede ver que  $D_1 f(x, y)$  no es continua en los puntos de la forma  $(0, y)$  con  $y \neq 0$ ).

b) El valor máximo de la derivada direccional se alcanza en la dirección del vector  $\nabla f(2/\pi, \pi) = (4, 4/\pi^2)$ . Sea  $g(x, y) = x^2 + \pi^2 y^2 - 1$ . Si  $(a, b)$  es un punto de la elipse, los vectores tangentes a la elipse en  $(a, b)$  son los ortogonales a  $\nabla g(a, b) = (2a, 2\pi^2 b)$ .

Buscamos un punto  $(a, b)$  de la elipse tal que  $(4, 4/\pi^2)$  sea ortogonal a  $(2a, 2\pi^2 b)$ , es decir, un punto que verifique  $0 = a + b$ ;  $a^2 + \pi^2 b^2 = 1$ ; . Como soluciones se obtienen los dos puntos de la elipse  $(a, -a)$  y  $(-a, a)$  con  $a = \sqrt{1/(1 + \pi^2)}$ .

Hay dos vectores unitarios  $\mathbf{u}, -\mathbf{u}$  tangentes a la elipse en  $(a, -a)$  (resp.  $(-a, a)$ ) y proporciona la derivada direccional máxima el que tiene la misma dirección que  $\nabla f(2/\pi, \pi)$ . ■

**Ejercicio 5.37** *Si  $\mathbf{a} \in \mathbb{R}^n$ , demuestre que la función  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , definida por*

$$f(\mathbf{x}) = \langle \mathbf{x} | \mathbf{a} \rangle e^{-\|\mathbf{x}\|_2}$$

*alcanza en  $\mathbb{R}^n$  un máximo y un mínimo absoluto. Calcule sus valores.*

SOLUCIÓN

Como el caso  $\mathbf{a} = \mathbf{0}$  es trivial suponemos en lo que sigue que  $\mathbf{a} \neq \mathbf{0}$ . En virtud de la desigualdad de Cauchy-Schwartz 2.2, para todo  $\mathbf{x} \in \mathbb{R}^n$  se cumple

$$|f(\mathbf{x})| \leq \|\mathbf{a}\|_2 \|\mathbf{x}\|_2 e^{-\|\mathbf{x}\|_2}, \quad \text{luego} \quad \lim_{\|\mathbf{x}\|_2 \rightarrow +\infty} f(\mathbf{x}) = 0$$

Como  $f(\mathbf{a}) > 0$ , existe  $R > 0$  tal que  $\|\mathbf{x}\|_2 > R \Rightarrow |f(\mathbf{x})| < f(\mathbf{a}) = \|\mathbf{a}\|_2 e^{-\|\mathbf{a}\|_2}$ .

Se sigue que  $\|\mathbf{a}\|_2 \leq R$  y que el máximo absoluto que alcanza  $|f|$  en un punto del compacto  $K = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_2 \leq R\}$  es realmente el máximo absoluto en  $\mathbb{R}^n$ .

Sea  $\mathbf{b} \in K$  tal que  $|f(\mathbf{b})| \geq |f(\mathbf{x})|$  para todo  $\mathbf{x} \in \mathbb{R}^n$ . Como  $f(-\mathbf{b}) = -f(\mathbf{b})$ , podemos suponer que  $|f(\mathbf{b})| = f(\mathbf{b})$ , luego para todo  $\mathbf{x} \in \mathbb{R}^n$  se cumple  $f(\mathbf{b}) \geq f(\mathbf{x})$ . Es inmediato que si  $f$  alcanza un máximo absoluto en  $\mathbf{b}$  entonces alcanza también un mínimo absoluto en  $-\mathbf{b}$ . Según la proposición 5.9,  $\mathbf{b}$  es punto estacionario de  $f$  y debe ser solución del sistema

$$\begin{aligned} 1) \quad 0 &= D_1 f(\mathbf{x}) \equiv \left( a_1 - \frac{x_1}{\|\mathbf{x}\|_2} \langle \mathbf{a} \mid \mathbf{x} \rangle \right) e^{-\|\mathbf{x}\|_2} \\ 2) \quad 0 &= D_2 f(\mathbf{x}) \equiv \left( a_2 - \frac{x_2}{\|\mathbf{x}\|_2} \langle \mathbf{a} \mid \mathbf{x} \rangle \right) e^{-\|\mathbf{x}\|_2} \\ &\dots\dots \\ n) \quad 0 &= D_n f(\mathbf{x}) \equiv \left( a_n - \frac{x_n}{\|\mathbf{x}\|_2} \langle \mathbf{a} \mid \mathbf{x} \rangle \right) e^{-\|\mathbf{x}\|_2} \end{aligned}$$

Multiplicando la ecuación de la línea  $j$  por  $x_j$  y sumando se obtiene que los puntos estacionarios de  $f$  deben cumplir

$$0 = \langle \mathbf{a} \mid \mathbf{x} \rangle - \|\mathbf{x}\|_2 \langle \mathbf{a} \mid \mathbf{x} \rangle \quad (*)$$

En particular, como el punto  $\mathbf{b}$  donde  $f$  alcanza el máximo satisface (\*) y cumple  $\langle \mathbf{a} \mid \mathbf{b} \rangle = f(\mathbf{b})e^{\|\mathbf{b}\|_2} > 0$ , se obtiene que  $\|\mathbf{b}\|_2 = 1$ . Como  $\mathbf{b}$  satisface el sistema de ecuaciones de los puntos estacionarios,

$$a_j = b_j \langle \mathbf{a} \mid \mathbf{b} \rangle, \quad 1 \leq j \leq n$$

Elevando al cuadrado y sumando para  $1 \leq j \leq n$  se llega a la igualdad

$$\|\mathbf{a}\|_2^2 = \|\mathbf{b}\|_2^2 \langle \mathbf{a} \mid \mathbf{b} \rangle^2$$

Como  $\|\mathbf{b}\|_2 = 1$  y  $\langle \mathbf{a} \mid \mathbf{b} \rangle > 0$  resulta  $\|\mathbf{a}\|_2 = \langle \mathbf{a} \mid \mathbf{b} \rangle$ . Si en las ecuaciones de los puntos estacionarios sustituimos  $\mathbf{x}$  por  $\mathbf{b}$  y luego  $\langle \mathbf{a} \mid \mathbf{b} \rangle$  por  $\|\mathbf{a}\|_2$  se obtiene

$$0 = a_j - b_j \langle \mathbf{a} \mid \mathbf{b} \rangle = a_j - b_j \|\mathbf{a}\|_2$$

es decir,  $\mathbf{b} = \mathbf{a} / \|\mathbf{a}\|_2$ , luego el máximo absoluto es  $f(\mathbf{a} / \|\mathbf{a}\|_2) = \|\mathbf{a}\|_2 / e$ , y el mínimo absoluto  $f(-\mathbf{a} / \|\mathbf{a}\|_2) = -\|\mathbf{a}\|_2 / e$ . ■

**Ejercicio 5.38** *Dados  $\mathbf{p}^1, \mathbf{p}^2, \dots, \mathbf{p}^m \in \mathbb{R}^n$  demuestre que existe  $\mathbf{p} \in \mathbb{R}^n$  donde la suma  $f(\mathbf{x}) = \sum_{k=1}^m \|\mathbf{x} - \mathbf{p}^k\|_2^2$  alcanza un mínimo absoluto y calcule  $\mathbf{p}$ .*

SOLUCIÓN

La desigualdad  $f(\mathbf{x}) \geq \sum_{k=1}^m (\|\mathbf{x}\|_2 - \|\mathbf{p}^k\|_2)^2$ , implica que  $f(\mathbf{x})$  tiende hacia  $+\infty$  cuando  $\|\mathbf{x}\|_2 \rightarrow +\infty$ , luego existe  $R > 0$  tal que  $\|\mathbf{x}\|_2 > R \Rightarrow f(\mathbf{x}) > f(0)$ .

En algún punto  $\mathbf{p}$  del compacto  $\{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_2 \leq R\}$  la función continua  $f|_K$  alcanza un mínimo absoluto. Si  $\|\mathbf{x}\|_2 > R$  se cumple  $f(\mathbf{x}) > f(0) \geq f(\mathbf{p})$ , luego  $\mathbf{f}(\mathbf{p})$  es realmente el mínimo absoluto de  $f$  en todo  $\mathbb{R}^n$ . según la proposición 5.9,  $\mathbf{p} = (p_1, p_2, \dots, p_n)$  es un punto estacionario de  $f$ , es decir, satisface las ecuaciones

$$D_j f(\mathbf{x}) = 0, \quad 1 \leq j \leq n \quad (*)$$

Para calcular las derivadas parciales  $D_j f$  escribimos  $f(\mathbf{x}) = \sum_{k=1}^m f_k(\mathbf{x})$ , donde  $f_k(x_1, x_2, \dots, x_n) = \sum_{j=1}^n (x_j - p_j^k)^2$ . Resulta

$$D_j f(\mathbf{x}) = \sum_{k=1}^m D_j f_k(\mathbf{x}) = 2 \sum_{k=1}^m (x_j - p_j^k) = 2(mx_j - \sum_{k=1}^m p_j^k); \quad 1 \leq j \leq n$$

Se sigue que  $\mathbf{p}$  es la única solución de (\*), luego  $p_j = \frac{1}{m} \sum_{k=1}^m p_j^k$ . Queda establecido así que  $f$  alcanza su mínimo absoluto en el punto  $\mathbf{p} = \frac{1}{m} \sum_{k=1}^m \mathbf{p}^k$ . ■

**Ejercicio 5.39** Demuestre que entre todas las cajas (con tapa) cuyo volumen es 1 litro hay una cuya superficie es mínima. Calcule sus dimensiones.

SOLUCIÓN

Si  $x > 0$ ,  $y > 0$ ,  $z > 0$  son las dimensiones de la caja, expresadas en cm., hay que demostrar que área de la caja  $S(x, y, z) = 2xy + 2xz + 2yz$  alcanza un mínimo absoluto cuando  $x, y, z$  cumplen la condición de ligadura  $xyz = 1000 \text{ cm}^3$ .

Basta demostrar que la función de dos variables

$$f(x, y) = S(x, y, 1000/(xy)) = 2xy + 2000/x + 2000/y$$

alcanza un mínimo absoluto en el abierto  $U = \{(x, y) \in \mathbb{R}^2 : x > 0, y > 0\}$ . Haremos la demostración buscando un compacto  $K \subset U$  donde podamos asegurar que el mínimo absoluto de  $f|_K$  es realmente el mínimo absoluto de  $f$  en  $U$ .

Vemos a continuación que este requisito lo cumple

$$K = \{(x, y) \in \mathbb{R}^2 : x \geq 1, y \geq 1, xy \leq 1000\} \subset U$$

que es compacto por ser cerrado y acotado (obsérvese que para  $(x, y) \in K$  se cumple  $0 \leq y \leq 1000$ ,  $0 \leq x \leq 1000$ ). Cuando  $(x, y) \in U \setminus K$  se verifica al menos una de las desigualdades  $x < 1$ ,  $y < 1$ ,  $xy > 1000$ , de donde se sigue que

$$f(x, y) = 2xy + 2000/x + 2000/y > 2000$$

Es fácil ver que existe  $\mathbf{q} \in K$  tal que  $f(\mathbf{q}) < 2000$  (por ejemplo  $\mathbf{q} = (5, 10)$ ), luego el mínimo absoluto de la función continua  $f$  en el compacto  $K$  también es el mínimo absoluto de  $f$  en  $U$ . El punto  $\mathbf{p} \in K$  donde  $f$  alcanza el mínimo absoluto debe satisfacer las ecuaciones

$$0 = D_1 f(x, y) = 2y - 2000/x^2; \quad 0 = D_2 f(x, y) = 2x - 2000/y^2$$

que sólo tienen una sola solución,  $x = y = 10$ , luego  $\mathbf{p} = (10, 10)$ , lo que significa que la caja de superficie mínima es el cubo de lado 10 cm. ■

**Ejercicio 5.40** Calcule la mínima distancia entre la superficie  $z = x^2 + y^2$  y el plano  $x + 2y - z = 4$ , justificando previamente que la mínima distancia se alcanza.

DEM: Es fácil ver que el paraboloides de revolución  $S = \{(x, y, z) : z = x^2 + y^2\}$  y el plano  $P = \{(x, y, z) : x + 2y - z = 4\}$  son disjuntos: Si hubiese algún punto  $(x, y)$ , en su intersección, cumpliría  $4 = x + 2y - (x^2 + y^2) = 5/4 - (x - 1/2)^2 - (y - 1)^2 \leq 5/4!$

Como los conjuntos  $S$  y  $P$  son cerrados, pero ninguno es acotado no disponemos de un resultado topológico de carácter general que nos garantice que la distancia

$$d(S, P) = \inf\{\|\mathbf{x} - \mathbf{y}\| : \mathbf{x} \in S, \mathbf{y} \in P\}$$

se alcanza en una pareja de puntos  $\mathbf{p} \in P$ ,  $\mathbf{q} \in S$ .

Sin embargo, en el caso que nos ocupa, lo podremos asegurar acudiendo a un resultado bien conocido de geometría elemental: La distancia de un punto  $(x, y, z)$  al plano  $x + 2y - z = 4$  viene dada por la fórmula  $D(x, y, z) = |x + 2y - z - 4|/\sqrt{6}$ .

El problema lo podremos resolver demostrando primero que sobre el paraboloides  $S$  la función  $D(x, y, z)$  alcanza un mínimo absoluto en algún punto  $\mathbf{q} = (x_0, y_0, z_0)$  y calculando ese punto. Como la distancia de  $\mathbf{q}$  al plano  $P$  se alcanza en algún  $\mathbf{p} \in P$  (porque  $P$  es cerrado), quedará justificado que la distancia  $d(S, P) = \|\mathbf{q} - \mathbf{p}\|_2$  se alcanza en los puntos  $\mathbf{p} \in P$ ,  $\mathbf{q} \in S$ .

La función  $D(x, y, z)$  alcanza un mínimo absoluto sobre el paraboloides  $z = x^2 + y^2$  si y sólo si la función de dos variables reales  $f(x, y) = |x + 2y - (x^2 + y^2) - 4|$  alcanza un mínimo absoluto en algún punto. Completando cuadrados se observa que  $x^2 + y^2 - x - 2y + 4 = (x - 1/2)^2 + (y - 1)^2 + 11/4 > 0$  para todo  $(x, y) \in \mathbb{R}^2$ , luego

$$f(x, y) = x^2 + y^2 - x - 2y + 4 = (x^2 + y^2) \left( \frac{4 - x - 2y}{x^2 + y^2} + 1 \right)$$

es una función continua en  $\mathbb{R}^2$  que cumple  $\lim_{\|(x,y)\|_2 \rightarrow +\infty} f(x, y) = +\infty$ , de donde se sigue, por un razonamiento estandar, que  $f$  alcanza un mínimo absoluto en algún  $(x_0, y_0)$ , que debe satisfacer las ecuaciones  $0 = D_1 f(x, y) = D_2 f(x, y)$ . Como estas ecuaciones sólo tienen la solución  $(1/2, 1)$ , se sigue que  $\mathbf{q} = (x_0, y_0, z_0) = (1/2, 1, 5/4)$  es el punto del paraboloides  $S$  que está más cerca del plano  $P$ . Finalmente, el cálculo del punto  $\mathbf{p} \in P$  que está más cerca de  $\mathbf{q}$  es un asunto elemental que dejamos al cuidado del lector (se puede hacer con los procedimientos usuales de geometría analítica, o con los métodos que estamos exponiendo). ■

**Ejercicio 5.41** *En cada uno de los siguientes casos calcule los extremos absolutos de  $f$  sobre el compacto que se indica:*

- a)  $f(x, y) = x - x^2 - y^2$   $K = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$  ;
- b)  $f(x, y) = \text{sen}(xy)$   $K = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$  ;
- c)  $f(x, y) = x^2 + 5y^3 - y$   $K = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1, y \geq 0\}$  ;

SOLUCIÓN

a) Un punto  $\mathbf{p} \in K$  donde  $f|_K$  alcanza un extremo absoluto puede estar en el interior de  $K$  o en su frontera,  $C = \{(x, y) : x^2 + y^2 = 1\}$ .

- Si ocurre lo primero,  $\mathbf{p}$  debe satisfacer las ecuaciones  $0 = D_1 f(x, y) = 1 - 2x$ ;

$0 = D_2f(x, y) = -2y$ , que sólo tienen una solución  $\mathbf{a} = (1/2, 0) \in K$ .

- Si ocurre lo segundo  $f|_C$  alcanza un extremo absoluto en  $\mathbf{p}$ . Los extremos absolutos de  $f|_C$  son inmediatos en el caso que nos ocupa: Cuando  $(x, y) \in C$  se cumple  $f(x, y) = x - 1$  y es obvio que el mínimo (resp. el máximo) de  $x - 1$  sobre la circunferencia  $C$  se alcanza en  $\mathbf{b} = (-1, 0)$ , (resp.  $\mathbf{c} = (1, 0)$ ).

En definitiva,  $f|_K$  alcanza sus extremos absolutos en algunos de los puntos  $\mathbf{a}, \mathbf{b}, \mathbf{c}$ . Como  $f(\mathbf{a}) = 1/4$ ,  $f(\mathbf{b}) = -2$ , y  $f(\mathbf{c}) = 0$ , concluimos que el máximo absoluto es  $1/4 = f(\mathbf{a})$ , y el mínimo absoluto es  $-2 = f(\mathbf{b})$ .

b) El problema de calcular los extremos absolutos de  $f(x, y) = \text{sen}(xy)$  sobre  $K$  se simplifica calculando previamente los extremos absolutos de  $g(x, y) = xy$  sobre  $K$

$$\alpha = g(\mathbf{p}) = \min\{g(x, y) : (x, y) \in K\}; \quad \beta = g(\mathbf{q}) = \max\{g(x, y) : (x, y) \in K\}$$

que se alcanzarán en sendos puntos  $\mathbf{p}, \mathbf{q} \in K$ . En efecto, como  $g(K)$  es conexo y compacto (por serlo  $K$ ) también lo será  $g(K)$ , luego  $g(K) = [\alpha, \beta]$ . Entonces los extremos absolutos de  $f(x, y) = \text{sen } g(x, y)$  sobre  $K$  serán los extremos absolutos de la función  $\text{sen } t$  en el intervalo  $[\alpha, \beta]$ .

Los puntos de  $K$  donde  $g|_K$  alcanza los extremos absolutos están en el interior de  $K$ , o en su frontera  $C = \{(x, y) : x^2 + y^2 = 1\}$ .

- Los que estén en el interior deben satisfacer las ecuaciones

$$0 = D_1g(x, y) = y; \quad 0 = D_2g(x, y) = x$$

Como su solución trivial  $\mathbf{a} = (0, 0)$  es un punto interior a  $K$ , este será el primer candidato a punto de extremo absoluto (para la función  $g|_K$ ).

- Los que estén en  $C$  serán puntos donde  $g|_C$  alcance extremos absolutos. En el caso que nos ocupa estos extremos se pueden calcular fácilmente usando la parametrización de  $C$ ,  $x = \cos t$ ,  $y = \text{sen } t$ . Con los recursos habituales del cálculo diferencial de funciones de una variable se obtiene que los extremos absolutos de  $\varphi(t) = g(\cos t, \text{sen } t) = \text{sen } t \cos t$  sobre  $[0, 2\pi]$  son

$$\alpha = \varphi(3\pi/4) = \varphi(7\pi/4) = -1/2; \quad \beta = \varphi(\pi/4) = \varphi(5\pi/4) = 1/2$$

Se sigue de esto que el máximo absoluto de  $g|_C$  es

$$\frac{1}{2} = g(\mathbf{b}) = g(-\mathbf{b}) \quad \text{donde } \mathbf{b} = (\sqrt{2}/2, \sqrt{2}/2)$$

y que el mínimo absoluto de  $g|_C$  es

$$-\frac{1}{2} = g(\mathbf{c}) = g(-\mathbf{c}), \quad \text{donde } \mathbf{c} = (+\sqrt{2}/2, -\sqrt{2}/2)$$

Como  $g|_K$  alcanza sus extremos absolutos en puntos de la terna  $\mathbf{a}, \mathbf{b}, \mathbf{c}$ , concluimos que su mínimo absoluto es  $-1/2 = g(\mathbf{c})$ , y su máximo absoluto es  $1/2 = g(\mathbf{b})$ .

Según lo indicado al principio, el mínimo y el máximo absoluto de  $f|_K$  son, respectivamente,  $-\text{sen}(1/2) = f(\mathbf{c})$ ,  $\text{sen}(1/2) = f(\mathbf{b})$ , pues estos son los valores

extremos de  $\sin t$  en el intervalo  $[-1/2, 1/2]$ .

c) Los puntos de  $K$  donde  $f|_K$  alcanza los extremos absolutos estarán en el interior de  $K$  o en su frontera,  $\partial K = S \cup L$  donde

$$L = \{(x, 0) : -1 \leq x \leq 1\}; \quad S = \{(x, y) : y \geq 0, x^2 + y^2 = 1\}$$

- De estos puntos, los que estén en el interior de  $K$ , deben satisfacer las ecuaciones

$$0 = D_1 f(x, y) = 2x; \quad 0 = D_2 f(x, y) = 15y^2 - 1$$

que sólo tienen una solución  $\mathbf{a} = (0, 1/\sqrt{15})$  en el interior de  $K$ .

- Los otros puntos donde  $f$  puede alcanzar extremos absolutos son los puntos donde  $f|_L$  y  $f|_S$  alcanzan sus extremos absolutos.

Los extremos absolutos de  $f|_L$ , coinciden con los de  $f(x, 0) = x^2$  sobre  $[-1, 1]$ , que son  $0 = f(0, 0)$ , y  $1 = f(1, 0) = f(-1, 0)$ .

Los extremos absolutos de  $f|_S$ , coinciden con los extremos absolutos del polinomio  $p(y) = f(1 - y^2, y) = 5y^3 - y^2 - y + 1$  sobre el intervalo  $[0, 1]$ . Con las técnicas habituales para funciones de una variable se obtienen los valores extremos de  $p$  en  $[0, 1]$ , que son  $p(1/3) = 20/27$  (mínimo), y  $p(1) = 4$  (máximo). En  $S$  hay dos puntos con  $y = 1/3$ , que son  $(\pm 2\sqrt{2}/3, 1/3)$ , y un punto con  $y = 1$ , que es  $(0, 1)$ .

En definitiva, los extremos absolutos de  $f|_K$  se alcanzan en algunos de los puntos de la siguiente lista de candidatos:

$$(0, 1/\sqrt{15}), (-1, 0), (0, 0), (1, 0), (2\sqrt{2}/3, 1/3), (-2\sqrt{2}/3, 1/3), (0, 1)$$

y evaluando  $f$  en estos puntos, se obtienen los extremos absolutos de  $f|_K$ :

$$-2/(3\sqrt{15}) = f(0, 1/\sqrt{15}); \quad 4 = f(0, 1).$$

■

**Ejercicio 5.42** Sean  $(E, \|\cdot\|), (F, \|\cdot\|)$  espacios normados. Una aplicación  $\mathbf{f} : \Omega \rightarrow F$ , definida en un abierto  $\Omega \subset E$ , es diferenciable en  $\mathbf{a} \in \Omega$  si y sólo si se cumplen las dos condiciones siguientes:

a) Para cada  $\mathbf{v} \in E$  existe  $D_{\mathbf{v}}\mathbf{f}(\mathbf{a})$  y la aplicación  $\mathbf{v} \rightarrow D_{\mathbf{v}}\mathbf{f}(\mathbf{a})$  es lineal continua.

b)  $\lim_{t \rightarrow 0} \frac{\mathbf{f}(\mathbf{a} + t\mathbf{v}) - \mathbf{f}(\mathbf{a})}{t} = D_{\mathbf{v}}\mathbf{f}(\mathbf{a})$  uniformemente en  $\{\mathbf{v} \in E : \|\mathbf{v}\| = 1\}$ .

(es decir, la condición b) es lo que le falta a una aplicación que cumple a) para ser diferenciable).

SOLUCIÓN

Si  $\mathbf{f}$  es diferenciable en  $\mathbf{a}$ , según 5.14, se cumple a). Sea  $L = d\mathbf{f}(\mathbf{a})$  y  $\boldsymbol{\alpha}(\mathbf{h}) = \mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a}) - L(\mathbf{h})$  el error que se comete al aproximar el incremento  $\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a})$



mediante el incremento de la diferencial  $L(\mathbf{h}) = L(\mathbf{a} + \mathbf{h}) - L(\mathbf{a})$ . Según la definición de diferencial  $\lim_{\mathbf{h} \rightarrow 0} \alpha(\mathbf{h}) / \|\mathbf{h}\| = 0$ , es decir, para cada  $\epsilon > 0$  existe  $\delta > 0$  tal que  $\|\mathbf{h}\| < \delta \Rightarrow \|\alpha(\mathbf{h})\| < \epsilon \|\mathbf{h}\|$ . Si  $|t| < \delta$ , para cada  $\mathbf{v} \in E$ , con  $\|\mathbf{v}\| = 1$ , se cumple  $\|\alpha(t\mathbf{v})\| < \epsilon|t|$ , luego  $\lim_{t \rightarrow 0} \alpha(t\mathbf{v})/t = 0$  uniformemente en  $\{\mathbf{v} \in E : \|\mathbf{v}\| = 1\}$ . Como  $\alpha(t\mathbf{v}) = \mathbf{f}(\mathbf{a} + t\mathbf{v}) - \mathbf{f}(\mathbf{a}) - tD_{\mathbf{v}}\mathbf{f}(\mathbf{a})$  se obtiene b).

El recíproco es inmediato: En virtud de b) la aplicación lineal continua  $L : E \rightarrow F$ ,  $L(\mathbf{v}) = D_{\mathbf{v}}\mathbf{f}(\mathbf{a})$  proporcionada por a) satisface la definición de diferencial.

■

## 5.7. Ejercicios propuestos

◇ 5.7.1 Calcule, en un punto genérico, las derivadas parciales de las funciones

a)  $x^{y^z}$ , definida en  $\{(x, y, z) : x > 0, y > 0\}$ .

b)  $\operatorname{sen}(x^y + y^z + z^x)$ , definida en  $\{(x, y, z) : x > 0, y > 0, z > 0\}$ .

c)  $\operatorname{sen}(x \operatorname{sen}(y \operatorname{sen} z))$ , definida en todo  $\mathbb{R}^3$ .

d)  $\int_0^{x^2 y^3} g(t) dt + \int_{x^2}^{x^2+z^2} g(t) dt$ , donde  $g : \mathbb{R} \rightarrow \mathbb{R}$  es continua.

◇ 5.7.2 Demuestre que la función

$$f(x, y) = \frac{x^3 + 2xy^2}{x^2 + y^2} \quad \text{si } (x, y) \neq (0, 0), \quad f(0, 0) = 0$$

es uniformemente continua en  $\mathbb{R}^2$ .

◇ 5.7.3 Demuestre que la función  $f(x, y) = 1/(1 + x^2 + y^2)$  es uniformemente continua en todo  $\mathbb{R}^2$ . Estudie si sus derivadas parciales alcanzan extremos absolutos en todo el plano y obtenga la mejor cota de  $|D_1 f(x, y)|$ ,  $|D_2 f(x, y)|$ .

◇ 5.7.4 Demuestre que la función

$$f(x, y) = \frac{xy}{(x+y)(1+x)(1+y)}$$

es uniformemente continua en el abierto  $\{(x, y) : x > 0, y > 0\}$  y alcanza un máximo absoluto en este abierto. Obténgalo.

◇ 5.7.5 Determine los valores de  $n \in \mathbb{N}$  para los que es uniformemente continua la función  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ , definida por

$$f(x, y) = \frac{y^n}{x^2 + y^2} \quad \text{si } (x, y) \neq (0, 0); \quad f(0, 0) = 0 \quad (n \in \mathbb{N})$$

◇ 5.7.6 Sea  $f : \Omega \rightarrow \mathbb{R}$  definida en el abierto  $\Omega = \mathbb{R}^2 \setminus \{(x, 0) : x \geq 0\}$  tal que en cada  $(x, y) \in \Omega$  existe y es nula la derivada parcial  $D_1 f(x, y) = 0$ . Demuestre que  $f$  no depende de la primera variable. ¿Se obtiene un resultado análogo para funciones  $f : \Omega \rightarrow \mathbb{R}$  tales que en cada  $(x, y) \in \Omega$  existe y es nula la derivada parcial  $D_2 f(x, y) = 0$ ?

◇ 5.7.7 Obtenga los vectores  $\mathbf{v} \in \mathbb{R}^3$  para los que existe la derivada  $D_{\mathbf{v}} f(1, -1, 0)$  de la función  $f(x, y, z) = |x + y + z|$ .

◇ 5.7.8 Sea  $A \subset \mathbb{R}^n$  un abierto acotado y  $f : \bar{A} \rightarrow \mathbb{R}$  una función continua, diferenciable en cada  $\mathbf{x} \in A$ , tal que  $f(\mathbf{x}) = 0$  para todo  $\mathbf{x} \in \partial A$ . Demuestre que existe  $\mathbf{a} \in A$  tal que  $\nabla f(\mathbf{a}) = \mathbf{0}$ .

◇ **5.7.9** Calcule los extremos absolutos de  $x^2 + y^2 - xy + x + y$  sobre el compacto

$$\{(x, y) \in \mathbb{R}^2 : x \leq 0, y \leq 0, x + y + 3 \geq 0\}$$

◇ **5.7.10** Obtenga la mínima distancia entre la recta  $y - x + 5 = 0$  y la parábola  $y = x^2$ .

◇ **5.7.11** Obtenga la distancia de  $(0, 0, 0)$  a cada una de las superficies

$$x^2 + y^2 - z^2 + 2xy = 16; \quad z^2 - xy - 1 = 0$$

◇ **5.7.12** Determine las dimensiones de una caja rectangular sin tapa, con superficie de  $16 \text{ m}^2$ , que encierre un volumen máximo. Justifique la existencia de la caja de volumen máximo.

◇ **5.7.13** Calcule el máximo de  $\log x + \log y + 3 \log z$ , sobre la porción de esfera  $x^2 + y^2 + z^2 = 5r^2$  en la que  $x > 0, y > 0, z > 0$ . Aplique el resultado para demostrar que si  $a, b, c > 0$  entonces

$$abc^3 \leq 27 \left( \frac{a + b + c}{5} \right)^5$$

◇ **5.7.14** Utilizando la definición compruebe que  $f(x, y, z) = (x - 1)^3 yz + x^2 + 2y^2$  es diferenciable en  $(1, 0, 0)$  y obtenga la diferencial  $df(1, 0, 0)$ .

◇ **5.7.15** Sea  $f : (a, b) \rightarrow \mathbb{R}$  derivable en  $x_0 \in (a, b)$ . Demuestre que la función de tres variables  $F(x, y, z) = f(x)$ , definida en  $A = \{(x, y, z) : a < x < b\}$ , es diferenciable en  $(x_0, y, z)$ . Obtenga  $dF(x_0, y, z)$ .

◇ **5.7.16** Se supone que  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  verifica  $|f(\mathbf{x})| \leq \|\mathbf{x}\|^p$  para todo  $\mathbf{x} \in \mathbb{R}^n$ , donde  $p > 1$ . Demuestre que  $f$  es diferenciable en  $0$  y obtenga  $df(0)$ .

◇ **5.7.17** Si  $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$  verifica  $\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| \leq \|\mathbf{x} - \mathbf{y}\|^2$  para todo  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ , demuestre que  $\mathbf{f}$  es constante.

◇ **5.7.18** Sea  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  diferenciable en  $(0, 0)$ . Obtenga  $df(0, 0)$  sabiendo que  $D_{\mathbf{u}}f(0, 0) = 1$  y  $D_{\mathbf{v}}f(0, 0) = 0$ , donde  $\mathbf{u} = (1, 1)$  y  $\mathbf{v} = (-1, 1)$ .

◇ **5.7.19** Se consideran las siguientes funciones  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ :

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

b)  $f(x, y) = x \operatorname{sen} \frac{1}{x^2 + y^2}$  si  $(x, y) \neq (0, 0)$ ,  $f(0, 0) = 0$ .

$$c) f(x, y) = \frac{x|y|}{\sqrt{x^2 + y^2}} \quad \text{si } (x, y) \neq (0, 0), \quad f(0, 0) = 0.$$

$$d) f(x, y) = (x^2 + y^2) \operatorname{sen} \frac{1}{\sqrt{x^2 + y^2}} \quad \text{si } (x, y) \neq (0, 0), \quad f(0, 0) = 0.$$

$$e) f(x, y) = xy^2 \operatorname{sen}(1/y) \quad \text{si } y \neq 0, \quad f(x, 0) = 0.$$

$$f) f(x, y) = (x + y)^n \operatorname{sen} \frac{1}{\sqrt{x^2 + y^2}} \quad \text{si } (x, y) \neq (0, 0), \quad f(0, 0) = 0.$$

$$g) f(x, y) = \frac{x^n}{x^2 + y^2} \quad \text{si } (x, y) \neq (0, 0), \quad f(0, 0) = 0. \quad (n \in \mathbb{N})$$

Compruebe las afirmaciones que se hacen sobre de cada una de ellas:

- $f$  no es diferenciable en  $(0, 0)$  porque no es continua en  $(0, 0)$ .
- $f$  no es diferenciable en  $(0, 0)$  porque no existe una derivada parcial en  $(0, 0)$ .
- $f$  es continua en  $(0, 0)$  y existe la derivada  $D_{\mathbf{u}}f(0, 0)$  según cualquier vector  $\mathbf{u} \in \mathbb{R}^2$  pero  $f$  no es diferenciable en  $(0, 0)$ .
- $f$  es diferenciable en  $(0, 0)$  y sus derivadas parciales no son continuas en  $(0, 0)$ .
- $f$  es diferenciable en  $(0, 0)$  porque se cumple la condición suficiente de diferenciability (una de las dos derivadas parciales es continua en  $(0, 0)$ ).
- $f$  es continua; es diferenciable si y sólo si  $n \geq 2$ ; tiene derivadas parciales continuas si y sólo si  $n \geq 3$ .
- $f$  es diferenciable en  $(0, 0)$  si y sólo si  $n > 3$ .

◇ **5.7.20** Estudie, según los valores de  $n \in \mathbb{N}$ , la continuidad y diferenciability en  $(0, 0)$  de la función  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  definida por

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

◇ **5.7.21** Sea  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  definida por  $f(x, y) = \frac{y(x^2 - y^2)}{x^2 + y^2}$  si  $(x, y) \neq (0, 0)$ ,  $f(0, 0) = 0$ .

¿Es  $f$  diferenciable en  $(0, 0)$ ? ¿Es de clase  $C^1(\mathbb{R}^2)$ ? ¿Es uniformemente continua?

◇ **5.7.22** Sea  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$  una función derivable de la forma  $\varphi(t) = \alpha t + t^2 A(t)$  donde  $\lim_{t \rightarrow 0} A(t) = 1$ . Se define  $f(x, y) = \varphi(xy)/(xy)$  si  $xy \neq 0$ ;  $f(x, y) = \alpha$  si  $xy = 0$ . Justifique que  $f$  es diferenciable en todo punto.

◇ **5.7.23** Sea  $f(x, y) = g(x^2 + y^2)$  donde  $g : [0, +\infty) \rightarrow \mathbb{R}$  es derivable dos veces en  $(0, +\infty)$  y  $g(0) = 0$ . Demuestre las siguientes afirmaciones

- Si existe  $D_1 f(0, 0) = A$  entonces para todo  $\mathbf{u} \in \mathbb{R}^2$  existe  $D_{\mathbf{u}}f(0, 0)$  y obtenga su valor en función de  $A$ .
- $f$  es diferenciable en  $(0, 0)$  si y sólo si existe y es nula la derivada  $D_1 f(0, 0)$ .

c)  $f$  es de clase  $C^1$  si y sólo si  $\lim_{r \rightarrow 0} rg'(r^2) = 0$ .

◇ **5.7.24** Determine los valores de  $p > 0$  para los que

$$f(x, y) = (x^2 + y^2)^p \operatorname{sen} \left( \frac{1}{x^2 + y^2} \right) \quad \text{si } (x, y) \neq (0, 0), \quad f(0, 0) = 0$$

es diferenciable. ¿Para qué valores de  $p$  es de clase  $C^1$ ?

◇ **5.7.25** Estudie, según los valores del número real  $\alpha > 0$  la diferenciable en  $(0, 0)$  de la función

$$f(x, y) = \frac{x}{(x^2 + y^2)^\alpha} \operatorname{sen}(\sqrt{x^2 + y^2}) \quad \text{si } (x, y) \neq (0, 0); \quad f(0, 0) = 0$$

◇ **5.7.26** Sea  $f : \mathbb{R} \rightarrow \mathbb{R}$  derivable dos veces en  $a \in \mathbb{R}$  y  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  definida por

$$F(x, y) = \frac{f(y) - f(x)}{y - x} \quad \text{si } x \neq y, \quad F(x, x) = f'(x)$$

Demuestre que  $F$  es diferenciable en  $(a, a)$  y obtenga  $dF(a, a)$ . (Indicación: Considere el desarrollo de Taylor de  $f$  en el punto  $a$ )

◇ **5.7.27** Sean  $f, g : \Omega \rightarrow \mathbb{R}$  definidas en un abierto  $\Omega \subset \mathbb{R}^n$ .

a) Si  $g$  es continua en  $\mathbf{a} \in \Omega$  y  $f$  es diferenciable en  $\mathbf{a}$  con  $f(\mathbf{a}) = 0$ , demuestre que  $fg$  es diferenciable en  $\mathbf{a}$  (aunque  $g$  no sea diferenciable en  $\mathbf{a}$ ).

b) Compruebe que es cierto lo que se afirma en el apartado anterior cuando  $\Omega = \mathbb{R}^2$ ,  $\mathbf{a} = (0, 0)$ ,

$$f(x, y) = e^x \operatorname{sen} y; \quad g(x, y) = x^3 / (x^2 + y^2) \quad \text{si } (x, y) \neq (0, 0), \quad g(0, 0) = 0$$

(Estudie la continuidad y diferenciable de  $g$  en  $(0, 0)$ ).

◇ **5.7.28** Sean  $(E, \|\cdot\|)$ ,  $(F, \|\cdot\|)$  espacios normados reales y  $\mathbf{f} : E \rightarrow F$  una aplicación homogénea:  $\mathbf{f}(t\mathbf{x}) = t\mathbf{f}(\mathbf{x})$  para todo  $\mathbf{x} \in E$  y todo  $t > 0$ . Demuestre que  $\mathbf{f}$  es diferenciable en  $0$  si y sólo si es una aplicación lineal continua.

◇ **5.7.29** Demuestre que en un espacio normado  $(E, \|\cdot\|)$  la norma nunca es diferenciable en  $\mathbf{0}$ . Si la norma procede de un producto escalar, demuestre que es diferenciable en cada  $\mathbf{a} \neq \mathbf{0}$  y obtenga su diferencial.

◇ **5.7.30** Sea  $(E, \|\cdot\|)$  un espacio normado real cuya norma procede de un producto escalar  $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x} | \mathbf{x} \rangle}$ . Si  $L : E \rightarrow E$  es una aplicación lineal continua, demuestre que  $f(\mathbf{x}) = \langle L(\mathbf{x}) | \mathbf{x} \rangle$  es diferenciable en todo  $\mathbf{a} \in E$  y obtenga su diferencial  $df(\mathbf{a})$ .

◇ **5.7.31** Dadas dos aplicaciones lineales  $A, B : \mathbb{R}^n \rightarrow \mathbb{R}^m$ , se considera la función  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  definida por  $f(\mathbf{x}) = \langle A(\mathbf{x}) | B(\mathbf{x}) \rangle$ . Demuestre que  $f$  es diferenciable y obtenga su diferencial  $df(\mathbf{a})$  en un punto genérico  $\mathbf{a} \in \mathbb{R}^n$ .

◇ **5.7.32** Sea  $\Omega$  un abierto de un espacio normado  $(E, \|\cdot\|)$  y  $\mathbf{f} : \Omega \rightarrow E$  una aplicación de la forma  $\mathbf{f}(\mathbf{x}) = \alpha(\mathbf{x})\mathbf{x}$  donde  $\alpha : \Omega \rightarrow \mathbb{R}$  es diferenciable en  $\mathbf{a} \in \Omega$ . Demuestre que  $\mathbf{f}$  también es diferenciable en  $\mathbf{a}$  y obtenga una fórmula para  $d\mathbf{f}(\mathbf{a})$ .

*Aplicación:* Si la norma de  $(E, \|\cdot\|)$  procede de un producto escalar, obtenga que  $f(\mathbf{x}) = \|\mathbf{x}\| \mathbf{x}$  es diferenciable en  $\mathbf{a} \in E \setminus \{\mathbf{0}\}$  y calcule  $df(\mathbf{a})$ .

◇ **5.7.33** En un espacio normado  $(E, \|\cdot\|)$  la norma procede de un producto escalar. Demuestre que  $f(\mathbf{x}) = \frac{\mathbf{x}}{\|\mathbf{x}\|}$  es diferenciable en cada  $\mathbf{a} \in E \setminus \{\mathbf{0}\}$  y calcule  $df(\mathbf{a})$ .

◇ **5.7.34** Se dice que  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es homogénea de grado  $m$  si  $f(t\mathbf{x}) = t^m f(\mathbf{x})$  para todo  $\mathbf{x} \in \mathbb{R}^n \setminus \{0\}$  y todo  $t > 0$ . Si  $f$  es diferenciable, demuestre que son equivalentes

i)  $f$  es homogénea de grado  $m$ .

ii)  $\langle \nabla f(\mathbf{x}) | \mathbf{x} \rangle = m f(\mathbf{x})$  para todo  $\mathbf{x} \in \mathbb{R}^n$ .

◇ **5.7.35** Si  $f : \Omega \rightarrow \mathbb{R}$  es diferenciable en  $\Omega = \mathbb{R}^n \setminus \{0\}$ , demuestre que son equivalentes:

i) Existe una función derivable  $g : (0, +\infty) \rightarrow \mathbb{R}$  tal que  $f(\mathbf{x}) = g(\|\mathbf{x}\|_2)$  para todo  $\mathbf{x} \in \Omega$ .

ii) Existe una función  $\alpha : \Omega \rightarrow \mathbb{R}$  tal que  $\nabla f(\mathbf{x}) = \alpha(\mathbf{x})\mathbf{x}$  para todo  $\mathbf{x} \in \Omega$ .

Si se cumple i) ¿qué relación hay entre  $g$  y  $\alpha$ ?

◇ **5.7.36** Sea  $\text{Arctg} : \mathbb{R} \rightarrow (-\pi/2, \pi/2)$  la rama principal de la función  $\arctg$ . En el abierto  $\Omega = \mathbb{R}^2 \setminus \{(0, 0)\}$  se define la función  $f : \Omega \rightarrow \mathbb{R}$

$$f(x, y) = \text{Arctg}(y/x^2) \quad \text{si } x \neq 0.$$

$$f(0, y) = -\pi/2 \quad \text{si } y < 0; \quad f(0, y) = \pi/2 \quad \text{si } y > 0$$

Demuestre que  $f$  es de clase  $C^1(\Omega)$  y estudie su continuidad uniforme en  $\Omega$ .

◇ **5.7.37** Si  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es de clase  $C^1$  y  $f(0) = 0$ , demuestre que existen funciones continuas,  $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $1 \leq i \leq n$ , tales que  $f(\mathbf{x}) = \sum_{i=1}^n x_i g_i(\mathbf{x})$  para todo  $\mathbf{x} \in \mathbb{R}^n$ .

◇ **5.7.38** Demuestre que es de clase  $C^1$  la función

$$f(x, y) = \frac{x^4 + \text{sen } y^4}{x^2 + y^2} \quad \text{si } (x, y) \neq (0, 0), \quad f(0, 0) = 0$$

◇ **5.7.39** Se supone que  $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ , es diferenciable en  $\mathbf{a}$ , que  $g : \mathbb{R}^m \rightarrow \mathbb{R}$  es diferenciable en  $\mathbf{b} = \mathbf{f}(\mathbf{a})$  y que  $\gamma : [-1, 1] \rightarrow \mathbb{R}^n$  es derivable en 0. Si  $\gamma'(0) = \mathbf{u}$  y  $\gamma(0) = \mathbf{a}$  demuestre que  $\varphi(t) = g(\mathbf{f}(\gamma(t)))$  es derivable en  $t = 0$  y que  $\varphi'(0) = \langle \nabla g(\mathbf{b}) | D_{\mathbf{u}} \mathbf{f}(\mathbf{a}) \rangle$

◇ **5.7.40** Escriba, en un punto genérico, la matriz jacobiana de las aplicaciones

$$a) \mathbf{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^3, \quad \mathbf{f}(x, y) = (\text{sen}(xy), \text{sen}(x \text{sen } y), x^4);$$

b)  $\mathbf{F} = f \circ \mathbf{g}$  donde  $f(x, y, z) = x^2 + y^2 + z^2$ ,  $\mathbf{g}(u, v) = (u - v, u - v, u^2 - v^2)$ .

◇ **5.7.41** Sea  $\mathbf{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ , la aplicación definida por

$$\mathbf{f}(x, y) = \left(x + \cos y, \frac{x^4}{x^2 + y^2}\right) \text{ si } (x, y) \neq (0, 0), \quad \mathbf{f}(0, 0) = (1, 0).$$

Justifique que  $\mathbf{F} = \mathbf{f} \circ \mathbf{f} + \mathbf{f}$  es diferenciable en  $(0, 0)$  y calcule  $d\mathbf{F}(0, 0)$ .

◇ **5.7.42** Sean  $\mathbf{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ ,  $\mathbf{g} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  definidas por

$$\mathbf{f}(x, y) = (xy, e^x \cos y, e^{xy}), \quad \mathbf{g}(u, v) = (u + e^v, u^3 v^2)$$

Obtenga que  $\mathbf{F} = \mathbf{f} \circ \mathbf{g}$  es diferenciable en todo punto y calcule  $d\mathbf{F}(0, 0)$ .

◇ **5.7.43** ¿En qué puntos es diferenciable la función  $f(x, y) = \sqrt{1 + |xy|}$ ? Demuestre que  $|D_{\mathbf{u}}f(1, -1)| \leq \|\mathbf{u}\|_2 / 2$  para todo  $\mathbf{u} \in \mathbb{R}^2$ .

◇ **5.7.44** Si  $f(x, y, z) = x^2 - y^2 + xyz^2 - zx$ , obtenga el máximo valor de las derivadas direccionales  $\{D_{\mathbf{u}}f(1, 2, 3) : \|\mathbf{u}\|_2 = 1\}$

◇ **5.7.45** Obtenga  $a, b, c \in \mathbb{R}$  con la condición de que la derivada direccional de  $f(x, y, z) = axy^2 + byz + cz^2x^3$  en el punto  $(1, 2, -1)$  alcance un valor máximo, igual a 64, según la dirección del vector  $(0, 0, 1)$ .

◇ **5.7.46** Sea  $f(x, y) = \frac{x^3 + 2xy^2}{x^2 + y^2}$  si  $(x, y) \neq (0, 0)$ ,  $f(0, 0) = 0$

Calcule el máximo valor de la derivada direccional  $D_{\mathbf{u}}f(0, 0)$ ,  $\|\mathbf{u}\|_2 = 1$ .

◇ **5.7.47** Sea  $\Omega \subset \mathbb{R}^n$  abierto y  $\mathbf{a} \in \Omega$ . Si  $f : \Omega \rightarrow \mathbb{R}$  es continua en  $\mathbf{a}$ , diferenciable en cada  $\mathbf{x} \in \Omega \setminus \{\mathbf{a}\}$ , y existen los límites  $\lim_{\mathbf{x} \rightarrow \mathbf{a}} D_k f(\mathbf{x}) = A_k$ ,  $1 \leq k \leq n$ , demuestre que  $f$  es diferenciable en  $\mathbf{a}$ .

(Indic. Considere  $\varphi(t) = f(\mathbf{a} + t\mathbf{h}) - t \sum_{k=1}^n A_k h_k$ )

◇ **5.7.48** Sean  $(E, \|\cdot\|)$ ,  $(F, \|\cdot\|)$  espacios normados,  $\mathbf{f} : \Omega \rightarrow F$  una función continua, definida en un abierto  $\Omega \subset E$  y  $\mathbf{a} \in \Omega$ . Si  $\mathbf{f}$  diferenciable en cada  $\mathbf{x} \in \Omega \setminus \{\mathbf{a}\}$ , y existe  $\lim_{\mathbf{x} \rightarrow \mathbf{a}} d\mathbf{f}(\mathbf{x}) = L$ , demuestre que  $\mathbf{f}$  es diferenciable en  $\mathbf{a}$  y que  $d\mathbf{f}(\mathbf{a}) = L$ .

◇ **5.7.49** Sean  $(E, \|\cdot\|)$ ,  $(F, \|\cdot\|)$  espacios normados y  $\mathbf{f} : \Omega \rightarrow F$  una función continua definida en un abierto  $\Omega \subset E$ . Se supone que para cada  $\mathbf{x} \in \Omega$  y cada  $\mathbf{u} \in E$  existen las derivadas  $D_{\mathbf{u}}\mathbf{f}(\mathbf{x}) = A(\mathbf{x})\mathbf{u}$  donde  $A : \Omega \rightarrow \mathcal{L}(E, F)$  es continua. Demuestre que  $f$  es diferenciable y que  $d\mathbf{f}(\mathbf{x}) = A(\mathbf{x})$  para todo  $\mathbf{x} \in \Omega$ .

◇ **5.7.50** Si  $\Omega \subset \mathbb{R}^n$  es un abierto convexo y  $f : \Omega \rightarrow \mathbb{R}$  es diferenciable con derivadas parciales acotadas demuestre que  $f$  es uniformemente continua.

◇ **5.7.51** Sea  $\Omega \subset \mathbb{R}^n$  abierto convexo y  $(E, \|\cdot\|)$  un espacio normado completo. Si  $\mathbf{f} : \Omega \rightarrow (E, \|\cdot\|)$  es de clase  $C^1(\Omega, E)$  y  $\|d\mathbf{f}(\mathbf{x})\| \leq M$  para todo  $\mathbf{x} \in \Omega$ , demuestre que  $\mathbf{f}$  admite una extensión continua  $\hat{\mathbf{f}} : \hat{\Omega} \rightarrow E$ .

◇ **5.7.52** Escriba las ecuaciones de los planos tangentes a las siguientes superficies en los puntos que se indican:

a) Superficie de ecuación  $x^2 + y^2 + 4z^2 = 12$ , en el punto  $\mathbf{p} = (2, 2, 1)$ ;

b) Superficie de ecuaciones paramétricas

$$x(r, t) = r \cos t, \quad y(r, t) = r \sin t, \quad z(r, t) = t, \quad 0 < r, \quad 0 < t < 2\pi$$

en el punto  $\mathbf{p} = (0, 2, \pi/2)$  imagen de  $(2, \pi/2)$ .

◇ **5.7.53** Se considera la función  $f : \mathbb{R}^3 \rightarrow \mathbb{R}$  definida por

$$f(x, y, z) = \frac{x}{\sqrt{x^2 + y^2}} \sin(x^2 + y^2) + e^z \quad \text{si } (x, y, z) \neq (0, 0, z), \quad f(0, 0, z) = e^z$$

Compruebe que  $f$  es diferenciable en  $(0, 0, 0)$  y obtenga la derivada  $D_{\mathbf{v}}f(0, 0, 0)$  según un vector unitario  $\mathbf{v}$  normal a la superficie  $x + y^2 + 2z = 0$  en el punto  $(0, 0, 0)$ .

◇ **5.7.54** Demuestre que la función  $f(x, y, z) = \int_y^z e^{-x^2 t^2} dt$  es diferenciable en todos los puntos. Calcule la derivada  $D_{\mathbf{v}}f(\mathbf{p})$  donde  $\mathbf{p} = (2, 2, 1)$  y  $\mathbf{v}$  es un vector unitario normal al elipsoide  $2x^2 + y^2 + z^2 = 13$ , en el punto  $\mathbf{p}$ , que apunta hacia el exterior del elipsoide.

◇ **5.7.55** Sea  $f : \mathbb{R}^3 \rightarrow \mathbb{R}$  definida por

$$f(x, y, z) = \frac{x^2 z^2}{x^2 + y^2} \quad \text{si } (x, y) \neq (0, 0), \quad f(0, 0, z) = 0.$$

Obtenga  $D_{\mathbf{v}}f(1, 1, \sqrt{2})$ , donde  $\mathbf{v}$  es un vector unitario tangente a la curva

$$x^2 + y^2 = 2x, \quad x^2 + y^2 + z^2 = 4$$

en el punto  $(1, 1, \sqrt{2})$ .

◇ **5.7.56** Sea  $f(x, y) = y^2 \sin(x/y)$  si  $y \neq 0$ ,  $f(x, 0) = 0$ .

i) Estudie la diferenciable de  $f$  en un punto genérico  $(a, b) \in \mathbb{R}^2$

ii) Estudie la continuidad y diferenciable de  $D_1 f(x, y)$  y  $D_2 f(x, y)$  en el punto  $(0, 0)$ .

iii) Determine un vector unitario  $u \in \mathbb{R}^2$  tal que  $D_u f(\pi, 1) = 1$ .

◇ **5.7.57** Estudie la diferenciable en  $(0, 0, 0)$  de la función  $f : \mathbb{R}^3 \rightarrow \mathbb{R}$  definida por

$$f(x, y, z) = z + y \sin \frac{x^2}{\sqrt{x^2 + y^2}} \quad \text{si } (x, y) \neq (0, 0), \quad f(0, 0, z) = z$$

Dado un punto  $\mathbf{p}$  del elipsoide  $E = \{(x, y, z) : x^2 + 2y^2 + 2z^2 = 2\}$ , obtenga un vector unitario  $\mathbf{v}$  tangente a  $E$  en  $\mathbf{p}$  que proporcione el mayor valor de la derivada  $D_{\mathbf{v}}f(\mathbf{p})$ .



# Capítulo 6

## Funciones dos veces diferenciables

*Derivadas parciales de segundo orden. Aplicaciones dos veces diferenciables: Teorema de Young sobre permutabilidad del orden de las derivaciones. Desarrollo de Taylor de orden 2. Extremos relativos. Funciones convexas.*

Este capítulo está dedicado a las funciones dos veces diferenciables y sus principales aplicaciones: La discusión de la naturaleza de los extremos locales y el estudio de las funciones convexas.

La noción de aplicación diferenciable dos veces la formulamos sólo en el caso de funciones de varias variables reales, tomando como base la diferenciabilidad de las derivadas parciales. El primer resultado básico es el teorema de Young 6.4 sobre la simetría de las derivadas parciales de segundo orden, con el que se obtiene que la diferencial segunda  $d^2f(\mathbf{a})(\mathbf{h}, \mathbf{k})$  es una aplicación bilineal simétrica que restringida a la diagonal produce una forma cuadrática  $Q_{\mathbf{a}}(\mathbf{h}) = d^2f(\mathbf{a})(\mathbf{h}, \mathbf{h})$  con la que se consigue una aproximación local de la función mejor que la proporcionada por la diferencial primera. Esta aproximación se establece de modo preciso mediante el desarrollo de Taylor de segundo orden (un anticipo de la versión general que se considerará en el capítulo 7) con el que se aborda la discusión de la naturaleza de los extremos locales y el estudio de las funciones convexas.

Como material complementario, en el apéndice ?? el lector interesado puede ver, como alternativa razonable al teorema de Young, el clásico teorema de Schwarz, sobre la igualdad de las derivadas mixtas. También puede encontrar allí una formulación equivalente de la noción de función dos veces diferenciable en la que no intervienen las derivadas parciales, lo que permite extender la definición al caso de funciones cuyo dominio no es finito dimensional. En el apéndice F se repasan y amplían los resultados básicos sobre las funciones convexas de una variable que se suelen estudiar en el curso de Análisis Matemático I. También se completa el estudio de las funciones convexas de varias variables demostrando que toda función convexa definida en un subconjunto abierto y convexo de  $\mathbb{R}^n$  es continua.

## 6.1. Funciones dos veces diferenciables.

Sea  $\mathbf{f} : \Omega \rightarrow F$  una aplicación definida en un abierto  $\Omega$  del espacio normado  $(E, \|\cdot\|)$ , con valores en el espacio normado  $(F, \|\cdot\|)$ . Supongamos que  $\mathbf{a} \in \Omega$  posee un entorno abierto  $V_{\mathbf{a}} \subset \Omega$  tal que para todo  $\mathbf{x} \in V_{\mathbf{a}}$  existe la derivada  $D_{\mathbf{v}}\mathbf{f}(\mathbf{x})$  según el vector  $\mathbf{v} \in E$ . Si la aplicación  $\mathbf{x} \rightarrow D_{\mathbf{v}}\mathbf{f}(\mathbf{x})$ , definida en  $V_{\mathbf{a}}$  con valores en  $F$ , es derivable en  $\mathbf{a}$  según el vector  $\mathbf{u} \in E$ , entonces la derivada  $D_{\mathbf{u}}(D_{\mathbf{v}}\mathbf{f})(\mathbf{a})$ , denotada  $D_{\mathbf{uv}}\mathbf{f}(\mathbf{a})$ , se llama *derivada segunda* de  $\mathbf{f}$  en  $\mathbf{a}$  según el par de vectores  $(\mathbf{u}, \mathbf{v}) \in E^2$ :

$$D_{\mathbf{uv}}\mathbf{f}(\mathbf{a}) = D_{\mathbf{u}}D_{\mathbf{v}}\mathbf{f}(\mathbf{a})$$

Si  $E = \mathbb{R}^n$ , cuando  $\mathbf{u} = \mathbf{e}_i$ , y  $\mathbf{v} = \mathbf{e}_j$  son vectores de la base canónica, como caso particular resulta la definición de las *derivadas parciales de segundo orden*,  $D_{\mathbf{e}_i\mathbf{e}_j}\mathbf{f}(\mathbf{a})$ , para las que se utilizan las notaciones

$$D_{ij}\mathbf{f}(\mathbf{a}) = \frac{\partial^2\mathbf{f}}{\partial x_i\partial x_j}(\mathbf{a}) \quad \text{si } i \neq j, \quad D_{ii}\mathbf{f}(\mathbf{a}) = \frac{\partial^2\mathbf{f}}{\partial x_i^2}(\mathbf{a}).$$

Habitualmente las derivadas parciales de segundo orden se pueden calcular fácilmente haciendo uso de las reglas del cálculo diferencial de funciones de una variable: Para obtener  $D_{ij}\mathbf{f}(\mathbf{a})$  basta derivar respecto a la variable  $x_i$  (en el punto  $\mathbf{a}$ ) la función de  $n$  variables reales  $(x_1, x_2, \dots, x_n) \rightarrow D_j\mathbf{f}(x_1, x_2, \dots, x_n)$ .

Si  $E = \mathbb{R}^3$  es costumbre que  $(x, y, z)$  designe un punto genérico de  $\mathbb{R}^3$ , y las derivadas parciales segundas de una función de tres variables reales  $\mathbf{f}(x, y, z)$  se escriben en la forma

$$\frac{\partial^2\mathbf{f}}{\partial x^2}, \quad \frac{\partial^2\mathbf{f}}{\partial y^2}, \quad \frac{\partial^2\mathbf{f}}{\partial z^2}, \quad \frac{\partial^2\mathbf{f}}{\partial x\partial y}, \quad \frac{\partial^2\mathbf{f}}{\partial y\partial x}, \quad \frac{\partial^2\mathbf{f}}{\partial x\partial z}, \quad \frac{\partial^2\mathbf{f}}{\partial z\partial x}, \quad \frac{\partial^2\mathbf{f}}{\partial y\partial z}, \quad \frac{\partial^2\mathbf{f}}{\partial z\partial y}.$$

donde, para simplificar la escritura, hemos omitido el punto donde se evalúan las derivadas. A veces es más cómodo utilizar la notación  $\mathbf{f}_x, \mathbf{f}_y, \mathbf{f}_z$  para las derivadas parciales primeras y  $\mathbf{f}_{xx}, \mathbf{f}_{xy}, \mathbf{f}_{xz}, \mathbf{f}_{yx}, \mathbf{f}_{yy}, \mathbf{f}_{yz}, \mathbf{f}_{zx}, \mathbf{f}_{zy}, \mathbf{f}_{zz}$  para las derivadas parciales segundas. Cuando se usa esta notación,  $\mathbf{f}_{xy}$  es una abreviatura de  $(\mathbf{f}_x)_y$ , luego

$$\mathbf{f}_{xy} = \frac{\partial^2\mathbf{f}}{\partial y\partial x}$$

de manera que se invierte el orden de la  $x$  y la  $y$  en las dos notaciones.

Más adelante veremos que bajo ciertas condiciones se puede asegurar la coincidencia de las llamadas *derivadas mixtas*  $\mathbf{f}_{xy} = \mathbf{f}_{yx}$ ,  $\mathbf{f}_{xz} = \mathbf{f}_{zx}$ ,  $\mathbf{f}_{yz} = \mathbf{f}_{zy}$ , de modo que en el caso de funciones de tres variables sólo habrá 3 derivadas mixtas distintas.

Se introducen notaciones análogas para funciones de distinto número de variables. Así por ejemplo, si  $\mathbf{f}(x, y, z, t)$  es una función de cuatro variables, hay 16 derivadas parciales segundas:  $\mathbf{f}_{xx}, \mathbf{f}_{xy}, \mathbf{f}_{xz}, \mathbf{f}_{xt}, \dots, \mathbf{f}_{tx}, \mathbf{f}_{ty}, \mathbf{f}_{tz}, \mathbf{f}_{tt}$ . y cuando haya coincidencia de las derivadas mixtas sólo habrá 6 derivadas mixtas distintas.

**Definición 6.1** Una aplicación  $\mathbf{f} : \Omega \rightarrow F$ , definida en un abierto  $\Omega \subset \mathbb{R}^n$ , es diferenciable dos veces en el punto  $\mathbf{a} \in \Omega$  cuando se verifican las condiciones:

- i)  $\mathbf{f}$  es diferenciable en un entorno abierto de  $\mathbf{a}$ ,  $V_{\mathbf{a}} \subset \Omega$ .
- ii) Las derivadas parciales  $D_i \mathbf{f}(\mathbf{x})$ ,  $1 \leq i \leq n$ , que están definidas en  $V_{\mathbf{a}}$ , son diferenciables en  $\mathbf{a}$ .

Si  $\mathbf{f}$  es diferenciable dos veces en cada  $\mathbf{x} \in \Omega$  se dice que  $\mathbf{f}$  es diferenciable dos veces en  $\Omega$ . Si en cada  $\mathbf{x} \in \Omega$  existen todas las derivadas parciales segundas y son continuas, se dice que  $\mathbf{f}$  es de clase  $C^2$  en  $\Omega$  y se escribe  $\mathbf{f} \in C^2(\Omega, F)$ .

OBSERVACIONES:

- a) Si en cada  $\mathbf{x} \in \Omega$  existen las derivadas parciales  $D_i \mathbf{f}(\mathbf{x})$ ,  $1 \leq i \leq n$ , y son diferenciables en  $\Omega$  entonces  $\mathbf{f}$  es diferenciable dos veces en  $\Omega$ : Basta tener en cuenta que la condición i) de 6.1 se cumple en todo punto de  $\Omega$  ya que  $\mathbf{f}$  tiene derivadas parciales continuas (porque son diferenciables).
- b) Toda función de clase  $C^2(\Omega, F)$  es diferenciable dos veces en  $\Omega$  (las derivadas parciales  $D_i \mathbf{f}(\mathbf{x})$ ,  $1 \leq i \leq n$ , son diferenciables porque tienen derivadas parciales continuas) y toda función diferenciable dos veces en  $\Omega$  es de clase  $C^1(\Omega, F)$  ya que las derivadas parciales  $D_i \mathbf{f}(\mathbf{x})$ ,  $1 \leq i \leq n$ , son continuas (por ser diferenciables).

Comenzamos obteniendo una condición suficiente, de tipo local, para que una función sea diferenciable dos veces en un punto, similar a la condición suficiente de diferenciability.

**Proposición 6.2** Sea  $\mathbf{f} : \Omega \rightarrow F$  definida en un abierto  $\Omega \subset \mathbb{R}^n$ , con valores en el espacio normado  $(F, \|\cdot\|)$ . Se supone que  $\mathbf{a} \in \Omega$  posee un entorno abierto  $V_{\mathbf{a}} \subset \Omega$  donde existen las derivadas parciales segundas  $D_{ij} \mathbf{f}(\mathbf{x})$ ,  $1 \leq i, j \leq n$ , y todas son continuas en  $\mathbf{a}$ . Entonces  $\mathbf{f}$  es diferenciable dos veces en  $\mathbf{a}$ .

DEM: Veamos en primer lugar que se cumple la condición i) de la definición 6.1), es decir, que existe un entorno abierto de  $\mathbf{a}$ ,  $U_{\mathbf{a}} \subset V_{\mathbf{a}}$  donde  $\mathbf{f}$  es diferenciable.

Como todas las derivadas parciales segundas  $D_{ij} \mathbf{f} = D_i(D_j \mathbf{f})$ , que están definidas en  $V_{\mathbf{a}}$ , son continuas en  $\mathbf{a}$ , existe un entorno abierto  $U_{\mathbf{a}} \subset V_{\mathbf{a}}$  donde todas ellas están acotadas. Como cada derivada primera  $D_j \mathbf{f}$  tiene derivadas parciales acotadas en  $U_{\mathbf{a}}$ , y las funciones con derivadas parciales acotadas son continuas (5.5) podemos asegurar que las derivadas parciales  $D_i \mathbf{f} : U_{\mathbf{a}} \rightarrow F$ ,  $1 \leq i \leq n$ , son continuas en  $U_{\mathbf{a}}$  y por lo tanto  $\mathbf{f}$  es diferenciable en  $U_{\mathbf{a}}$  (condición suficiente de diferenciability 5.16).

Según la hipótesis, las funciones  $D_i \mathbf{f}(\mathbf{x})$ ,  $1 \leq i \leq n$ , que están definidas en  $U_{\mathbf{a}}$ , tienen derivadas parciales continuas  $\mathbf{a}$ , y por lo tanto son diferenciables en  $\mathbf{a}$ , de modo que también se cumple la condición ii) de la definición 6.1. ■

El siguiente ejemplo muestra que en general no se puede asegurar la igualdad  $D_{ij} \mathbf{f}(\mathbf{a}) = D_{ji} \mathbf{f}(\mathbf{a})$ .

**Ejemplo 6.3** La función  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ , definida por

$$f(x, y) = xy \frac{x^2 - y^2}{x^2 + y^2} \quad \text{si } (x, y) \neq (0, 0), \quad f(0, 0) = 0$$

es diferenciable en todo punto pero  $D_{12}f(0, 0) \neq D_{21}f(0, 0)$ .

DEM: Es fácil ver que las derivadas parciales  $D_1f(x, y)$ ,  $D_2f(x, y)$  existen en todo punto, y con un cálculo elemental se obtiene el valor

$$D_1f(x, y) = \frac{y(x^4 + 4x^2y^2 - y^4)}{(x^2 + y^2)^2} \quad \text{si } (x, y) \neq 0, \quad D_1f(0, 0) = 0$$

La función  $D_1f$  es continua en todo punto: La continuidad en los puntos  $(x, y) \neq (0, 0)$  es inmediata y la continuidad en  $(0, 0)$  se obtiene fácilmente considerando coordenadas polares,  $x = r \cos \theta$ ,  $y = r \sin \theta$ , con las que se establece la acotación  $|D_1f(x, y)| \leq 5r = 5\sqrt{x^2 + y^2}$ , de donde se sigue que

$$\lim_{(x,y) \rightarrow (0,0)} D_1f(x, y) = 0 = D_1f(0, 0)$$

Análogamente  $D_2f(x, y)$  es continua en cada  $(x, y) \in \mathbb{R}^2$  y con el teorema 5.16 se concluye que  $f$  es diferenciable en todo punto. Como  $D_1f(0, y) = -y$  se sigue que  $D_{21}f(0, 0) = -1$ , y análogamente se obtiene que  $D_{12}f(0, 0) = 1$ . ■

El siguiente teorema proporciona una condición suficiente para que se cumpla la igualdad  $D_{ij}\mathbf{f}(\mathbf{a}) = D_{ji}\mathbf{f}(\mathbf{a})$ .

**Teorema 6.4** [Young] Sea  $\mathbf{f} : \Omega \rightarrow F$  definida en un abierto  $\Omega \subset \mathbb{R}^n$ . Se supone que  $\mathbf{a} \in \Omega$  posee un entorno  $V_{\mathbf{a}} \subset \Omega$  donde existen las derivadas parciales  $D_i\mathbf{f}$ ,  $D_j\mathbf{f}$  ( $i \neq j$ ) y ambas son diferenciables en  $\mathbf{a}$ . Entonces se verifica  $D_{ij}\mathbf{f}(\mathbf{a}) = D_{ji}\mathbf{f}(\mathbf{a})$ .

En particular, si  $\mathbf{f}$  es diferenciable dos veces en  $\mathbf{a}$ , para cada par de índices  $1 \leq i < j \leq n$ , se cumple  $D_{ij}\mathbf{f}(\mathbf{a}) = D_{ji}\mathbf{f}(\mathbf{a})$ .

DEM: Como el teorema sólo involucra dos variables, basta considerar el caso  $n = 2$ . Para simplificar la notación suponemos  $\mathbf{a} = (0, 0)$  y  $V_{\mathbf{a}} = (-r, r) \times (-r, r)$ . Tomando  $0 < h < r$  podemos asegurar que el cuadrado de vértices  $(0, 0)$ ,  $(h, 0)$ ,  $(0, h)$ ,  $(h, h)$  está contenido en  $V$ . Demostraremos que  $\Delta(h) = \mathbf{f}(h, h) - \mathbf{f}(h, 0) - \mathbf{f}(0, h) + \mathbf{f}(0, 0)$  verifica

$$D_{12}\mathbf{f}(0, 0) = \lim_{h \rightarrow 0^+} \frac{\Delta(h)}{h^2} = D_{21}\mathbf{f}(0, 0)$$

Obsérvese que  $\Delta(h) = \mathbf{g}(h) - \mathbf{g}(0)$ , donde  $\mathbf{g}(x) = \mathbf{f}(x, h) - \mathbf{f}(x, 0)$  es derivable en cada  $x \in [0, h]$ , con  $\mathbf{g}'(x) = D_1\mathbf{f}(x, h) - D_1\mathbf{f}(x, 0)$ .

Como  $D_1\mathbf{f}$  es diferenciable en  $(0, 0)$  se tiene

- i)  $D_1\mathbf{f}(x, h) = D_1\mathbf{f}(0, 0) + D_{11}\mathbf{f}(0, 0)x + D_{21}\mathbf{f}(0, 0)h + \|(x, h)\|_{\infty} \boldsymbol{\rho}(x, h)$ .
- ii)  $D_1\mathbf{f}(x, 0) = D_1\mathbf{f}(0, 0) + D_{11}\mathbf{f}(0, 0)x + \|(x, 0)\|_{\infty} \boldsymbol{\rho}(x, 0)$ .

donde  $\boldsymbol{\rho}(x, y)$  tiende hacia 0 cuando  $\|(x, y)\|_\infty \rightarrow 0$ .

Sustituyendo i) y ii) en la expresión  $\mathbf{g}'(x) = D_1\mathbf{f}(x, h) - D_1\mathbf{f}(x, 0)$  resulta

$$\mathbf{g}'(x) = hD_{21}\mathbf{f}(0, 0) + \|(x, h)\|_\infty \boldsymbol{\rho}(x, h) - \|(x, 0)\|_\infty \boldsymbol{\rho}(x, 0)$$

La función  $\boldsymbol{\varphi}(x) = \mathbf{g}(x) - D_{21}\mathbf{f}(0, 0)hx$  está definida para  $0 \leq x \leq h$ , y

$$\boldsymbol{\varphi}'(x) = \mathbf{g}'(x) - D_{21}\mathbf{f}(0, 0)h = \|(x, h)\|_\infty \boldsymbol{\rho}(x, h) - \|(x, 0)\|_\infty \boldsymbol{\rho}(x, 0)$$

Dado  $\epsilon > 0$ , usando la definición de límite, podemos encontrar  $0 < r' < r$  tal que

$$\|(x, y)\|_\infty < r' \Rightarrow \|\boldsymbol{\rho}(x, y)\| < \epsilon/2$$

Si  $0 \leq x \leq h < r'$  se cumple  $\|(x, h)\|_\infty < r'$ ,  $\|(x, 0)\|_\infty < r'$ , luego

$$\|\boldsymbol{\varphi}'(x)\| < h\epsilon/2 + h\epsilon/2 = \epsilon h$$

y aplicando el teorema del incremento finito resulta  $\|\boldsymbol{\varphi}(h) - \boldsymbol{\varphi}(0)\| \leq \epsilon h^2$ , es decir

$$\begin{aligned} \|(\mathbf{f}(h, h) - \mathbf{f}(h, 0) - D_{21}\mathbf{f}(0, 0)h^2) - (\mathbf{f}(0, h) - \mathbf{f}(0, 0))\| &< \epsilon h^2 \\ \|\Delta(h) - D_{21}\mathbf{f}(0, 0)h^2\| &< \epsilon h^2 \end{aligned}$$

Hemos demostrado así que

$$0 < h < r' \Rightarrow \left\| \frac{\Delta(h)}{h^2} - D_{21}\mathbf{f}(0, 0) \right\| < \epsilon$$

es decir:

$$\lim_{h \rightarrow 0^+} \frac{\Delta(h)}{h^2} = D_{21}\mathbf{f}(0, 0)$$

Análogamente, considerando la función auxiliar  $\mathbf{g}(y) = \mathbf{f}(h, y) - \mathbf{f}(0, y)$ , que cumple  $\Delta(h) = \mathbf{g}(h) - \mathbf{g}(0)$ , usando ahora que  $D_2\mathbf{f}$  es diferenciable en  $(0, 0)$ , y estimando la derivada  $\mathbf{g}'(y) = D_2\mathbf{f}(h, y) - D_2\mathbf{f}(0, y)$ , con un razonamiento paralelo se llega a

$$\lim_{h \rightarrow 0^+} \frac{\Delta(h)}{h^2} = D_{12}\mathbf{f}(0, 0)$$

Finalmente, si  $\mathbf{f}$  es diferenciable dos veces en  $\mathbf{a}$ , según la definición 6.1 existe  $V_{\mathbf{a}} \subset \Omega$ , entorno abierto de  $\mathbf{a}$ , donde  $\mathbf{f}$  es diferenciable y además todas las derivadas parciales  $D_i\mathbf{f}(x)$ ,  $1 \leq i \leq n$ , son diferenciables en  $\mathbf{a}$ . Por lo tanto podemos aplicar el resultado demostrado a todas las derivadas mixtas  $D_{ij}\mathbf{f}(\mathbf{a})$ . ■

**Corolario 6.5** Sea  $\mathbf{f} : \Omega \rightarrow F$  definida en un abierto  $\Omega \subset \mathbb{R}^n$ . Si  $\mathbf{a} \in \Omega$  posee un entorno abierto  $V_{\mathbf{a}} \subset \Omega$  donde existen todas las derivadas parciales  $D_{ij}\mathbf{f}(\mathbf{x})$ ,  $1 \leq i, j \leq n$ , y son continuas en  $\mathbf{a}$ , entonces  $D_{ij}\mathbf{f}(\mathbf{a}) = D_{ji}\mathbf{f}(\mathbf{a})$ .

DEM: Es consecuencia directa de la proposición 6.2 y del teorema de Young 6.4. ■

**Corolario 6.6** Sea  $\mathbf{f} : \Omega \rightarrow F$  definida en un abierto  $\Omega \subset \mathbb{R}^2$ . Si  $\mathbf{a} \in \Omega$  posee un entorno  $V_{\mathbf{a}} \subset \Omega$  donde existen las cuatro derivadas parciales  $D_{ij}\mathbf{f}$ ,  $1 \leq i, j \leq 2$  y  $D_{12}\mathbf{f}$  y  $D_{21}\mathbf{f}$  son continuas en  $\mathbf{a}$  entonces  $D_{12}\mathbf{f}(\mathbf{a}) = D_{21}\mathbf{f}(\mathbf{a})$ .

DEM:  $D_1\mathbf{f}$  es diferenciable en  $\mathbf{a}$  porque en un entorno de  $\mathbf{a}$  existen sus dos derivadas parciales y una de ellas es continua en  $\mathbf{a}$ . Lo mismo se puede decir de  $D_2\mathbf{f}$  y por lo tanto se puede aplicar el teorema de Young 6.4. ■

**Proposición 6.7** Sea  $\mathbf{f} : \Omega \rightarrow F$  definida en un abierto  $\Omega \subset \mathbb{R}^n$ . Si  $\mathbf{f}$  es diferenciable dos veces en  $\mathbf{a} \in \Omega$  entonces para cada par  $(\mathbf{u}, \mathbf{v}) \in \mathbb{R}^n \times \mathbb{R}^n$  existen y son iguales las derivadas segundas  $D_{\mathbf{u}\mathbf{v}}\mathbf{f}(\mathbf{a})$ ,  $D_{\mathbf{v}\mathbf{u}}\mathbf{f}(\mathbf{a})$ , que vienen dadas por

$$D_{\mathbf{u}\mathbf{v}}\mathbf{f}(\mathbf{a}) = D_{\mathbf{v}\mathbf{u}}\mathbf{f}(\mathbf{a}) = \sum_{i,j=1}^n D_{ij}\mathbf{f}(\mathbf{a})u_i v_j$$

DEM: Según la definición 6.1  $\mathbf{f}$  es diferenciable en un entorno abierto de  $\mathbf{a}$ ,  $V_{\mathbf{a}} \subset \Omega$ , luego en cada  $\mathbf{x} \in V_{\mathbf{a}}$  existe la derivada  $D_{\mathbf{v}}\mathbf{f}(\mathbf{x}) = d\mathbf{f}(\mathbf{x})\mathbf{v}$ , que viene dada por

$$D_{\mathbf{v}}\mathbf{f}(\mathbf{x}) = D_1\mathbf{f}(\mathbf{x})v_1 + D_2\mathbf{f}(\mathbf{x})v_2 + \cdots + D_n\mathbf{f}(\mathbf{x})v_n$$

La función  $\mathbf{g}(\mathbf{x}) = D_{\mathbf{v}}\mathbf{f}(\mathbf{x})$ , definida en  $V_{\mathbf{a}}$ , es diferenciable en  $\mathbf{a}$  (por ser suma de funciones diferenciables en  $\mathbf{a}$ ) y por lo tanto existe su derivada según el vector  $\mathbf{u}$

$$D_{\mathbf{u}}\mathbf{g}(\mathbf{a}) = \sum_{i=1}^n D_i\mathbf{g}(\mathbf{a})u_i = \sum_{i=1}^n \left( \sum_{j=1}^n D_i(D_j\mathbf{f})(\mathbf{a})v_j \right) u_i$$

Es decir

$$D_{\mathbf{u}}(D_{\mathbf{v}}\mathbf{f})(\mathbf{a}) = \sum_{i,j=1}^n D_{ij}\mathbf{f}(\mathbf{a})u_i v_j$$

Usando esta fórmula y la simetría de las derivadas segundas  $D_{ij}\mathbf{f}(\mathbf{a}) = D_{ji}\mathbf{f}(\mathbf{a})$  (teorema 6.4) se concluye que  $D_{\mathbf{u}\mathbf{v}}\mathbf{f}(\mathbf{a}) = D_{\mathbf{v}\mathbf{u}}\mathbf{f}(\mathbf{a})$ . ■

**Definición 6.8** Si  $\mathbf{f} : \Omega \rightarrow F$ , definida en un abierto  $\Omega \subset \mathbb{R}^n$ , es diferenciable dos veces en  $\mathbf{a} \in \Omega$  la diferencial segunda de  $\mathbf{f}$  en  $\mathbf{a}$  es la aplicación bilineal simétrica  $d^2\mathbf{f}(\mathbf{a}) : \mathbb{R}^n \times \mathbb{R}^n \rightarrow F$ , definida por

$$d^2\mathbf{f}(\mathbf{a})(\mathbf{u}, \mathbf{v}) = D_{\mathbf{u}\mathbf{v}}\mathbf{f}(\mathbf{a}) = \sum_{i,j=1}^n D_{ij}\mathbf{f}(\mathbf{a})u_i v_j$$

En las condiciones de 6.8, en virtud de la simetría, de las  $n^2$  derivadas segundas  $D_{ij}\mathbf{f}(\mathbf{a})$  hay muchas que son iguales; hay a lo más  $n(n-1)/2$  derivadas parciales segundas que pueden ser diferentes. Cuando  $f : \Omega \rightarrow \mathbb{R}$  es una función con valores reales diferenciable dos veces en  $\mathbf{a}$ , la matriz simétrica formada con las derivadas parciales segundas,  $(D_{ij}f(\mathbf{a}))$ , se llama matriz *Hessiana* de  $f$  en el punto  $\mathbf{a}$ .

Si  $\mathbf{f}$  es diferenciable dos veces en  $\mathbf{a} \in \Omega$ , la expresión

$$Q_{\mathbf{a}}(\mathbf{h}) = d^2\mathbf{f}(\mathbf{a})(\mathbf{h}, \mathbf{h}) = \sum_{i,j=1}^n D_{ij}\mathbf{f}(\mathbf{a})h_i h_j$$

que a veces se suele denotar  $d^2\mathbf{f}(\mathbf{a})\mathbf{h}^2$  es un polinomio homogéneo de segundo grado en las variables  $(h_1, h_2, \dots, h_n)$  (cuyos coeficientes son vectores de  $F$ ). Para una función de dos variables reales se tiene:

$$Q(\mathbf{h}) = D_{11}\mathbf{f}(\mathbf{a})h_1^2 + 2D_{12}\mathbf{f}(\mathbf{a})h_1h_2 + D_{22}\mathbf{f}(\mathbf{a})h_2^2$$

y para una función de tres variables reales

$$D_{11}\mathbf{f}(\mathbf{a})h_1^2 + D_{22}\mathbf{f}(\mathbf{a})h_2^2 + D_{33}\mathbf{f}(\mathbf{a})h_3^2 + 2D_{12}\mathbf{f}(\mathbf{a})h_1h_2 + 2D_{13}\mathbf{f}(\mathbf{a})h_1h_3 + 2D_{23}\mathbf{f}(\mathbf{a})h_2h_3$$

**Teorema 6.9** Sea  $\mathbf{f} : \Omega \rightarrow F$ , definida en un abierto  $\Omega \subset \mathbb{R}^n$ , y diferenciable dos veces en  $\mathbf{a}$ . Entonces

$$\mathbf{f}(\mathbf{a} + \mathbf{h}) = \mathbf{f}(\mathbf{a}) + d\mathbf{f}(\mathbf{a})\mathbf{h} + \frac{1}{2}d^2\mathbf{f}(\mathbf{a})\mathbf{h}^2 + \mathbf{R}(\mathbf{h})$$

donde  $\mathbf{R}(\mathbf{h}) = o(\|\mathbf{h}\|^2)$  (es decir,  $\mathbf{R}(\mathbf{h}) = \|\mathbf{h}\|^2 \mathbf{r}(\mathbf{h})$ , con  $\lim_{\mathbf{h} \rightarrow 0} \mathbf{r}(\mathbf{h}) = 0$ ),

DEM: Como  $\mathbf{f}$  es diferenciable en una bola  $B(\mathbf{a}, r) \subset \Omega$ , podemos asegurar que

$$\mathbf{R}(\mathbf{h}) = \mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a}) - \sum_{i=1}^n D_i\mathbf{f}(\mathbf{a})h_i - \frac{1}{2} \sum_{i,j=1}^n D_{ij}\mathbf{f}(\mathbf{a})h_i h_j$$

es diferenciable en  $B(\mathbf{0}, r)$ , donde existirán sus derivadas parciales. Para calcularlas observemos que, en virtud de la simetría  $D_{ij}\mathbf{f}(\mathbf{a}) = D_{ji}\mathbf{f}(\mathbf{a})$ , las derivadas parciales de  $Q(\mathbf{h}) = \sum_{i,j=1}^n D_{ij}\mathbf{f}(\mathbf{a})h_i h_j$  vienen dadas por  $D_k Q(\mathbf{h}) = 2 \sum_{i=1}^n D_{ik}\mathbf{f}(\mathbf{a})h_i$ , luego

$$D_k\mathbf{R}(\mathbf{h}) = D_k\mathbf{f}(\mathbf{a} + \mathbf{h}) - D_k\mathbf{f}(\mathbf{a}) - \sum_{i=1}^n D_{ik}\mathbf{f}(\mathbf{a})h_i$$

Cada  $D_k\mathbf{f}$  es diferenciable en  $\mathbf{a}$ , luego  $D_k\mathbf{R}(\mathbf{h}) = o(\|\mathbf{h}\|)$  es decir

$$D_k\mathbf{R}(\mathbf{h}) = \|\mathbf{h}\| \boldsymbol{\rho}_k(\mathbf{h}) \text{ con } \lim_{\mathbf{h} \rightarrow 0} \boldsymbol{\rho}_k(\mathbf{h}) = 0$$

Dado  $\epsilon > 0$  existe  $\delta > 0$  tal que, para todo  $k \in \{1, \dots, n\}$  se cumple

$$\|\mathbf{h}\| < \delta \Rightarrow \|\boldsymbol{\rho}_k(\mathbf{h})\| < \epsilon/n$$

$\varphi(t) = \mathbf{R}(t\mathbf{h})$  es derivable en  $[0, 1]$  y su derivada  $\varphi'(t) = \sum_{j=1}^n D_j\mathbf{R}(t\mathbf{h})h_j$  verifica

$$\|\varphi'(t)\| \leq \sum_{j=1}^n \|\boldsymbol{\rho}_j(t\mathbf{h})\| \|t\mathbf{h}\| |h_j| \leq \epsilon \|\mathbf{h}\|^2$$

Aplicando el teorema del incremento finito a la función  $\varphi$  se obtiene

$$\|\varphi(1) - \varphi(0)\| \leq \epsilon \|\mathbf{h}\|^2, \text{ si } \|\mathbf{h}\| < \delta$$

es decir

$$\|\mathbf{R}(\mathbf{h})\| \leq \epsilon \|\mathbf{h}\|^2, \text{ si } \|\mathbf{h}\| < \delta.$$

■

## 6.2. Extremos relativos

Sea  $f : \Omega \rightarrow \mathbb{R}$  una función diferenciable de  $n$  variables reales definida en un abierto  $\Omega \subset \mathbb{R}^n$ . Ya hemos visto en el capítulo 5 que si  $f$  presenta un extremo relativo en  $\mathbf{a} \in \Omega$  entonces  $D_j f(\mathbf{a}) = 0$  para todo  $j \in \{1, 2, \dots, n\}$ , lo que se expresa brevemente diciendo que  $\mathbf{a}$  es un punto estacionario de  $f$ .

Los siguientes ejemplos muestran que, en lo referente a extremos locales, las funciones de varias variables pueden tener comportamientos que no ocurren en el caso de las funciones de una variable (la comprobación de las afirmaciones que en ellos se hacen se hacen se dejan al cuidado del lector).

### Ejemplos 6.10

a) El único punto estacionario de la función  $f(x, y) = x^2(1+y)^3 + y^2$  es  $(0, 0)$  donde  $f$  presenta un mínimo relativo que no es mínimo absoluto.

b) La función  $g(x, y) = (x^2y - x - 1)^2 + (x^2 - 1)^2$  sólo tiene dos puntos estacionarios  $(1, 2)$  y  $(-1, 0)$ , y ambos son puntos de mínimo absoluto.

En esta sección proseguimos con el estudio de los extremos locales de las funciones dos veces diferenciables, obteniendo condiciones necesarias y condiciones suficientes de segundo orden (en términos de las derivadas parciales segundas) para la existencia de máximo relativo, o mínimo relativo, en un punto estacionario.

**Proposición 6.11** Sea  $f : \Omega \rightarrow \mathbb{R}$ , definida en un abierto  $\Omega \subset \mathbb{R}^n$ , diferenciable dos veces en  $\mathbf{a} \in \Omega$ . Si  $f$  presenta en  $\mathbf{a} \in \Omega$  un mínimo relativo se cumple

$$d^2 f(\mathbf{a})\mathbf{h}^2 \geq 0 \quad \text{para todo } \mathbf{h} \in \mathbb{R}^n$$

Si  $f$  presenta en  $\mathbf{a} \in \Omega$  un máximo relativo se cumple

$$d^2 f(\mathbf{a})\mathbf{h}^2 \leq 0 \quad \text{para todo } \mathbf{h} \in \mathbb{R}^n$$

DEM: Por hipótesis  $\mathbf{a}$  posee un entorno abierto  $V_{\mathbf{a}} \subset \Omega$  donde  $f$  es diferenciable. Dado  $\mathbf{h} \in \mathbb{R}^n$ , sea  $r > 0$  tal que  $|t| < r \Rightarrow \mathbf{a} + t\mathbf{h} \in V_{\mathbf{a}}$ . En virtud de la regla de la cadena,  $\varphi(t) = f(\mathbf{a} + t\mathbf{h})$  está definida y es derivable en  $(-r, r)$ . Su derivada viene dada por

$$\varphi'(t) = df(\mathbf{a} + t\mathbf{h})\mathbf{h} = \sum_{j=1}^n D_j f(\mathbf{a} + t\mathbf{h})h_j$$

Por hipótesis, las derivadas parciales  $D_j f : V_{\mathbf{a}} \rightarrow \mathbb{R}$  son diferenciables en  $\mathbf{a}$ , y aplicando otra vez la regla de la cadena, obtenemos que las funciones  $\alpha_j(t) = D_j f(\mathbf{a} + t\mathbf{h})$  son derivables en  $t = 0$  y que la derivada vale  $\alpha_j'(0) = \sum_{i=1}^n D_i D_j f(\mathbf{a})h_i$ . Entonces  $\varphi'(t)$  es derivable en  $t = 0$  y

$$\varphi''(0) = \sum_{j=1}^n \left( \sum_{i=1}^n D_i D_j f(\mathbf{a})h_i \right) h_j = \sum_{i,j=1}^n D_{ij} f(\mathbf{a})h_i h_j = d^2 f(\mathbf{a})(\mathbf{h}, \mathbf{h})$$



Si  $f$  presenta en  $\mathbf{a}$  un mínimo (resp. máximo) relativo entonces  $\varphi$  presenta en  $t = 0$  un extremo local del mismo tipo y se concluye que

$$d^2 f(\mathbf{a})(\mathbf{h}, \mathbf{h}) = \varphi''(0) \geq 0, \quad (\text{resp. } \leq 0).$$

■

**Corolario 6.12** *Sea  $f : \Omega \rightarrow \mathbb{R}$  definida en un abierto  $\Omega \subset \mathbb{R}^n$  y diferenciable dos veces en  $\mathbf{a} \in \Omega$ . Si existen  $\mathbf{h}, \mathbf{k} \in \mathbb{R}^n$  tales que  $d^2 f(\mathbf{a})\mathbf{k}^2 < 0 < d^2 f(\mathbf{a})\mathbf{h}^2$  entonces  $f$  no presenta en  $\mathbf{a}$  ni máximo relativo ni mínimo relativo.*

DEM: Es consecuencia directa de la proposición 6.11

■

El siguiente ejemplo muestra que la condición  $d^2 f(\mathbf{a})(\mathbf{h}, \mathbf{h}) \geq 0$  para todo  $\mathbf{h} \in \mathbb{R}^n$ , no garantiza que  $f$  presente en  $\mathbf{a}$  un mínimo relativo. (Si  $d^2 f(\mathbf{a})(\mathbf{h}, \mathbf{h})$  no es idénticamente nula, esta condición sólo permite asegurar que  $f$  no presenta máximo relativo).

**Ejemplo 6.13** *En el punto  $(0, 0)$ , que es estacionario para  $f(x, y) = x^2 y + y^2$ , se cumple la condición*

$$d^2 f(0, 0)(\mathbf{h}, \mathbf{h}) = 2h_2^2 \geq 0 \quad \text{para todo } \mathbf{h} \in \mathbb{R}^2$$

y sin embargo  $f$  no presenta en este punto un extremo relativo:  $f(0, 0) = 0$  y es claro que en todo entorno de  $(0, 0)$  hay puntos donde la función  $f(x, y) = y(x^2 + y)$  es positiva y puntos donde es negativa.

Una forma cuadrática real de  $n$  variables reales es una aplicación  $Q : \mathbb{R}^n \rightarrow \mathbb{R}$  de la forma

$$Q(\mathbf{x}) = Q(x_1, x_2, \dots, x_n) = \sum_{i,j=1}^n \alpha_{ij} x_i x_j$$

Si  $Q(\mathbf{x}) > 0$  (resp.  $< 0$ ) para todo  $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$  se dice que  $Q$  es una forma cuadrática *definida positiva* (resp. *negativa*). Si  $Q(\mathbf{x}) \geq 0$  (resp.  $\leq 0$ ) para todo  $\mathbf{x} \in \mathbb{R}^n$  se dice que  $Q$  es una forma cuadrática *definida no negativa* (resp. *no positiva*). Finalmente, si existen  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$  tales que  $Q(\mathbf{x}) < 0 < Q(\mathbf{y})$  se dice que  $Q$  es una forma cuadrática *indefinida*.

**Lema 6.14** *Si la forma cuadrática  $Q : \mathbb{R}^n \rightarrow \mathbb{R}$  es definida positiva (resp. negativa) entonces existe  $m > 0$  (resp.  $m < 0$ ) tal que  $Q(\mathbf{x}) \geq m \|\mathbf{x}\|^2$  (resp.  $Q(\mathbf{x}) \leq m \|\mathbf{x}\|^2$ ) para todo  $\mathbf{x} \in \mathbb{R}^n$ .*

DEM: Como  $Q$  es continua y  $S = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| = 1\}$  es compacto (cerrado y acotado) existe  $\mathbf{a} \in S$  tal que  $Q(\mathbf{a}) = \min\{Q(\mathbf{z}) : \mathbf{z} \in S\}$ . Si  $Q$  es definida positiva podemos asegurar que  $Q(\mathbf{a}) > 0$  y tomando  $m = Q(\mathbf{a})$  se cumple la condición requerida: Es evidente si  $\mathbf{x} = 0$  y cuando  $\mathbf{x} \neq 0$  basta considerar  $\mathbf{z} = \mathbf{x}/\|\mathbf{x}\| \in S$ , y tener en cuenta que  $Q$  es homogénea de grado 2, para obtener

$$Q(\mathbf{x}) = Q(\|\mathbf{x}\| \mathbf{z}) = \|\mathbf{x}\|^2 Q(\mathbf{z}) \geq m \|\mathbf{x}\|^2$$

Análogamente se razona, cuando  $Q$  es definida negativa, considerando el máximo de  $Q$  sobre  $S$ , que es negativo. ■

**Teorema 6.15** *Sea  $f : \Omega \rightarrow \mathbb{R}$ , definida en un abierto  $\Omega \subset \mathbb{R}^n$ , y diferenciable dos veces en  $\mathbf{a} \in \Omega$ . Si  $\mathbf{a}$  es punto estacionario de  $f$  y la forma cuadrática*

$$d^2 f(\mathbf{a})\mathbf{h}^2 = \sum_{i,j=1}^n D_{ij}f(\mathbf{a})h_i h_j$$

*es definida positiva (resp. negativa) entonces  $f$  presenta en  $\mathbf{a}$  un mínimo (resp. un máximo) relativo estricto*

DEM: Sea  $Q_{\mathbf{a}}(\mathbf{h}) = \frac{1}{2}d^2 f(\mathbf{a})\mathbf{h}^2$ . En virtud del teorema 6.9

$$f(\mathbf{a} + \mathbf{h}) = f(\mathbf{a}) + df(\mathbf{a})\mathbf{h} + Q_{\mathbf{a}}(\mathbf{h}) + \mathbf{r}(\mathbf{h}) \|\mathbf{h}\|^2$$

donde  $\lim_{\mathbf{h} \rightarrow \mathbf{0}} \mathbf{r}(\mathbf{h}) = \mathbf{0}$ . Según el lema 6.14 existe  $m > 0$  tal que  $Q_{\mathbf{a}}(\mathbf{h}) \geq m \|\mathbf{h}\|^2$  y teniendo en cuenta que  $df(\mathbf{a}) = 0$  resulta

$$f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) \geq (m + \mathbf{r}(\mathbf{h})) \|\mathbf{h}\|^2$$

Por la definición de límite, existe  $B(\mathbf{a}, \delta) \subset \Omega$  tal que  $\|\mathbf{h}\| < \delta \Rightarrow |\mathbf{r}(\mathbf{h})| < m/2$ , luego

$$\|\mathbf{h}\| < \delta \Rightarrow f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) \geq (m - m/2) \|\mathbf{h}\|^2 > 0$$

Hemos demostrado así que  $\mathbf{x} \in B(\mathbf{a}, \delta) \Rightarrow f(\mathbf{x}) > f(\mathbf{a})$ , y por lo tanto  $f$  presenta en  $\mathbf{a}$  un mínimo relativo estricto. Análogamente se razona, en el caso alternativo, para el máximo relativo estricto. ■

Si  $f$  es diferenciable dos veces en un punto estacionario  $\mathbf{a}$  y la forma cuadrática  $d^2 f(\mathbf{a})\mathbf{h}^2 = \sum_{i,j=1}^n D_{ij}f(\mathbf{a})h_i h_j$  es indefinida, según el corolario 6.12,  $f$  no presenta ni mínimo ni máximo relativo en este punto. Utilizando el lema 6.14 se consigue una descripción más precisa de lo que ocurre en este caso

**Proposición 6.16** *Se supone que  $f : \Omega \rightarrow \mathbb{R}$ , definida en un abierto  $\Omega \subset \mathbb{R}^n$ , es diferenciable dos veces en  $\mathbf{a} \in \Omega$ . Si  $\mathbf{a}$  es punto estacionario de  $f$  y la forma cuadrática*

$$d^2 f(\mathbf{a})\mathbf{h}^2 = \sum_{i,j=1}^n D_{ij}f(\mathbf{a})h_i h_j$$

*es indefinida entonces existen dos rectas que pasan por  $\mathbf{a}$  de modo que a lo largo de una de ellas  $f$  presenta en  $\mathbf{a}$  un mínimo relativo estricto y a lo largo de la otra un máximo relativo estricto.*

DEM: Razonando como en la demostración del teorema 6.9, existe  $B(\mathbf{a}, r) \subset \Omega$  tal que si  $\mathbf{x} \in B(\mathbf{a}, r)$  y  $\mathbf{h} = \mathbf{x} - \mathbf{a}$  se cumple  $f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) = Q_{\mathbf{a}}(\mathbf{h}) + \epsilon(\mathbf{h}) \|\mathbf{h}\|^2$ ,

donde  $Q_{\mathbf{a}}(\mathbf{h}) = \frac{1}{2}d^2f(\mathbf{a})\mathbf{h}^2$ , y  $\lim_{\mathbf{h} \rightarrow \mathbf{0}} \epsilon(\mathbf{h}) = \mathbf{0}$ . Por hipótesis existen  $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$  tales que  $Q(\mathbf{u}) < 0 < Q(\mathbf{v})$ . Cuando  $\mathbf{h} = t\mathbf{u}$ , la diferencia

$$f(\mathbf{a} + t\mathbf{u}) - f(\mathbf{a}) = Q_{\mathbf{a}}(t\mathbf{u}) + \epsilon(t\mathbf{u}) \|t\mathbf{u}\|^2 = t^2(Q_{\mathbf{a}}(\mathbf{u}) + \epsilon(t\mathbf{u}) \|\mathbf{u}\|^2)$$

tiene el mismo signo que  $Q_{\mathbf{a}}(\mathbf{u}) + \epsilon(t\mathbf{u}) \|\mathbf{u}\|^2$ . Cuando  $t \rightarrow 0$  esta expresión tiene límite  $Q_{\mathbf{a}}(\mathbf{u}) < 0$ , luego existe  $\rho > 0$  tal que  $Q_{\mathbf{a}}(\mathbf{u}) + \epsilon(t\mathbf{u}) \|\mathbf{u}\|^2 < 0$  si  $|t| < \rho$ . Vemos así que a lo largo de la recta  $\mathbf{a} + t\mathbf{u}$  la función  $f$  presenta en  $\mathbf{a}$  un máximo relativo estricto:  $|t| < \rho \Rightarrow f(\mathbf{a} + t\mathbf{u}) - f(\mathbf{a}) < 0$ .

Con un razonamiento similar se obtiene que a lo largo de la recta  $\mathbf{a} + t\mathbf{v}$  la función  $f$  presenta en  $\mathbf{a}$  un mínimo relativo estricto. ■

A la vista de los resultados expuestos en los teoremas 6.15 y 6.16, es claro el interés que tienen, para el asunto que nos ocupa, los criterios que permiten asegurar que una forma cuadrática concreta es definida positiva, definida negativa o indefinida. Un resultado de esta naturaleza es el siguiente.

**Proposición 6.17** Sea  $Q(\mathbf{x}) = \sum_{i,j=1}^n \alpha_{ij}x_i x_j$  la forma cuadrática asociada a una matriz simétrica  $\alpha_{ij} = \alpha_{ji}$ , y  $\Delta_k$  el determinante de la matriz  $\{\alpha_{ij} : 1 \leq i, j \leq k\}$ .

$Q$  es definida positiva si y sólo si  $\Delta_k > 0$  para todo  $k \in \{1, 2, \dots, n\}$   
 $Q$  es definida negativa si y sólo si  $(-1)^k \Delta_k > 0$  para todo  $k \in \{1, 2, \dots, n\}$ .

DEM: De momento damos una demostración elemental para el caso  $n = 2$ . El lector interesado puede consultar más adelante, en la sección I.1, una demostración del caso general. En el caso particular que nos ocupa escribimos

$$Q(x, y) = Ax^2 + 2Bxy + Cy^2$$

Si  $AC - B^2 > 0$  debe ser  $A \neq 0$ , y completando cuadrados se obtiene

$$Q(x, y) = \frac{1}{A}[(Ax + By)^2 + (AC - B^2)y^2]$$

Cuando  $(x, y) \neq (0, 0)$ , teniendo en cuenta que  $x \neq 0$  si  $y = 0$ , se obtiene que  $Q$  es definida positiva si  $A > 0$  y definida negativa si  $A < 0$ .

Si  $AC - B^2 < 0$ , y  $A \neq 0$ , completando cuadrados se puede expresar  $Q(x, y)$  como producto de dos factores lineales distintos:

$$Q(x, y) = \frac{1}{A}[(Ax + By)^2 - M^2y^2] = \frac{1}{A}[Ax + (B + M)y][Ax + (B - M)y]$$

donde  $M = \sqrt{B^2 - AC}$ . Si consideremos los semiplanos

$$U^+ = \{(x, y) : Ax + (B + M)y > 0\}, \quad U^- = \{(x, y) : Ax + (B + M)y < 0\}$$

$$V^+ = \{(x, y) : Ax + (B - M)y > 0\}, \quad V^- = \{(x, y) : Ax + (B - M)y < 0\}$$

es claro que sobre las regiones angulares  $U^+ \cap V^+$ ,  $U^- \cap V^-$  la forma cuadrática tiene el signo de  $A$ , mientras que sobre las regiones angulares  $U^+ \cap V^-$ ,  $U^- \cap V^+$  tiene el signo de  $-A$ , luego se trata de una forma cuadrática indefinida.

Cuando  $AC - B^2 < 0$ , y  $A = 0$ , debe ser  $B \neq 0$ . En este caso  $Q(x, y)$  también se expresa como producto de dos factores lineales distintos,  $Q(x, y) = y(2Bx + Cy)$  y razonando en forma similar se concluye que la forma cuadrática es indefinida. ■

En el ejercicio 9.17 volveremos obtener, con las interpretaciones geométricas oportunas, los resultados anteriores sobre formas cuadráticas de dos variables. Combinando el teorema 6.15 con la proposición 6.17 se llega al siguiente resultado

**Corolario 6.18** *Se supone que  $f : \Omega \rightarrow \mathbb{R}$  es diferenciable dos veces en  $\mathbf{a} \in \Omega \subset \mathbb{R}^2$  y que  $D_1f(\mathbf{a}) = D_2f(\mathbf{a}) = 0$ ,  $D_{11}f(\mathbf{a}) = A$ ,  $D_{12}f(\mathbf{a}) = B$ ,  $D_{22}f(\mathbf{a}) = C$ . Se verifica:*

$AC - B^2 > 0$ ,  $A > 0 \Rightarrow \mathbf{a}$  es punto de mínimo relativo estricto.

$AC - B^2 > 0$ ,  $A < 0 \Rightarrow \mathbf{a}$  es punto de máximo relativo estricto.

$AC - B^2 < 0, \Rightarrow \mathbf{a}$  es punto de silla (ni máximo ni mínimo)

*Cuando  $AC - B^2 = 0$ , la función  $f$  puede presentar en  $\mathbf{a}$  un máximo relativo, un mínimo relativo o ninguna de las dos cosas (caso dudoso).*

DEM: Para completar la demostración sólo queda mostrar ejemplos de lo que puede ocurrir en el caso dudoso  $AC - B^2 = 0$ :

El origen  $(0, 0)$  es punto estacionario de  $f(x, y) = x^2 + y^4$ , (resp.  $-(x^2 + y^4)$ ) donde se cumple  $AC - B^2 = 0$ , y es inmediato que  $f$  presenta en  $(0, 0)$  un mínimo (resp. un máximo) relativo.

El origen  $(0, 0)$  también es punto estacionario de  $g(x, y) = x^2y + y^2$ . En este caso  $AC - B^2 = 0$  y en  $(0, 0)$  no hay ni máximo ni mínimo relativo porque  $(x^2 + y)y$  cambia de signo en todo entorno de  $(0, 0)$ . ■

### 6.3. Funciones convexas

En esta sección se extienden, al contexto de las funciones diferenciables de varias variables, las caracterizaciones habituales de las funciones convexas derivables que se suelen estudiar en el curso de Análisis Matemático I. El lector interesado puede acudir al apéndice F donde se repasan y amplían los resultados básicos referentes a funciones convexas de una variable.

Recordemos que un conjunto  $\Omega \subset \mathbb{R}^n$  es convexo cuando, para cada par de puntos  $\mathbf{x}, \mathbf{y} \in \Omega$ , el segmento  $[\mathbf{x}, \mathbf{y}] := \{(1 - t)\mathbf{x} + t\mathbf{y} : 0 \leq t \leq 1\}$  está contenido en  $\Omega$ .

**Definición 6.19** *Una función  $f : \Omega \rightarrow \mathbb{R}$ , definida en un conjunto convexo  $\Omega \subset \mathbb{R}^n$  se dice que es convexa cuando para todo segmento  $[\mathbf{x}, \mathbf{y}] \subset \Omega$  y todo  $t \in (0, 1)$  se cumple:*

$$f((1 - t)\mathbf{x} + t\mathbf{y}) \leq (1 - t)f(\mathbf{x}) + tf(\mathbf{y})$$

Es inmediato que cada bola (abierta o cerrada) de centro  $\mathbf{a} \in \mathbb{R}^n$  y radio  $r > 0$ , es convexa y que en esa bola la función  $f(\mathbf{x}) = \|\mathbf{x} - \mathbf{a}\|$  es convexa. Combinando este hecho con la siguiente proposición, de carácter elemental, se obtienen buenos ejemplos de funciones convexas:

**Proposición 6.20** Si  $f : \Omega \rightarrow \mathbb{R}$  es convexa en un conjunto convexo  $\Omega \subset \mathbb{R}^n$  y  $\varphi : I \rightarrow \mathbb{R}$  es creciente y convexa en un intervalo  $I \supset f(\Omega)$ , entonces la función  $\varphi \circ f : \Omega \rightarrow \mathbb{R}$  es convexa en  $\Omega$ . (Brevemente, la composición de una función convexa con una función creciente y convexa es convexa.)

DEM: Es una comprobación rutinaria que se deja al cuidado del lector ■

**Ejemplos 6.21** Las siguientes funciones son convexas en los conjuntos convexos que se indican:

- a)  $f(\mathbf{x}) = \|\mathbf{x}\|^p$ , en  $\mathbb{R}^n$ , con  $p \geq 1$ .
- b)  $f(\mathbf{x}) = (1 + \|\mathbf{x}\|^2)^{\|\mathbf{x}\|^2}$ , en  $\mathbb{R}^n$ .
- c)  $f(\mathbf{x}) = e^{-\sqrt{1-\|\mathbf{x}\|^2}}$ , en la bola  $B(\mathbf{0}, 1)$ .

DEM: a) Basta aplicar la proposición 6.20 con  $f(\mathbf{x}) = \|\mathbf{x}\|$ , pues  $\varphi(t) = t^p$  es creciente y convexa en  $I = [0, +\infty)$ , ya que para todo  $t \in I$  se cumple  $\varphi'(t) = pt^{p-1} \geq 0$  y  $\varphi''(t) = p(p-1)t^{(p-1)(p-2)} \geq 0$ .

b)  $f(\mathbf{x}) = e^{g(\mathbf{x})}$  donde  $g(\mathbf{x}) = \|\mathbf{x}\|^2 \log(1 + \|\mathbf{x}\|^2)$ . Como la función exponencial es creciente y convexa, según la proposición 6.20, basta ver que  $g$  es convexa en  $\mathbb{R}^n$ . Esto también se puede justificar con la proposición 6.20: Según a) la función  $\|\mathbf{x}\|^2$  es convexa en  $\mathbb{R}^n$ , y la función  $\varphi(t) = t \log(1+t)$  es creciente y convexa en  $[0, +\infty)$  pues, para todo  $t \geq 0$ , se cumple  $\varphi'(t) = \log(1+t) + t/(1+t) \geq 0$ ,  $\varphi''(t) = t/(1+t) \geq 0$ .

c) Razonando como en b) basta ver que la función  $g(\mathbf{x}) = -\sqrt{1 - \|\mathbf{x}\|^2}$  es convexa en la bola abierta  $B(\mathbf{0}, 1)$ , donde está definida. Esto también se puede justificar con la proposición 6.20 considerando la función  $\varphi(t) = -\sqrt{1-t}$  que es creciente y convexa en  $[0, +\infty)$ , ya que para todo  $t \in [0, 1)$  se cumple  $\varphi'(t) = \frac{1}{2}(1-t)^{-1/2} \geq 0$ , y  $\varphi''(t) = \frac{1}{4}(1-t)^{-3/2} \geq 0$ . ■

Si  $\Omega \subset \mathbb{R}^n$  es convexo, dados  $\mathbf{a} \in \Omega$  y  $\mathbf{h} \in \mathbb{R}$  el conjunto

$$I_{\mathbf{a}, \mathbf{h}} = \{t \in \mathbb{R} : \mathbf{a} + t\mathbf{h} \in \Omega\}$$

es un intervalo porque tiene la propiedad de que dados  $x, y \in I_{\mathbf{a}, \mathbf{h}}$ , cualquier punto intermedio  $z = \alpha x + \beta y$ , con  $\alpha > 0, \beta > 0, \alpha + \beta = 1$ , está en  $I_{\mathbf{a}, \mathbf{h}}$ : Basta observar que  $\mathbf{a} + z\mathbf{h} = \alpha(\mathbf{a} + x\mathbf{h}) + \beta(\mathbf{a} + y\mathbf{h})$  pertenece al conjunto convexo  $\Omega$  porque  $\mathbf{a} + x\mathbf{h}$  y  $\mathbf{a} + y\mathbf{h}$  son puntos de  $\Omega$ . Cuando  $\Omega \subset \mathbb{R}^n$  sea un abierto convexo, el intervalo  $I_{\mathbf{a}, \mathbf{h}}$  será abierto por ser la preimagen de  $\Omega$  mediante la función continua  $t \rightarrow \mathbf{a} + t\mathbf{h}$ .

**Proposición 6.22** Sea  $\Omega \subset \mathbb{R}^n$  un conjunto convexo. Una función  $f : \Omega \rightarrow \mathbb{R}$ , es convexa si y sólo si para cada  $\mathbf{a} \in \Omega$  y cada  $\mathbf{h} \in \mathbb{R}^n$  la función  $\varphi_{\mathbf{a}, \mathbf{h}}(t) = f(\mathbf{a} + t\mathbf{h})$ , es convexa en el intervalo  $I_{\mathbf{a}, \mathbf{h}} = \{t \in \mathbb{R} : \mathbf{a} + t\mathbf{h} \in \Omega\}$ .

DEM: Una comprobación rutinaria permite establecer que si  $f$  es convexa entonces las funciones  $\varphi = \varphi_{\mathbf{a}, \mathbf{h}}$  también cumplen la condición de convexidad, es decir

$$\varphi(\alpha x + \beta y) \leq \alpha\varphi(x) + \beta\varphi(y)$$

siempre que  $x, y \in I_{\mathbf{a}, \mathbf{h}}$ ,  $\alpha > 0$ ,  $\beta > 0$ , y  $\alpha + \beta = 1$ .

Para demostrar el recíproco, consideramos un segmento arbitrario  $[\mathbf{x}, \mathbf{y}] \subset \Omega$ , al que asociamos la función  $\varphi(t) = \varphi_{\mathbf{x}, \mathbf{h}}(t) = f(\mathbf{x} + t\mathbf{h})$ , con  $\mathbf{h} = \mathbf{y} - \mathbf{x}$ , que está definida en un intervalo  $I \supset [0, 1]$ . Esta función es convexa por hipótesis, luego dados  $\alpha > 0$  y  $\beta > 0$ , con  $\alpha + \beta = 1$ , se cumple:

$$\begin{aligned} f(\alpha\mathbf{x} + \beta\mathbf{y}) &= f(\mathbf{x} + \beta\mathbf{h}) = \varphi(\beta) = \\ &= \varphi(\alpha \cdot 0 + \beta \cdot 1) \leq \alpha\varphi(0) + \beta\varphi(1) = \alpha f(\mathbf{x}) + \beta f(\mathbf{y}) \end{aligned}$$

y queda demostrado que  $f$  cumple la condición de convexidad.  $\blacksquare$

Como en el caso de las funciones de una variable se cumple que toda función convexa definida en un abierto convexo  $\Omega \subset \mathbb{R}^n$  es continua. El lector interesado en la demostración puede consultar el apéndice F, o [13] (Teor.3.5, pág.110).

Aquí sólo nos ocuparemos de la caracterización de las funciones convexas diferenciables definidas en un abierto convexo  $\Omega \subset \mathbb{R}^n$ . La razón de considerar dominios abiertos se debe a que para  $n > 1$  no hay una alternativa razonable a la noción de derivada lateral que interviene, cuando  $n = 1$ , en los extremos del intervalo.

Como en el caso de las funciones de una variable la convexidad se puede caracterizar por la condición de que la gráfica quede siempre por encima del plano tangente.

**Proposición 6.23** *Si  $f : \Omega \rightarrow \mathbb{R}$  es diferenciable en un abierto convexo  $\Omega \subset \mathbb{R}^n$ , son equivalentes*

a)  $f$  es convexa.

b)  $f(\mathbf{y}) \geq f(\mathbf{x}) + df(\mathbf{x})(\mathbf{y} - \mathbf{x})$  para todo par de puntos  $\mathbf{x}, \mathbf{y} \in \Omega$ .

DEM: a)  $\Rightarrow$  b): Dados  $\mathbf{x}, \mathbf{y} \in \Omega$ , sea  $\mathbf{h} = \mathbf{y} - \mathbf{x}$ . Si  $f$  es convexa para cada  $t \in [0, 1]$  se verifica:

$$f(\mathbf{x} + t\mathbf{h}) = f((1-t)\mathbf{x} + t\mathbf{y}) \leq (1-t)f(\mathbf{x}) + tf(\mathbf{y})$$

luego  $f(\mathbf{x} + t\mathbf{h}) - f(\mathbf{x}) \leq t(f(\mathbf{y}) - f(\mathbf{x}))$ . Dividiendo por  $t > 0$  se conserva la desigualdad y se obtiene

$$\frac{f(\mathbf{x} + t\mathbf{h}) - f(\mathbf{x})}{t} \leq f(\mathbf{y}) - f(\mathbf{x})$$

Como  $f$  es diferenciable en  $\mathbf{x}$ , existe la derivada  $D_{\mathbf{h}}f(\mathbf{x})$ , y pasando al límite cuando  $t \rightarrow 0+$ , resulta

$$df(\mathbf{x})\mathbf{h} = D_{\mathbf{h}}f(\mathbf{x}) \leq f(\mathbf{y}) - f(\mathbf{x})$$

es decir  $f(\mathbf{y}) \geq f(\mathbf{x}) + df(\mathbf{x})(\mathbf{y} - \mathbf{x})$ .

b)  $\Rightarrow$  a): Si se cumple b), dados  $\mathbf{x}, \mathbf{y} \in \Omega$  y  $t \in [0, 1]$ , consideremos el punto  $\mathbf{x}_t = (1-t)\mathbf{x} + t\mathbf{y}$ . Aplicando la desigualdad b) a los puntos  $\mathbf{x}, \mathbf{x}_t$  y a los puntos  $\mathbf{y}, \mathbf{x}_t$  se obtiene

$$\text{i) } f(\mathbf{x}) \geq f(\mathbf{x}_t) + df(\mathbf{x}_t)(\mathbf{x} - \mathbf{x}_t)$$

$$\text{ii) } f(\mathbf{y}) \geq f(\mathbf{x}_t) + df(\mathbf{x}_t)(\mathbf{y} - \mathbf{x}_t)$$

Multiplicando la primera desigualdad por  $(1-t) \geq 0$ , la segunda por  $t \geq 0$  sumando miembro a miembro, y teniendo en cuenta que  $(1-t)(\mathbf{x} - \mathbf{x}_t) + t(\mathbf{y} - \mathbf{x}_t) = 0$  resulta  $(1-t)f(\mathbf{x}) + tf(\mathbf{y}) \geq f(\mathbf{x}_t)$ , es decir, se cumple la condición de convexidad. ■

**Corolario 6.24** *Sea  $f : \Omega \rightarrow \mathbb{R}$  una función convexa y diferenciable en el abierto convexo  $\Omega \subset \mathbb{R}^n$ . Si  $\mathbf{a} \in \Omega$  es un punto estacionario de  $f$  (e.d.  $D_k f(\mathbf{a}) = 0$  para  $1 \leq k \leq n$ ) entonces  $f$  presenta en  $\mathbf{a}$  un mínimo absoluto.*

DEM: Según 6.23, para todo  $\mathbf{x} \in \Omega$  se cumple  $f(\mathbf{x}) \geq f(\mathbf{a}) + df(\mathbf{a})(\mathbf{x} - \mathbf{a}) = f(\mathbf{a})$ . ■

**Proposición 6.25** *Si  $f : \Omega \rightarrow \mathbb{R}$  es diferenciable dos veces en un abierto convexo  $\Omega \subset \mathbb{R}^n$ , son equivalentes*

a)  $f$  es convexa.

b)  $\sum_{i,j=1}^n D_{ij} f(\mathbf{x}) u_i u_j \geq 0$ , para todo  $\mathbf{x} \in \Omega$  y todo  $\mathbf{u} \in \mathbb{R}^n$ .

DEM: a)  $\Rightarrow$  b): A cada  $\mathbf{x} \in \Omega$  le asociamos la forma cuadrática

$$Q_{\mathbf{x}}(\mathbf{h}) = \frac{1}{2} \sum_{i,j=1}^n D_{ij} f(\mathbf{x}) h_i h_j$$

Según 6.9,  $f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) - df(\mathbf{x})\mathbf{h} = Q_{\mathbf{x}}(\mathbf{h}) + \|\mathbf{h}\|^2 \epsilon(\mathbf{h})$ , donde  $\lim_{\mathbf{h} \rightarrow \mathbf{0}} \epsilon(\mathbf{h}) = 0$ . Dado  $\mathbf{u} \in \mathbb{R}^n$  existe  $r > 0$  tal que  $0 < t < \delta \Rightarrow \mathbf{x} + t\mathbf{u} \in \Omega$ , y aplicando 6.23 b) con  $\mathbf{y} = \mathbf{x} + t\mathbf{u}$ , se obtiene

$$Q_{\mathbf{x}}(t\mathbf{u}) + \|t\mathbf{u}\|^2 \epsilon(t\mathbf{u}) = f(\mathbf{x} + t\mathbf{u}) - f(\mathbf{x}) - df(\mathbf{x})(t\mathbf{u}) \geq 0$$

es decir

$$t^2 Q_{\mathbf{x}}(\mathbf{u}) + t^2 \|\mathbf{u}\|^2 \epsilon(t\mathbf{u}) \geq 0 \text{ para todo } 0 < t < \delta.$$

Dividiendo por  $t^2 > 0$  se conserva la desigualdad, y pasando al límite cuando  $t \rightarrow 0+$  resulta  $Q_{\mathbf{x}}(\mathbf{u}) \geq 0$ , para todo  $\mathbf{x} \in \Omega$  y todo  $\mathbf{u} \in \mathbb{R}^n$ .

b)  $\Rightarrow$  a) Recíprocamente, si se cumple b), dados  $\mathbf{x} \in \Omega$ ,  $\mathbf{h} \in \mathbb{R}^n$ , la función real de variable real  $\varphi(t) = f(\mathbf{x} + t\mathbf{h})$ , que está definida en un intervalo abierto  $I \subset \mathbb{R}$ , es derivable dos veces en cada  $t \in I$  y su derivada vale

$$\varphi''(t) = \sum_{i,j=1}^n D_{ij} f(\mathbf{x} + t\mathbf{h}) h_i h_j \geq 0$$

Un resultado bien conocido de la teoría de funciones reales de una variable (véase F.5) permite asegurar que  $\varphi$  es convexa sobre el intervalo  $I$ , y con la proposición 6.22 se concluye que  $f$  es convexa. ■

## 6.4. Ejercicios resueltos

**Ejercicio 6.26** Sea  $\Omega = \mathbb{R}^n \setminus \{\mathbf{0}\}$ . Obtenga las funciones  $f : \Omega \rightarrow \mathbb{R}$  que son de la forma  $f(\mathbf{x}) = g(\|\mathbf{x}\|)$  con  $g : (0, +\infty) \rightarrow \mathbb{R}$  de clase  $C^2$ , y que satisfacen la ecuación de Laplace  $\sum_{j=1}^n D_{jj}f = 0$ . ( $\|\mathbf{x}\|$  es la norma euclídea).

SOLUCIÓN

$$\frac{\partial f(\mathbf{x})}{\partial x_j} = g'(\|\mathbf{x}\|) \frac{\partial \|\mathbf{x}\|}{\partial x_j} = g'(\|\mathbf{x}\|) \frac{x_j}{\|\mathbf{x}\|}$$

Derivando otra vez respecto a  $x_j$ :

$$\frac{\partial^2 f(\mathbf{x})}{\partial x_j^2} = g''(\|\mathbf{x}\|) \frac{x_j^2}{\|\mathbf{x}\|^2} + \left(1 - \frac{x_j^2}{\|\mathbf{x}\|^2}\right) \frac{g'(\|\mathbf{x}\|)}{\|\mathbf{x}\|}$$

Sumando para  $j = 1, 2, \dots, n$  resulta:

$$\Delta f(\mathbf{x}) = g''(\|\mathbf{x}\|) + (n-1) \frac{g'(\|\mathbf{x}\|)}{\|\mathbf{x}\|}$$

Si  $r = \|\mathbf{x}\|$  la ecuación  $\Delta f(\mathbf{x}) = 0$  se escribe en la forma

$$g''(r) + (n-1) \frac{g'(r)}{r} = 0$$

La derivada de  $g'(r)e^{(n-1)\log r}$  es idénticamente nula luego existe una constante  $K$  tal que  $g'(r) = Ke^{(1-n)\log r} = Kr^{1-n}$ , y se concluye que

$$g(r) = A + \frac{K}{(2-n)r^{n-2}} = A + \frac{B}{r^{n-2}} \quad \text{si } n \geq 3.$$

$$g(r) = A + K \log r \quad \text{si } n = 2.$$

■

**Ejercicio 6.27** Sea  $F = \varphi \circ f : \Omega \rightarrow \mathbb{R}$  definida en un abierto  $\Omega \subset \mathbb{R}^n$ , donde  $f : \Omega \rightarrow \mathbb{R}$  es diferenciable dos veces en  $\mathbf{a} \in \Omega$  y  $\varphi : V \rightarrow \mathbb{R}$ , definida en un abierto  $V \subset \mathbb{R}$ , con  $V \supset f(\Omega)$ , es derivable dos veces en  $b = f(\mathbf{a})$ .

Demuestre que es  $F$  diferenciable dos veces en  $\mathbf{a}$  y obtenga, para cada  $\mathbf{h} \in \mathbb{R}$ , el valor de  $d^2F(\mathbf{a})\mathbf{h}^2$  en términos de las funciones  $f$  y  $\varphi$ .

Si  $\varphi'(b) > 0$  y  $\varphi''(b) \geq 0$ , obtenga que  $d^2f(\mathbf{a})\mathbf{h}^2 > 0 \Rightarrow d^2F(\mathbf{a})\mathbf{h}^2 > 0$ .

SOLUCIÓN

La existencia de la derivada segunda  $\varphi''(b)$  implica que hay un entorno de  $b$ ,  $V_b \subset V$  donde  $\varphi$  es derivable. Como  $f$  es continua en  $\mathbf{a}$ , hay un entorno abierto de  $\mathbf{a}$ ,  $U_{\mathbf{a}} \subset \Omega$ , tal que  $f(U_{\mathbf{a}}) \subset V_b$ . De acuerdo con la definición de función dos veces diferenciable



en  $\mathbf{a}$  podemos suponer también que  $f$  es diferenciable en  $U_{\mathbf{a}}$ , donde estarán definidas las derivadas parciales  $D_j f$ ,  $1 \leq j \leq n$  que, por hipótesis, son diferenciables en  $\mathbf{a}$ .

Según la regla de la cadena la composición  $F = \varphi \circ f$  es diferenciable en  $U_{\mathbf{a}}$  donde sus derivadas parciales vienen dadas por

$$D_j F(\mathbf{x}) = \varphi'(f(\mathbf{x}))D_j f(\mathbf{x}), \quad 1 \leq j \leq n$$

Como las dos funciones  $\varphi'(f(\mathbf{x}))$ ,  $D_j f(\mathbf{x})$  son diferenciables en  $\mathbf{x} = \mathbf{a}$ , también lo es su producto y por lo tanto cada derivada parcial  $D_j F$  es diferenciable en  $\mathbf{a}$ . Queda demostrado así que  $F$  es diferenciable dos veces en  $\mathbf{a}$ . Las derivadas parciales  $D_{ij} F(\mathbf{a})$  se pueden calcular con la regla para derivar un producto:

$$D_{ij} F(\mathbf{a}) = D_i D_j F(\mathbf{a}) = \varphi''(f(\mathbf{a}))D_i f(\mathbf{a})D_j f(\mathbf{a}) + \varphi'(f(\mathbf{a}))D_i D_j f(\mathbf{a})$$

Entonces, para cada  $\mathbf{h} \in \mathbb{R}^n$  se verifica

$$\sum_{i,j=1}^n D_{ij} F(\mathbf{a})h_i h_j = \varphi''(b) \sum_{i,j=1}^n D_i f(\mathbf{a})D_j f(\mathbf{a})h_i h_j + \varphi'(b) \sum_{i,j=1}^n D_{ij} f(\mathbf{a})h_i h_j$$

es decir

$$\sum_{i,j=1}^n D_{ij} F(\mathbf{a})h_i h_j = \varphi''(b) \left[ \sum_{i=1}^n D_i f(\mathbf{a})h_i \right]^2 + \varphi'(b) \sum_{i,j=1}^n D_{ij} f(\mathbf{a})h_i h_j$$

Si  $\varphi'(b) > 0$  y  $\varphi''(b) \geq 0$ , manejando la última igualdad es inmediato que

$$d^2 f(\mathbf{a})\mathbf{h}^2 = \sum_{i,j=1}^n D_{ij} f(\mathbf{a})h_i h_j > 0 \quad \Rightarrow \quad d^2 F(\mathbf{a})\mathbf{h}^2 = \sum_{i,j=1}^n D_{ij} F(\mathbf{a})h_i h_j > 0$$

■

**Ejercicio 6.28** Determine los valores del parámetro  $a \in \mathbb{R}$  para los que la función

$$f_a(x, y) = a(2xy + y^2 + yx^2 + \cos(x + y)) + x^2(a^2 - y)$$

presenta un extremo relativo en el punto  $(0, 0)$ .

SOLUCIÓN

Se comprueba fácilmente que para cualquier valor de  $a \in \mathbb{R}$  el punto  $(0, 0)$  es estacionario:  $D_1 f(0, 0) = D_2 f(0, 0) = 0$ . Un cálculo rutinario permite obtener la matriz Hessiana

$$\begin{pmatrix} D_{11} f(0, 0) & D_{12} f(0, 0) \\ D_{21} f(0, 0) & D_{22} f(0, 0) \end{pmatrix} = \begin{pmatrix} 2a^2 - a & a \\ a & a \end{pmatrix}$$

cuyo determinante vale  $\Delta(a) = 2a^2(a - 1)$ .

Si  $a > 1$  se cumple  $\Delta(a) > 0$  y  $2a^2 - a > 0$  luego, en virtud de 6.18,  $f$  presenta en  $(0, 0)$  un mínimo relativo.

Si  $a < 1$  y  $a \neq 0$  se cumple  $\Delta(a) < 0$  luego, en virtud de 6.18,  $f$  no presenta en  $(0, 0)$  ni máximo ni mínimo relativo (punto de silla).

Finalmente, cuando  $a = 1$  ó  $a = 0$  se cumple  $\Delta(a) = 0$  por lo que no es aplicable el criterio 6.18 (caso dudoso) y debemos estudiar directamente la función.

Cuando  $a = 1$  la función toma la forma

$$f(x, y) = (x + y)^2 + \cos(x + y) = \varphi(x + y)$$

donde  $\varphi(t) = t^2 + \cos t$  presenta un mínimo relativo en  $t = 0$ , luego  $f$  también presenta un mínimo relativo en  $(0, 0)$ .

Cuando  $a = 0$  la función se reduce a  $f(x, y) = -x^2y$  y es obvio que en  $(0, 0)$  no hay ni máximo ni mínimo relativo.

Resumiendo:  $f$  presenta extremo relativo en  $(0, 0)$  si y sólo si  $a \geq 1$ , y el extremo siempre es un mínimo relativo. ■

**Ejercicio 6.29** *Estudie la existencia de extremos relativos de la función de tres variables  $f(x, y, z) = xy + yz + zx$ .*

SOLUCIÓN

Los puntos estacionarios, donde  $f$  puede presentar extremo local, son las soluciones del sistema  $D_k f(x, y, z) = 0$ ,  $k = 1, 2, 3$ , es decir:  $y + z = 0$ ;  $x + z = 0$ ;  $y + x = 0$ .

La única solución de este sistema lineal es  $x = 0$ ,  $y = 0$ ,  $z = 0$ . En este punto las derivadas parciales segundas valen

$$D_{ij}f(0, 0, 0) = 1 \text{ si } i \neq j, \quad D_{ii}f(0, 0, 0) = 0$$

El criterio de la proposición 6.17 no sirve aquí para decidir si  $f$  presenta en  $(0, 0, 0)$  un extremo local, por lo que debemos estudiar directamente la función. Para simplificar empezamos restringiendo  $f$  al plano  $y = z$  (que contiene al punto  $(0, 0, 0)$ ). Obtenemos así una función de dos variables

$$h(x, y) = f(x, y, y) = y(y + 2x)$$

en la que es fácil apreciar cambio de signo en todo entorno de  $(0, 0)$ : Para todo  $t > 0$ , se cumple  $h(t, t) = f(t, t, t) = 3t^2 > 0$  y  $h(-t, t) = f(-t, t, t) = -2t^2 < 0$ , luego  $f$  no presenta un extremo local en  $(0, 0, 0)$ , pues  $f(0, 0, 0) = 0$ .

**Ejercicio 6.30** *Encuentre los extremos relativos y absolutos de la función de tres variables  $g(x, y, z) = (x + z^2)e^{x(1+y^2+z^2)}$ .*

SOLUCIÓN

Empecemos calculando los puntos estacionarios, donde  $g$  puede presentar extremos locales. Para ello debemos resolver el sistema de las tres ecuaciones  $D_k f(x, y, z) = 0$ ,  $k = 1, 2, 3$ , es decir:

$$\text{i) } [(1 + y^2 + z^2)(x + z^2) + 1]e^{x(1+y^2+z^2)} = 0;$$

$$\text{ii) } 2xy(x + z^2)e^{x(1+y^2+z^2)} = 0;$$

$$\text{iii) } [2z + 2xz(x + z^2)]e^{x(1+y^2+z^2)} = 0;$$

El sistema equivale al siguiente

$$\text{a) } (1 + y^2 + z^2)(x + z^2) + 1 = 0;$$

$$\text{b) } xy(x + z^2) = 0;$$

$$\text{c) } z[1 + x(x + z^2)] = 0;$$

La ecuación b) conduce a tres alternativas

$$x = 0; \quad x + z^2 = 0; \quad y = 0$$

La ecuación a) permite descartar la dos primeras. Sustituyendo la tercera alternativa  $y = 0$  en la otras dos ecuaciones obtenemos

$$\text{a) } (1 + z^2)(x + z^2) + 1 = 0;$$

$$\text{c') } z(1 + x(x + z^2)) = 0$$

Ahora la ecuación c') conduce a dos posibilidades  $z = 0$  ó  $x(x + z^2) = -1$ .

Con  $z = 0$  obtenemos la solución  $x = -1, y = 0, z = 0$ .

La ecuación  $x(x + z^2) = -1$  no proporciona soluciones: Si multiplicamos a) por  $x$  resulta  $x + x(x + z^2)(1 + z^2) = 0$ , luego  $x - (1 + z^2) = 0$ , y sustituyendo  $x = 1 + z^2$  en a) se obtiene una ecuación sin soluciones:  $1 + (1 + z^2)[1 + z^2 + z^2] = 0$ .

Hemos obtenido así que  $\mathbf{p} = (-1, 0, 0)$  es el un único punto estacionario de  $g$ . Un cálculo rutinario proporciona la matriz Hessiana

$$\begin{pmatrix} D_{11}g(\mathbf{p}) & D_{12}g(\mathbf{p}) & D_{13}g(\mathbf{p}) \\ D_{21}g(\mathbf{p}) & D_{22}g(\mathbf{p}) & D_{23}g(\mathbf{p}) \\ D_{31}g(\mathbf{p}) & D_{32}g(\mathbf{p}) & D_{33}g(\mathbf{p}) \end{pmatrix} = \begin{pmatrix} 1/e & 0 & 0 \\ 0 & 2/e & 0 \\ 0 & 0 & 4/e \end{pmatrix}$$

Como los determinantes de los sucesivos menores cumplen la condición

$$\Delta_1 > 0; \quad \Delta_2 > 0; \quad \Delta_3 > 0;$$

en virtud de 6.15 y 6.17  $g$  presenta un mínimo relativo en  $\mathbf{p} = (-1, 0, 0)$ .

Para averiguar si este mínimo relativo  $g(-1, 0, 0) = -1/e$  es un mínimo absoluto efectuamos un estudio directo sobre la función. Para simplificar la cuestión podemos restringir el estudio al abierto donde  $g$  es negativa  $A := \{(x, y, z) : x + z^2 < 0\}$ . En un punto de este abierto siempre es  $x < 0$ , luego  $e^x \geq e^{x(1+y^2+z^2)}$ . Multiplicando por  $x$  ( $< 0$ ) resulta

$$xe^x \leq xe^{x(1+y^2+z^2)} \leq (x + z^2)e^{x(1+y^2+z^2)}$$

es decir,  $g(x, 0, 0) \leq g(x, y, z)$  para todo  $(x, y, z) \in A$ .

Con los métodos del cálculo de funciones de una variable se obtiene fácilmente que la función  $\varphi(x) = g(0, 0, x) = xe^x$  alcanza un mínimo absoluto para  $x = -1$  que vale  $\varphi(-1) = -1/e$ , es decir,

$$-1/e = g(-1, 0, 0) \leq g(x, y, z) \text{ para todo } (x, y, z) \in A$$

Como en  $\mathbb{R}^3 \setminus A$  los valores de  $g$  son positivos, se sigue que  $g(-1, 0, 0) = -1/e$  es el mínimo absoluto de  $g$  en todo  $\mathbb{R}^3$ . ■

**Ejercicio 6.31** Demuestre que la forma cuadrática asociada a una matriz simétrica:

$$Q(\mathbf{x}) = \sum_{i,j=1}^n \alpha_{ij} x_i x_j \text{ con } \alpha_{ij} = \alpha_{ji}, \quad 1 \leq i, j \leq n$$

es convexa si y sólo si es definida no negativa e.d.  $Q(\mathbf{x}) \geq 0$  para todo  $\mathbf{x} \in \mathbb{R}^n$ .

SOLUCIÓN

Es consecuencia inmediata de 6.25 ya que para todo  $\mathbf{x} \in \mathbb{R}$  se verifica  $D_{ij}Q(\mathbf{x}) = 2\alpha_{ij}$ , luego  $\sum_{i,j=1}^n D_{ij}Q(\mathbf{x})u_i u_j = 2Q(\mathbf{u})$ . ■

**Ejercicio 6.32** Sea  $\alpha < 0 < \beta$  y  $g : (\alpha, \beta) \rightarrow \mathbb{R}$  una función de clase  $C^2$  tal que  $g''(t) \leq 0$  para todo  $t \in [0, \beta)$ . Se considera la función  $f(x, y) = g(x^2 + y^2)$ , que está definida en  $\{(x, y) : x^2 + y^2 < \beta\}$ .

Si  $0 < R^2 < \beta$ , demuestre que  $f$  es convexa sobre el disco  $x^2 + y^2 < R^2$  si y sólo si  $g'(t) + 2tg''(t) \geq 0$  para todo  $t \in [0, R^2]$ .

Para cada una de las funciones  $\log(1 + x^2 + y^2)$ ,  $\sin(x^2 + y^2)$  obtenga el mayor disco  $D_R = \{(x, y) : x^2 + y^2 < R^2\}$  sobre el que la función es convexa.

DEM: Como  $f$  es de clase  $C^2$  para estudiar su convexidad podemos utilizar la proposición 6.25 donde interviene la forma cuadrática

$$Q_{\mathbf{p}}(\mathbf{h}) = D_{11}f(\mathbf{p})h_1^2 + 2D_{12}f(\mathbf{p})h_1 h_2 + D_{22}f(\mathbf{p})h_2^2$$

asociada a un punto genérico  $\mathbf{p}$  de su dominio. En lo que sigue conviene expresar  $\mathbf{p} = (a, b)$  en coordenadas polares,  $a = r \cos \theta$ ,  $b = r \sin \theta$ .

Calculando previamente las derivadas parciales

$$D_{11}f(\mathbf{p}) = 2g'(r^2) + 4a^2g''(r^2); \quad D_{12}f(\mathbf{p}) = 4abg''(r^2); \quad D_{22}f(\mathbf{p}) = 2g'(r^2) + 4b^2g''(r^2).$$

resulta

$$\begin{aligned} Q_{\mathbf{p}}(\mathbf{h}) &= 2g'(r^2)(h_1^2 + h_2^2) + 4g''(r^2)(ah_1 + bh_2)^2 = \\ &= 2g'(r^2)(h_1^2 + h_2^2) + 4r^2g''(r^2)(h_1 \cos \theta + h_2 \sin \theta)^2 \end{aligned}$$

Si suponemos que para todo  $t \in [0, R^2]$  se verifica  $g'(t) + 2tg''(t) \geq 0$ , entonces  $2g'(r^2) \geq -4r^2g''(r^2)$  para cada  $r \in [0, R]$ , luego para cada  $\mathbf{p} \in D_R$  y cada  $\mathbf{h} \in \mathbb{R}^2$  se cumple la desigualdad

$$Q_{\mathbf{p}}(\mathbf{h}) \geq -4r^2g''(r^2)(h_1^2 + h_2^2) + 4r^2g''(r^2)(h_1 \cos \theta + h_2 \sin \theta)^2 =$$

$$= -4r^2 g''(r^2)(h_1 \sin \theta + h_2 \cos \theta)^2$$

Por hipótesis  $g''(r^2) \leq 0$ , luego  $Q_{\mathbf{p}}(\mathbf{h}) \geq 0$  para cada  $\mathbf{p} \in D_R$  y cada  $\mathbf{h} \in \mathbb{R}^2$ , y con la proposición 6.25 concluimos que  $f$  es convexa en  $D_R$

Recíprocamente, si  $f$  es convexa en  $D_R$ , según la proposición 6.25, podemos asegurar que para todo  $r \in [0, R]$  y todo  $\mathbf{h} \in \mathbb{R}^2$  se cumple

$$Q_{\mathbf{p}}(\mathbf{h}) = 2g'(r^2)(h_1^2 + h_2^2) + 4g''(r^2)(ah_1 + bh_2)^2 \geq 0$$

Aplicando esta desigualdad cuando  $a = r < R$ ,  $b = 0$ ,  $h_1 > 0$ , y  $h_2 = 0$  se obtiene que  $g'(r^2)h_1^2 + 2g''(r^2)r^2h_1^2 \geq 0$ , luego  $g'(r^2) + 2g''(r^2)r^2 \geq 0$  para todo  $r \in [0, R]$ .

De acuerdo con lo que se acaba de establecer, para determinar el mayor disco  $D_R = \{(x, y) : x^2 + y^2 < R^2\}$  donde  $\log(1 + x^2 + y^2)$  es convexa consideramos la función  $g(t) = \log(1 + t)$ , sobre el intervalo  $(-1, +\infty)$ , donde cumple  $g''(t) = -1/(1 + t)^2 \leq 0$ . En este caso la desigualdad  $g'(t) + 2tg''(t) = (1 - t)/(1 + t)^2 \geq 0$  se satisface si y sólo si  $t \leq 1$ , luego la función  $\log(1 + x^2 + y^2)$  es convexa en el disco  $D_1$ , y este es el mayor disco centrado en  $(0, 0)$  donde es convexa.

Análogamente, para obtener el mayor disco  $D_R$  sobre el que  $\sin(x^2 + y^2)$  es convexa consideramos la función  $g(t) = \sin t$  en el intervalo  $(-\pi/2, \pi/2)$ , donde cumple que  $g''(t) = -\sin t \leq 0$ . Ahora  $g'(t) + 2tg''(t) = \cos t - 2t \sin t$ , y el mayor intervalo  $[0, R^2]$  donde se cumple la desigualdad  $\cos t - 2t \sin t \geq 0$  se consigue cuando  $R^2$  es el único cero que tiene en  $(0, \pi/2)$  la función  $\varphi(t) = \cos t - 2t \sin t$  (obsérvese que  $\varphi$  es decreciente en  $[0, \pi/2]$ , y que  $\varphi(0) = 1$ ,  $\varphi(\pi/2) = -\pi$ ). ■

## 6.5. Ejercicios propuestos

◇ **6.5.1** Sea  $g : \mathbb{R}^2 \rightarrow \mathbb{R}$  definida por

$$g(x, y) = \frac{x \operatorname{sen} y - y \operatorname{sen} x}{x^2 + y^2} \quad \text{si } (x, y) \neq (0, 0); \quad g(0, 0) = 0$$

Compruebe que  $g$  es diferenciable en  $(0, 0)$ , y que existen y son distintas las derivadas parciales  $D_{12}g(0, 0) \neq D_{21}g(0, 0)$ .

◇ **6.5.2** Sea  $f(x, y) = xyg(x, y)$  donde  $g : \mathbb{R}^2 \rightarrow \mathbb{R}$  es una función acotada diferenciable en  $\mathbb{R}^2 \setminus \{(0, 0)\}$ . Justifique las afirmaciones siguientes:

i) Si  $\lim_{x \rightarrow 0} g(x, 0) \neq \lim_{y \rightarrow 0} g(0, y)$  entonces  $f$  es diferenciable en  $(0, 0)$  pero no es diferenciable dos veces en  $(0, 0)$ .

ii) Si  $g$  es continua en  $(0, 0)$  y sus derivadas parciales están acotadas entonces  $f$  es diferenciable dos veces en  $(0, 0)$ .

◇ **6.5.3** En cada caso estudie si  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  es dos veces diferenciable en  $(0, 0)$ :

i)  $f(x, y) = y^2 \operatorname{sen}(x/y^2)$  si  $y \neq 0$ ,  $f(x, 0) = 0$ .

ii)  $f(x, y) = \frac{x}{\sqrt{x^2 + y^2}} \operatorname{sen}(x^2 + y^2)$  si  $(x, y) \neq 0$ ,  $f(0, 0) = 0$

iii)  $f(x, y) = \frac{\operatorname{sen} x^2 \operatorname{sen} y^2}{x^2 + y^2}$  si  $(x, y) \neq (0, 0)$   $f(0, 0) = 0$ .

◇ **6.5.4** Demuestre que la función

$$f(x, y) = y^4/(x^2 + y^2) \quad \text{si } (x, y) \neq (0, 0); \quad f(0, 0) = 0$$

es de clase  $C^1$  en  $\mathbb{R}^2$ . Estudie si  $f$  es diferenciable dos veces en  $(0, 0)$ .

◇ **6.5.5** Sea  $f_\alpha : \mathbb{R}^2 \rightarrow \mathbb{R}$  definida por

$$f_\alpha(x, y) = \frac{x^\alpha}{x^2 + y^2} \quad \text{si } (x, y) \neq (0, 0) \quad \text{y } f_\alpha(0, 0) = 0.$$

Estudie los valores  $\alpha \in \mathbb{R}$ ,  $\alpha > 2$ , para los que  $f_\alpha$  es dos veces diferenciable en  $(0, 0)$ .

◇ **6.5.6** Estudie los puntos  $(a, b) \in \mathbb{R}^2$  donde es dos veces diferenciable la función  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  definida por  $f(x, y) = yx^2 \operatorname{sen}(1/x)$  si  $x \neq 0$ ;  $f(0, y) = 0$ .

◇ **6.5.7** Sean  $u, v : \mathbb{R}^2 \rightarrow \mathbb{R}$  de clase  $C^2$  verificando

$$u_x(x, y) = v_y(x, y), \quad u_y(x, y) = -v_x(x, y) \quad \text{para todo } (x, y) \in \mathbb{R}^2.$$

Si  $g : \mathbb{R}^2 \rightarrow \mathbb{R}$  es de clase  $C^2$  y satisface la ecuación de Laplace  $D_{11}g + D_{22}g = 0$  compruebe que  $G(x, y) = g(u(x, y), v(x, y))$  también la satisface.

◇ **6.5.8** Sea  $f : A \rightarrow \mathbb{R}$  de clase  $C^2(A)$  en un abierto conexo  $A \subset \mathbb{R}^n$ , tal que  $\nabla f(\mathbf{x}) \neq 0$  y  $\sum_{i=1}^n D_{ii}f(\mathbf{x}) = 0$  para todo  $\mathbf{x} \in A$ . Demuestre que  $B = f(A)$  es un intervalo abierto y obtenga la forma general de las funciones de clase  $C^2(B)$ ,  $g : B \rightarrow \mathbb{R}$  tales que  $F = g \circ f$  verifica  $\sum_{i=1}^n D_{ii}F(\mathbf{x}) = 0$ .

◇ **6.5.9** Compruebe que la función  $f(x, y, z) = \int_x^{y^2} \frac{\operatorname{sen} zt}{t} dt$  está bien definida y es de clase  $C^2$  en todo  $\mathbb{R}^3$ . Calcule el máximo valor de la derivada direccional  $D_{\mathbf{u}}f(1, 1, \pi/2)$ ,  $\|\mathbf{u}\|_2 = 1$ .

◇ **6.5.10** Justifique que la integral  $F(x, y, z) = \int_y^z \frac{\operatorname{sen} tx}{t} dt$  define en  $\mathbb{R}^3$  una función de clase  $C^2$  que verifica  $zF_z(x, y, z) = xF_x(x, y, z) + yF_y(x, y, z)$ .

◇ **6.5.11** Sean  $A, \Omega \subset \mathbb{R}^n$  abiertos y  $\mathbf{g} : A \rightarrow \Omega$  un homeomorfismo. Dada  $f : \Omega \rightarrow \mathbb{R}$  compruebe que  $F = f \circ \mathbf{g} : A \rightarrow \mathbb{R}$  tiene un extremo relativo en  $\mathbf{a} \in A$  si y sólo si  $f$  tiene un extremo relativo (del mismo tipo) en  $\mathbf{b} = \mathbf{g}(\mathbf{a})$ . Aplique lo anterior para determinar los extremos relativos de la función

$$f(x, y) = \left( \operatorname{arctg} \frac{y}{x} - 1 \right)^2 + \sqrt{(x^2 + y^2)}(x^2 + y^2 - 3)$$

definida en  $\Omega = \{(x, y) \in \mathbb{R}^2 : x > 0, y > 0\}$ .

◇ **6.5.12** Sea  $f : \Omega \rightarrow \mathbb{R}$  definida en un abierto  $\Omega \subset \mathbb{R}^n$  y  $\varphi : I \rightarrow \mathbb{R}$  una función creciente definida en un intervalo  $I \supset f(\Omega)$ . Compruebe que  $f$  y  $F = \varphi \circ f$  tienen los mismos extremos relativos. ¿Qué ocurre cuando  $\varphi$  es decreciente?. Aplique lo anterior para determinar los extremos relativos de las funciones

a)  $F(x, y) = e^{x((\log x)^2 + y^2)}$ , definida en  $\{(x, y) : x > 0\}$ ;

b)  $F(x, y) = \frac{x^3 y^3}{(x-a)(y-b)}$ , definida en  $\mathbb{R}^2 \setminus \{(a, b)\}$ ;

◇ **6.5.13** Obtenga los extremos relativos de los polinomios

$$x^3 + 3xy^2 - 15x - 12y; \quad 3x^4 - 4x^2y + y^2;$$

$$(3-x)(3-y)(x+y-3) \quad x^2 - y^2 + x^3 + x^2y + y^3/3$$

Estudie en cada caso la existencia de extremos absolutos.

◇ **6.5.14** Obtenga los extremos relativos de las funciones de dos variables

$$(ax^2 + by^2)e^{-(x^2+y^2)}; \quad \frac{x-y}{x^2+y^2+1}; \quad 1 - (x^2 + y^2)^{2/3}$$

Estudie en cada caso la existencia de extremos absolutos.

◇ **6.5.15** Obtenga los extremos relativos de la función de dos variables

$$f(x, y) = \log(1 + x^2 + y^2) - \int_0^x \frac{2t}{1 + t^4} dt$$

◇ **6.5.16** Demuestre que hay un entorno del punto  $(0, 0, 1)$  en el que la gráfica de la función  $z = y^2 + y \cos x + e^{2x} \cos y$  queda por encima de su plano tangente en este punto.

◇ **6.5.17** Estudie (según los valores de  $a, b \in \mathbb{R}$ ) cuando la superficie

$$z = e^{ax+y^2} + b \cos(x^2 + y^2)$$

queda por encima o por debajo de su plano tangente en un entorno de  $(0, 0)$  y cuando no ocurre ni una cosa ni la otra.

◇ **6.5.18** Sea  $\Omega \subset \mathbb{R}^n$  un abierto convexo. Se dice que  $f : \Omega \rightarrow \mathbb{R}$  es estrictamente convexa si

$$f(t\mathbf{x} + (1-t)\mathbf{y}) < tf(\mathbf{x}) + (1-t)f(\mathbf{y})$$

para todo par de puntos distintos  $\mathbf{x}, \mathbf{y} \in \Omega$  y todo  $t \in (0, 1)$ . Demuestre:

i) Una función diferenciable  $f : \Omega \rightarrow \mathbb{R}$  es estrictamente convexa si y sólo si

$$f(\mathbf{a}) + df(\mathbf{a})(\mathbf{x} - \mathbf{a}) < f(\mathbf{x}) \quad \text{para todo par de puntos distintos } \mathbf{a}, \mathbf{x} \in \Omega$$

ii) Una función diferenciable estrictamente convexa presenta a lo más un único extremo relativo que necesariamente es un mínimo absoluto.

iii) Una función de clase  $C^2$ ,  $f : \Omega \rightarrow \mathbb{R}$  que verifica  $d^2f(\mathbf{x})\mathbf{h}^2 > 0$  para todo  $\mathbf{h} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$  es estrictamente convexa.

◇ **6.5.19** Compruebe que la función  $x^3 + y^3 + 6xy$  es convexa en el abierto  $A = \{(x, y) : xy > 1, x > 0\}$ .

◇ **6.5.20** Compruebe que la función  $f(x, y) = x^2 + y(y^3 - 4)$  es convexa en  $\mathbb{R}^2$  y estudie la existencia de mínimo absoluto global.

◇ **6.5.21** Obtenga una bola centrada en  $(0, 0, 0)$  sobre la que sea convexa la función  $\log(1 + x^2 + y^2 + z^2)$ .

◇ **6.5.22** Sea  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$  una función derivable con  $\varphi'(0) > 0$ . Demuestre que  $f(x, y) = \varphi(x^2 + y^2)$  presenta un mínimo relativo en  $(0, 0)$ . Indique una clase de funciones  $\varphi$  para las que se pueda asegurar que  $f$  es convexa en  $\mathbb{R}^2$ .

◇ **6.5.23** a) Demuestre que las siguientes funciones son convexas en  $\Omega = \{(x, y, z) : x > 0, y > 0, z > 0\}$ .

$$f(x, y, z) = -\log(xyz^3); \quad g(x, y, z) = 1/(xyz^3)$$

b) Demuestre que la función  $\log(xyz^3)$  alcanza un máximo relativo sobre el trozo de esfera  $M = \{(x, y, z) \in \Omega : x^2 + y^2 + z^2 = 5r^2\}$ . Justifique que este máximo es absoluto (para ello se recomienda considerar la función  $xyz^3$ ).



# Capítulo 7

## Desarrollo de Taylor

*Funciones diferenciables  $m$  veces y funciones de clase  $C^m$ . Polinomio de Taylor. Serie de Taylor de una función de clase  $C^\infty$*

En este capítulo seguimos considerando funciones de  $n$  variables reales con valores en un espacio vectorial normado. Como esto involucra, en los resultados centrales relativos al desarrollo de Taylor, la consideración de polinomios con coeficientes vectoriales, el lector que no se encuentre cómodo considerando estos polinomios generalizados puede suponer que las funciones toman valores reales.

Si  $\mathbf{f}$  es diferenciable en  $\mathbf{a}$ , la definición de diferencial asegura que  $\mathbf{P}_1(\mathbf{x} - \mathbf{a}) = \mathbf{f}(\mathbf{a}) + \sum_{j=1}^n D_j \mathbf{f}(\mathbf{a})(x_j - a_j)$  es un polinomio de primer grado en las  $n$  variables reales  $x_1, x_2, \dots, x_n$ , con coeficientes en  $F$ , con el que se consigue una aproximación local de primer orden de  $\mathbf{f}$  en el punto  $\mathbf{a}$ , es decir  $\mathbf{f}(\mathbf{x}) = \mathbf{P}_1(\mathbf{x} - \mathbf{a}) + \mathbf{R}_1(\mathbf{x} - \mathbf{a})$  donde  $\mathbf{R}_1(\mathbf{x} - \mathbf{a}) = o(\|\mathbf{x} - \mathbf{a}\|)$ . En el capítulo 6 ya se ha visto que si  $\mathbf{f}$  es diferenciable 2 veces en  $\mathbf{a} \in \Omega$ , entonces hay un polinomio de grado 2, en las variables  $x_1, x_2, \dots, x_n$ , que proporciona una aproximación local de orden 2, de la función  $\mathbf{f}$ , en el punto  $\mathbf{a}$ .

En este capítulo, después de considerar la noción de función  $m$  veces diferenciable y de función de clase  $C^m$ , extendemos el resultado al caso de una función diferenciable  $m$  veces en un punto. En este caso hay un polinomio de grado  $m$ , en las variables  $x_1, x_2, \dots, x_n$ , con coeficientes en  $F$ , llamado polinomio de Taylor de  $\mathbf{f}$  en  $\mathbf{a}$ , que proporciona una aproximación local de orden  $m$ , de la función  $\mathbf{f}$  en el punto  $\mathbf{a}$ . Esto significa que el polinomio de Taylor  $\mathbf{P}_m(\mathbf{x} - \mathbf{a}) = \mathbf{P}_m(x_1 - a_1, x_2 - a_2, \dots, x_n - a_n)$ , (que conviene escribir usando potencias de  $x_j - a_j$ ) cumple la condición

$$\mathbf{f}(\mathbf{x}) = \mathbf{P}_m(\mathbf{x} - \mathbf{a}) + \mathbf{R}_m(\mathbf{x} - \mathbf{a}) \text{ donde } \mathbf{R}_m(\mathbf{x} - \mathbf{a}) = o(\|\mathbf{x} - \mathbf{a}\|^m)$$

Cuando  $\mathbf{f}$  es de clase  $C^{m+1}$  se puede conseguir una fórmula explícita, en forma de integral, para el término complementario  $\mathbf{R}_m(\mathbf{x} - \mathbf{a})$  y acotaciones útiles de este término. Esta fórmula integral requiere la consideración de la integral de una función continua  $\mathbf{f} : [a, b] \rightarrow F$  con valores en un espacio normado completo  $(F, \|\cdot\|)$ , que el lector interesado puede consultar en el apéndice D.

Los resultados de este capítulo se pueden completar con la breve incursión a la teoría de las funciones analíticas de varias variables reales que se ofrece, como

material de carácter complementario, en el apéndice G. Allí se presenta una caracterización útil de estas funciones y se mencionan, a título informativo, algunos resultados de carácter básico.

## 7.1. Funciones diferenciables $m$ veces

Para  $m > 2$  las derivadas parciales de orden  $m$  se definen análogamente a como se han definido las derivadas parciales de segundo orden y se usan notaciones análogas. Para una función de  $n$ -variables reales  $\mathbf{f}(x_1, x_2, \dots, x_n)$ , con valores en un espacio normado  $(F, \|\cdot\|)$  la derivada tercera  $D_{ijk}\mathbf{f}(\mathbf{a})$ , si existe, es la que resulta al derivar respecto a la variable  $x_i$  la función de  $n$  variables  $D_{jk}\mathbf{f}(x_1, x_2, \dots, x_n)$ . También se usa la notación

$$\frac{\partial^3 \mathbf{f}}{\partial x_i \partial x_j \partial x_k}(\mathbf{a}); \quad \frac{\partial^3 \mathbf{f}}{\partial x_i^3}(\mathbf{a}) \text{ si } i = j = k.$$

Si hay dos índices repetidos p.e.  $i = j \neq k$ , cuando se pueda asegurar que el orden de derivación es indiferente (véase el corolario 6.5) el valor común de las derivadas  $D_{iik}\mathbf{f}(\mathbf{a}) = D_{iki}\mathbf{f}(\mathbf{a}) = D_{kii}\mathbf{f}(\mathbf{a})$ , se designará con la notación habitual

$$\frac{\partial^3 \mathbf{f}}{\partial x_i^2 \partial x_k}(\mathbf{a}) = \frac{\partial^3 \mathbf{f}}{\partial x_k \partial x_i^2}(\mathbf{a})$$

En el caso de funciones de tres variables,  $\mathbf{f}(x, y, z)$ , para las derivadas parciales de tercer orden también se usan las notaciones  $\mathbf{f}_{xxx}$ ,  $\mathbf{f}_{xxy}$ ,  $\mathbf{f}_{xxz}$ ,  $\mathbf{f}_{xyx}$ ,  $\mathbf{f}_{xyz}$ , etc. cuyo significado es claro.

Según las observaciones que siguen a la definición 6.1 toda función  $\mathbf{f} : \Omega \rightarrow F$  de clase  $C^2(\Omega, F)$  es diferenciable dos veces en  $\Omega$ , lo que significa que las derivadas parciales  $D_j\mathbf{f}$ ,  $1 \leq j \leq n$ , están definidas y son diferenciables en  $\Omega$ . En este caso todas las derivadas parciales de segundo orden  $D_{ij}\mathbf{f}$  están definidas en  $\Omega$  y en virtud de 6.4 en todo punto  $\mathbf{x} \in \Omega$  se cumple la condición de simetría:

$$D_{ij}\mathbf{f}(\mathbf{x}) = D_{ji}\mathbf{f}(\mathbf{x}) \text{ para todo par } i, j \in \{1, 2, \dots, n\}$$

Si  $\mathbf{f} : \Omega \rightarrow F$  es diferenciable dos veces en un abierto  $V \subset \Omega$  y todas las derivadas parciales de segundo orden  $D_{ij}\mathbf{f} : V \rightarrow F$  son diferenciables en  $\mathbf{a} \in V$ , se dice que  $\mathbf{f}$  es diferenciable tres veces en  $\mathbf{a}$ . En este caso existen las derivadas parciales de tercer orden  $D_{kij}\mathbf{f}(\mathbf{a}) = D_k D_{ij}\mathbf{f}(\mathbf{a})$ ,  $1 \leq k, i, j \leq n$ , y se cumple

$$\text{i) } D_{kij}\mathbf{f}(\mathbf{a}) = D_{kji}\mathbf{f}(\mathbf{a}) \quad \text{ii) } D_{kij}\mathbf{f}(\mathbf{a}) = D_{ikj}\mathbf{f}(\mathbf{a})$$

La igualdad i) es consecuencia de la igualdad  $D_{ij}\mathbf{f}(\mathbf{x}) = D_{ji}\mathbf{f}(\mathbf{x})$  válida en todo  $\mathbf{x} \in V$ . La igualdad ii) se cumple porque  $D_j\mathbf{f}(\mathbf{x})$  es dos veces diferenciable en  $\mathbf{a}$  y por lo tanto  $D_{ki}(D_j\mathbf{f})(\mathbf{a}) = D_{ik}(D_j\mathbf{f})(\mathbf{a})$ . Combinando i) y ii) se concluye que  $D_{j_1 j_2 j_3}\mathbf{f}(\mathbf{a}) = D_{j_{\sigma(1)} j_{\sigma(2)} j_{\sigma(3)}}\mathbf{f}(\mathbf{a})$  para cada permutación  $\sigma$  de  $\{1, 2, 3\}$ .

La definición de función diferenciable  $m$  veces se hace por recurrencia: Se suponen definidas las funciones diferenciables  $m - 1$  veces en un abierto  $V \subset \Omega$ , lo que lleva consigo la existencia en  $V$  de todas las derivadas parciales de orden  $m - 1$ .

**Definición 7.1** Se dice que  $\mathbf{f} : \Omega \rightarrow F$  es diferenciable  $m$  veces en  $\mathbf{a} \in \Omega$  si hay un entorno abierto de  $\mathbf{a}$ ,  $V_{\mathbf{a}} \subset \Omega$ , donde  $\mathbf{f}$  es diferenciable  $m - 1$  veces y todas sus derivadas parciales de orden  $m - 1$  son diferenciables en  $\mathbf{a}$ .

Evidentemente,  $\mathbf{f}$  es diferenciable  $m$  veces en  $\mathbf{a}$  si y sólo si existe un entorno abierto de  $\mathbf{a}$ ,  $V_{\mathbf{a}} \subset \Omega$ , donde  $\mathbf{f}$  es diferenciable y todas las derivadas parciales  $D_1\mathbf{f}$ ,  $D_2\mathbf{f}$ , ...,  $D_n\mathbf{f}$ , que están definidas en  $V_{\mathbf{a}}$ , son diferenciables  $m - 1$  veces en  $\mathbf{a}$ .

**Proposición 7.2** Sea  $\Omega \subset \mathbb{R}$  abierto y  $\mathbf{f} : \Omega \rightarrow F$  una función diferenciable  $m$  veces en  $\mathbf{a} \in \Omega$  ( $m \geq 2$ ). Entonces existen todas las derivadas parciales de orden  $m$ ,  $D_{j_1 j_2 \dots j_m} \mathbf{f}(\mathbf{a})$ ,  $1 \leq j_1, j_2, \dots, j_m \leq n$ , y se verifica la condición de simetría:  $D_{j_1 j_2 \dots j_m} \mathbf{f}(\mathbf{a}) = D_{j_{\sigma(1)} j_{\sigma(2)} \dots j_{\sigma(m)}} \mathbf{f}(\mathbf{a})$ , para toda permutación  $\sigma$  de  $\{1, 2, \dots, m\}$ .

DEM: Ya sabemos que el resultado es cierto para  $m = 2$  y  $m = 3$ . La demostración general se puede hacer por inducción sobre  $m$ . Para ello suponemos que el resultado es cierto para  $m - 1 \geq 1$ . Si  $\mathbf{f} : \Omega \rightarrow F$  es diferenciable  $m$  veces en  $\mathbf{a}$ , sus derivadas de orden  $m - 1$  están definidas en un entorno abierto de  $\mathbf{a}$ ,  $V_{\mathbf{a}} \subset \Omega$ , y en todo  $\mathbf{x} \in V_{\mathbf{a}}$  se cumple la condición de simetría de orden  $m - 1$ : Para toda permutación  $\sigma$  de  $\{2, 3, \dots, m\}$  y todo  $\mathbf{x} \in V_{\mathbf{a}}$  se verifica  $D_{j_2 j_3 \dots j_m} \mathbf{f}(\mathbf{x}) = D_{j_{\sigma(2)} j_{\sigma(3)} \dots j_{\sigma(m)}} \mathbf{f}(\mathbf{x})$ , luego

$$D_{j_1 j_2 j_3 \dots j_m} \mathbf{f}(\mathbf{a}) = D_{j_1 j_{\sigma(2)} j_{\sigma(3)} \dots j_{\sigma(m)}} \mathbf{f}(\mathbf{a})$$

es decir, se cumple la condición de simetría de orden  $m$  cuando  $\sigma$  es una permutación de  $\{1, 2, \dots, m\}$  que cumple  $\sigma(1) = 1$ .

Por otra parte, como  $D_{j_3 \dots j_m} \mathbf{f}(\mathbf{x})$  es diferenciable dos veces en  $\mathbf{a}$ , en virtud del teorema de Young 6.4, se cumple

$$D_{i j j_3 \dots j_m} \mathbf{f}(\mathbf{a}) = D_{j i j_3 \dots j_m} \mathbf{f}(\mathbf{a})$$

luego también se cumple la condición de simetría de orden  $m$  cuando  $\sigma$  es la permutación de  $\{1, 2, \dots, m\}$  definida por  $\sigma(1) = 2$ ,  $\sigma(2) = 1$ ,  $\sigma(k) = k$  si  $3 \leq k \leq m$ .

Para terminar la demostración basta tener en cuenta que cualquier permutación  $\sigma$  de  $\{1, 2, \dots, m\}$  se puede obtener componiendo permutaciones de los dos tipos para los que hemos demostrado la validez de la condición de simetría. ■

Si  $\mathbf{f} : \Omega \rightarrow F$  es diferenciable  $m$  veces en  $\mathbf{a} \in \Omega$ , en virtud de la simetría de las derivadas  $D_{j_1 j_2 \dots j_m} \mathbf{f}(\mathbf{a})$  no es preciso tener en cuenta el orden en que se han realizado las derivaciones; basta conocer el número de veces que se deriva respecto a cada variable. Si en la sucesión finita  $(j_1, j_2, \dots, j_m)$  el subíndice  $i$  aparece  $p_i$  veces es decir si se deriva  $p_i$  veces respecto a la variable  $x_i$ ,  $1 \leq i \leq n$ , se introduce la notación

$$\frac{\partial^m \mathbf{f}(\mathbf{a})}{\partial x_1^{p_1} \partial x_2^{p_2} \partial x_n^{p_n}} := D_{j_1 j_2 \dots j_m} \mathbf{f}(\mathbf{a})$$

Esto motiva la utilización de los índices múltiples: Si  $\mathbf{p} = (p_1, p_2, \dots, p_n)$ , con  $p_i \in \{0, 1, 2, \dots\}$ , y se define  $|\mathbf{p}| := p_1 + p_2 + \dots + p_n$  entonces la derivada parcial de orden  $m$  considerada arriba se escribe con la notación más breve

$$D^{\mathbf{p}} \mathbf{f}(\mathbf{a}) := \frac{\partial^{|\mathbf{p}|} \mathbf{f}(\mathbf{a})}{\partial x_1^{p_1} \partial x_2^{p_2} \partial x_n^{p_n}} \quad \text{con } |\mathbf{p}| = m.$$

**Funciones de clase  $C^m$ .** Análogamente a como se hizo para  $m = 2$  se definen las funciones de clase  $C^m(\Omega, F)$ . Estas funciones son diferenciables  $m$  veces en  $\Omega$ , y las funciones diferenciables  $m$  veces en  $\Omega$  son de clase  $C^{m-1}(\Omega, F)$ .

Cuando  $F = \mathbb{R}^n$  es inmediato que  $\mathbf{f} : \Omega \rightarrow \mathbb{R}^n$ ,  $\mathbf{f} = (f_1, f_2 \cdots f_n)$  es de clase  $C^m$  si y sólo si cada componente  $f_j$  es de clase  $C^m$ .

**Proposición 7.3** *Sea  $\Omega \subset \mathbb{R}^n$  abierto y  $(F, \|\cdot\|)$  un espacio normado.*

$$i) \mathbf{f}, \mathbf{g} \in C^m(\Omega, F) \Rightarrow \mathbf{f} + \mathbf{g} \in C^m(\Omega, F)$$

$$ii) f \in C^m(\Omega, \mathbb{R}), \mathbf{g} \in C^m(\Omega, F) \Rightarrow f\mathbf{g} \in C^m(\Omega, F)$$

DEM: i) es inmediato y ii) se demuestra fácilmente por inducción sobre  $m$  usando que las derivadas parciales del producto  $\varphi = f\mathbf{g}$  vienen dadas por la fórmula

$$D_j\varphi(\mathbf{x}) = D_j f(\mathbf{x})\mathbf{g}(\mathbf{x}) + f(\mathbf{x})D_j\mathbf{g}(\mathbf{x}), \quad 1 \leq j \leq n.$$

Con ella se obtiene que el resultado es cierto para  $m = 1$  y que si se supone cierto para funciones de clase  $C^m$  también lo es para funciones de clase  $C^{m+1}$ . ■

**Proposición 7.4** *Sean  $\mathbf{g} : U \rightarrow \mathbb{R}^n$ ,  $\mathbf{f} : \Omega \rightarrow \mathbb{R}^p$  funciones de clase  $C^m$ , donde  $U \subset \mathbb{R}^k$ ,  $\Omega \subset \mathbb{R}^n$  son abiertos y  $\mathbf{g}(U) \subset \Omega$ . Entonces  $\varphi = \mathbf{f} \circ \mathbf{g} : U \rightarrow \mathbb{R}^p$  es de clase  $C^m$  (brevemente, la composición de dos aplicaciones de clase  $C^m$  es de clase  $C^m$ ).*

DEM: Lo demostraremos por inducción sobre  $m$ . En virtud de la regla de la cadena,  $\varphi = \mathbf{f} \circ \mathbf{g}$  es diferenciable en  $U$  y según la fórmula 5.1 de la sección 5.3, para  $1 \leq j \leq k$ , la derivada parcial  $D_j\varphi_i$  de la componente  $\varphi_i$  viene dada por

$$D_j\varphi_i(\mathbf{x}) = D_1 f_i(\mathbf{g}(\mathbf{x}))D_j g_1(\mathbf{x}) + D_2 f_i(\mathbf{g}(\mathbf{x}))D_j g_2(\mathbf{x}) + \cdots + D_n f_i(\mathbf{g}(\mathbf{x}))D_j g_n(\mathbf{x})$$

En el caso  $m = 1$  las funciones  $D_r f_i(\mathbf{g}(\mathbf{x}))$ ,  $D_j g_r(\mathbf{x})$  son continuas, luego las derivadas parciales  $D_j\varphi_i(\mathbf{x})$  son continuas en  $U$  y por lo tanto  $\varphi$  es de clase  $C^1$ .

Si suponemos que el resultado es cierto para funciones de clase  $m \geq 1$ , dadas  $\mathbf{f}$ ,  $\mathbf{g}$  de clase  $C^{m+1}$ , las derivadas parciales  $D_j g_r$ ,  $D_r f_i$  son de clase  $C^m$  y por hipótesis de inducción  $D_r f_i(\mathbf{g}(\mathbf{x}))$  también es de clase  $C^m$ . Según la proposición 7.3 la suma y el producto de funciones de clase  $C^m$  es de la misma clase, luego todas las derivadas parciales

$$D_j\varphi_i(\mathbf{x}) = \sum_{r=1}^n D_r f_i(\mathbf{g}(\mathbf{x}))D_j g_r(\mathbf{x})$$

son de clase  $C^m$  lo que significa que  $\varphi$  es de clase  $C^{m+1}$ . ■

Cuando  $\mathbf{f}$  es diferenciable  $m$  veces en  $\mathbf{a}$ , según la proposición 7.2, entre las derivadas parciales  $D_{j_1 j_2 \cdots j_m} \mathbf{f}(\mathbf{a})$ , hay muchas repetidas que conviene contar: Los subíndices  $(j_1, j_2, \cdots, j_m)$  recorren las permutaciones con repetición de  $(1, 2, \cdots, n)$  y aparecerán repetidas las derivadas que corresponden a las permutaciones en las

que el 1 aparece  $p_1$  veces, el 2 aparece  $p_2$  veces... y el  $n$  aparece  $p_n$  veces (con  $p_1 + p_2 + \dots + p_n = m$ ). El número de estas permutaciones es

$$\frac{m!}{p_1!p_2!p_3! \cdots p_n!}$$

lo que motiva la introducción de la notación  $\mathbf{p}!$  para designar el producto de los factoriales de las componentes:  $\mathbf{p}! := p_1!p_2! \cdots p_n!$ . Con esta notación el número de derivadas parciales de orden  $m$ , con índice de derivación  $\mathbf{p} = (p_1, p_2, \dots, p_n)$ , que se repiten en virtud del principio de simetría, es exactamente  $m!/\mathbf{p}! = |\mathbf{p}|/\mathbf{p}!$ . Para lo que sigue, si  $\mathbf{h} = (h_1, h_2, \dots, h_n) \in \mathbb{R}^n$  y  $\mathbf{p} = (p_1, p_2, \dots, p_n)$  también resultará cómoda la notación abreviada:  $\mathbf{h}^{\mathbf{p}} = h_1^{p_1} h_2^{p_2} \cdots h_n^{p_n}$ .

**Lema 7.5** *Sea  $\mathbf{f} : \Omega \rightarrow F$  diferenciable  $m > 1$  veces en cada punto del segmento  $[\mathbf{a}, \mathbf{a} + \mathbf{h}] = \{\mathbf{a} + t\mathbf{h} : 0 \leq t \leq 1\} \subset \Omega$ . Entonces la función  $\mathbf{v} : [0, 1] \rightarrow F$ , definida por  $\mathbf{v}(t) = \mathbf{f}(\mathbf{a} + t\mathbf{h})$ , es derivable  $m$  veces en cada  $t \in [0, 1]$  y sus derivadas sucesivas vienen dadas por*

$$\mathbf{v}^{(k)}(t) = \sum_{j_1, j_2, \dots, j_k=1}^n D_{j_1 j_2 \dots j_k} \mathbf{f}(\mathbf{a} + t\mathbf{h}) h_{j_1} h_{j_2} \cdots h_{j_k} = k! \sum_{|\mathbf{p}|=k} \frac{D^{\mathbf{p}} \mathbf{f}(\mathbf{a} + t\mathbf{h})}{\mathbf{p}!} \mathbf{h}^{\mathbf{p}}$$

DEM: En virtud de la regla de la cadena  $\mathbf{v}(t) = \mathbf{f}(\mathbf{a} + t\mathbf{h})$  es derivable en cada  $t \in [0, 1]$  y su derivada es

$$\mathbf{v}'(t) = \sum_{j=1}^n D_j \mathbf{f}(\mathbf{a} + t\mathbf{h}) h_j$$

Como  $m > 1$  esta función vuelve a ser derivable con derivada

$$\mathbf{v}''(t) = \sum_{j=1}^n \left( \sum_{i=1}^n D_i D_j \mathbf{f}(\mathbf{a} + t\mathbf{h}) h_i \right) h_j = \sum_{i,j=1}^n D_{ij} \mathbf{f}(\mathbf{a} + t\mathbf{h}) h_i h_j$$

y de modo recurrente se obtiene que  $\mathbf{v}$  es derivable  $m$  veces en  $[0, 1]$  y que

$$\mathbf{v}^{(k)}(t) = \sum_{j_1, j_2, \dots, j_k=1}^n D_{j_1 j_2 \dots j_k} \mathbf{f}(\mathbf{a} + t\mathbf{h}) h_{j_1} h_{j_2} \cdots h_{j_k}, \quad \text{con } 1 \leq k \leq m.$$

En virtud del principio de simetría en la última suma hay  $k!/\mathbf{p}!$  sumandos repetidos para cada índice de derivación  $\mathbf{p} = (p_1, p_2, \dots, p_n)$ , con  $|\mathbf{p}| = k$ . Agrupando los sumandos que se repiten resulta

$$\mathbf{v}^{(k)}(t) = \sum_{|\mathbf{p}|=k} \frac{k!}{\mathbf{p}!} D^{\mathbf{p}} \mathbf{f}(\mathbf{a} + t\mathbf{h}) \mathbf{h}^{\mathbf{p}}$$

■

### Ejemplo 7.6

Veamos, en un caso concreto, como se agrupan en el sumatorio del lema 7.5, los términos repetidos de cuarto orden: Si  $f(x_1, x_2, x_3)$  es una función de 3 variables, diferenciable cuatro veces en  $\mathbf{a}$ , aparecen  $4!/2 = 12$  términos iguales a  $D_{1123}f(\mathbf{a})h_1h_1h_2h_3$ , que corresponden a las 12 ordenaciones posibles de los símbolos 1, 1, 2, 3:

$$D_{1123}f(\mathbf{a})h_1h_1h_2h_3, \quad D_{1213}f(\mathbf{a})h_1h_2h_1h_3, \quad D_{2113}f(\mathbf{a})h_2h_1h_1h_3, \dots$$

En cada término interviene dos veces la derivada respecto a  $x_1$ , una vez la derivada respecto a  $x_2$  y una vez la derivada respecto a  $x_3$ , luego  $\mathbf{p} = (2, 1, 1)$ ,  $\mathbf{p}! = 2$ , y la suma de los 12 términos iguales vale

$$12 \frac{\partial^4 f(\mathbf{a})}{\partial x_1^2 \partial x_2 \partial x_3} = \frac{4!}{\mathbf{p}!} D^{\mathbf{p}} f(\mathbf{a})$$

## 7.2. Desarrollo de Taylor

El polinomio Taylor de una función vectorial de  $n$  variables reales  $\mathbf{f} : \Omega \rightarrow F$  es un polinomio en  $n$  variables reales con coeficientes en  $F$  y para escribirlo cómodamente conviene utilizar los índices múltiples, es decir  $n$ -plas  $\mathbf{p} = (p_1, p_2, \dots, p_n)$  donde  $p_1, p_2, \dots, p_n$  son números enteros mayores o iguales que 0. Un polinomio en las  $n$  variables reales  $x_1, x_2, \dots, x_n$ , de grado  $\leq m$ , con coeficientes en  $F$ , se escribirá en la forma abreviada

$$\mathbf{P}(\mathbf{x}) = \sum_{|\mathbf{p}| \leq m} \mathbf{a}_{\mathbf{p}} \mathbf{x}^{\mathbf{p}}$$

donde  $\mathbf{a}_{\mathbf{p}} \in F$ ,  $\mathbf{x}^{\mathbf{p}} = x_1^{p_1} x_2^{p_2} \dots x_n^{p_n}$  y  $|\mathbf{p}| = p_1 + p_2 + \dots + p_n$ . En esta suma interviene el índice múltiple  $\mathbf{0} = (0, 0, \dots, 0)$  con  $|\mathbf{0}| = 0$ , para el cual se obtiene el término independiente  $\mathbf{a}_{\mathbf{0}}$ . Agrupando los términos de grado 0, de grado 1, de grado 2, etc. el polinomio se expresa en la forma

$$\mathbf{P}(\mathbf{x}) = \mathbf{A}_0 + \mathbf{A}_1(\mathbf{x}) + \mathbf{A}_2(\mathbf{x}) + \mathbf{A}_3(\mathbf{x}) + \dots + \mathbf{A}_m(\mathbf{x})$$

donde  $\mathbf{A}_0 = \mathbf{a}_0$  y para  $k \geq 1$ , y  $\mathbf{A}_k(\mathbf{x}) = \sum_{|\mathbf{p}|=k} \mathbf{a}_{\mathbf{p}} \mathbf{x}^{\mathbf{p}}$ , es un polinomio homogéneo de grado  $k$ , es decir,  $\mathbf{A}(t\mathbf{x}) = t^k \mathbf{A}(\mathbf{x})$  para todo  $\mathbf{x} \in \mathbb{R}^n$  y todo  $t \in \mathbb{R}$ .

Para motivar la definición del polinomio de Taylor de una función de varias variables conviene comenzar con la siguiente versión del desarrollo de Taylor que sólo es válida para funciones con valores reales. Esta versión es consecuencia directa del desarrollo de Taylor para funciones reales de una variable real con el término complementario en la forma de Lagrange:

**Teorema 7.7** *Sea  $f : \Omega \rightarrow \mathbb{R}$  de clase  $C^{m+1}(\Omega)$  en un abierto  $\Omega \subset \mathbb{R}^n$ . Si el segmento  $[\mathbf{a}, \mathbf{a} + \mathbf{h}] = \{\mathbf{a} + t\mathbf{h} : 0 \leq t \leq 1\}$  está contenido en  $\Omega$ , existe  $\theta \in (0, 1)$  tal que*

$$f(\mathbf{a} + \mathbf{h}) = \sum_{|\mathbf{p}| \leq m} \frac{D^{\mathbf{p}} f(\mathbf{a})}{\mathbf{p}!} \mathbf{h}^{\mathbf{p}} + \sum_{|\mathbf{p}|=m+1} \frac{D^{\mathbf{p}} f(\mathbf{a} + \theta\mathbf{h})}{\mathbf{p}!} \mathbf{h}^{\mathbf{p}}$$

DEM: Según la proposición 7.4 la función real de variable real  $v(t) = f(\mathbf{a} + t\mathbf{h})$ , está definida y es de clase  $C^{m+1}$  en el abierto  $\{t \in \mathbb{R} : \mathbf{a} + t\mathbf{h} \in \Omega\} \supset [0, 1]$ . Basta considerar su desarrollo de Taylor con el resto en la forma de Lagrange:

$$v(1) = v(0) + v'(0) + \cdots + \frac{1}{k!}v^{(k)}(0) + \cdots + \frac{1}{m!}v^{(m)}(0) + \frac{1}{(m+1)!}v^{(m+1)}(\theta)$$

y sustituir los valores  $v^{(k)}$  calculados en el lema 7.5 para obtener el resultado. ■

El polinomio que interviene en 7.7, escrito en términos de  $\mathbf{x} = \mathbf{a} + \mathbf{h}$

$$P_m(\mathbf{x} - \mathbf{a}) = \sum_{|\mathbf{p}| \leq m} \frac{D^{\mathbf{p}}f(\mathbf{a})}{\mathbf{p}!}(\mathbf{x} - \mathbf{a})^{\mathbf{p}}$$

se llama polinomio de Taylor de orden  $m$  de la función  $f$  en el punto  $\mathbf{a}$  y a la diferencia  $R_m(\mathbf{x} - \mathbf{a}) = f(\mathbf{x}) - P_m(\mathbf{x} - \mathbf{a})$  se le suele llamar término complementario. La fórmula que adopta el término complementario en el teorema 7.7, llamada forma de Lagrange, no es posible conseguirla para el caso de funciones con valores en un espacio de dimensión  $\geq 2$ , pues en el caso  $m = 1$  ya vimos en el capítulo 5 que era imposible expresar el incremento  $\mathbf{R}_1(\mathbf{h}) = \mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a})$  en términos de la diferencial primera evaluada en un punto intermedio  $\mathbf{a} + \theta\mathbf{h}$ , con  $\theta \in (0, 1)$ .

La definición del polinomio de Taylor se extiende en forma natural al caso de funciones con valores vectoriales:

**Definición 7.8** Si  $\mathbf{f} : \Omega \rightarrow F$  es diferenciable  $m$  veces en  $\mathbf{a} \in \Omega$ , se llama polinomio de Taylor de  $\mathbf{f}$  en  $\mathbf{a}$ , de orden  $\leq m$ , al polinomio (con coeficientes en  $F$ )

$$\mathbf{P}_m(\mathbf{x} - \mathbf{a}) = \sum_{|\mathbf{p}| \leq m} \frac{D^{\mathbf{p}}\mathbf{f}(\mathbf{a})}{\mathbf{p}!}(\mathbf{x} - \mathbf{a})^{\mathbf{p}}$$

En el sumatorio anterior interviene el índice múltiple  $\mathbf{0} = (0, 0, \dots, 0)$ , que da lugar, con los convenios habituales, al término independiente del polinomio

$$\frac{D^{\mathbf{0}}\mathbf{f}(\mathbf{a})}{\mathbf{0}!}(\mathbf{x} - \mathbf{a})^{\mathbf{0}} = \mathbf{f}(\mathbf{a})$$

Para una función con valores vectoriales que sólo se supone diferenciable  $m$  veces en el punto  $\mathbf{a}$  se puede conseguir un desarrollo de Taylor en ese punto con término complementario en forma infinitesimal. En este caso, si  $\mathbf{f}$  es diferenciable  $m$  veces en  $\mathbf{a} \in \Omega$ , existe un entorno abierto de  $\mathbf{a}$  donde están definidas todas las derivadas parciales  $D_j\mathbf{f}(\mathbf{x})$ ,  $1 \leq j \leq m$ , y son diferenciables  $m - 1$  veces en  $\mathbf{a}$ . Necesitamos el siguiente lema, según el cual el desarrollo de Taylor, de orden  $m - 1$ , de estas derivadas parciales es el que cabe esperar.

**Lema 7.9** En las condiciones de la definición 7.8 el polinomio de Taylor de grado  $\leq m - 1$ , en el punto  $\mathbf{a}$ , de la derivada parcial  $D_j\mathbf{f}(\mathbf{x})$  es  $D_j\mathbf{P}_m(\mathbf{x} - \mathbf{a})$ , es decir

$$D_j\mathbf{P}_m(\mathbf{x} - \mathbf{a}) = \sum_{|\mathbf{q}| \leq m-1} \frac{D^{\mathbf{q}}D_j\mathbf{f}(\mathbf{a})}{\mathbf{q}!}(\mathbf{x} - \mathbf{a})^{\mathbf{q}}$$

DEM: Para simplificar la escritura se supone  $j = 1$ . Al derivar  $\mathbf{P}_m(\mathbf{x} - \mathbf{a})$  respecto a la variable  $x_1$  los únicos términos del polinomio que tienen derivada no nula son los de la forma

$$\frac{1}{\mathbf{p}!} D^{\mathbf{p}} \mathbf{f}(\mathbf{a})(\mathbf{x} - \mathbf{a})^{\mathbf{p}} = \frac{1}{\mathbf{p}!} D^{\mathbf{p}} \mathbf{f}(\mathbf{a})(x_1 - a_1)^{p_1} (x_2 - a_2)^{p_2} \cdots (x_n - a_n)^{p_n}$$

con  $p_1 \geq 1$ . Si  $\mathbf{q} = (p_1 - 1, p_2, \dots, p_n)$  entonces  $D^{\mathbf{p}} \mathbf{f} = D^{\mathbf{q}} D_1 \mathbf{f}$  y la derivada respecto a  $x_1$  de estos términos se expresa en la forma

$$\frac{1}{p_1! p_2! \cdots p_n!} D^{\mathbf{q}} D_1 \mathbf{f}(\mathbf{a}) p_1 (x_1 - a_1)^{p_1 - 1} (x_2 - a_2)^{p_2} \cdots (x_n - a_n)^{p_n} = \frac{D^{\mathbf{q}} D_1 \mathbf{f}(\mathbf{a})}{\mathbf{q}!} (\mathbf{x} - \mathbf{a})^{\mathbf{q}}$$

La suma de todos estos términos da el resultado

$$D_1 \mathbf{P}_m(\mathbf{x} - \mathbf{a}) = \sum_{|\mathbf{q}| \leq m-1} \frac{D^{\mathbf{q}} D_1 \mathbf{f}(\mathbf{a})}{\mathbf{q}!} (\mathbf{x} - \mathbf{a})^{\mathbf{q}}$$

■

El siguiente teorema es una versión para funciones de varias variables del correspondiente resultado referente a funciones de una variable. Su demostración por inducción, esencialmente la misma que se hizo en el caso  $m = 2$  (6.9), se basa en el teorema de incremento finito. Antes de enunciar el teorema hay que introducir la siguiente notación que interviene en el mismo: Si  $E, F$  son espacios normados y  $\mathbf{f}, \mathbf{g} : \Omega \rightarrow F$  son funciones definidas en un abierto  $\Omega \subset E$  se dice que  $\mathbf{g}$  es una aproximación local de orden  $m \geq 1$  de  $\mathbf{f}$  en el punto  $\mathbf{a} \in \Omega$  si

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{\mathbf{f}(\mathbf{x}) - \mathbf{g}(\mathbf{x})}{\|\mathbf{x} - \mathbf{a}\|^m} = 0$$

En este caso se escribe  $\mathbf{f}(\mathbf{x}) - \mathbf{g}(\mathbf{x}) = o(\|\mathbf{x} - \mathbf{a}\|^m)$  y se dice que  $\mathbf{f}$  y  $\mathbf{g}$  tienen en el punto  $\mathbf{a}$  un contacto o tangencia de orden  $m$ .

**Teorema 7.10** *Si  $\mathbf{f} : \Omega \rightarrow F$  es diferenciable  $m$  veces en  $\mathbf{a} \in \Omega$ , y  $\mathbf{P}_m(\mathbf{x} - \mathbf{a})$  es su polinomio de Taylor en  $\mathbf{a}$ , de grado  $\leq m$ , se verifica*

$$\mathbf{f}(\mathbf{x}) = \mathbf{P}_m(\mathbf{x} - \mathbf{a}) + \mathbf{R}_m(\mathbf{x} - \mathbf{a}) \quad \text{donde} \quad \mathbf{R}_m(\mathbf{x} - \mathbf{a}) = o(\|\mathbf{x} - \mathbf{a}\|^m)$$

DEM: En lo que sigue  $\mathbf{h} = \mathbf{x} - \mathbf{a}$ . El resultado es cierto para  $m = 1$  ya que

$$\mathbf{P}_1(\mathbf{h}) = \mathbf{f}(\mathbf{a}) + \sum_{j=1}^n D_j \mathbf{f}(\mathbf{a}) \mathbf{h}_j = \mathbf{f}(\mathbf{a}) + d\mathbf{f}(\mathbf{a})\mathbf{h}$$

y sabemos que, según la definición de diferencial,  $\mathbf{R}_1(\mathbf{h}) = \mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{P}_1(\mathbf{h})$  cumple la condición requerida,  $\mathbf{R}_1(\mathbf{h}) = o(\|\mathbf{h}\|)$ .

Supongamos el resultado cierto para funciones diferenciables  $m - 1$  veces en  $\mathbf{a}$ , con  $m \geq 2$ . Si  $\mathbf{f} : \Omega \rightarrow F$  es diferenciable  $m$  veces en  $\mathbf{a} \in \Omega$ , existe  $B(\mathbf{a}, r) \subset \Omega$  tal que en todos los puntos de  $B(\mathbf{a}, r)$  la función  $\mathbf{f}$  es diferenciable  $m - 1$  veces.



En lo que sigue en  $\mathbb{R}^n$  se considera la norma  $\|\cdot\|_1$ . Para cada  $1 \leq j \leq n$  la función  $D_j \mathbf{f}(\mathbf{x})$  es diferenciable  $m-1$  veces en  $\mathbf{a}$  y en virtud del lema 7.9 su polinomio de Taylor de grado  $\leq m-1$  es  $D_j \mathbf{P}_m(\mathbf{x} - \mathbf{a})$  por lo que, en virtud de la hipótesis de inducción, se cumple

$$D_j \mathbf{f}(\mathbf{a} + \mathbf{h}) = D_j \mathbf{P}_m(\mathbf{h}) + o(\|\mathbf{h}\|_1^{m-1})$$

Si cambiamos de notación y escribimos  $\mathbf{g}(\mathbf{h}) = \mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{P}_m(\mathbf{h})$ , tenemos que para  $1 \leq j \leq n$  se cumple  $D_j \mathbf{g}(\mathbf{h}) = o(\|\mathbf{h}\|_1^{m-1})$  luego, dado  $\epsilon > 0$  existe  $0 < \delta < r$  tal que para todo  $j \in \{1, 2, \dots, n\}$  podemos asegurar que

$$\mathbf{y} \in B(0, \delta) \Rightarrow \|D_j \mathbf{g}(\mathbf{y})\| \leq \epsilon \|\mathbf{y}\|_1^{m-1}$$

Fijado  $\mathbf{h} \in \mathbb{R}^n$  con  $\|\mathbf{h}\|_1 < \delta$ , la función auxiliar  $\varphi(t) = \mathbf{g}(t\mathbf{h}) = \mathbf{f}(\mathbf{a} + t\mathbf{h}) - \mathbf{P}_m(t\mathbf{h})$ , está definida y es derivable en  $[0, 1]$  (ya que  $[\mathbf{a}, \mathbf{a} + \mathbf{h}] \subset B(\mathbf{a}, r)$ ). Su derivada viene dada por  $\varphi'(t) = \sum_{j=1}^n D_j \mathbf{g}(t\mathbf{h}) h_j$ , luego para  $0 \leq t \leq 1$  se verifica

$$\|\varphi'(t)\| \leq \sum_{j=1}^n \|D_j \mathbf{g}(t\mathbf{h})\| |h_j| \leq \sum_{j=1}^n \epsilon \|\mathbf{h}\|_1^{m-1} |h_j| = \epsilon \|\mathbf{h}\|_1^m$$

En virtud del teorema del incremento finito,  $\|\varphi(1) - \varphi(0)\| \leq \epsilon \|\mathbf{h}\|_1^m$ , es decir

$$\|\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{P}_m(\mathbf{h})\| \leq \epsilon \|\mathbf{h}\|_1^m \quad \text{si} \quad \|\mathbf{h}\| < \delta$$

y queda demostrado así que  $\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{P}_m(\mathbf{h}) = o(\|\mathbf{h}\|_1^m)$ . ■

Nuestro siguiente objetivo es obtener el llamado recíproco del desarrollo de Taylor. Este resultado, recogido en la proposición 7.12, suele resultar útil a la hora de obtener desarrollos de Taylor de funciones concretas. La demostración que ofrecemos se basa en el siguiente lema que extiende al contexto  $n$ -dimensional un resultado bien conocido para los polinomios reales de una variable real

**Lema 7.11** *Si  $\mathbf{Q}(\mathbf{x}) = \sum_{|\mathbf{p}| \leq m} \mathbf{a}_{\mathbf{p}} \mathbf{x}^{\mathbf{p}}$  es un polinomio de grado  $\leq m$  con coeficientes  $\mathbf{a}_{\mathbf{p}} \in F$ , son equivalentes*

- a)  $\mathbf{Q}(\mathbf{x}) = o(\|\mathbf{x}\|^m)$ .
- b)  $\mathbf{Q}(\mathbf{x}) = 0$  para cada  $\mathbf{x} \in \mathbb{R}^n$ ;
- c)  $\mathbf{a}_{\mathbf{p}} = 0$  para cada  $|\mathbf{p}| \leq m$ .

DEM: a)  $\Rightarrow$  b): Es fácil ver esta implicación se cumple para polinomios de una variable real: Si  $\mathbf{Q}(t) = \mathbf{a}_0 + \mathbf{a}_1 t + \mathbf{a}_2 t^2 + \dots + \mathbf{a}_m t^m = o(|t|^m)$  debe ser  $\mathbf{a}_0 = 0$ , luego  $\mathbf{a}_1 + \mathbf{a}_2 t + \dots + \mathbf{a}_m t^{m-1} = o(|t|^{m-1})$  de donde se sigue que  $\mathbf{a}_1 = 0$ , etc ... y así se obtiene que  $\mathbf{a}_k = 0$  para  $0 \leq k \leq m$  y por lo tanto  $\mathbf{Q}(t) = 0$  para todo  $t \in \mathbb{R}$ .

La demostración para polinomios de  $n$  variables se reduce al caso de una variable: Fijado  $\mathbf{x} \neq 0$  se considera el polinomio de la variable real  $t$

$$\mathbf{Q}_{\mathbf{x}}(t) = \mathbf{Q}(t\mathbf{x}) = a_0 + \sum_{k=1}^m \left( \sum_{|\mathbf{p}|=k} a_{\mathbf{p}} \mathbf{x}^{\mathbf{p}} \right) t^k$$

Si se cumple a) es fácil ver que  $\mathbf{Q}_x(t) = o(|t|^m)$ , luego el polinomio de una variable  $\mathbf{Q}_x(t)$  es idénticamente nulo y en particular  $\mathbf{Q}(\mathbf{x}) = \mathbf{Q}_x(1) = 0$ .

Para demostrar que b)  $\Rightarrow$  c) se puede razonar por inducción sobre el número  $n$  de variables. El resultado es inmediato cuando  $n = 1$ . Supongamos que es cierto para polinomios de  $n - 1$  variables y sea  $\mathbf{Q}(\mathbf{x})$  un polinomio idénticamente nulo de  $n$  variables. Si para cada  $1 \leq k \leq n$  se agrupan los términos donde figura  $x_1^k$ , y se saca este factor común podemos escribir  $\mathbf{Q}(\mathbf{x})$  en la forma

$$\mathbf{Q}(\mathbf{x}) = \mathbf{c}_0 x_1^m + \mathbf{c}_1(x_2, x_3, \dots, x_n) x_1^{m-1} + \dots + \mathbf{c}_m(x_2, x_3, \dots, x_n) = 0$$

donde cada  $\mathbf{c}_k(x_2, x_3, \dots, x_n)$  es un polinomio de grado  $k$  en  $m - 1$  variables.

Manteniendo fijos  $x_2, x_3, \dots, x_n$ , y considerando sólo la variable  $x_1$ , obtenemos un polinomio idénticamente nulo en esta variable. Como el resultado que queremos demostrar es cierto para polinomios de una variable concluimos que son nulos todos sus coeficientes, es decir,  $\mathbf{c}_k(x_2, x_3, \dots, x_n) = 0$ ,  $0 \leq k \leq m$ . Esta afirmación es cierta para cualquier  $(x_2, x_3, \dots, x_n) \in \mathbb{R}^{n-1}$ , aplicando la hipótesis de inducción obtenemos que los coeficientes de los polinomios  $\mathbf{c}_k$  son nulos, lo que significa que todos los coeficientes de  $\mathbf{Q}$  son nulos.

La implicación c)  $\Rightarrow$  a) es evidente. ■

**Proposición 7.12** *Sea  $\mathbf{f} : \Omega \rightarrow F$  diferenciables  $m$  veces en  $\mathbf{a} \in \Omega$  y  $\mathbf{Q}(\mathbf{x} - \mathbf{a})$  un polinomio de grado  $\leq m$  que verifica  $\mathbf{f}(\mathbf{x}) = \mathbf{Q}(\mathbf{x} - \mathbf{a}) + o(\|\mathbf{x} - \mathbf{a}\|^m)$ . Entonces  $\mathbf{Q}(\mathbf{x} - \mathbf{a})$  es el polinomio de Taylor de orden  $m$ , de la función  $\mathbf{f}$  en el punto  $\mathbf{a}$ .*

DEM: Si  $\mathbf{P}_m(\mathbf{x} - \mathbf{a})$  es el polinomio de Taylor de  $\mathbf{f}$  en  $\mathbf{a}$  (de orden  $m$ ) en virtud de la hipótesis y del teorema 7.10 podemos asegurar que la diferencia

$$\mathbf{S}(\mathbf{x} - \mathbf{a}) = \mathbf{P}_m(\mathbf{x} - \mathbf{a}) - \mathbf{Q}(\mathbf{x} - \mathbf{a})$$

es un polinomio que cumple  $\mathbf{S}(\mathbf{x} - \mathbf{a}) = o(\|\mathbf{x} - \mathbf{a}\|^m)$  y aplicando el lema 7.11 se concluye que  $\mathbf{Q} = \mathbf{P}_m$ . ■

**Proposición 7.13** *Si  $\mathbf{f}, \mathbf{g} : \Omega \rightarrow F$  son diferenciables  $m$  veces en  $\mathbf{a} \in \Omega \subset \mathbb{R}^n$ , son equivalentes*

$$a) \mathbf{f}(\mathbf{x}) - \mathbf{g}(\mathbf{x}) = o(\|\mathbf{x} - \mathbf{a}\|^m);$$

$$b) \mathbf{f}(\mathbf{a}) = \mathbf{g}(\mathbf{a}) \text{ y } D^p \mathbf{f}(\mathbf{a}) = D^p \mathbf{g}(\mathbf{a}) \text{ cuando } |\mathbf{p}| \leq m.$$

DEM: Sean  $\mathbf{P}_m(\mathbf{x} - \mathbf{a})$  y  $\mathbf{Q}_m(\mathbf{x} - \mathbf{a})$ , respectivamente, los polinomios de Taylor (de orden  $m$ ) de las funciones  $\mathbf{f}$  y  $\mathbf{g}$ , en el punto  $\mathbf{a}$ . Según el teorema 7.10 se verifica

$$\mathbf{f}(\mathbf{x}) - \mathbf{P}_m(\mathbf{x} - \mathbf{a}) = o(\|\mathbf{x} - \mathbf{a}\|^m), \quad \mathbf{g}(\mathbf{x}) - \mathbf{Q}_m(\mathbf{x} - \mathbf{a}) = o(\|\mathbf{x} - \mathbf{a}\|^m) \quad (*)$$

a)  $\Rightarrow$  b): Si se cumple a), con los desarrollos anteriores se obtiene

$$\mathbf{P}_m(\mathbf{x} - \mathbf{a}) - \mathbf{Q}_m(\mathbf{x} - \mathbf{a}) = o(\|\mathbf{x} - \mathbf{a}\|^m)$$

y con el lema 7.11 se concluye que los coeficientes de  $\mathbf{P}_m(\mathbf{x} - \mathbf{a}) - \mathbf{Q}_m(\mathbf{x} - \mathbf{a})$  son idénticamente nulos, luego  $D^{\mathbf{p}}\mathbf{f}(\mathbf{a}) - D^{\mathbf{p}}\mathbf{g}(\mathbf{a}) = \mathbf{0}$  siempre que  $|\mathbf{p}| \leq m$ .

b)  $\Rightarrow$  a): Si se cumple b) entonces  $\mathbf{P}_m = \mathbf{Q}_m$  y usando los desarrollos de Taylor (\*) se obtiene que  $\mathbf{f}(\mathbf{x}) - \mathbf{g}(\mathbf{x}) = o(\|\mathbf{x} - \mathbf{a}\|^m)$ . ■

### 7.3. Serie de Taylor de una función de clase $C^\infty$

Dada una función  $\mathbf{f}$  de clase  $C^\infty$  en un abierto  $\Omega \subset \mathbb{R}^n$  y un punto  $\mathbf{a} \in \Omega$  podemos considerar en este punto polinomios de Taylor de  $\mathbf{f}$  arbitrariamente largos, que generan una serie de potencias en varias variables reales, llamada *serie de Taylor* de la función en el punto  $\mathbf{a}$ . En el teorema 7.15 se obtendrá, para una función de clase  $C^\infty$ , una condición suficiente para que su serie de Taylor en un punto sea convergente y represente a la función en un entorno de ese punto.

Comenzamos obteniendo una acotación útil del error que se comete cuando se utiliza el polinomio de Taylor para aproximar a la función en un entorno del punto donde se efectúa el desarrollo.

**Teorema 7.14** *Sea  $\mathbf{f} : \Omega \rightarrow F$  una función de clase  $C^{m+1}(\Omega, F)$  definida en un abierto  $\Omega \subset \mathbb{R}^n$ , con valores en un espacio normado  $(F, \|\cdot\|)$ . Si el segmento  $[\mathbf{a}, \mathbf{a} + \mathbf{h}]$  está contenido en  $\Omega$  y  $\mathbf{P}_m(\mathbf{x} - \mathbf{a})$  es el polinomio de Taylor de  $\mathbf{f}$  en  $\mathbf{a}$ , de orden  $m$ , entonces para el término complementario  $\mathbf{R}_m(\mathbf{x} - \mathbf{a}) = \mathbf{f}(\mathbf{x}) - \mathbf{P}_m(\mathbf{x} - \mathbf{a})$  vale la siguiente acotación*

$$\|\mathbf{R}_m(\mathbf{h})\| \leq \frac{M}{(m+1)!} \|\mathbf{h}\|_1^{m+1}$$

donde

$$M = \sup\{\|D_{j_1 j_2 \dots j_{m+1}} \mathbf{f}(\mathbf{z})\| : 1 \leq j_1, j_2, \dots, j_{m+1} \leq n, \mathbf{z} \in [\mathbf{a}, \mathbf{a} + \mathbf{h}]\}$$

DEM: En  $[0, 1]$  están definidas las funciones  $\mathbf{v}(t) = \mathbf{f}(\mathbf{a} + t\mathbf{h})$ , y

$$\mathbf{g}(t) = \mathbf{v}(t) + (1-t)\mathbf{v}'(t) + \frac{1}{2!}(1-t)^2\mathbf{v}''(t) + \dots + \frac{1}{m!}(1-t)^m\mathbf{v}^{(m)}(t)$$

cuya derivada

$$\mathbf{g}'(t) = \frac{(1-t)^m}{m!} \mathbf{v}^{(m+1)}(t) = \sum_{j_1 j_2 \dots j_{m+1}=1}^n D_{j_1 j_2 \dots j_{m+1}} \mathbf{f}(\mathbf{a} + t\mathbf{h}) h_{j_1} h_{j_2} \dots h_{j_{m+1}}$$

cumple la desigualdad

$$\|\mathbf{g}'(t)\| \leq \frac{M \|\mathbf{h}\|_1^{m+1}}{m!} (1-t)^m$$

es decir

$$\|\mathbf{g}'(t)\| \leq \alpha'(t) \quad \text{con} \quad \alpha(t) = -\frac{M \|\mathbf{h}\|_1^{m+1}}{(m+1)!} (1-t)^{m+1}$$

Utilizando el teorema del incremento finito 4.7 se concluye que

$$\|\mathbf{R}_m(\mathbf{h})\| = \|\mathbf{g}(1) - \mathbf{g}(0)\| \leq \alpha(1) - \alpha(0) = \frac{M}{(m+1)!} \|\mathbf{h}\|_1^{m+1}$$

■

**Teorema 7.15** *Sea  $\mathbf{f} : \Omega \rightarrow \mathbf{F}$  de clase  $C^\infty$  en un abierto  $\Omega \subset \mathbb{R}^n$ , con valores en un espacio normado  $(F, \|\cdot\|)$  y  $B(\mathbf{a}, \delta) \subset \Omega$  una bola tal que existen constantes  $M > 0$  y  $R > 0$  que verifican*

$$\mathbf{x} \in B(\mathbf{a}, \delta), \quad |\mathbf{p}| = k \Rightarrow |D^{\mathbf{p}}\mathbf{f}(\mathbf{x})| \leq Mk!R^k$$

Entonces, si  $\mathbf{x} = \mathbf{a} + \mathbf{h} \in B(\mathbf{a}, \delta)$  y  $\|\mathbf{h}\|_1 < 1/R$  se verifica

$$\mathbf{f}(\mathbf{x}) = \sum_{k=1}^{\infty} \sum_{|\mathbf{p}|=k} \frac{D^{\mathbf{p}}\mathbf{f}(\mathbf{a})}{\mathbf{p}!} \mathbf{h}^{\mathbf{p}}, \quad y \quad \sum_{k=1}^{\infty} \sum_{|\mathbf{p}|=k} \left\| \frac{D^{\mathbf{p}}\mathbf{f}(\mathbf{a})}{\mathbf{p}!} \mathbf{h}^{\mathbf{p}} \right\| < +\infty$$

DEM: Observemos en primer lugar que

$$\|\mathbf{h}\|_1^k = (|h_1| + |h_2| + \cdots + |h_n|)^k = \sum_{j_1 j_2 \cdots j_k=1}^n |h_{j_1}| |h_{j_2}| \cdots |h_{j_k}| = \sum_{|\mathbf{p}|=k} \frac{k!}{\mathbf{p}!} |\mathbf{h}|^{\mathbf{p}}$$

luego

$$\sum_{|\mathbf{p}|=k} \left\| \frac{D^{\mathbf{p}}\mathbf{f}(\mathbf{a})}{\mathbf{p}!} \mathbf{h}^{\mathbf{p}} \right\| \leq MR^k \sum_{|\mathbf{p}|=k} \frac{k!}{\mathbf{p}!} |\mathbf{h}|^{\mathbf{p}} = M(R \|\mathbf{h}\|_1)^k$$

Como  $R \|\mathbf{h}\|_1 < 1$ , se obtiene la convergencia absoluta de la serie

$$\sum_{k=0}^{\infty} \sum_{|\mathbf{p}|=k} \frac{D^{\mathbf{p}}\mathbf{f}(\mathbf{a})}{\mathbf{p}!} \mathbf{h}^{\mathbf{p}}$$

Para terminar debemos demostrar su suma es  $\mathbf{f}(\mathbf{x})$ , y para ello basta ver que el término complementario del desarrollo de Taylor

$$\mathbf{R}_m(\mathbf{h}) = \mathbf{f}(\mathbf{x}) - \sum_{k=0}^m \sum_{|\mathbf{p}|=k} \frac{D^{\mathbf{p}}\mathbf{f}(\mathbf{a})}{\mathbf{p}!} \mathbf{h}^{\mathbf{p}}$$

converge hacia  $\mathbf{0}$  cuando  $m \rightarrow \infty$ .

Por hipótesis, para todo  $\mathbf{z} \in [\mathbf{a}, \mathbf{a} + \mathbf{h}] \subset B(\mathbf{a}, \delta)$  todas las derivadas  $D^{\mathbf{p}}(\mathbf{z})$  con  $|\mathbf{p}| = m+1$  cumplen la desigualdad  $\|D^{\mathbf{p}}\mathbf{f}(\mathbf{z})\| \leq M(m+1)! R^{m+1}$  y teniendo en cuenta el teorema 7.14 se obtiene la desigualdad

$$\|\mathbf{R}_m(\mathbf{h})\| \leq \frac{M(m+1)! R^{m+1}}{(m+1)!} \|\mathbf{h}\|_1^{m+1} = M(R \|\mathbf{h}\|_1)^{m+1}$$

y con ella el resultado deseado, porque  $R \|\mathbf{h}\|_1 < 1$ . ■

## 7.4. Fórmula integral para el resto

En el caso de funciones con valores vectoriales se puede conseguir una fórmula integral para el resto o término complementario del desarrollo de Taylor. Como en ella interviene la integral de una función continua con valores en  $F$ , para este resultado hay que suponer que el espacio normado  $(F, \| \cdot \|)$  es completo (véase el apéndice D). Para mayor simplicidad, el lector puede considerar sólo funciones con valores en  $F = \mathbb{R}^k$ , ya que en este caso la integral de una función continua se puede definir componente a componente y los resultados de integración vectorial requeridos son consecuencia directa de los referentes a funciones escalares (véase 4.2).

**Teorema 7.16** *Sea  $\mathbf{f} : \Omega \rightarrow F$  una función de clase  $C^{m+1}(\Omega, F)$  definida en un abierto  $\Omega \subset \mathbb{R}^n$ , con valores en un espacio normado completo  $(F, \| \cdot \|)$ . Si el segmento  $[\mathbf{a}, \mathbf{a} + \mathbf{h}]$  está contenido en  $\Omega$  y  $\mathbf{P}_m(\mathbf{x} - \mathbf{a})$  es el polinomio de Taylor de  $\mathbf{f}$  en  $\mathbf{a}$ , de orden  $m$ , entonces para el término complementario  $\mathbf{R}_m(\mathbf{x} - \mathbf{a}) = \mathbf{f}(\mathbf{x}) - \mathbf{P}_m(\mathbf{x} - \mathbf{a})$  vale la siguiente fórmula integral:*

$$\mathbf{R}_m(\mathbf{h}) = (m+1) \int_0^1 (1-t)^m \mathbf{r}_{m+1}(\mathbf{h}, t) dt$$

donde

$$\mathbf{r}_{m+1}(\mathbf{h}, t) = \sum_{|\mathbf{p}|=m+1} \frac{D^{\mathbf{p}}\mathbf{f}(\mathbf{a} + t\mathbf{h})}{\mathbf{p}!} \mathbf{h}^{\mathbf{p}} = \frac{1}{(m+1)!} \sum_{j_1, \dots, j_{m+1}=1}^n D_{j_1, \dots, j_{m+1}} \mathbf{f}(\mathbf{a} + t\mathbf{h}) h_{j_1} \cdots h_{j_{m+1}}$$

DEM: La función  $\mathbf{v}(t) = \mathbf{f}(\mathbf{a} + t\mathbf{h})$  está definida y es de clase  $C^{m+1}$  en el abierto  $\{t \in \mathbb{R} : \mathbf{a} + t\mathbf{h} \in \Omega\} \supset [0, 1]$ , y según el lema 7.5,

$$\begin{aligned} \mathbf{v}^{(m+1)}(t) &= \sum_{j_1, \dots, j_{m+1}=1}^n D_{j_1 j_2 \dots j_{m+1}} \mathbf{f}(\mathbf{a} + t\mathbf{h}) h_{j_1} h_{j_2} \cdots h_{j_{m+1}} = \\ &= (m+1)! \sum_{|\mathbf{p}|=m+1} \frac{D^{\mathbf{p}}\mathbf{f}(\mathbf{a} + t\mathbf{h})}{\mathbf{p}!} \mathbf{h}^{\mathbf{p}} = (m+1)! \mathbf{r}_{m+1}(\mathbf{h}, t) \end{aligned}$$

Se comprueba fácilmente que

$$\mathbf{R}_m(\mathbf{h}) = \mathbf{v}(1) - \left[ \mathbf{v}(0) + \mathbf{v}'(0) + \cdots + \frac{1}{m!} \mathbf{v}^{(m)}(0) \right]$$

y aplicando el teorema 4.16 a la función  $\mathbf{v}(t)$  en el intervalo  $[0, 1]$ , se obtiene

$$\mathbf{R}_m(\mathbf{h}) = \frac{1}{m!} \int_0^1 (1-t)^m \mathbf{v}^{(m+1)}(t) dt = (m+1) \int_0^1 (1-t)^m \mathbf{r}_{m+1}(\mathbf{h}, t) dt$$

■

NOTA: Usando el teorema 7.16 se puede dar otra demostración del teorema 7.14 bajo la hipótesis de que el espacio normado  $(F, \|\cdot\|)$  es completo:

Aplicando la desigualdad triangular a la suma

$$\mathbf{r}_m(\mathbf{h}, t) = \frac{1}{(m+1)!} \sum_{j_1, \dots, j_{m+1}=1}^n D_{j_1, \dots, j_{m+1}} \mathbf{f}(\mathbf{a} + t\mathbf{h}) h_{j_1} \cdots h_{j_{m+1}}$$

se obtiene

$$\|\mathbf{r}_m(\mathbf{h}, t)\| \leq \frac{M}{(m+1)!} \sum_{j_1, \dots, j_{m+1}=1}^n |h_{j_1}| |h_{j_2}| \cdots |h_{j_{m+1}}|$$

Teniendo en cuenta que

$$\sum_{j_1, j_2, \dots, j_{m+1}=1}^n |h_{j_1}| |h_{j_2}| \cdots |h_{j_{m+1}}| = (|h_1| + |h_2| + \cdots + |h_n|)^{m+1} = \|\mathbf{h}\|_1^{m+1}$$

resulta

$$\|\mathbf{r}_m(\mathbf{h}, t)\| \leq \frac{M}{(m+1)!} \|\mathbf{h}\|_1^{m+1}$$

y utilizando la desigualdad D.10 a) se deduce

$$\begin{aligned} \|\mathbf{R}_m(\mathbf{h})\| &\leq (m+1) \int_0^1 (1-t)^m \|\mathbf{r}(\mathbf{h}, t)\| dt \leq \\ &\leq \frac{M}{(m+1)!} \|\mathbf{h}\|_1^{m+1} \int_0^1 (m+1)(1-t)^m dt = \frac{M}{(m+1)!} \|\mathbf{h}\|_1^{m+1} \end{aligned}$$

■

## 7.5. Ejercicios resueltos

**Ejercicio 7.17** Demuestre que la función

$$f(x, y) = \frac{\operatorname{sen} y - \operatorname{sen} x}{y - x} \quad \text{si } x \neq y; \quad f(x, x) = \cos x$$

es de clase  $C^\infty(\mathbb{R}^2)$ . Obtenga su desarrollo de Taylor en  $(0, 0)$  y compruebe que converge en todo punto  $(x, y)$  hacia el valor de la función  $f(x, y)$ . Utilícelo para demostrar que  $f$  presenta en  $(0, 0)$  un máximo relativo.

SOLUCIÓN

Con la sustitución  $t = y - x$  se obtiene que para  $t \neq 0$  el valor de la función se expresa en la forma

$$f(x, y) = \frac{\operatorname{sen}(x+t) - \operatorname{sen} x}{t} = \left( \frac{\cos t - 1}{t} \right) \operatorname{sen} x + \left( \frac{\operatorname{sen} t}{t} \right) \cos x$$

luego

$$f(x, y) = \alpha(y - x) \operatorname{sen} x + \beta(y - x) \operatorname{cos} x \quad (7.1)$$

donde las funciones  $\alpha(t) = (\cos t - 1)/t$ ,  $\beta(t) = (\operatorname{sen} t)/t$ , se pueden suponer definidas en todo  $\mathbb{R}$  mediante los desarrollos en serie de potencias

$$\alpha(t) = -\frac{1}{2!}t + \frac{1}{4!}t^3 - \dots; \quad \beta(t) = 1 - \frac{1}{3!}t^2 + \frac{1}{5!}t^4 - \dots$$

con lo cual la fórmula (7.1) sirve incluso para el caso  $x = y$  haciendo intervenir los valores  $\alpha(0) = 0$  y  $\beta(0) = 1$ .

Como las funciones  $\alpha, \beta$  son de clase  $C^\infty(\mathbb{R})$  con la fórmula 7.1 se pone de manifiesto que  $f$  es de clase  $C^\infty(\mathbb{R}^2)$ . Restando los desarrollos en serie de potencias

$$\operatorname{sen} x = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \dots; \quad \operatorname{sen} y = y - \frac{1}{3!}y^3 + \frac{1}{5!}y^5 - \dots$$

que son válidos para todo  $x \in \mathbb{R}$  y todo  $y \in \mathbb{R}$ , y utilizando la igualdad

$$y^n - x^n = (y - x)(y^{n-1} + xy^{n-2} + x^2y^{n-3} + \dots + x^{n-2}y + x^{n-1})$$

se consigue la siguiente representación de la función  $f$  mediante un desarrollo en serie convergente en todo punto  $(x, y) \in \mathbb{R}^2$ :

$$\begin{aligned} f(x, y) &= 1 - \frac{1}{3!}(x^2 + xy + y^2) + \frac{1}{5!}(x^4 + x^3y + x^2y^2 + xy^3 + y^4) - \dots = \\ &= 1 + A_2(x, y) + A_4(x, y) + \dots + A_{2m}(x, y) + R_{2m}(x, y) \end{aligned}$$

donde  $A_{2k}(x, y)$  es un polinomio homogéneo de grado  $2k$ . Si  $r = \|(x, y)\|_2$ , se verifica

$$|A_{2k}(x, y)| \leq \frac{1}{(2k+1)!}(2k+1)r^{2k} = \frac{r^{2k}}{(2k)!}$$

luego

$$|R_{2m}(x, y)| = |A_{2m+2}(x, y) + A_{2m+4}(x, y) + \dots| \leq \sum_{k=1}^{\infty} \frac{r^{2(m+k)}}{(2m+2k)!} = \varphi(r)$$

Como

$$\varphi(r) = r^{2m+1} \sum_{k=1}^{\infty} \frac{r^{2k-1}}{(2m+2k)!} = o(r^{2m+1})$$

se sigue que  $|R_{2m}(x, y)| = o(\|x, y\|^{2m+1})$  y la proposición 7.12 permite asegurar que  $1 + A_2(x, y) + A_4(x, y) + \dots + A_{2m}(x, y)$  es el polinomio de Taylor de orden  $2m$  (incluso de orden  $2m+1$ ) de la función  $f$  en  $(0, 0)$ . En particular, el polinomio de Taylor de grado 2 es

$$1 - \frac{1}{2}(x^2 + xy + y^2)$$

Por consiguiente:

$$D_1f(0, 0) = D_2f(0, 0) = 0$$

$$D_{11}f(0, 0) = D_{22}f(0, 0) = -1/2; \quad D_{12}f(0, 0) = D_{21}f(0, 0) = -1/4$$

Con el criterio usual del Hessiano se obtiene que  $f$  presenta un máximo relativo en el punto  $(0, 0)$ . ■

**Ejercicio 7.18** Si  $\mathbf{Q}(\mathbf{x})$  es un polinomio de grado  $m$  con coeficientes reales (o más generalmente en un espacio normado  $(F, \|\cdot\|)$ ) demuestre las afirmaciones:

a)  $|\mathbf{p}| > m \Rightarrow D^{\mathbf{p}}\mathbf{Q}(\mathbf{x}) = 0$  para todo  $\mathbf{x} \in \mathbb{R}^n$ .

b) Para cada  $\mathbf{a} \in \mathbb{R}^n$ ,  $\mathbf{Q}$  coincide con su polinomio de Taylor en  $\mathbf{a}$ , de orden  $m$ :

$$\mathbf{Q}(\mathbf{x}) = \sum_{|\mathbf{p}| \leq m} \frac{D^{\mathbf{p}}\mathbf{Q}(\mathbf{a})}{\mathbf{p}!} (\mathbf{x} - \mathbf{a})^{\mathbf{p}} \quad \text{para todo } \mathbf{x} \in \mathbb{R}^n$$

SOLUCIÓN

a) Se demuestra fácilmente por inducción sobre  $m$ : El resultado es evidente para  $m = 1$ . Si se supone cierto para el valor  $m - 1$  y  $\mathbf{Q}(\mathbf{x})$  es un polinomio de grado  $m$ , entonces  $D_j\mathbf{Q}(\mathbf{x})$ ,  $1 \leq j \leq n$ , son polinomios de grado  $\leq m - 1$ , y por la hipótesis de inducción,  $D^{\mathbf{q}}D_j\mathbf{Q}(\mathbf{x}) = 0$  para todo  $\mathbf{q}$  con  $|\mathbf{q}| = m$  y todo  $j \in \{1, 2, \dots, n\}$ , luego  $D^{\mathbf{p}}\mathbf{Q}(\mathbf{x}) = 0$  para todo  $\mathbf{p}$  con  $|\mathbf{p}| = m + 1$ .

b) Sea  $\mathbf{P}_m(\mathbf{x} - \mathbf{a})$  el polinomio de Taylor de  $\mathbf{Q}$  en  $\mathbf{a}$ , de grado  $\leq m$ , y  $\mathbf{R}_m(\mathbf{x} - \mathbf{a}) = \mathbf{Q}(\mathbf{x}) - \mathbf{P}_m(\mathbf{x} - \mathbf{a})$  el término complementario correspondiente. En virtud de a) podemos aplicar el teorema 7.14 a la función  $\mathbf{Q}$ , con  $M = 0$ , y así se obtiene que el término complementario  $\mathbf{R}_m$  es idénticamente nulo. ■

**Ejercicio 7.19** Sea  $f : \Omega \rightarrow F$  una función de clase  $C^{m+1}$  con todas sus derivadas parciales de orden  $m + 1$  idénticamente nulas. Si el abierto  $\Omega \subset \mathbb{R}^n$  es conexo demuestre que  $f$  es la restricción a  $\Omega$  de un polinomio de grado  $\leq m$ .

SOLUCIÓN

En lo que sigue  $\mathbf{P}_{\mathbf{a}}(\mathbf{x}) = \mathbf{P}_m(\mathbf{x} - \mathbf{a})$  denotará al polinomio de Taylor de orden  $m$  de la función  $\mathbf{f}$  en el punto  $\mathbf{a} \in \Omega$ . Aplicando el teorema 7.14 con  $M = 0$  se obtiene que las funciones  $\mathbf{f}$  y  $\mathbf{P}_{\mathbf{a}}$  coinciden en cada bola  $B(\mathbf{a}, r) \subset \Omega$ . Fijado  $\mathbf{b} \in \Omega$  el conjunto  $\Omega_{\mathbf{b}} := \{\mathbf{a} \in \Omega : \mathbf{P}_{\mathbf{a}} = \mathbf{P}_{\mathbf{b}}\}$  no es vacío.

Mostraremos en primer lugar que  $\Omega_{\mathbf{b}}$  es abierto: Dado  $\mathbf{a} \in \Omega_{\mathbf{b}}$  existe  $r > 0$  tal que  $B(\mathbf{a}, r) \subset \Omega$ . Fijado un punto  $\mathbf{a}' \in B(\mathbf{a}, r)$ , como las funciones  $\mathbf{f}$  y  $\mathbf{P}_{\mathbf{a}}$  coinciden en la bola  $B(\mathbf{a}, r)$  tienen el mismo polinomio de Taylor en  $\mathbf{a}'$ , luego  $\mathbf{P}_{\mathbf{a}'} = \mathbf{P}_{\mathbf{a}}$  (ya que, según el ejercicio 7.18, el polinomio de Taylor de  $\mathbf{P}_{\mathbf{a}}$  en  $\mathbf{a}'$  es  $\mathbf{P}_{\mathbf{a}}$ ). Teniendo en cuenta que  $\mathbf{P}_{\mathbf{a}} = \mathbf{P}_{\mathbf{b}}$  (porque  $\mathbf{a} \in \Omega_{\mathbf{b}}$ ) se obtiene que  $\mathbf{P}_{\mathbf{a}'} = \mathbf{P}_{\mathbf{b}}$ , luego  $\mathbf{a}' \in \Omega_{\mathbf{b}}$  y queda demostrado así que  $B(\mathbf{a}, r) \subset \Omega_{\mathbf{b}}$ , y con ello que  $\Omega_{\mathbf{b}}$  es abierto.

Ahora veremos que  $\Omega_{\mathbf{b}}$  es un conjunto cerrado relativo en espacio conexo  $\Omega$ :

Si  $\mathbf{a} \in \Omega \cap \overline{\Omega_{\mathbf{b}}}$  existe una sucesión  $\mathbf{a}_k \in \Omega_{\mathbf{b}}$  tal que  $\mathbf{a} = \lim_k \mathbf{a}_k$ . Elegimos  $\rho > 0$  tal que  $B(\mathbf{a}, 2\rho) \subset \Omega$ , y también un término de la sucesión  $\mathbf{a}_k \in B(\mathbf{a}, \rho)$ . Como



$\mathbf{a} \in B(\mathbf{a}_k, \rho) \subset B(\mathbf{a}, 2\rho) \subset \Omega$ , según lo que se indicó al principio las funciones  $\mathbf{f}$  y  $\mathbf{P}_{\mathbf{a}_k}$  coinciden en la bola  $B(\mathbf{a}_k, \rho)$  por lo que tienen el mismo polinomio de Taylor en  $\mathbf{a}$ . Razonando como antes se obtiene  $\mathbf{P}_{\mathbf{a}} = \mathbf{P}_{\mathbf{a}_k} = \mathbf{P}_{\mathbf{b}}$ , luego  $\mathbf{a} \in \Omega_{\mathbf{b}}$  y queda demostrado que  $\Omega_{\mathbf{b}}$  es cerrado en la topología relativa de  $\Omega$ .

Como  $\Omega_{\mathbf{b}} \neq \emptyset$  es un subconjunto abierto y cerrado del espacio conexo  $\Omega$  se concluye que  $\Omega = \Omega_{\mathbf{b}}$ . Esto significa que para todo  $\mathbf{a} \in \Omega$  se cumple  $\mathbf{P}_{\mathbf{a}} = \mathbf{P}_{\mathbf{b}}$ , luego  $\mathbf{f}(\mathbf{a}) = P_{\mathbf{a}}(\mathbf{a}) = P_{\mathbf{b}}(\mathbf{a})$ , es decir, todos los polinomios  $\mathbf{P}_{\mathbf{a}}$  coinciden con  $\mathbf{P}_{\mathbf{b}}$  y  $\mathbf{f} = \mathbf{P}_{\mathbf{b}}|_{\Omega}$ . ■

## 7.6. Ejercicios propuestos

◇ **7.6.1** Utilice la fórmula de Taylor para desarrollar  $x^3 + y^3 + xy$  en potencias de  $(x - 1)$  y de  $(y - 2)$ .

◇ **7.6.2** Calcule el polinomio de Taylor de grado 2 de la función

$$F(x, y) = \int_0^{f(x,y)} e^{-t^2} dt$$

en el punto  $(0, 0)$ , donde  $f(x, y) = y + x^2 + 2y^2 + 2xy + \text{sen}(x^4 + y^4)$

◇ **7.6.3** Escriba los desarrollos de Taylor de orden 3 de las siguientes funciones, en el punto  $(0, 0)$  con el término complementario en forma integral.

i)  $\log(1 + x + y)$ .

ii)  $x \text{sen } y + y \text{sen } x$ .

◇ **7.6.4** Si  $\Omega \subset \mathbb{R}^n$  es abierto, una función  $f : \Omega \rightarrow \mathbb{R}$  de clase  $C^\infty$  se dice que es analítica en  $\Omega$  si para cada  $\mathbf{a} \in \Omega$  existe una bola  $B(\mathbf{a}, r) \subset \Omega$  tal que para cada  $\mathbf{x} \in B(\mathbf{a}, r)$  se cumple

$$f(\mathbf{x}) = \lim_m P_m(\mathbf{x} - \mathbf{a})$$

donde  $P_m(\mathbf{x} - \mathbf{a})$  es el polinomio de Taylor de  $f$  en  $\mathbf{a}$ , de grado  $\leq m$ .

Si existe  $M > 0$  tal que  $|D_{j_1 j_2 \dots j_k}^k f(\mathbf{x})| \leq M^k$  para cada  $\mathbf{x} \in \Omega$ ,  $k \in \mathbb{N}$  y  $1 \leq j_1, \dots, j_k \leq n$ , demuestre que  $f$  es analítica en  $\Omega$ .

◇ **7.6.5** Demuestre que las siguientes funciones son analíticas en los abiertos que se indican

i)  $x^3 + y^2 - xy^7$ , en  $\mathbb{R}^2$ .

ii)  $\log(1 + x + y)$ , en  $\{(x, y) \in \mathbb{R}^2 : x + y + 1 > 0\}$ .

◇ **7.6.6** Demuestre que las siguientes funciones son analíticas en los abiertos que se indican

i)  $\log(x + y)$ , en  $A = \{(x, y) \in \mathbb{R}^2 : x + y > 0\}$ .

ii)  $x \text{sen } y + y \text{sen } x$ , en  $\mathbb{R}^2$ .

iii)  $xe^{x+y+z}$ , en  $\mathbb{R}^3$ .

iv)  $f(ax + by)$  en  $\mathbb{R}^2$ , siendo  $f : \mathbb{R} \rightarrow \mathbb{R}$  analítica.

v)  $\text{sen}(ax^2 + by)$  en  $\mathbb{R}^2$ .

# Capítulo 8

## Función inversa y función implícita

*Aplicaciones localmente inyectivas y aplicaciones abiertas. Inversión local de aplicaciones diferenciables. Funciones implícitas. Cambio de variable y técnicas de cálculo con funciones inversas e implícitas.*

El objetivo del cálculo diferencial es el estudio del comportamiento local de una función en el entorno de un punto. Si la aproximación local de primer orden proporcionada por la diferencial tiene cierta propiedad, cabe esperar que la función también tenga esa propiedad localmente. Resultados de este tipo son los que se tratan en este capítulo al estudiar la inversión local de aplicaciones diferenciables y la existencia de funciones definidas implícitamente.

Si  $\mathbf{f} : \Omega \rightarrow F$  es diferenciable en  $\mathbf{a} \in \Omega$  y su diferencial  $d\mathbf{f}(\mathbf{a}) : E \rightarrow F$  es una aplicación lineal invertible cabe esperar que  $\mathbf{f}$  sea localmente invertible en  $\mathbf{a}$ , lo que significa que existe algún entorno  $U \subset \Omega$  de  $\mathbf{a}$  tal que  $\mathbf{f}|_U$  tiene inversa. Los resultados de naturaleza local que conciernen al teorema de la función inversa los presentamos desdoblados en dos tipos de resultados: Los que garantizan la inyectividad local y los que aseguran que la aplicación es abierta y con ello la continuidad de la inversa.

Después de estudiar la diferenciabilidad de las funciones inversas se introducen los  $C^m$ -difeomorfismos, que son los cambios de variable naturales en problemas de cálculo diferencial donde intervienen funciones de clase  $C^m$ .

Los problemas de existencia de funciones definidas implícitamente se enmarcan en el siguiente planteamiento: Dado un sistema de  $m = n - k$  ecuaciones con  $n$  incógnitas

$$g_1(x_1, \dots, x_n) = 0; \quad g_2(x_1, \dots, x_n) = 0; \quad g_m(x_1, \dots, x_n) = 0;$$

se trata de resolverlo localmente con el fin de expresar, en el entorno de un punto, a las  $m$  variables  $x_{k+1}, x_{k+2}, \dots, x_n$  en función de las restantes variables  $x_1, x_2, \dots, x_k$ .

Los resultados que sobre este asunto se exponen aquí son versiones no lineales de resultados bien conocidos en el ámbito lineal. En ellos se asume que la diferencial de cierta aplicación cumple la hipótesis del caso lineal y se demuestra, bajo las condiciones naturales, que esta propiedad se transmite localmente a la función.

Este capítulo finaliza exponiendo con detalle las técnicas de cálculo con funciones definidas implícitamente, y en particular con funciones inversas. Finalmente se

consideran los problemas de cambio de variable en el contexto del cálculo diferencial. Con varios ejemplos se explica la técnica sistemática para realizarlos y se aplica para resolver algunas ecuaciones funcionales sencillas.

Los teoremas de la función inversa y de la función implícita intervienen en el siguiente capítulo para establecer la equivalencia de las diferentes formas de definir las subvariedades diferenciables de  $\mathbb{R}^n$ . Como tema complementario directamente relacionado con el material de este capítulo el lector interesado puede ver en el apéndice H las nociones de dependencia e independencia funcional, otro asunto interesante para el que son esenciales los teoremas de la función inversa y de la función implícita.

Las demostraciones que se ofrecen aquí para estos teoremas son de naturaleza finito dimensional. No obstante, hay otras técnicas más generales que permiten extenderlos al caso de aplicaciones entre espacios normados completos arbitrarios.

## 8.1. Aplicaciones con inversa local

Para que una función  $\mathbf{f} : \Omega \rightarrow \mathbb{R}^n$ , definida en un abierto  $\Omega \subset \mathbb{R}^n$ , se pueda invertir localmente en  $\mathbf{a} \in \Omega$ , hay que encontrar un entorno abierto  $U$  de  $\mathbf{a}$  tal que  $\mathbf{f}|_U$  sea inyectiva y abierta. Esta segunda propiedad garantizará que  $\mathbf{f}(U) = V$  es abierto y la continuidad de la inversa  $(\mathbf{f}|_U)^{-1} : V \rightarrow U$ . Comenzamos obteniendo condiciones suficientes para la inyectividad local.

**Aplicaciones localmente inyectivas.** El siguiente ejemplo muestra que la hipótesis de que la diferencial  $d\mathbf{f}(\mathbf{a})$  sea inyectiva no garantiza que  $\mathbf{f}|_U$  sea inyectiva en algún entorno  $U$  de  $\mathbf{a}$ :

**Ejemplo 8.1** *La función  $f : \mathbb{R} \rightarrow \mathbb{R}$ , definida por  $f(x) = x/2 + x^2 \operatorname{sen}(1/x)$  si  $x \neq 0$ ,  $f(0) = 0$ , es derivable en todo  $x \in \mathbb{R}$  y  $f'(0) \neq 0$ , luego su diferencial  $h \rightarrow f'(0)h$  es inyectiva. Sin embargo  $f$  no es inyectiva en los intervalos de la forma  $(-\epsilon, \epsilon)$  porque en ellos no es estrictamente monótona (ya que  $f'(1/(n\pi)) = 1/2 - (-1)^n$ ) y por lo tanto  $f'$  cambia de signo en estos intervalos)*

En este ejemplo se aprecia que la discontinuidad de  $f'$  en  $x = 0$  es la que permite que  $f$  no sea inyectiva en los entornos de 0. Para impedir situaciones como esta, en la proposición 8.2 y en el teorema de la función inversa 8.13 interviene la hipótesis de que la diferencial  $d\mathbf{f}(\mathbf{x})$  exista en un entorno de  $\mathbf{a}$ , y sea continua en  $\mathbf{a}$ . Veremos que con esta hipótesis, en el caso finito dimensional, el problema de la inversión local tiene solución satisfactoria y se consigue una inversa local, definida y continua en un entorno de  $\mathbf{b} = \mathbf{f}(\mathbf{a})$ , que resulta de clase  $C^m$  si  $\mathbf{f}$  es de clase  $C^m$ .

En lo que sigue, si  $\mathbf{f} : \Omega \rightarrow \mathbb{R}^n$  está definida en un abierto  $\Omega \subset \mathbb{R}^n$  y en el punto  $\mathbf{a} \in \Omega$  existen todas las derivadas parciales  $D_i f_j(\mathbf{a})$ ,  $1 \leq i, j \leq n$ , escribiremos

$$\det \mathbf{f}'(\mathbf{a}) = \det [D_i f_j(\mathbf{a})]$$

para denotar el valor del determinante de la matriz Jacobiana en el punto  $\mathbf{a}$ .

**Proposición 8.2** Sea  $f : \Omega \rightarrow \mathbb{R}^n$  diferenciable en un abierto  $\Omega \subset \mathbb{R}^n$  con derivadas parciales continuas en  $\mathbf{a} \in \Omega$ . Si  $\det \mathbf{f}'(\mathbf{a}) \neq 0$  entonces existe  $B(\mathbf{a}, r) \subset \Omega$  tal que  $f|_{B(\mathbf{a}, r)}$  es inyectiva.

DEM: La función  $h(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n) = \det[D_i f_k(\mathbf{z}_k)]$ , está definida en

$$\Omega^n \subset \mathbb{R}^n \times \dots \times \mathbb{R}^n = \mathbb{R}^{n^2}$$

Como las funciones  $D_i f_k$  son continuas en  $\mathbf{a}$  se sigue que  $h$  es continua en  $(\mathbf{a}, \mathbf{a}, \dots, \mathbf{a})$ . Es claro que  $h(\mathbf{a}, \mathbf{a}, \dots, \mathbf{a}) = \det \mathbf{f}'(\mathbf{a}) \neq 0$ , y si suponemos que  $\det \mathbf{f}'(\mathbf{a}) \neq 0$ , la continuidad de  $h$  en  $(\mathbf{a}, \mathbf{a}, \dots, \mathbf{a})$  permite asegurar que existe un entorno de  $(\mathbf{a}, \mathbf{a}, \dots, \mathbf{a})$ , de la forma

$$V = B(\mathbf{a}, r) \times \dots \times B(\mathbf{a}, r) \subset \Omega^n$$

tal que  $h(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n) \neq 0$  si  $(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n) \in V$ .

Demostramos a continuación que si  $\mathbf{x}, \mathbf{y} \in B(\mathbf{a}, r)$ , y  $\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{y})$ , entonces  $\mathbf{x} = \mathbf{y}$ . Como el segmento  $[\mathbf{x}, \mathbf{y}] = \{t\mathbf{x} + (1-t)\mathbf{y} : 0 \leq t \leq 1\}$  está contenido en  $B(\mathbf{a}, r) \subset \Omega$ , para  $k \in \{1, 2, \dots, n\}$ , están definidas las funciones reales de variable real  $\varphi_k(t) = f_k(\mathbf{z}(t))$  donde  $\mathbf{z}(t) = t\mathbf{x} + (1-t)\mathbf{y}$ . En virtud de la regla de la cadena estas funciones son derivables en  $[0, 1]$ , con derivada

$$\varphi'_k(t) = df_k(\mathbf{z}(t))\mathbf{z}'(t) = df_k(\mathbf{z}(t))(\mathbf{x} - \mathbf{y}) = \sum_{j=1}^n D_j f_k(\mathbf{z}(t))(x_j - y_j)$$

Según el teorema del valor medio existe  $\theta_k \in (0, 1)$  tal que

$$0 = f_k(\mathbf{x}) - f_k(\mathbf{y}) = \varphi_k(1) - \varphi_k(0) = \varphi'_k(\theta_k)$$

es decir, los puntos  $\mathbf{z}_k = \mathbf{z}(\theta_k)$  verifican

$$\sum_{j=1}^n D_j f_k(\mathbf{z}_k)(x_j - y_j) = \varphi'_k(\theta_k) = 0$$

Como  $\mathbf{z}_k = \mathbf{z}(\theta_k) \in [\mathbf{x}, \mathbf{y}] \subset B(\mathbf{a}, r)$ , podemos asegurar que  $\det[D_i f_k(\mathbf{z}_k)] \neq 0$ . Entonces, considerando

$$\sum_{j=1}^n D_j f_k(\mathbf{z}_k)(x_j - y_j) = 0, \quad 1 \leq k \leq n$$

como un sistema lineal en las incógnitas  $x_j - y_j$ , cuyo determinante no es nulo, se concluye que  $x_j - y_j = 0$  para  $1 \leq j \leq n$ , es decir,  $\mathbf{x} = \mathbf{y}$ . ■

Obsérvese que el resultado que se obtiene con la proposición 8.2 es de naturaleza local. La razón de esto se debe a que las propiedades de la diferencial, que aproxima localmente a la función en un punto, sólo pueden propiciar propiedades de la función de tipo local. El siguiente ejemplo muestra que, en general, con la proposición 8.2 no se pueden conseguir un resultado de tipo global

**Ejemplo 8.3** La aplicación  $\mathbf{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ , definida por  $\mathbf{f}(x, y) = e^x(\cos y, \operatorname{sen} y)$ , es localmente inyectiva en todo punto, pero no es globalmente inyectiva aunque su diferencial  $d\mathbf{f}(\mathbf{a})$  lo es en todo punto  $\mathbf{a} \in \mathbb{R}^2$ .

El lector que sólo esté interesado en la demostración del teorema de la función inversa puede omitir el siguiente resultado que completa el que acabamos de obtener.

**Teorema 8.4** Sea  $\mathbf{f} : \Omega \rightarrow \mathbb{R}^n$  diferenciable en un abierto  $\Omega \subset \mathbb{R}^k$ , con derivadas parciales continuas en  $\mathbf{a} \in \Omega$ . Si  $k \leq n$  y  $d\mathbf{f}(\mathbf{a}) : \mathbb{R}^k \rightarrow \mathbb{R}^n$  es inyectiva (e.d si la matriz jacobiana  $\mathbf{f}'(\mathbf{a}) = (D_i f_j(\mathbf{a}))_{1 \leq i \leq k, 1 \leq j \leq n}$  tiene rango  $k$ ) entonces existe una bola abierta  $B(\mathbf{a}, r) \subset \Omega$  tal que  $\mathbf{f}|_{B(\mathbf{a}, r)}$  es inyectiva.

DEM: Después de la proposición 8.2 sólo tenemos que considerar el caso  $k < n$ . Como la matriz  $\mathbf{f}'(\mathbf{a}) = (D_i f_j(\mathbf{a}))_{1 \leq i \leq k, 1 \leq j \leq n}$  tiene rango  $k$  suponemos, para simplificar la notación, que no es nulo el determinante de la matriz  $(D_i f_j(\mathbf{a}))_{1 \leq i \leq k, 1 \leq j \leq k}$ . Entonces la función  $\mathbf{g}(\mathbf{x}) = (f_1, f_2, \dots, f_k)$  cumple que  $\det \mathbf{g}'(\mathbf{a}) \neq 0$ , y con la proposición 8.2 se obtiene una bola abierta  $B(\mathbf{a}, r) \subset \Omega$  tal que  $\mathbf{g}|_{B(\mathbf{a}, r)}$  es inyectiva lo que implica que  $\mathbf{f}|_{B(\mathbf{a}, r)}$  también lo es ■

### Ejemplo 8.5

La aplicación  $\mathbf{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  definida por  $\mathbf{f}(x, y, z) = (x + y, x^2 - y, y^4)$  es localmente inyectiva en cada punto  $(x, y) \neq (-1/2, 0)$  ya que la matriz Jacobiana

$$\begin{pmatrix} 1 & 2x & 0 \\ 1 & -1 & 4y^3 \end{pmatrix}$$

tiene rango dos en todo  $(x, y) \neq (-1/2, 0)$ . El punto  $(-1/2, 0)$ , donde el rango de la matriz es 1, no tiene ningún entorno sobre el que  $\mathbf{f}$  sea inyectiva: Basta observar que para todo  $\epsilon > 0$  se cumple  $\mathbf{f}(-1/2 + \epsilon, -\epsilon) = \mathbf{f}(-1/2 - \epsilon, +\epsilon)$ . ■

**Aplicaciones abiertas.** Recordemos que una transformación espacios topológicos se dice que es abierta cuando transforma abiertos en abiertos, lo que equivale a que transforma cada entorno de un punto de su dominio en un entorno del punto imagen. El siguiente lema proporciona un ingrediente básico para la demostración del teorema de la aplicación abierta.

**Lema 8.6** Sea  $\mathbf{f} : \overline{B} \rightarrow \mathbb{R}^n$  una aplicación continua, donde  $B = B(\mathbf{a}, r) \subset \mathbb{R}^n$  una bola abierta para la norma euclídea. Se supone que

- i) Para cada  $\mathbf{x} \in B(\mathbf{a}, r)$  existen las derivadas  $D_i f_k(\mathbf{x})$ ,  $1 \leq i, k \leq n$ , y  $\det \mathbf{f}'(\mathbf{x}) \neq 0$ .
- ii)  $f(\mathbf{a}) \neq f(\mathbf{x})$  si  $\|\mathbf{x} - \mathbf{a}\|_2 = r$ .

Entonces existe  $\rho > 0$  tal que  $B(\mathbf{f}(\mathbf{a}), \rho) \subset \mathbf{f}(B(\mathbf{a}, r))$ .

DEM: La función continua  $\mathbf{x} \rightarrow \|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a})\|_2$  alcanza un mínimo absoluto sobre el compacto  $S_r = \{\mathbf{x} : \|\mathbf{x} - \mathbf{a}\|_2 = r\}$ , luego existe  $\mathbf{z} \in S_r$  tal que

$$M = \min\{\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a})\|_2 : \|\mathbf{x} - \mathbf{a}\|_2 = r\} = \|\mathbf{f}(\mathbf{z}) - \mathbf{f}(\mathbf{a})\|_2$$

Según la hipótesis ii),  $\mathbf{f}(\mathbf{z}) \neq \mathbf{f}(\mathbf{a})$ , y así podemos asegurar que  $M > 0$ . Vamos a demostrar que con  $\rho = M/2$  se cumple la inclusión  $B(\mathbf{f}(\mathbf{a}), \rho) \subset \mathbf{f}(B(\mathbf{a}, r))$ .

Dado  $\mathbf{y} \in B(\mathbf{f}(\mathbf{a}), \rho)$ , la función continua  $h(\mathbf{x}) = \|\mathbf{f}(\mathbf{x}) - \mathbf{y}\|_2$  alcanza en un punto  $\mathbf{e}$  del compacto  $\overline{B}$  el mínimo absoluto

$$\min\{h(\mathbf{x}) : \|\mathbf{x} - \mathbf{a}\|_2 \leq r\} = h(\mathbf{e})$$

Si  $\|\mathbf{x} - \mathbf{a}\|_2 = r$  se cumple  $h(\mathbf{x}) > h(\mathbf{e})$ , ya que

$$h(\mathbf{x}) = \|\mathbf{f}(\mathbf{x}) - \mathbf{y}\|_2 = \|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a}) - (\mathbf{y} - \mathbf{f}(\mathbf{a}))\|_2 \geq$$

$$\geq \|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a})\|_2 - \|\mathbf{y} - \mathbf{f}(\mathbf{a})\|_2 \geq M - \rho = \rho > \|\mathbf{f}(\mathbf{a}) - \mathbf{y}\|_2 = h(\mathbf{a}) \geq h(\mathbf{e})$$

y podemos asegurar así que  $\mathbf{e} \in B(\mathbf{a}, r)$ . En definitiva, en la bola abierta  $B(\mathbf{a}, r)$ , la función diferenciable  $h(\mathbf{x})^2 = \sum_{j=1}^n (f_j(\mathbf{x}) - y_j)^2$  alcanza un mínimo (absoluto) en  $\mathbf{e} \in B(\mathbf{a}, r)$ , y por lo tanto  $D_k h^2(\mathbf{e}) = 0$  para  $1 \leq k \leq n$ , es decir

$$2 \sum_{j=1}^n (f_j(\mathbf{e}) - y_j) D_k f_j(\mathbf{e}) = 0$$

Por hipótesis  $\det[D_k f_j(\mathbf{e})] = \det \mathbf{f}'(\mathbf{e}) \neq 0$ , luego el sistema homogéneo de ecuaciones lineales asociado a la matriz  $D_k f_j(\mathbf{e})$  sólo tiene la solución trivial. Por lo tanto  $f_j(\mathbf{e}) - y_j = 0$  para cada  $1 \leq j \leq n$ , y queda demostrado que  $\mathbf{y} = \mathbf{f}(\mathbf{e}) \in \mathbf{f}(B(\mathbf{a}, r))$ . ■

**Proposición 8.7** *Sea  $\mathbf{f} : \Omega \rightarrow \mathbb{R}^n$  continua e inyectiva en un abierto  $\Omega \subset \mathbb{R}^n$  tal que en cada  $\mathbf{x} \in \Omega$  existen las derivadas parciales  $D_i f_k(\mathbf{x})$ ,  $1 \leq i, k \leq n$ , y  $\det \mathbf{f}'(\mathbf{x}) \neq 0$ . Entonces  $\mathbf{f}$  es abierta, es decir,  $\mathbf{f}(V)$  es abierto para cada abierto  $V \subset \Omega$ .*

DEM: Dado  $\mathbf{a} \in V$ , sea  $r > 0$  tal que  $\overline{B(\mathbf{a}, r)} \subset V$ . Sobre la bola cerrada  $\overline{B(\mathbf{a}, r)}$  se satisfacen las hipótesis del lema 8.6, luego existe  $\rho > 0$  tal que

$$B(\mathbf{f}(\mathbf{a}), \rho) \subset \mathbf{f}(B(\mathbf{a}, r)) \subset \mathbf{f}(V)$$

■

Obsérvese que en la proposición 8.7 no se ha supuesto que  $\mathbf{f}$  sea de clase  $C^1$ . Para funciones de clase  $C^1$  el resultado proporcionado por esta proposición queda cubierto por el siguiente teorema, donde no se supone que la función sea inyectiva.

**Teorema 8.8** [Aplicación abierta] *Si  $\mathbf{f} : \Omega \rightarrow \mathbb{R}^n$  es de clase  $C^1(\Omega)$  en un abierto  $\Omega \subset \mathbb{R}^n$  y  $\det \mathbf{f}'(\mathbf{x}) \neq 0$  para todo  $\mathbf{x} \in \Omega$  entonces  $\mathbf{f}$  es abierta.*

DEM: Si  $U \subset \Omega$  es abierto y  $\mathbf{a} \in U$ , aplicando la proposición 8.2, se obtiene una bola abierta  $B_{\mathbf{a}} \subset U$ , centrada en  $\mathbf{a}$ , tal que  $\mathbf{f}|_{B_{\mathbf{a}}}$  es inyectiva. En virtud de la proposición 8.7 cada  $\mathbf{f}(B_{\mathbf{a}})$  es abierto, luego  $\mathbf{f}(U) = \bigcup_{\mathbf{a} \in U} \mathbf{f}(B_{\mathbf{a}})$  es abierto. ■

NOTA: En el ejercicio resuelto 8.19 se muestra que el teorema 8.8 sigue valiendo con

una hipótesis más débil: Basta suponer que  $\mathbf{f}$  es diferenciable en cada  $\mathbf{x} \in \Omega$  y que  $\det \mathbf{f}'(\mathbf{x}) \neq 0$  para todo  $\mathbf{x} \in \Omega$ .

El lector que sólo esté interesado en el teorema de la función inversa también puede omitir el siguiente teorema y su corolario, que completan el resultado anterior.

**Teorema 8.9** *Sea  $\mathbf{f} : \Omega \rightarrow \mathbb{R}^m$  de clase  $C^1$  en un abierto  $\Omega \subset \mathbb{R}^n$ , donde  $m \leq n$  y  $\mathbf{a} \in \Omega$ . Si  $d\mathbf{f}(\mathbf{a}) : \mathbb{R}^n \rightarrow \mathbb{R}^m$  es sobreyectiva (e.d. si la matriz jacobiana  $\mathbf{f}'(\mathbf{a})$  es de rango  $m$ ) y  $A \subset \Omega$  es entorno de  $\mathbf{a}$  entonces  $\mathbf{f}(A)$  es entorno de  $\mathbf{f}(\mathbf{a})$ . Más aún,  $\mathbf{a}$  posee un entorno abierto  $U \subset \Omega$  tal que  $\mathbf{f}|_U$  es abierta.*

DEM: Como el caso  $n = m$  ya ha sido demostrado en el teorema 8.8, sólo tenemos que considerar el caso  $m < n$ . Suponemos, para simplificar la notación, que no se anula el determinante  $\Delta(\mathbf{a})$  de la matriz cuadrada  $(D_i f_j(\mathbf{a}))_{1 \leq i, j \leq m}$ . Como el determinante  $\Delta(\mathbf{x})$  de la matriz  $(D_i f_j(\mathbf{x}))_{1 \leq i, j \leq m}$  es una función continua de  $\mathbf{x}$  podemos asegurar que existe  $B(\mathbf{a}, r) \subset A$  tal que  $\Delta(\mathbf{x}) \neq 0$  para cada  $\mathbf{x} \in B(\mathbf{a}, r)$ . Si consideramos la función auxiliar  $\mathbf{g} : \Omega \rightarrow \mathbb{R}^n$ , definida por

$$\mathbf{g}(x_1, x_2, \dots, x_n) = (f_1(x_1, \dots, x_n), f_2(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n), x_{m+1}, \dots, x_n)$$

es fácil comprobar que  $\det \mathbf{g}'(\mathbf{x}) = \Delta(\mathbf{x})$ . Como la función  $\mathbf{g}$  es de clase  $C^1(\Omega)$  y  $\det \mathbf{g}'(\mathbf{x}) \neq 0$  para todo  $\mathbf{x} \in B(\mathbf{a}, r)$ , con el teorema 8.8 obtenemos que  $\mathbf{g}$  es abierta sobre  $U = B(\mathbf{a}, r)$ . Como la proyección

$$\pi : \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad \pi(x_1, x_2, \dots, x_m, x_{m+1}, \dots, x_n) = (x_1, x_2, \dots, x_m)$$

transforma abiertos en abiertos se sigue que  $\mathbf{f}|_U = \pi \circ \mathbf{g}|_U : U \rightarrow \mathbb{R}^m$  es abierta, luego  $\mathbf{f}(U)$  es un subconjunto abierto de  $\mathbb{R}^m$  contenido en  $\mathbf{f}(A)$ , y por lo tanto  $\mathbf{f}(A)$  es entorno de  $\mathbf{f}(\mathbf{a})$ . ■

**Corolario 8.10** *Sea  $\mathbf{f} : \Omega \rightarrow \mathbb{R}^m$  de clase  $C^1$  en un abierto  $\Omega \subset \mathbb{R}^n$ , donde  $m \leq n$ . Si  $d\mathbf{f}(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{R}^m$  es sobreyectiva en todo punto  $\mathbf{x} \in \Omega$  (e.d. si la matriz jacobiana  $\mathbf{f}'(\mathbf{x})$  es de rango  $m$  en cada  $\mathbf{x} \in \Omega$ ) entonces  $\mathbf{f}$  es abierta.*

DEM: Es consecuencia directa del teorema 8.9 ya que se cumplen sus hipótesis en cada punto  $\mathbf{a} \in \Omega$ . ■

**Funciones inversas. Teorema de inversión local.** Antes de demostrar el teorema de inversión local demostraremos dos resultados preliminares que nos dicen que, bajo las hipótesis naturales, cuando una función  $\mathbf{f}$  tiene inversa continua las propiedades de diferenciabilidad de  $\mathbf{f}$  las hereda la inversa.

**Teorema 8.11** *Sea  $\mathbf{f} : A \rightarrow B$  una biyección entre dos abiertos  $A, B \subset \mathbb{R}^n$  tal que  $\mathbf{f}$  es diferenciable en  $\mathbf{a} \in A$  y su inversa  $\mathbf{g} : B \rightarrow A$  es continua en  $\mathbf{b} = \mathbf{f}(\mathbf{a})$ .*

*Si  $\det \mathbf{f}'(\mathbf{a}) \neq 0$  entonces  $\mathbf{g} = \mathbf{f}^{-1} : B \rightarrow A$  es diferenciable en  $\mathbf{b} = \mathbf{f}(\mathbf{a})$  y su diferencial  $d\mathbf{g}(\mathbf{b})$  es la inversa de la aplicación lineal  $d\mathbf{f}(\mathbf{a})$ .*



DEM: La diferencial  $L = df(\mathbf{a})$  es una aplicación lineal  $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$  que satisface

$$\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a}) = L(\mathbf{h}) + \|\mathbf{h}\| \epsilon(\mathbf{h}) \text{ donde } \lim_{\mathbf{h} \rightarrow \mathbf{0}} \epsilon(\mathbf{h}) = \mathbf{0}$$

En lo que sigue suponemos definido  $\epsilon(\mathbf{0}) = \mathbf{0}$  de modo que  $\epsilon(\mathbf{h})$  es continua en  $\mathbf{h} = \mathbf{0}$  y cumple la igualdad anterior incluso en el caso  $\mathbf{h} = \mathbf{0}$ .

Si  $\mathbf{k} \in \mathbb{R}^n$ , con  $\mathbf{b} + \mathbf{k} \in B$ , entonces el incremento  $\mathbf{h} = \mathbf{g}(\mathbf{b} + \mathbf{k}) - \mathbf{g}(\mathbf{b}) \in \mathbb{R}^n$  cumple que  $\mathbf{a} + \mathbf{h} \in A$  y además  $\mathbf{k} = \mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a})$ , luego

$$\mathbf{k} = L(\mathbf{h}) + \|\mathbf{h}\| \epsilon(\mathbf{h})$$

La aplicación lineal  $L$  tiene inversa porque  $\det(L) = \det \mathbf{f}'(\mathbf{a}) \neq 0$ , y aplicando  $L^{-1}$  a los dos miembros de la última igualdad resulta

$$L^{-1}(\mathbf{k}) = \mathbf{h} + \|\mathbf{h}\| L^{-1}(\epsilon(\mathbf{h}))$$

y sustituyendo  $\mathbf{h} = \mathbf{g}(\mathbf{b} + \mathbf{k}) - \mathbf{g}(\mathbf{b})$  se obtiene

$$\mathbf{g}(\mathbf{b} + \mathbf{k}) - \mathbf{g}(\mathbf{b}) = L^{-1}(\mathbf{k}) - \|\mathbf{h}\| L^{-1}(\epsilon(\mathbf{h}))$$

En la fórmula anterior, y en lo que sigue, se considera siempre que  $\mathbf{h}$  es función de  $\mathbf{k}$ , es decir, se supone efectuada la sustitución  $\mathbf{h} = \mathbf{g}(\mathbf{b} + \mathbf{k}) - \mathbf{g}(\mathbf{b})$ . Obsérvese que el incremento  $\mathbf{h} = \mathbf{g}(\mathbf{b} + \mathbf{k}) - \mathbf{g}(\mathbf{b})$  tiende hacia  $\mathbf{0}$  cuando  $\mathbf{k} \rightarrow \mathbf{0}$  porque  $\mathbf{g}$  es continua en  $\mathbf{b}$ . Se sigue que  $\epsilon(\mathbf{h})$  también tiende hacia  $\epsilon(\mathbf{0}) = \mathbf{0}$  cuando  $\mathbf{k}$  tiende hacia  $\mathbf{0}$ .

Para demostrar que  $\mathbf{g}$  es diferenciable en  $\mathbf{b}$ , con  $d\mathbf{g}(\mathbf{b}) = L^{-1}$ , basta ver que

$$\|\mathbf{h}\| L^{-1}(\epsilon(\mathbf{h})) = o(\|\mathbf{k}\|)$$

Como  $L^{-1}$  es continua y  $\epsilon(\mathbf{h})$  tiende hacia  $\mathbf{0}$  cuando  $\mathbf{k}$  tiende hacia  $\mathbf{0}$ , existe  $\delta > 0$  tal que  $\|\mathbf{k}\| < \delta \Rightarrow \|L^{-1}(\epsilon(\mathbf{h}))\| < 1/2$ , luego

$$\|L^{-1}(\mathbf{k})\| = \|\mathbf{h} + \|\mathbf{h}\| L^{-1}(\epsilon(\mathbf{h}))\| \geq \|\mathbf{h}\| - \|\mathbf{h}\| \|L^{-1}(\epsilon(\mathbf{h}))\| \geq \frac{1}{2} \|\mathbf{h}\|$$

Entonces, cuando  $\|\mathbf{k}\| < \delta$ , se cumple

$$\|\mathbf{h}\| \|L^{-1}(\epsilon(\mathbf{h}))\| \leq 2 \|L^{-1}(\mathbf{k})\| \|L^{-1}(\epsilon(\mathbf{h}))\| \leq \|\mathbf{k}\| \|L^{-1}\|^2 \|\epsilon(\mathbf{h})\|$$

y se sigue de esta desigualdad que  $\|\mathbf{h}\| \|L^{-1}(\epsilon(\mathbf{h}))\| / \|\mathbf{k}\|$  tiende hacia 0 cuando  $\mathbf{k}$  tiende hacia  $\mathbf{0}$ . ■

OBSERVACIÓN. En el teorema anterior la hipótesis  $\det \mathbf{g}'(\mathbf{a}) \neq 0$  es crucial para conseguir la diferenciable de  $\mathbf{g}$  en  $\mathbf{b} = \mathbf{f}(\mathbf{a})$ : La función  $f(x) = x^3$  es derivable en  $a = 0$  pero su inversa  $g(x) = \sqrt[3]{x}$ , que es continua en  $b = 0$ , no es derivable en este punto.

**Teorema 8.12** *Sea  $\mathbf{f} : A \rightarrow B$  una biyección entre dos abiertos  $A, B \subset \mathbb{R}^n$ , diferenciable en cada  $\mathbf{x} \in A$  con  $\det \mathbf{f}'(\mathbf{x}) \neq 0$ . Si la inversa  $\mathbf{g} = \mathbf{f}^{-1} : B \rightarrow A$  es continua, entonces es diferenciable en cada  $\mathbf{y} \in B$ . Si  $\mathbf{f}$  es de clase  $C^m(A)$  entonces  $\mathbf{g}$  es de clase  $C^m(B)$ .*

DEM: Por lo que se acaba de demostrar en el teorema 8.11 la inversa  $\mathbf{g} : B \rightarrow A$  es diferenciable en cada  $\mathbf{y} \in B$  y su diferencial es  $d\mathbf{g}(\mathbf{y}) = [d\mathbf{f}(\mathbf{g}(\mathbf{y}))]^{-1}$ , es decir, la matriz jacobiana de  $\mathbf{g}'(\mathbf{y})$  se obtiene invirtiendo la matriz jacobiana  $\mathbf{f}'(\mathbf{x})$  y sustituyendo luego  $\mathbf{x} = \mathbf{g}(\mathbf{y})$ .

Para demostrar que  $\mathbf{g}$  es de clase  $C^m(B)$  cuando  $\mathbf{f}$  es de clase  $C^m(A)$  consideramos el espacio  $\mathcal{M}$  formado por las matrices cuadradas  $n \times n$  de números reales, que se supone identificado con  $\mathbb{R}^{n^2}$  dotado de la topología usual. Con esta topología la aplicación  $\det : \mathcal{M} \rightarrow \mathbb{R}$  que asocia a cada matriz  $M \in \mathcal{M}$  su determinante es continua y por lo tanto el conjunto de las matrices invertibles

$$\Gamma = \{M \in \mathcal{M} : \det(M) \neq 0\}$$

es un subconjunto abierto en  $\mathcal{M}$ . La aplicación  $\text{Inv} : \Gamma \rightarrow \Gamma$  que asocia a cada matriz  $M = (m_{ij}) \in \Gamma$  su matriz inversa  $\text{Inv}(M) = M^{-1}$  es de clase  $C^\infty$  (basta tener en cuenta que cada elemento de la matriz inversa  $M^{-1}$  es una función racional de las variables  $m_{ij}$  cuyo denominador  $\det[m_{ij}] \neq 0$  no se anula). Si  $\mathbf{f}$  es de clase  $C^1$ , como  $\mathbf{g}'$  se obtiene componiendo las aplicaciones continuas

$$\mathbf{g}' : B \xrightarrow{\mathbf{g}} A \xrightarrow{\mathbf{f}'} \Gamma \xrightarrow{\text{Inv}} \Gamma$$

obtenemos que  $\mathbf{g}'$  es continua, lo que significa que  $\mathbf{g}$  es de clase  $C^1$ . Razonado por inducción sobre  $m$  se demuestra que  $\mathbf{g}$  es de clase  $C^m$  si  $\mathbf{f}$  lo es:

Ya hemos visto que el resultado es cierto para  $m = 1$ . Si  $\mathbf{f}$  es de clase  $C^m$  y el resultado se supone cierto para funciones de clase  $C^{m-1}$ , esta hipótesis de inducción conduce a que  $\mathbf{g}$  es de clase  $C^{m-1}$ . Como  $\mathbf{f}'$  también es de clase  $C^{m-1}$ , y lo mismo le ocurre a  $\text{Inv}$  resulta que  $\mathbf{g}'$  es la composición de tres aplicaciones de clase  $C^{m-1}$ . La proposición 7.4 permite concluir que  $\mathbf{g}'$  es de clase  $C^{m-1}$ , lo que significa que  $\mathbf{g}$  es de clase  $C^m$ . ■

**Teorema 8.13** [Función inversa] *Sea  $\mathbf{f} : \Omega \rightarrow \mathbb{R}^n$  diferenciable en un abierto  $\Omega \subset \mathbb{R}^n$  y  $\mathbf{a} \in \Omega$  un punto donde las derivadas  $D_i f_j$ ,  $1 \leq i, j \leq n$  son continuas.*

*Si  $\det \mathbf{f}'(\mathbf{a}) \neq 0$ , existen abiertos  $A \subset \Omega$ ,  $B \subset \mathbb{R}^n$ , con  $\mathbf{a} \in A$ ,  $\mathbf{b} = \mathbf{f}(\mathbf{a}) \in B$ , verificando*

*i)  $\mathbf{f}|_A$  es inyectiva y  $\mathbf{f}(A) = B$ .*

*ii)  $\mathbf{g} = (\mathbf{f}|_A)^{-1} : B \rightarrow A$  es diferenciable en  $B$ .*

*Si  $\mathbf{f}$  es de clase  $C^m(A)$  entonces  $\mathbf{g}$  sea de clase  $C^m(B)$ .*

DEM: La hipótesis sobre las derivadas parciales garantiza que la función  $\mathbf{x} \rightarrow \det \mathbf{f}'(\mathbf{x})$  es continua en  $\mathbf{a}$ , luego la condición  $\det \mathbf{f}'(\mathbf{a}) \neq 0$  permite asegurar que existe  $B(\mathbf{a}, \rho) \subset \Omega$  tal que  $\det \mathbf{f}'(\mathbf{x}) \neq 0$  para todo  $\mathbf{x} \in B(\mathbf{a}, \rho)$ . Según la proposición 8.2 existe una bola abierta  $A = B(\mathbf{a}, r) \subset B(\mathbf{a}, \rho)$  tal que  $\mathbf{f}|_A$  es inyectiva. En el abierto  $A$  la función  $\mathbf{f}$  cumple las hipótesis de la proposición 8.7 luego  $\mathbf{f}|_A$  es abierta, es decir  $B = \mathbf{f}(A)$  es abierto y  $\mathbf{f}$  transforma cada abierto  $U \subset A$  en un abierto  $V = \mathbf{f}(U) \subset B$ , lo que significa que la inversa  $\mathbf{g} = (\mathbf{f}|_A)^{-1} : B \rightarrow A$  de la aplicación inyectiva  $\mathbf{f}|_A$  es continua en  $B$ . Queda establecido así que los abiertos  $A = B(\mathbf{a}, r)$ ,

$B = \mathbf{f}(A)$  cumplen i), y que la inversa  $g : B \rightarrow A$  es continua. Como  $\det \mathbf{f}'(\mathbf{x}) \neq 0$  para todo  $\mathbf{x} \in A \subset B(\mathbf{a}, \rho)$ , con el teorema 8.12 se concluye que  $\mathbf{g}$  es diferenciable en cada  $\mathbf{y} \in B$ , y de clase  $C^m(B)$  si  $\mathbf{f}$  es de clase  $C^m(A)$ . ■

Cuando  $\mathbf{f} : U \rightarrow V$  es una biyección de clase  $C^m$  entre dos abiertos  $U, V \subset \mathbb{R}^n$  y su inversa  $\mathbf{g} = \mathbf{f}^{-1} : V \rightarrow U$  también es de clase  $C^m$  se dice que  $\mathbf{f}$  es un *difeomorfismo* de clase  $C^m$  (o un  *$C^m$ -difeomorfismo*) entre los abiertos  $U$  y  $V$ . En este caso se dice que  $U$  y  $V$  son  $C^m$ -difeomorfos.

**Corolario 8.14** *Sea  $\mathbf{f} : U \rightarrow V$  una biyección de clase  $C^m$ , entre dos abiertos,  $U, V \subset \mathbb{R}^n$ . Una condición necesaria y suficiente para que  $\mathbf{f}$  sea un  $C^m$ -difeomorfismo es que para todo  $\mathbf{x} \in U$  sea  $\det \mathbf{f}'(\mathbf{x}) \neq 0$ .*

DEM: La condición es suficiente: En un entorno  $B$  de cada  $\mathbf{b} = \mathbf{f}(\mathbf{a}) \in V$  la inversa  $\mathbf{g}$  de  $\mathbf{f}$  coincide con la inversa local de  $\mathbf{f}$  proporcionada por el teorema de la función inversa 8.13, que es de clase  $C^m$ , luego  $\mathbf{g}$  es de clase  $C^m(V)$ .

La condición es necesaria: Dado  $\mathbf{x} \in U$ , por hipótesis  $\mathbf{f}$  es diferenciable en  $\mathbf{x}$  y  $\mathbf{g}$  es diferenciable en  $\mathbf{y} = \mathbf{f}(\mathbf{x})$ . Como  $\mathbf{g}(\mathbf{f}(\mathbf{x})) = \mathbf{x}$  para todo  $\mathbf{x} \in U$ , en virtud de la regla de la cadena  $d\mathbf{g}(\mathbf{y}) \circ d\mathbf{f}(\mathbf{x}) = I$  (identidad), luego la aplicación lineal  $d\mathbf{f}(\mathbf{x})$  es invertible y por lo tanto  $\det \mathbf{f}'(\mathbf{x}) \neq 0$ . ■

## 8.2. Funciones implícitas

El problema de la función implícita, en su forma más simple, se plantea en los siguientes términos: Si  $g$  es una función real de dos variables reales definida en un abierto  $\Omega \subset \mathbb{R}^2$ , se considera la ecuación  $g(x, y) = 0$ . Si no es vacío el conjunto de sus soluciones  $S = \{(x, y) \in \Omega : g(x, y) = 0\}$ , se trata de decidir cuándo y en qué sentido esta ecuación determina a la variable  $y$  como función de la variable  $x$ , es decir, bajo qué condiciones queda definida una función  $y = f(x)$  tal que  $g(x, f(x)) = 0$ .

Esto ocurrirá con seguridad cuando  $S = \{(x, y) \in \Omega : g(x, y) = 0\}$  sea la gráfica de una función  $f : A \rightarrow \mathbb{R}$ , definida en  $A = \pi_1(\Omega)$ . En este caso, la función  $f$ , que asigna a cada  $x \in A$  la única solución  $y$  de la ecuación  $g(x, y) = 0$  se dice que está definida implícitamente por dicha ecuación. Esto es lo que ocurre, por ejemplo, con  $g(x, y) = x^2 + y^2 - 1$ , que define en el abierto  $\Omega = \{(x, y) \in \mathbb{R}^2 : y > 0\}$ , la función implícita  $f(x) = \sqrt{1 - x^2}$ , con dominio  $A = (-1, +1)$ .

Es raro que se presente una situación tan sencilla como la anterior pues puede ocurrir que la ecuación  $g(x, y) = 0$  no tenga solución, o que el conjunto de sus soluciones se reduzca a un punto (p.e. si  $g(x, y) = x^2 + y^2$ ). Aún suponiendo que este no es el caso, puede ocurrir que para algunos valores  $x \in A = \pi_1(S)$  la ecuación  $g(x, y) = 0$  tenga varias o infinitas soluciones. En estos casos, para que quede determinada una función implícita  $f : A \rightarrow \mathbb{R}$ , será preciso considerar alguna condición adicional que garantice que para cada  $x \in A$  hay un único  $y$  que satisface la ecuación  $g(x, y) = 0$  y la condición propuesta.

Las condición adicional que se suele proponer para determinar una función implícita es de tipo local: Dado un punto  $(a, b) \in S$  se trata de encontrar un entorno  $A \times B \subset \Omega$  de  $(a, b)$  con la propiedad de que para cada  $x \in A$  haya un único  $y \in B$  que sea solución de la ecuación  $g(x, y) = 0$  de modo que  $S \cap (A \times B)$  es la gráfica de una función  $f : A \rightarrow B$ . En este caso se suele decir que  $f$  es la función implícita que  $A \times B$  determina en la ecuación  $g(x, y) = 0$ .

En el caso particular que estamos considerando la interpretación geométrica de la hipótesis que habrá que considerar para garantizar la existencia de función implícita es sencilla y natural: En el plano se tiene una curva  $S$  de ecuación  $g(x, y) = 0$  que se pretende expresar, en un entorno  $A \times B$  de  $(a, b) \in S$ , en la forma explícita habitual, como gráfica de una función real de variable real  $f : A \rightarrow B$ . Para ello debemos descartar los puntos  $(a, b)$  de la curva donde hay tangente vertical porque en ellos puede ocurrir que sea imposible despejar localmente a la variable  $y$  como función implícita de la variable  $x$ . Un ejemplo típico de esta situación lo proporciona la circunferencia  $C = \{(x, y) : x^2 + y^2 - 1 = 0\}$  cuando se considera el punto  $(a, b) = (1, 0)$ : En todo entorno  $A \times B$  de  $(1, 0)$ , por pequeño que sea, es imposible expresar  $C \cap (A \times B)$  como la gráfica de una función  $f : A \rightarrow B$ . Este ejemplo revela, en el caso de funciones de dos variables reales, el papel que desempeña la hipótesis  $D_2g(a, b) \neq 0$  para conseguir que la ecuación  $g(x, y) = 0$  defina, en un entorno de  $(a, b)$ , a la variable  $y$  como función implícita de la variable  $x$ .

Más generalmente, dado un sistema de  $m$  ecuaciones con  $n > m$  incógnitas

$$g_1(x_1, x_2, \dots, x_k, x_{k+1}, \dots, x_n) = 0$$

$$g_2(x_1, x_2, \dots, x_k, x_{k+1}, \dots, x_n) = 0$$

.....

$$g_m(x_1, x_2, \dots, x_k, x_{k+1}, \dots, x_n) = 0$$

del que se conoce una solución particular  $(a_1, a_2, \dots, a_n)$ , se plantea el problema análogo de estudiar cuando es posible despejar localmente  $m = n - k$  variables, por ejemplo  $(x_{k+1}, \dots, x_n)$  en función de las restantes  $(x_1, x_2, \dots, x_k)$ .

En lo que sigue, si  $k \in \{1, 2, \dots, n - 1\}$ , y  $m = n - k$ , conviene considerar  $\mathbb{R}^n$  identificado con  $\mathbb{R}^k \times \mathbb{R}^m$ , y así un elemento genérico de  $\mathbb{R}^n$  se escribirá en la forma  $(\mathbf{x}, \mathbf{y})$  con  $\mathbf{x} = (x_1, x_2, \dots, x_k) \in \mathbb{R}^k$ ,  $\mathbf{y} = (y_1, y_2, \dots, y_m)$ . Con estos convenios de notación, usando notación vectorial,  $\mathbf{g} = (g_1, g_2, \dots, g_m)$ ,  $\mathbf{x} = (x_1, \dots, x_k)$ ,  $\mathbf{y} = (x_{k+1}, \dots, x_n)$ ,  $\mathbf{a} = (a_1, \dots, a_k)$ ,  $\mathbf{b} = (a_{k+1}, \dots, a_n)$ , el problema de existencia de función implícita se plantea así: Dada la ecuación  $\mathbf{g}(\mathbf{x}, \mathbf{y}) = 0$ , de la que se conoce una solución particular  $\mathbf{g}(\mathbf{a}, \mathbf{b}) = 0$ , se trata de estudiar cuando es posible despejar localmente, en un entorno de  $(\mathbf{a}, \mathbf{b})$ , a la variable vectorial  $\mathbf{y}$  en función de la variable vectorial  $\mathbf{x}$ .

**Definición 8.15** Si  $\mathbf{g} : \Omega \rightarrow \mathbb{R}^m$  está definida en un abierto  $\Omega \subset \mathbb{R}^k \times \mathbb{R}^m$  y  $(\mathbf{a}, \mathbf{b}) \in \Omega$  es un punto que cumple  $\mathbf{g}(\mathbf{a}, \mathbf{b}) = 0$ , se dice que la ecuación  $\mathbf{g}(\mathbf{x}, \mathbf{y}) = 0$  define, en un entorno de  $(\mathbf{a}, \mathbf{b})$ , a la variable  $\mathbf{y} \in \mathbb{R}^m$  como función implícita de la variable  $\mathbf{x} \in \mathbb{R}^k$  si ocurre lo siguiente: Existe un entorno abierto  $A$  de  $\mathbf{a}$ , y un entorno abierto

$B$  de  $\mathbf{b}$ , tales que  $A \times B \subset \Omega$ , y para cada  $\mathbf{x} \in A$  hay un único  $\mathbf{y} = \mathbf{f}(\mathbf{x}) \in B$  que verifica  $\mathbf{g}(\mathbf{x}, \mathbf{f}(\mathbf{x})) = 0$ . En este caso se dice que  $\mathbf{f} : A \rightarrow B$  es la función implícita que  $A \times B$  determina en la ecuación  $\mathbf{g}(\mathbf{x}, \mathbf{y}) = 0$ .

Según esta definición existirá función implícita cuando se pueda garantizar la existencia de un entorno abierto  $A \times B \subset \Omega$  de  $(\mathbf{a}, \mathbf{b})$  tal que  $\{(\mathbf{x}, \mathbf{y}) \in A \times B : \mathbf{g}(\mathbf{x}, \mathbf{y}) = 0\}$  sea la gráfica de alguna función  $\mathbf{f} : A \rightarrow B$ .

Antes de formular y demostrar el teorema de existencia de funciones implícitas 8.16 una sencilla reflexión preliminar revelará el papel que desempeña la hipótesis central de este teorema (la no anulación, en el punto correspondiente, del jacobiano de  $\mathbf{g}$  respecto a las variables que se desean despejar).

En el caso particularmente simple de que las ecuaciones sean lineales

$$g_i(x_1, x_2, \dots, x_k, x_{k+1}, \dots, x_n) = \sum_{j=1}^k a_{ij}x_j + \sum_{j=k+1}^n a_{ij}x_j, \quad 1 \leq i \leq n - k$$

para poder resolver el sistema respecto a las variables  $x_{k+1}, x_{k+2}, \dots, x_n$  sabemos que hay que requerir que la matriz cuadrada  $\{a_{i,k+j} : 1 \leq i, j \leq m\}$  tenga determinante no nulo. En este caso, si  $\mathbf{g} : \mathbb{R}^k \times \mathbb{R}^m \rightarrow \mathbb{R}^m$  es la aplicación lineal de componentes  $(g_1 \cdots g_m)$ , fijado  $\mathbf{x} \in \mathbb{R}^k$ , la diferencial de la aplicación afín  $\mathbf{y} \rightarrow g_{\mathbf{x}}(\mathbf{y}) = g(\mathbf{x}, \mathbf{y})$  es una aplicación lineal invertible  $L : \mathbb{R}^m \rightarrow \mathbb{R}^m$  ya que el determinante de su matriz  $(a_{ik+j})_{1 \leq i, j \leq m}$  no es nulo.

Cuando el sistema no es lineal, para que quede defina una función implícita en un entorno de  $(\mathbf{a}, \mathbf{b})$  se deberá reemplazar la condición anterior por la condición de que la matriz cuadrada  $(D_{k+i}\mathbf{g}_j(\mathbf{a}, \mathbf{b}))_{1 \leq i, j \leq m}$  tenga determinante no nulo, lo que significa que la función parcial  $\mathbf{g}_{\mathbf{a}} : \mathbf{y} \rightarrow \mathbf{g}(\mathbf{a}, \mathbf{y})$  tiene en  $\mathbf{b}$  una diferencial invertible.

De un modo heurístico e intuitivo podemos pensar así: Si  $\mathbf{g}$  es diferenciable en  $(\mathbf{a}, \mathbf{b})$  y su diferencial  $L = d\mathbf{g}(\mathbf{a}, \mathbf{b})$  tiene la propiedad de que en la ecuación  $L(\mathbf{x}, \mathbf{y}) = 0$  se puede despejar a  $\mathbf{y}$  en función  $\mathbf{x}$ , cabe esperar que ocurra lo mismo en la ecuación original  $\mathbf{g}(\mathbf{x}, \mathbf{y}) = 0$ , después de restringir esta ecuación a un entorno suficientemente pequeño de  $(\mathbf{a}, \mathbf{b})$ . Como la ecuación lineal  $L(\mathbf{x}, \mathbf{y}) = 0$  se escribe en forma de sistema lineal

$$\sum_{i=1}^n D_i g_j(\mathbf{a}, \mathbf{b}) x_i = 0; \quad 1 \leq j \leq m$$

la no anulación del determinante de la matriz  $(D_{k+i}g_j(\mathbf{a}, \mathbf{b}))_{1 \leq i, j \leq m}$  es la hipótesis que permite despejar en este sistema lineal a las variables  $(x_{k+1}, \dots, x_n)$  en función de las variables  $(x_1, x_2, \dots, x_k)$ . Esta es la hipótesis crucial que interviene en el teorema de la función implícita.

**Teorema 8.16** *Sea  $\mathbf{g} : \Omega \rightarrow \mathbb{R}^m$  una función de clase  $C^p(\Omega)$ , ( $p \geq 1$ ) definida en un abierto  $\Omega \subset \mathbb{R}^k \times \mathbb{R}^m$  y  $(\mathbf{a}, \mathbf{b}) \in \Omega$  un punto que satisface  $\mathbf{g}(\mathbf{a}, \mathbf{b}) = 0$ . Si el determinante jacobiano de las componentes de  $\mathbf{g}$  respecto a las variables  $(y_1, y_2, \dots, y_m)$  no se anula en el punto  $(\mathbf{a}, \mathbf{b})$*

$$\frac{D(g_1, g_2, \dots, g_m)}{D(y_1, y_2, \dots, y_m)}(\mathbf{a}, \mathbf{b}) \neq 0$$

entonces la ecuación  $\mathbf{g}(\mathbf{x}, \mathbf{y}) = 0$  define, en un entorno de  $(\mathbf{a}, \mathbf{b})$ , a la variable  $\mathbf{y} = (y_1, y_2, \dots, y_m)$  como función implícita de la variable  $\mathbf{x} = (x_1, x_2, \dots, x_k)$ , es decir existe  $A \times B \subset \Omega$ , entorno abierto de  $(\mathbf{a}, \mathbf{b})$ , tal que para cada  $\mathbf{x} \in A$  hay un único  $\mathbf{y} = \mathbf{f}(\mathbf{x}) \in B$  que verifica  $\mathbf{g}(\mathbf{x}, \mathbf{y}) = 0$ . La función implícita  $\mathbf{f} : A \rightarrow B$  es de clase  $C^p(A)$

DEM: La función  $\mathbf{G} : \Omega \rightarrow \mathbb{R}^k \times \mathbb{R}^m$ , definida por  $\mathbf{G}(\mathbf{x}, \mathbf{y}) = (\mathbf{x}, \mathbf{g}(\mathbf{x}, \mathbf{y}))$  es de clase  $C^p(\Omega)$  y verifica

$$\det \mathbf{G}'(\mathbf{a}, \mathbf{b}) = \frac{D(g_1, g_2, \dots, g_m)}{D(y_1, y_2, \dots, y_m)}(\mathbf{a}, \mathbf{b}) \neq 0$$

Aplicando a la función  $\mathbf{G}$  el teorema de la función inversa 8.13, se obtienen abiertos  $U, V \subset \mathbb{R}^n$  con  $(\mathbf{a}, \mathbf{b}) \in U$  y  $\mathbf{G}(\mathbf{a}, \mathbf{b}) = (\mathbf{a}, \mathbf{0}) \in V$ , tales que  $\mathbf{G}|_U : U \rightarrow V$  es una biyección con inversa  $\mathbf{F} : V \rightarrow U$  de clase  $C^p$ . No hay inconveniente suponer que  $U = A_0 \times B_0$ , donde  $A_0$  es un entorno abierto de  $\mathbf{a}$ , y  $B_0$  un entorno abierto de  $\mathbf{b}$ .

En lo que sigue, para cada  $(\mathbf{u}, \mathbf{v}) \in V$  escribimos  $\mathbf{F}(\mathbf{u}, \mathbf{v}) = (\mathbf{F}_1(\mathbf{u}, \mathbf{v}), \mathbf{F}_2(\mathbf{u}, \mathbf{v}))$  con  $\mathbf{F}_1(\mathbf{u}, \mathbf{v}) \in \mathbb{R}^k$ ,  $\mathbf{F}_2(\mathbf{u}, \mathbf{v}) \in \mathbb{R}^m$ . Como  $\mathbf{G}$  deja fijas las primeras variables  $(x_1, x_2, \dots, x_k)$ , lo mismo le ocurre a su inversa  $\mathbf{F}$ , luego  $\mathbf{F}_1(\mathbf{u}, \mathbf{v}) = \mathbf{u}$ .

Para cada  $(\mathbf{u}, \mathbf{v}) \in V$  se cumple

$$(\mathbf{u}, \mathbf{v}) = \mathbf{G}(\mathbf{F}(\mathbf{u}, \mathbf{v})) = \mathbf{G}(\mathbf{u}, \mathbf{F}_2(\mathbf{u}, \mathbf{v})) = (\mathbf{u}, \mathbf{g}(\mathbf{u}, \mathbf{F}_2(\mathbf{u}, \mathbf{v})))$$

luego  $\mathbf{g}(\mathbf{u}, \mathbf{F}_2(\mathbf{u}, \mathbf{v})) = \mathbf{v}$  para cada  $(\mathbf{u}, \mathbf{v}) \in V$ .

Entonces la función  $\mathbf{f} : A \rightarrow \mathbb{R}^m$ , definida en  $A = \{\mathbf{x} \in \mathbb{R}^k : (\mathbf{x}, \mathbf{0}) \in V\}$  por  $\mathbf{f}(\mathbf{x}) = \mathbf{F}_2(\mathbf{x}, \mathbf{0})$  cumple que  $\mathbf{g}(\mathbf{x}, \mathbf{f}(\mathbf{x})) = 0$  para todo  $\mathbf{x} \in A$ . Obsérvese que  $A$  es un entorno abierto de  $\mathbf{a}$  (porque la función  $\mathbf{x} \rightarrow (\mathbf{x}, 0)$  es continua y  $(\mathbf{a}, \mathbf{0}) \in V$ ). Además,  $A \subset A_0$  y  $\mathbf{f}(A) \subset B_0$ , ya que

$$\mathbf{x} \in A \Rightarrow (\mathbf{x}, \mathbf{0}) \in V \Rightarrow \mathbf{F}(\mathbf{x}, \mathbf{0}) = (\mathbf{x}, \mathbf{f}(\mathbf{x})) \in U = A_0 \times B_0$$

Con  $B = B_0$  se cumple que  $\mathbf{f}(A) \subset B$ , y  $\mathbf{g}(\mathbf{x}, \mathbf{f}(\mathbf{x})) = 0$  para todo  $\mathbf{x} \in A$ . Para concluir la demostración basta ver que para cada  $\mathbf{x} \in A$  hay un único  $\mathbf{y} \in B$  que verifica  $\mathbf{g}(\mathbf{x}, \mathbf{y}) = \mathbf{0}$ . Efectivamente, si  $\mathbf{y} \in B$  y  $\mathbf{g}(\mathbf{x}, \mathbf{y}) = 0$  se cumple

$$\mathbf{G}(\mathbf{x}, \mathbf{y}) = (\mathbf{x}, \mathbf{g}(\mathbf{x}, \mathbf{y})) = (\mathbf{x}, \mathbf{0}), \quad \mathbf{G}(\mathbf{x}, \mathbf{f}(\mathbf{x})) = (\mathbf{x}, \mathbf{g}(\mathbf{x}, \mathbf{f}(\mathbf{x}))) = (\mathbf{x}, \mathbf{0})$$

Entonces, teniendo en cuenta que  $\mathbf{G}|_U$  es inyectiva, y que los puntos  $(\mathbf{x}, \mathbf{y})$ ,  $(\mathbf{x}, \mathbf{f}(\mathbf{x}))$  pertenecen a  $A \times B \subset A_0 \times B_0 = U$  se concluye que  $\mathbf{y} = \mathbf{f}(\mathbf{x})$ . Finalmente  $\mathbf{f}$  es de clase  $C^p(A)$  porque  $\mathbf{F}$  es de clase  $C^p(V)$ . ■

NOTA: En el teorema 8.16 se ha supuesto para simplificar la notación, la hipótesis apropiada para obtener a las últimas  $m$  variables como funciones implícitas de las  $k$  primeras. Análogamente, si  $(i_1, i_2, \dots, i_m)$  es un subconjunto de  $\{1, 2, \dots, n\}$ , con  $m$  elementos y se supone que

$$\frac{D(g_1, g_2, \dots, g_m)}{D(x_{i_1}, x_{i_2}, \dots, x_{i_m})}(\mathbf{p}) \neq 0$$

entonces la ecuación vectorial  $\mathbf{g}(x_1, x_2, \dots, x_n) = \mathbf{0}$  define, en un entorno de  $\mathbf{p}$ , a las variables  $(x_{i_1}, x_{i_2}, \dots, x_{i_m})$  como funciones implícitas de las restantes.

### 8.3. Cálculo con funciones implícitas e inversas

**Derivadas parciales de funciones implícitas.** En las condiciones del teorema 8.16 cuando  $\mathbf{g}$  es de clase  $C^p$ , aunque no se conozca una fórmula explícita para la función implícita  $\mathbf{f}$  es posible calcular sus derivadas parciales sucesivas (hasta las de orden  $p$ ) en el punto concreto  $(\mathbf{a}, \mathbf{b})$ . Para simplificar la exposición del método consideremos el caso particular  $n = 4$  y  $k = 2$ , denotando  $(x, y, u, v)$  a las variables de la función  $\mathbf{g}$ . Si  $f_1, f_2$  son las componentes de la función implícita:  $\mathbf{f} : A \rightarrow B$  en lo que sigue resultará cómodo designarlas con la notación más flexible  $u(x, y) = f_1(x, y)$ ,  $v(x, y) = f_2(x, y)$ . Para todo  $(x, y) \in A$  se cumple

$$g_1(x, y, u(x, y), v(x, y)) = 0, \quad g_2(x, y, u(x, y), v(x, y)) = 0.$$

Derivando las dos ecuaciones respecto a la variable  $x$  y respecto a la variable  $y$  se obtienen las identidades i), ii), iii), iv).

$$\begin{array}{ll} \text{i)} & \frac{\partial g_1}{\partial x} + \frac{\partial g_1}{\partial u} u_x + \frac{\partial g_1}{\partial v} v_x = 0; \\ \text{ii)} & \frac{\partial g_2}{\partial x} + \frac{\partial g_2}{\partial u} u_x + \frac{\partial g_2}{\partial v} v_x = 0; \\ \text{iii)} & \frac{\partial g_1}{\partial y} + \frac{\partial g_1}{\partial u} u_y + \frac{\partial g_1}{\partial v} v_y = 0; \\ \text{iv)} & \frac{\partial g_2}{\partial y} + \frac{\partial g_2}{\partial u} u_y + \frac{\partial g_2}{\partial v} v_y = 0; \end{array}$$

donde las derivadas parciales de  $g_1$  y  $g_2$  se suponen evaluadas  $(x, y, u(x, y), v(x, y))$ , y las derivadas parciales  $u_x, u_y, v_x, v_y$  en el punto  $(x, y)$ . Cuando  $(x, y) = \mathbf{a}$  se tiene  $(u(\mathbf{a}), v(\mathbf{a})) = \mathbf{b}$  y si las ecuaciones i), ii) se particularizan en el punto  $\mathbf{p} = (\mathbf{a}, \mathbf{b})$  resulta

$$\begin{array}{l} \text{i')} \quad \frac{\partial g_1}{\partial x}(\mathbf{p}) + \frac{\partial g_1}{\partial u}(\mathbf{p}) u_x(\mathbf{a}) + \frac{\partial g_1}{\partial v}(\mathbf{p}) v_x(\mathbf{a}) = 0; \\ \text{ii')} \quad \frac{\partial g_2}{\partial x}(\mathbf{p}) + \frac{\partial g_2}{\partial u}(\mathbf{p}) u_x(\mathbf{a}) + \frac{\partial g_2}{\partial v}(\mathbf{p}) v_x(\mathbf{a}) = 0; \end{array}$$

Este sistema lineal de dos ecuaciones con dos incógnitas  $u_x(\mathbf{a}), v_x(\mathbf{a})$ , con determinante  $\frac{D(g_1, g_2)}{D(u, v)}(\mathbf{p}) \neq 0$ , permite calcular los valores  $u_x(\mathbf{a}), v_x(\mathbf{a})$ . Análogamente, usando las ecuaciones iii) y iv), se pueden calcular  $u_y(\mathbf{a})$  y  $v_y(\mathbf{a})$ .

Si suponemos que  $p \geq 2$  podemos seguir derivando respecto a  $x$  y respecto a  $y$  las identidades i), ii), iii) y iv). Para calcular, en el punto  $\mathbf{a}$ , las derivadas parciales segundas  $u_{xx}, u_{xy}, u_{yy}, v_{xx}, v_{xy}, v_{yy}$ , se derivan las identidades i), ii), iii) y iv) respecto a las variables  $x$  e  $y$ , se particulariza el resultado para  $(x, y) = \mathbf{a}$ ,  $(u, v) = \mathbf{b}$  y se obtienen sistemas de ecuaciones lineales que permiten calcular los valores particulares de estas derivadas segundas. Así por ejemplo, derivando i) y ii) respecto a  $x$

$$\begin{aligned} \text{a)} \quad & \frac{\partial^2 g_1}{\partial x^2} + \frac{\partial^2 g_1}{\partial u \partial x} u_x + \frac{\partial^2 g_1}{\partial v \partial x} v_x + \left[ \frac{\partial^2 g_1}{\partial x \partial u} + \frac{\partial^2 g_1}{\partial^2 u} u_x + \frac{\partial^2 g_1}{\partial v \partial u} v_x \right] u_x + \frac{\partial g_1}{\partial u} u_{xx} + \\ & + \left[ \frac{\partial^2 g_1}{\partial x \partial v} + \frac{\partial^2 g_1}{\partial u \partial v} u_x + \frac{\partial^2 g_1}{\partial^2 v} v_x \right] v_x + \frac{\partial g_1}{\partial v} v_{xx} = 0 \end{aligned}$$

$$\begin{aligned}
 \text{b) } & \frac{\partial^2 g_2}{\partial x^2} + \frac{\partial^2 g_2}{\partial u \partial x} u_x + \frac{\partial^2 g_2}{\partial v \partial x} v_x + \left[ \frac{\partial^2 g_2}{\partial x \partial u} + \frac{\partial^2 g_2}{\partial^2 u} u_x + \frac{\partial^2 g_2}{\partial v \partial u} v_x \right] u_x + \frac{\partial g_2}{\partial u} u_{xx} + \\
 & + \left[ \frac{\partial^2 g_2}{\partial x \partial v} + \frac{\partial^2 g_2}{\partial u \partial v} u_x + \frac{\partial^2 g_2}{\partial^2 v} v_x \right] v_x + \frac{\partial g_2}{\partial v} v_{xx} = 0
 \end{aligned}$$

donde las derivadas parciales de  $g_1$  y  $g_2$  están evaluadas en  $(x, y, u(x, y), v(x, y))$ , y las derivadas parciales de  $u$  y  $v$  están evaluadas en  $(x, y)$ .

Sustituyendo  $(x, y) = \mathbf{a}$ ,  $(u, v) = \mathbf{b}$ ,  $(x, y, u, v) = \mathbf{p}$ , y utilizando los valores ya calculados  $u_x(\mathbf{a})$ ,  $v_x(\mathbf{a})$  se llega a un sistema lineal de dos ecuaciones con dos incógnitas  $u_{xx}(\mathbf{a})$ ,  $v_{xx}(\mathbf{a})$ , que tiene determinante  $\frac{D(g_1, g_2)}{D(u, v)}(\mathbf{p}) \neq 0$ , y su solución proporciona las derivadas segundas  $u_{xx}(\mathbf{a})$ ,  $v_{xx}(\mathbf{a})$ .

Procediendo en forma similar, derivando las identidades i) y ii) respecto a la variable  $y$ , al particularizar el resultado en el punto  $\mathbf{p}$ , se llega a otro sistema lineal, con el mismo determinante no nulo, que permite calcular  $u_{xy}(\mathbf{a})$ ,  $v_{xy}(\mathbf{a})$  (estos valores también se pueden calcular derivando respecto a  $x$  las identidades iii) y iv), sustituyendo luego  $(x, y, u, v) = \mathbf{p}$  y resolviendo el correspondiente sistema lineal). Finalmente  $u_{yy}(\mathbf{b})$ ,  $v_{yy}(\mathbf{b})$  se calculan con el mismo método, derivando respecto a  $y$  las identidades iii) y iv).

Cuando  $p \geq 3$  se puede continuar con el procedimiento: Se deriva respecto a  $x$  y respecto a  $y$  cada una de las identidades obtenidas en la etapa anterior, se particulariza el resultado en el punto  $\mathbf{p}$  y se resuelven luego los correspondientes sistemas lineales de dos ecuaciones con dos incógnitas, cuyo determinante siempre es el mismo,  $\frac{D(g_1, g_2)}{D(u, v)}(\mathbf{p}) \neq 0$ . ■

**Derivadas parciales de funciones inversas.** En las condiciones del teorema 8.12, cuando  $\mathbf{f}$  es de clase  $C^m$ , aunque no se conozca una fórmula explícita para la función inversa  $\mathbf{g}$  también es posible calcular sus derivadas parciales sucesivas (hasta las de orden  $m$ ) en un punto concreto  $\mathbf{b} = \mathbf{f}(\mathbf{a})$ . La técnica que acabamos de exponer para calcular derivadas parciales de funciones implícitas también se puede utilizar ahora ya que una inversa local de la función  $\mathbf{y} = \mathbf{f}(\mathbf{x})$  se puede considerar como función implícita definida por la ecuación  $\mathbf{g}(\mathbf{x}, \mathbf{y}) = \mathbf{0}$ , donde  $\mathbf{g}(\mathbf{x}, \mathbf{y}) = \mathbf{f}(\mathbf{x}) - \mathbf{y}$ .

Como conviene adquirir destreza en estas técnicas de cálculo insistimos con ella en el contexto de las funciones inversas. Para simplificar la exposición lo hacemos en el caso particular  $n = 2$ .

Si  $f_1, f_2$  son las componentes de  $\mathbf{f}$ , se suele decir que las ecuaciones  $u = f_1(x, y)$ ,  $v = f_2(x, y)$  establecen una transformación de un abierto  $A$  del plano de las variables  $(x, y)$  sobre un abierto  $B$  del plano de las variables  $(u, v)$ . La función inversa  $\mathbf{g}$ , transforma cada  $(u, v) \in B$  en el punto  $(x, y) \in A$  dado por  $x = g_1(u, v)$ ,  $y = g_2(u, v)$ .

En lo que sigue designamos las componentes de  $\mathbf{g}$  con la notación más cómoda  $x(u, v) = g_1(u, v)$ ,  $y(u, v) = g_2(u, v)$ . Para todo  $(u, v) \in B$  se cumple

$$f_1(x(u, v), y(u, v)) = u, \quad f_2(x(u, v), y(u, v)) = v.$$

Derivando las dos ecuaciones respecto a la variable  $u$  y respecto a la variable  $v$  se obtienen las identidades i), ii), iii), iv).



$$\begin{array}{ll} \text{i)} & \frac{\partial f_1}{\partial x}x_u + \frac{\partial f_1}{\partial y}y_u = 1; \\ \text{ii)} & \frac{\partial f_2}{\partial x}x_u + \frac{\partial f_2}{\partial y}y_u = 0; \\ \text{iii)} & \frac{\partial f_1}{\partial x}x_v + \frac{\partial f_1}{\partial y}y_v = 0; \\ \text{iv)} & \frac{\partial f_2}{\partial x}x_v + \frac{\partial f_2}{\partial y}y_v = 1; \end{array}$$

donde las derivadas parciales de  $f_1$  y  $f_2$  se suponen evaluadas en  $(x(u, v), y(u, v))$ , y las derivadas parciales  $x_u, y_u$  en  $(u, v)$ . Sustituyendo los valores  $(u, v) = \mathbf{b}$ ,  $(x(u, v), y(u, v)) = \mathbf{a}$  las ecuaciones i), ii) se concretan en

$$\text{i')} \quad \frac{\partial f_1}{\partial x}(\mathbf{a})x_u(\mathbf{b}) + \frac{\partial f_1}{\partial y}(\mathbf{a})y_u(\mathbf{b}) = 1$$

$$\text{ii')} \quad \frac{\partial f_2}{\partial x}(\mathbf{a})x_u(\mathbf{b}) + \frac{\partial f_2}{\partial y}(\mathbf{a})y_u(\mathbf{b}) = 0$$

Este sistema lineal de dos ecuaciones con dos incógnitas  $x_u(\mathbf{b}), y_u(\mathbf{b})$ , con determinante  $\det \mathbf{f}'(\mathbf{a}) \neq 0$ , permite calcular los valores  $x_u(\mathbf{b}), y_u(\mathbf{b})$ . Análogamente, usando las ecuaciones iii) y iv), se pueden calcular  $x_v(\mathbf{b})$  y  $y_v(\mathbf{b})$ .

Si suponemos que  $m \geq 2$  podemos seguir derivando respecto a  $u$  y respecto a  $v$  las identidades i), ii), iii) y iv). Para calcular las derivadas parciales segundas  $x_{uu}(\mathbf{b}), y_{uu}(\mathbf{b})$  se derivan las identidades i) y ii) respecto a la variable  $u$  y se obtiene

$$\text{a)} \quad \left[ \frac{\partial^2 f_1}{\partial x^2}x_u + \frac{\partial^2 f_1}{\partial y \partial x}y_u \right] x_u + \frac{\partial f_1}{\partial x}x_{uu} + \left[ \frac{\partial^2 f_1}{\partial x \partial y}x_u + \frac{\partial^2 f_1}{\partial y^2}y_u \right] y_u + \frac{\partial f_1}{\partial y}y_{uu} = 0$$

$$\text{b)} \quad \left[ \frac{\partial^2 f_2}{\partial x^2}x_u + \frac{\partial^2 f_2}{\partial y \partial x}y_u \right] x_u + \frac{\partial f_2}{\partial x}x_{uu} + \left[ \frac{\partial^2 f_2}{\partial x \partial y}x_u + \frac{\partial^2 f_2}{\partial y^2}y_u \right] y_u + \frac{\partial f_2}{\partial y}y_{uu} = 0$$

donde las derivadas parciales de  $f_1$  y  $f_2$  están evaluadas en  $(x(u, v), y(u, v))$ , y las derivadas parciales  $x_u, y_u, x_{uu}, y_{uu}$  están evaluadas en  $(u, v)$ .

Sustituyendo  $(u, v) = \mathbf{b}$ ,  $(x(u, v), y(u, v)) = \mathbf{a}$ , y utilizando los valores ya calculados  $x_u(\mathbf{b}), y_u(\mathbf{b})$  se obtiene un sistema lineal de determinante  $\det \mathbf{f}'(\mathbf{a}) \neq 0$  cuya solución proporciona las derivadas segundas  $x_{uu}(\mathbf{b}), y_{uu}(\mathbf{b})$ .

Análogamente, si se derivan las identidades i) y ii) respecto a la variable  $v$ , y se concreta el resultado en el punto  $\mathbf{b}$ , resulta otro sistema lineal, con determinante  $\det \mathbf{f}'(\mathbf{a}) \neq 0$ , que permite calcular  $x_{uv}(\mathbf{b}), y_{uv}(\mathbf{b})$  (los valores  $x_{vu}(\mathbf{b}) = x_{uv}(\mathbf{b}), y_{vu}(\mathbf{b}) = y_{uv}(\mathbf{b})$  también se pueden obtener derivando respecto a  $u$  las identidades iii) y iv), sustituyendo luego  $(u, v) = \mathbf{b}$  y resolviendo el correspondiente sistema lineal).

Finalmente  $x_{vv}(\mathbf{b}), y_{vv}(\mathbf{b})$  se calculan usando el mismo método, empezando con las derivadas parciales respecto a  $v$  de las identidades iii) y iv).

Cuando  $m \geq 3$  se puede continuar con el procedimiento. Así por ejemplo, para calcular las derivadas terceras

$$x_{uuv}(\mathbf{b}) = \frac{\partial^3 x}{\partial v \partial u^2}(\mathbf{b}), \quad y_{uuv}(\mathbf{b}) = \frac{\partial^3 y}{\partial v \partial u^2}(\mathbf{b})$$

habría que derivar las identidades a), b) respecto a la variable  $v$ , sustituir en el resultado los valores, en el punto  $\mathbf{b}$ , de las funciones  $x(u, v), y(u, v)$  y de sus derivadas parciales primeras y segundas (calculados en las etapas anteriores) y resolver finalmente un sistema lineal, cuyo determinante seguirá siendo  $\det \mathbf{f}'(\mathbf{a}) \neq 0$ .

## 8.4. Cambio de variable en el cálculo diferencial

Dado un  $C^m$ -difeomorfismo  $\mathbf{P} : U \rightarrow \Omega$  entre dos abiertos  $U, \Omega \subset \mathbb{R}^n$  cada punto  $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \Omega$  es la imagen de un único  $\mathbf{u} = (u_1, u_2, \dots, u_n)$  y se suele decir que  $(u_1, u_2, \dots, u_n)$  son las coordenadas curvilíneas de  $\mathbf{x} = \mathbf{P}(\mathbf{u})$  en el sistema de coordenadas curvilíneas asociado a  $\mathbf{P}$ .

Ejemplos típicos son las coordenadas polares en el plano, y también las coordenadas cilíndricas y las coordenadas esféricas del espacio  $\mathbb{R}^3$ .

En el plano  $\mathbb{R}^2$  las *coordenadas polares* son las asociadas a  $\mathbf{P}(r, \theta) = (r \cos \theta, r \sin \theta)$ .

En el espacio  $\mathbb{R}^3$  las *coordenadas cilíndricas* son las asociadas a

$$\mathbf{P}(r, \theta, z) = (r \cos \theta, r \sin \theta, z)$$

y las *coordenadas esféricas* las asociadas a

$$\mathbf{P}(\rho, \theta, \varphi) = (\rho \cos \varphi \cos \theta, \rho \cos \varphi \sin \theta, \rho \sin \varphi)$$

En cada caso se supone  $\mathbf{P}$  definida en un abierto  $U$  sobre el que es inyectiva.

Dada una función  $f : \Omega \rightarrow \mathbb{R}$  de la variable  $\mathbf{x} \in \Omega$  si efectuamos la sustitución  $\mathbf{x} = \mathbf{P}(\mathbf{u})$  se obtiene la expresión  $F = f \circ \mathbf{P}$  de la función  $f$  en términos de las coordenadas curvilíneas  $(u_1, u_2, \dots, u_n) \in U$ . Aunque desde el punto de vista formal  $F$  y  $f$  son funciones distintas, sin embargo, desde el punto de vista de las posibles aplicaciones e interpretaciones conviene considerarlas como diferentes expresiones analíticas de la misma aplicación, que surgen al adoptar distintos sistemas de coordenadas curvilíneas para representar numéricamente los puntos de su dominio. Así, de acuerdo con este convenio se suele decir que  $F(u_1, u_2, \dots, u_n)$  es la expresión analítica de  $f$  en el sistema de coordenadas curvilíneas asociado a  $\mathbf{P}$ .

En esta situación, un problema que se plantea con frecuencia en el cálculo diferencial es el de calcular las derivadas parciales sucesivas de  $f$ , en un punto genérico, en términos de las derivadas parciales de la nueva función  $F$ . Cuando conocemos explícitamente las ecuaciones (fórmulas) para la función inversa  $\mathbf{P}^{-1}$  el cálculo se puede hacer fácilmente usando la regla de la cadena ya que  $f = F \circ \mathbf{P}^{-1}$ . Sin embargo, ocurre a menudo que no se conocen fórmulas explícitas para la inversa, o se conocen pero son engorrosas de manejar.

Para abordar este asunto exponemos a continuación, en el caso particular de los cambios de variable a coordenadas polares, una técnica sistemática que se puede aplicar, con modificaciones obvias, a otros casos.

Sea  $f(x, y)$  una función real de dos variables reales definida en un abierto  $\Omega \subset \mathbb{R}^2$ . Efectuando la sustitución  $x = r \cos \theta$ ,  $y = r \sin \theta$ , se obtiene la expresión de  $f$  en coordenadas polares:

$$F(r, \theta) = f(r \cos \theta, r \sin \theta)$$

es decir  $F = f \circ \mathbf{P}$ , donde  $\mathbf{P} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  es la transformación  $\mathbf{P}(r, \theta) = (r \cos \theta, r \sin \theta)$ .

Se supone que  $\Omega = \mathbf{P}(U)$ , donde  $U \subset \mathbb{R}^2$  es un abierto tal que  $\mathbf{g} = \mathbf{P}|_U$  es inyectiva (lo que lleva implícito que  $(0, 0) \notin \Omega$ ). Así se puede asegurar que las coordenadas polares  $(r, \theta)$  de un punto  $(x, y) \in \Omega$  quedan unívocamente determinadas

por la condición  $(r, \theta) \in U$ , lo que permite recuperar  $f$  a partir de  $F$  con el cambio de variable inverso  $(r, \theta) = \mathbf{P}^{-1}(x, y)$ .

Si suponemos que  $f$  es diferenciable en  $\Omega$ , en virtud de la regla de la cadena aplicada a la composición  $F = f \circ \mathbf{P}$ , se tendrá que las derivadas parciales  $f_x, f_y$  deben satisfacer el sistema lineal

$$F_r = \cos \theta f_x + \operatorname{sen} \theta f_y, \quad F_\theta = -r \operatorname{sen} \theta f_x + r \cos \theta f_y$$

cuyo determinante es  $r > 0$  (ya que  $(0, 0) \notin \Omega$ ). Resolviendo el sistema se obtienen

$$f_x = \cos \theta F_r - \frac{\operatorname{sen} \theta}{r} F_\theta, \quad f_y = \operatorname{sen} \theta F_r + \frac{\cos \theta}{r} F_\theta$$

Vemos así que la regla para calcular  $f_x$  consiste en aplicar el operador diferencial

$$A = \cos \theta \frac{\partial}{\partial r} - \frac{\operatorname{sen} \theta}{r} \frac{\partial}{\partial \theta}$$

a la función que resulta de expresar  $f$  en coordenadas polares. Análogamente, la regla para calcular  $f_y$  consiste en aplicar el operador diferencial

$$B = \operatorname{sen} \theta \frac{\partial}{\partial r} + \frac{\cos \theta}{r} \frac{\partial}{\partial \theta}$$

a la función que resulta de expresar  $f$  en coordenadas polares.

### **Ejemplo 8.17** *El operador de Laplace en coordenadas polares*

Se trata de expresar en coordenadas polares la laplaciana  $\Delta f = f_{xx} + f_{yy}$  de una función  $f$  que es dos veces diferenciable en un abierto  $\Omega \subset \mathbb{R}^2$ . Para obtener  $f_{xx}$  en coordenadas polares debemos aplicar el operador  $A$  a la expresión de  $f_x$  en coordenadas polares obtenida arriba, es decir

$$f_{xx} = \left[ \cos \theta \frac{\partial}{\partial r} - \frac{\operatorname{sen} \theta}{r} \frac{\partial}{\partial \theta} \right] \left( \cos \theta F_r - \frac{\operatorname{sen} \theta}{r} F_\theta \right)$$

Efectuando las operaciones indicadas se obtiene

$$f_{xx} = (\cos^2 \theta) F_{rr} + \frac{\operatorname{sen}^2 \theta}{r^2} F_{\theta\theta} - \alpha F_{\theta r} + \beta F_\theta + \frac{\operatorname{sen}^2 \theta}{r} F_r$$

donde

$$\alpha = \frac{2 \operatorname{sen} \theta \cos \theta}{r}, \quad \beta = \frac{2 \operatorname{sen} \theta \cos \theta}{r^2}.$$

Aplicando el operador  $B$  a la expresión de  $f_y$  en coordenadas polares se obtiene:

$$f_{yy} = \left[ \operatorname{sen} \theta \frac{\partial}{\partial r} + \frac{\cos \theta}{r} \frac{\partial}{\partial \theta} \right] \left( \operatorname{sen} \theta F_r + \frac{\cos \theta}{r} F_\theta \right)$$

y efectuando los cálculos se llega a

$$f_{yy} = (\operatorname{sen}^2 \theta) F_{rr} + \frac{\cos^2 \theta}{r^2} F_{\theta\theta} + \alpha F_{\theta r} - \beta F_{\theta} + \frac{\cos^2 \theta}{r} F_r$$

donde  $\alpha$  y  $\beta$  son los mismos valores que aparecieron en el cálculo de  $f_{xx}$ .

Sumando las expresiones obtenidas para  $f_{xx}$  y  $f_{yy}$  se obtiene la fórmula para el operador de Laplace  $\Delta f$  en coordenadas polares:

$$\Delta(f) = \frac{\partial^2 F}{\partial r^2} + \frac{1}{r} \frac{\partial F}{\partial r} + \frac{1}{r^2} \frac{\partial^2 F}{\partial \theta^2}$$

■

Una función  $f : \Omega \rightarrow \mathbb{R}$  que satisface la ecuación de Laplace  $\Delta f = 0$  se dice que es *armónica*. Usando el cálculo anterior es fácil ver que son armónicas las funciones que en coordenadas polares se expresan en la forma  $r^n \cos n\theta$ ,  $r^n \operatorname{sen} n\theta$ .

### Ejemplo 8.18 Ecuación de la cuerda vibrante

La ecuación en derivadas parciales

$$\frac{\partial^2 f}{\partial x^2} = \frac{1}{\alpha^2} \frac{\partial^2 f}{\partial t^2}$$

llamada ecuación de onda unidimensional fue considerada por John Bernoulli alrededor de 1727 y varios años después por Jean Le Ron d'Alembert al estudiar el movimiento de una cuerda vibrante:  $f(x, t)$  representa el desplazamiento vertical de un punto de abscisa  $x$  en el instante  $t$  de una cuerda que vibra. Con un cambio de variable lineal  $x = Au + Bv$ ,  $t = Cu + Dv$  la función  $f(x, t)$  se transforma en  $F(u, v) = f(Au + Bv, Cu + Dv)$ . Usando la regla de la cadena se obtiene

$$F_{uv} = ABf_{xx} + (AD + BC)f_{xt} + CDf_{tt}$$

La expresión anterior se simplifica eligiendo  $A, B, C, D$  de modo que se anule el coeficiente de  $f_{xt}$ . Esto se consigue con  $A = B = 1/2$ , y  $C = -D = 1/(2\alpha)$ . Así, con el cambio de variable

$$x = \frac{1}{2}(u + v), \quad y = \frac{1}{2\alpha}(u - v)$$

se obtiene

$$4F_{uv} = f_{xx} - \frac{1}{\alpha^2} \frac{\partial^2 f}{\partial t^2}$$

y la ecuación del enunciado se transforma en  $F_{uv} = 0$  cuyas soluciones son las funciones de la forma  $F(u, v) = \varphi(u) + \psi(v)$ , donde  $\varphi, \psi$  son funciones de una variable real dos veces derivables. Deshaciendo el cambio de variable con la sustitución  $u = x + \alpha t$ ,  $v = x - \alpha t$ , se llega a que las soluciones de la ecuación original son las funciones de la forma

$$f(x, t) = \varphi(x + \alpha t) + \psi(x - \alpha t)$$

## 8.5. Ejercicios resueltos

Con el siguiente ejercicio se mejora el teorema de la aplicación abierta expuesto en 8.8:

**Ejercicio 8.19** Sea  $\mathbf{f} : \Omega \rightarrow \mathbb{R}^n$  diferenciable en  $\mathbf{x} \in \Omega$  con  $\det \mathbf{f}'(\mathbf{x}) \neq 0$ , y  $\delta(\mathbf{x}) = \frac{1}{2} \|\mathbf{df}(\mathbf{x})^{-1}\|^{-1}$ .

Demuestre que existe  $\rho(\mathbf{x}) > 0$  tal que  $\overline{B(\mathbf{x}, \rho(\mathbf{x}))} \subset \Omega$  y

$$\|\mathbf{y} - \mathbf{x}\| < \rho(\mathbf{x}) \Rightarrow \|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})\| \geq \delta(\mathbf{x}) \|\mathbf{y} - \mathbf{x}\|$$

Deduzca de ello que si  $\mathbf{f} : \Omega \rightarrow \mathbb{R}^n$  es diferenciable en un abierto  $\Omega \subset \mathbb{R}^n$  y  $\det \mathbf{f}'(\mathbf{x}) \neq 0$  para todo  $\mathbf{x} \in \Omega$  entonces  $\mathbf{f}$  es abierta. Demuestre también que  $f^{-1}(\mathbf{b})$  es un conjunto discreto para cada  $\mathbf{b} \in \mathbb{R}^n$ .

SOLUCIÓN

Por la diferenciabilidad de  $f$  en  $\mathbf{x}$  existe  $\rho(\mathbf{x}) > 0$  tal que  $\overline{B(\mathbf{x}, \rho(\mathbf{x}))} \subset \Omega$  y

$$\|\mathbf{y} - \mathbf{x}\| < \rho(\mathbf{x}) \Rightarrow \|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x}) - \mathbf{df}(\mathbf{x})(\mathbf{y} - \mathbf{x})\| < \delta(\mathbf{x}) \|\mathbf{y} - \mathbf{x}\|$$

Como  $T = \mathbf{df}(\mathbf{x})$  es lineal y  $\det T \neq 0$ , dado  $\mathbf{v} = T(\mathbf{u})$  se verifica

$$\|\mathbf{u}\| \leq \|T^{-1}(\mathbf{v})\| \leq \|T^{-1}\| \|\mathbf{v}\|, \text{ luego } \|T(\mathbf{u})\| \geq \|T^{-1}\|^{-1} \|\mathbf{u}\|$$

Así se obtiene que

$$\begin{aligned} \|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})\| &\geq \|\mathbf{df}(\mathbf{x})(\mathbf{y} - \mathbf{x})\| - \|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x}) - \mathbf{df}(\mathbf{x})(\mathbf{y} - \mathbf{x})\| \geq \\ &\geq \|\mathbf{df}(\mathbf{x})^{-1}\|^{-1} \|\mathbf{y} - \mathbf{x}\| - \delta(\mathbf{x}) \|\mathbf{y} - \mathbf{x}\| = \delta(\mathbf{x}) \|\mathbf{y} - \mathbf{x}\| \end{aligned}$$

Sea  $V \subset \Omega$  abierto y  $\mathbf{a} \in V$ . Según a) cualquier bola  $\overline{B(\mathbf{a}, r)} \subset V$ , de radio  $0 < r < \rho(\mathbf{a})$  cumple la condición  $\mathbf{f}(\mathbf{a}) \notin \mathbf{f}(\partial[B(\mathbf{a}, r)])$  y con el lema 8.6 se obtiene que  $\mathbf{f}(V) \supset \mathbf{f}(B(\mathbf{a}, r))$  es entorno de  $\mathbf{f}(\mathbf{a})$ . Finalmente, si  $\mathbf{a} \in \mathbf{f}^{-1}(\mathbf{b})$ , en virtud de la implicación establecida  $0 < \|\mathbf{x} - \mathbf{a}\| < \rho(\mathbf{a}) \Rightarrow \mathbf{f}(\mathbf{x}) \neq \mathbf{b}$ , es decir,  $B(\mathbf{a}, \rho(\mathbf{a})) \cap \mathbf{f}^{-1}(\mathbf{b}) = \{\mathbf{a}\}$  luego  $\mathbf{f}^{-1}(\mathbf{b})$  es un conjunto discreto. ■

**Ejercicio 8.20** Sea  $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^n$  una aplicación diferenciable tal que existe  $C > 0$  verificando:  $\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| \geq C \|\mathbf{x} - \mathbf{y}\|$  para cada  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ .

Demuestre que  $\det \mathbf{f}'(\mathbf{x}) \neq 0$  para todo  $\mathbf{x} \in \mathbb{R}^n$  y deduzca de ello que  $\mathbf{f}(\mathbb{R}^n) = \mathbb{R}^n$ .

SOLUCIÓN

Demostraremos la primera afirmación por reducción al absurdo. Si para algún  $\mathbf{a} \in \mathbb{R}^n$  ocurriese que  $\det \mathbf{f}'(\mathbf{a}) = 0$ , entonces la diferencial  $\mathbf{df}(\mathbf{a})$  no sería inyectiva y existiría  $\mathbf{u} \in \mathbb{R}^n$ ,  $\mathbf{u} \neq 0$  con  $\mathbf{df}(\mathbf{a})\mathbf{u} = 0$ . Según la definición de diferencial

$$\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a}) = \mathbf{df}(\mathbf{a})\mathbf{h} + \|\mathbf{h}\| \rho(\mathbf{h})$$

donde  $\lim_{\mathbf{h} \rightarrow 0} \rho(\mathbf{h}) = 0$ . Como  $d\mathbf{f}(\mathbf{a})$  se anula sobre los vectores  $\mathbf{h} \in \{t\mathbf{u} : t \in \mathbb{R}\}$ , para estos vectores se cumpliría la desigualdad

$$C \|\mathbf{h}\| \leq \|\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a})\| = \|\mathbf{h}\| \|\rho(\mathbf{h})\|$$

luego  $C \leq \|\rho(t\mathbf{u})\|$ , lo que es absurdo porque  $\lim_{t \rightarrow 0} \rho(t\mathbf{u}) = 0$ .

La condición del enunciado implica que  $\mathbf{f}$  es inyectiva, y con la proposición 8.7 se obtiene que  $\mathbf{f}$  es abierta, luego  $\mathbf{f}(\mathbb{R}^n)$  es un subconjunto abierto no vacío de  $\mathbb{R}^n$ . Entonces, en virtud de la conexión de  $\mathbb{R}^n$ , basta demostrar que el conjunto  $F = \mathbf{f}(\mathbb{R}^n)$  es cerrado para concluir que  $\mathbf{f}(\mathbb{R}^n) = \mathbb{R}^n$ .

Dado  $\mathbf{y} \in \overline{F}$  existe una sucesión  $\mathbf{y}_k \in \mathbf{f}(\mathbb{R}^n)$  convergente hacia  $\mathbf{y}$ . Cada  $\mathbf{y}_k$  es de la forma  $\mathbf{y}_k = \mathbf{f}(\mathbf{x}_k)$  para algún  $\mathbf{x}_k \in \mathbb{R}^n$ . La desigualdad

$$C \|\mathbf{x}_p - \mathbf{x}_q\| \leq \|\mathbf{f}(\mathbf{x}_p) - \mathbf{f}(\mathbf{x}_q)\| = \|\mathbf{y}_p - \mathbf{y}_q\|$$

válida para cada  $p, q \in \mathbb{N}$  implica que  $(\mathbf{x}_k)$  es una sucesión de Cauchy en  $\mathbb{R}^n$  y por lo tanto convergente hacia un punto  $\mathbf{x} = \lim_k \mathbf{x}_k$ . En virtud de la continuidad

$$\mathbf{f}(\mathbf{x}) = \lim_k \mathbf{f}(\mathbf{x}_k) = \lim_k \mathbf{y}_k = \mathbf{y}$$

luego  $\mathbf{y} \in \mathbf{f}(\mathbb{R}^n)$ , y queda demostrado que  $\mathbf{f}(\mathbb{R}^n)$  es cerrado.

Un procedimiento alternativo para demostrar que  $\mathbf{f}(\mathbb{R}^n) = \mathbb{R}^n$ , sin utilizar la proposición 8.7, es el siguiente: Se comienza demostrando que  $F = \mathbf{f}(\mathbb{R}^n)$  es cerrado y luego se usa un resultado bien conocido de la topología de  $\mathbb{R}^n$  que afirma que la distancia de un punto  $\mathbf{p} \in \mathbb{R}^n$  a un cerrado  $F = \mathbf{f}(\mathbb{R}^n)$ ,

$$d(\mathbf{p}, F) = \inf\{d(\mathbf{p}, \mathbf{y}) : \mathbf{y} \in F\}$$

se alcanza en algún  $\mathbf{b} = \mathbf{f}(\mathbf{a}) \in F$  (véase el ejercicio 3.8.6 d)). Aplicando esta propiedad con la distancia euclídea  $d_2(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_2$  podemos afirmar que  $\|\mathbf{y} - \mathbf{p}\|_2^2$  alcanza en  $F$  un mínimo absoluto en algún punto  $\mathbf{b} = \mathbf{f}(\mathbf{a})$ . Esto significa que

$$g(\mathbf{x}) = \|\mathbf{f}(\mathbf{x}) - \mathbf{p}\|_2^2 = \sum_{j=1}^n (f_j(\mathbf{x}) - p_j)^2$$

alcanza un mínimo absoluto en  $\mathbf{a} \in \mathbb{R}^n$ , luego para  $k \in \{1, 2, \dots, n\}$  se cumple

$$0 = D_k g(\mathbf{a}) = 2 \sum_{j=1}^n (f_j(\mathbf{a}) - p_j) D_k f_j(\mathbf{a})$$

Es decir  $f_j(\mathbf{a}) - p_j$ ,  $1 \leq j \leq n$ , es solución del sistema lineal homogéneo cuya matriz  $D_k f_j(\mathbf{a})$  tiene determinante no nulo. La única solución de este sistema lineal es la trivial, luego  $\mathbf{p} = \mathbf{f}(\mathbf{a}) \in \mathbf{f}(\mathbb{R}^n)$ . Como  $\mathbf{p} \in \mathbb{R}^n$  era arbitrario, queda demostrado que  $\mathbf{f}(\mathbb{R}^n) = \mathbb{R}^n$ . ■

**Ejercicio 8.21** *Mediante un cambio de variable a coordenadas polares encuentre las funciones  $f : \Omega \rightarrow \mathbb{R}$ , diferenciables en  $\Omega = \mathbb{R}^2 \setminus \{(0, 0)\}$  que satisfacen la ecuación  $xf_x(x, y) + yf_y(x, y) + 2(x^2 + y^2) = 0$*

## SOLUCIÓN

Si  $F(r, \theta) = f(r \cos \theta, r \sin \theta)$ , la ecuación se transforma en  $rF_r - 2r^2 = 0$ . En los puntos de  $\Omega$  es  $r = \sqrt{x^2 + y^2} > 0$ , y la última ecuación es equivalente a  $F_r = 2r$ , luego  $F(r, \theta) = r^2 + g(\theta)$  donde  $g$  es cualquier función periódica de periodo  $2\pi$ . Se sigue que las soluciones de la ecuación propuesta son las funciones de la forma

$$f(x, y) = x^2 + y^2 + G(x, y)$$

donde  $G : \Omega \rightarrow \mathbb{R}$  es una función diferenciable que se mantiene constante sobre cada semirecta que surge de 0 (pues  $G(r \cos \theta, r \sin \theta) = g(\theta)$  sólo depende de  $\theta$ ). Obsérvese que si  $G$  tiene esta propiedad, para todo  $t > 0$  y todo  $(x, y) \in \Omega$  se cumple  $G(tx, ty) = G(x, y)$ . Derivando respecto a  $t$  se obtiene que para todo  $t > 0$  se verifica  $x D_1 G(tx, ty) + y D_2 G(tx, ty) = 0$ , y en particular,  $x D_1 G(x, y) + y D_2 G(x, y) = 0$ . Con esta igualdad, que se satisface en todo  $(x, y) \in \Omega$ , es inmediato comprobar que  $f(x, y) = x^2 + y^2 + G(x, y)$  satisface la ecuación propuesta. ■

**Ejercicio 8.22** Sea  $A = \{(u, v) : v > 0\}$ ,  $B = \{(x, y) : y > 0\}$  y  $\mathbf{g} : A \rightarrow B$  la aplicación definida por  $\mathbf{g}(u, v) = (x, y)$  donde  $x = (u^2 + v^2)/2$ ,  $y = u/v$ . Compruebe que  $\mathbf{g}$  establece un  $C^\infty$ -difeomorfismo entre  $A$  y  $B$  y obtenga las ecuaciones de la transformación inversa  $\mathbf{g}^{-1} : B \rightarrow A$ . Utilice el cambio de variable  $(x, y) = \mathbf{g}(u, v)$  para encontrar las funciones diferenciables  $f : B \rightarrow \mathbb{R}$  que satisfacen la ecuación  $2xf_x - y(1 + y^2)f_y = 0$ .

## SOLUCIÓN

$\mathbf{g}$  es inyectiva sobre  $A$ : Si  $\mathbf{g}(u, v) = \mathbf{g}(s, t)$  con  $v > 0$  y  $t > 0$  se cumple,  $u/v = s/t$ ,  $u^2 + v^2 = s^2 + t^2$ . Si  $c = u/v$  resulta  $(1 + c^2)v^2 = (1 + c^2)t^2$ , luego  $v = t$ , y  $u = s$ .  
 $\mathbf{g}(A) = B$ : Basta ver que para cada  $(x, y) \in B$  el sistema de dos ecuaciones  $u^2 + v^2 = 2x$ ,  $u/v = y$  tiene una única solución  $(u, v) \in A$ . Como  $2x = v^2(1 + y^2)$  resulta

$$v = \sqrt{\frac{2x}{1 + y^2}}, \quad u = y \sqrt{\frac{2x}{1 + y^2}}.$$

Como  $\mathbf{g}$  es de clase  $C^\infty$  y en todo  $(u, v) \in A$  es  $\det \mathbf{g}'(u, v) = -(u^2 + 1)/v^2 \neq 0$ , en virtud del corolario 8.14 podemos asegurar que  $\mathbf{g}$  es un  $C^\infty$ -difeomorfismo.

Con el cambio de variable propuesto  $f$  se transforma en  $F(u, v) = f(\mathbf{g}(u, v))$  y usando la regla de la cadena para el cálculo de las derivadas parciales

$$F_u = f_x x_u + f_y y_u = u f_x + \frac{1}{v} f_y; \quad F_v = f_x x_v + f_y y_v = v f_x - \frac{u}{v^2} f_y.$$

Resolviendo el sistema se obtiene

$$f_x = \frac{u F_u + v F_v}{u^2 + v^2}, \quad f_y = \frac{v F_u - u F_v}{u^2 + v^2} v^2$$

luego

$$2x f_x = u F_u + v F_v; \quad y(1 + y^2) f_y = \frac{u}{v} \left( 1 + \frac{u^2}{v^2} \right) f_y = u F_u - \frac{u^2}{v} F_v;$$

Restando miembro estas igualdades

$$2xf_x - y(1 + y^2)f_y = \frac{u^2 + v^2}{v}F_v$$

Obtenemos así que la ecuación del enunciado equivale a  $F_v = 0$  cuyas soluciones son las funciones de la forma  $F(u, v) = \varphi(u)$  donde  $\varphi$  es una función derivable de una variable real. Des haciendo el cambio de variable se concluye que las soluciones de la ecuación propuesta son las funciones de la forma

$$f(x, y) = \varphi\left(y\sqrt{\frac{2x}{y^2 + 1}}\right)$$

donde  $\varphi$  es una función derivable. Se deja al cuidado del lector la comprobación de que las funciones de este tipo son soluciones de la ecuación propuesta. ■

**Ejercicio 8.23** Sea  $g : \mathbb{R}^3 \rightarrow \mathbb{R}$  una función de clase  $C^2$  tal que  $g(1, 1, 0) = 0$  y  $\nabla g(1, 1, 0) = (2, 0, 0)$ , y sea  $x = \varphi(y, z)$  la función implícita de clase  $C^2$  definida por la ecuación  $g(x, y, z) = 0$  en un entorno del punto  $(1, 1, 0)$ . Compruebe que  $(1, 0)$  es un punto estacionario de la función  $F(y, z) = \varphi(y, z)^2 + y^2z^2$  y obtenga valores de  $A = D_{11}\varphi(1, 0)$ ,  $B = D_{12}\varphi(1, 0)$  y  $C = D_{22}\varphi(1, 0)$  para los que se pueda asegurar que  $F$  presenta en  $(1, 0)$  un mínimo relativo.

SOLUCIÓN

La función implícita  $x = \varphi(y, z)$  está definida en un entorno de  $(1, 0)$  y verifica  $\varphi(1, 0) = 1$ . Comprobemos en primer lugar que  $(1, 0)$  es un punto estacionario de  $F$ : Derivando respecto a las variables  $y, z$ , en la identidad  $g(\varphi(y, z), y, z) \equiv 0$  se obtiene

$$D_1g(\varphi(y, z), y, z)\varphi_y(y, z) + D_2g(\varphi(y, z), y, z) \equiv 0;$$

$$D_1g(\varphi(y, z), y, z)\varphi_z(y, z) + D_3g(\varphi(y, z), y, z) \equiv 0$$

Sustituyendo  $y = 1, z = 0, x = \varphi(1, 0) = 1$ , y teniendo en cuenta que  $D_1g(1, 1, 0) = 2, D_2g(1, 1, 0) = 0, D_3g(1, 1, 0) = 0$ , resulta  $\varphi_y(1, 0) = \varphi_z(1, 0) = 0$ . Ahora podemos calcular las derivadas parciales primeras y segundas de  $F$ :

$$F_y(y, z) = 2\varphi(y, z)\varphi_y(y, z) + 2yz^2; \quad F_z(y, z) = 2\varphi(y, z)\varphi_z(y, z) + 2y^2z; \quad [*]$$

Sustituyendo  $y = 1, z = 0, \varphi(1, 0) = 1, \varphi_y(1, 0) = \varphi_z(1, 0) = 0$ , se obtiene que  $F_y(1, 0) = 0, F_z(1, 0) = 0$ , luego  $(1, 0)$  es un punto estacionario de  $F$ .

Volviendo a derivar en [\*] respecto a las variables  $y, z$  se llega a

$$F_{yy}(y, z) = 2\varphi_y(y, z)^2 + 2\varphi(y, z)\varphi_{yy}(y, z) + 2z^2 = 0;$$

$$F_{zy}(y, z) = 2\varphi_y(y, z)\varphi_z(y, z) + 2\varphi(y, z)\varphi_{zy}(y, z) + 4yz = 0$$

$$F_{zz}(y, z) = 2\varphi_z(y, z)^2 + 2\varphi(y, z)\varphi_{zz}(y, z) + 2y^2 = 0;$$

Sustituyendo  $y = 1, z = 0, \varphi(1, 0) = 1, \varphi_y(1, 0) = \varphi_z(1, 0) = 0$ , se llega a

$$F_{yy}(1, 0) = 2A, \quad F_{yz}(1, 0) = 2B, \quad F_{zz}(1, 0) = 2C + 2$$



luego  $F$  presentará un mínimo relativo en el punto  $(1, 0)$  cuando se verifique

$$A > 0; \quad \text{y} \quad \begin{vmatrix} 2A & 2B \\ 2B & 2C + 2 \end{vmatrix} = 4[A(1 + C) - B^2] > 0$$

■

**Ejercicio 8.24** [Unicidad de funciones implícitas] Sea  $\mathbf{g} : \Omega \rightarrow \mathbb{R}^m$  continua en un abierto  $\Omega \subset \mathbb{R}^k \times \mathbb{R}^m$ , tal que  $S = \{(\mathbf{x}, \mathbf{y}) \in \Omega : \mathbf{g}(\mathbf{x}, \mathbf{y}) = 0\}$  no es vacío y cada  $(\mathbf{x}_0, \mathbf{y}_0) \in S$  posee un entorno abierto  $U \times V$  donde la ecuación  $\mathbf{g}(\mathbf{x}, \mathbf{y}) = 0$  define una función implícita  $\mathbf{f} : U \rightarrow V$ .

Se supone que en un abierto conexo  $G \subset \mathbb{R}^k$  hay definidas dos funciones implícitas continuas  $\varphi, \Psi : G \rightarrow \mathbb{R}^m$ , (es decir, dos funciones continuas con gráfica contenida en  $S$ ). Si  $\mathbf{a} \in G$  y  $\Psi(\mathbf{a}) = \varphi(\mathbf{a})$  demuestre que  $\Psi = \varphi$ . Muestre con un ejemplo que el resultado es falso cuando  $G$  no es conexo.

SOLUCIÓN

El conjunto  $G_0 = \{\mathbf{x} \in G : \varphi(\mathbf{x}) = \Psi(\mathbf{x})\}$  no es vacío, pues  $\mathbf{a} \in G_0$ . Como  $G$  es conexo basta demostrar que  $G_0$  es un subconjunto abierto y cerrado de  $G$ , con su topología relativa, para concluir que  $G = G_0$ . En virtud de la continuidad de  $\varphi$  y  $\Psi$ , el conjunto  $G_0$  es cerrado relativo a  $G$ . Veamos que también es abierto:

Por hipótesis, para cada  $\mathbf{x}_0 \in G_0$  el punto  $(\mathbf{x}_0, \varphi(\mathbf{x}_0)) \in S$  posee un entorno  $U \times V$  donde la ecuación  $\mathbf{g}(\mathbf{x}, \mathbf{y}) = 0$  define una función implícita  $\mathbf{f} : U \rightarrow V$ .

Como  $V$  es un entorno de  $\varphi(\mathbf{x}_0) = \Psi(\mathbf{x}_0)$ , y las funciones  $\varphi, \Psi$  son continuas en  $\mathbf{x}_0$ , podemos suponer que el entorno  $U$  de  $\mathbf{x}_0$  se ha elegido de modo que  $\varphi(U) \subset V$  y  $\Psi(U) \subset V$ . Necesariamente  $\varphi|_U$  y  $\Psi|_U$  coinciden en  $U$  con la función implícita  $\mathbf{f}$ , luego  $U$  es un entorno abierto de  $\mathbf{x}_0$  contenido en  $G_0$ .

El siguiente ejemplo muestra que el resultado es falso cuando  $G_0$  no es conexo:  $g(x, y) = x^2 + y^2 - 1$ ,  $G = (-1, 0) \cup (0, 1)$ , Es claro que las funciones  $\varphi, \psi : G \rightarrow \mathbb{R}$  definidas por

$$- \varphi(x) = \sqrt{1 - x^2} \quad \text{si } x \in G.$$

$$- \psi(x) = \sqrt{1 - x^2} \quad \text{si } -1 < x < 0, \quad \psi(x) = -\sqrt{1 - x^2} \quad \text{si } 0 < x < 1,$$

son distintas y ambas satisfacen las condiciones del enunciado. ■

## 8.6. Ejercicios propuestos

◇ **8.6.1** Sea  $f : \Omega \rightarrow \mathbb{R}^n$  una aplicación diferenciable definida en un abierto convexo  $\Omega \subset \mathbb{R}^n$ . Demuestre que la condición

$$\sum_{i,j=1}^n D_i f_j(\mathbf{x}) h_i h_j > 0 \quad \text{para todo } \mathbf{x} \in \Omega \quad \text{y todo } \mathbf{h} \in \mathbb{R}^n \setminus \{0\}$$

implica que  $f$  es inyectiva.

(Indic: Si  $f(\mathbf{x}) = f(\mathbf{y})$  con  $\mathbf{h} = \mathbf{y} - \mathbf{x} \neq 0$ , aplique el teorema del valor medio a  $\varphi(t) = \langle f(\mathbf{x} + t\mathbf{h}) - f(\mathbf{x}) \mid \mathbf{h} \rangle$ , en  $[0, 1]$ ).

◇ **8.6.2** Demuestre que la función  $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ , definida por

$$f(x, y, z) = (z \cos xy, z \sin xy, x + z)$$

admite inversa local en un entorno abierto  $A$  de  $\mathbf{a} = (1, 0, 1)$ . Si  $\mathbf{g} = (f|_A)^{-1}$  y  $\mathbf{b} = f(\mathbf{a})$ , obtenga  $d\mathbf{g}(\mathbf{b})$ .

◇ **8.6.3** Demuestre que la aplicación  $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $f(x, y, z) = (x^3, y^2 - z^2, yz)$  es abierta pero no es inyectiva.

◇ **8.6.4** Estudie los puntos donde es localmente inyectiva la función  $\mathbf{g} : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  definida por  $\mathbf{g}(x, y) = (x + y, (x + y)^2, (x + y)^3)$ .

◇ **8.6.5** Para las transformaciones consideradas en los ejercicios 1.5.6, 1.5.7, 1.5.8, 1.5.9 estudie su comportamiento local y global: Sobre qué abiertos son inyectivas, abiertas, difeomorfismos..)

◇ **8.6.6** Sea  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  definida por  $f(x, y) = (x + h(y), y + h(x))$  donde  $h : \mathbb{R} \rightarrow \mathbb{R}$  es derivable con derivada continua y  $|h'(t)| \leq c < 1$  para todo  $t \in \mathbb{R}$ . Demuestre que  $f(\mathbb{R}^2) = \mathbb{R}^2$  y que  $f$  establece un  $C^1$ -difeomorfismo de  $\mathbb{R}^2$  sobre  $\mathbb{R}^2$ .

◇ **8.6.7** Demuestre que  $\mathbf{g}(x, y) = (x - a \cos y, y - a \cos x)$ , con  $0 < a < 1$ , es inyectiva y abierta.

◇ **8.6.8** Sea  $f : A \rightarrow \mathbb{R}^k$  de clase  $C^2$  en un abierto  $A \subset \mathbb{R}^n$ , con  $n \geq k$ , tal que para cada  $\mathbf{x} \in A$  los vectores  $\nabla f_1(\mathbf{x}), \nabla f_2(\mathbf{x}), \dots, \nabla f_k(\mathbf{x})$  son linealmente independientes y sea  $\varphi : B \rightarrow \mathbb{R}$  definida y continua en  $B = f(A)$ .

Demuestre que  $B$  es abierto y que  $\varphi$  presenta un extremo relativo en  $\mathbf{b} = f(\mathbf{a})$  si y sólo si  $F = \varphi \circ f$  presenta un extremo relativo del mismo tipo en  $\mathbf{a} \in A$ .

◇ **8.6.9** Sea  $\mathbf{F}(x) = \nabla f(x)$  donde  $f : \Omega \rightarrow \mathbb{R}$  es de clase  $C^k$  ( $k \geq 2$ ) en un abierto  $\Omega \subset \mathbb{R}^n$ .

a) Dado  $\mathbf{a} \in \Omega$ , obtenga una condición (C) que garantice que  $\mathbf{F}$  posee una inversa local  $\mathbf{G}$ , de clase  $C^1$ , definida en un entorno abierto de  $\mathbf{b} = \mathbf{F}(\mathbf{a})$ .

b) Obtenga  $dh(\mathbf{x})$ , donde  $h(\mathbf{x}) = \langle \mathbf{x} | \mathbf{F}(\mathbf{x}) \rangle - f(\mathbf{x})$ .

c) Si se cumple (C) demuestre que  $\mathbf{G}$  se deduce de  $\mathbf{g} = h \circ \mathbf{G} : B \rightarrow \mathbb{R}^n$  de la misma forma que  $\mathbf{F}$  se deduce de  $f$ .

◇ **8.6.10** Sea  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  de clase  $C^2$  de la que se conoce  $f(0,0) = 0$ ,  $D_1f(0,0) = 0$ ,  $D_2f(0,0) = 1$ ,  $D_{11}f(0,0) = 1$ ,  $D_{12}f(0,0) = 1$ ,  $D_{22}f(0,0) = 2$ .

Demuestre que la ecuación  $\int_0^{f(x,y)} e^{t^2} dt = 0$  define, en un entorno de  $(0,0)$ , a la variable  $y$  como función implícita de la variable  $x$ .

◇ **8.6.11** Compruebe que la siguiente función es de clase  $C^1$ .

$$f(x,y) = xy \log(x^2 + y^2) + y \quad \text{si } (x,y) \neq (0,0), \quad f(0,0) = 0$$

Demuestre que existe un entorno abierto  $U$  de  $(1,0)$  y una función  $g : U \rightarrow \mathbb{R}$ , de clase  $C^1(U)$ , tal que  $g(1,0) = 2$  y

$$f(xy, g(x,y) - 2x) = 0; \quad xD_1g(x,y) - yD_2g(x,y) = 2x \quad \text{para todo } (x,y) \in U$$

◇ **8.6.12** Sea  $g : \mathbb{R}^2 \rightarrow \mathbb{R}$  de clase  $C^2$  tal que  $g(b/a, c/a) = 0$  y  $D_2g(b/a, c/a) \neq 0$ , ( $a \neq 0$ ). Demuestre que existe un entorno abierto  $A \subset \mathbb{R}^2$  de  $(a,b)$ , donde hay definida una única función  $f : A \rightarrow \mathbb{R}$ , de clase  $C^2(A)$ , verificando:

$$f(a,b) = c, \quad \text{y } g(y/x, f(x,y)/x) = 0 \quad \text{para todo } (x,y) \in A.$$

Demuestre que  $f$  satisface la ecuación:  $x D_1 f(x,y) + y D_2 f(x,y) = f(x,y)$ .

Con un cambio de variable a coordenadas polares obtenga que  $f(x,y)/\sqrt{x^2 + y^2}$  sólo depende de  $y/x$ .

◇ **8.6.13** Sea  $\psi : \mathbb{R} \rightarrow \mathbb{R}$  una función de clase  $C^1$  tal que  $\psi(0) = 0$  y  $\psi'(0) = 1$ . Dados tres números reales  $a, b$  y  $c \neq 0$ , demuestre que existe un entorno  $U \subset \mathbb{R}^2$  de  $(0,0)$  y una función  $g : U \rightarrow \mathbb{R}$  de clase  $C^1$  que para todo  $(x,y) \in U$  se cumplen las dos condiciones siguientes:

$$a) \quad x^2 + y^2 + (g(x,y))^2 = \psi(ax + by + cg(x,y))$$

$$b) \quad (cy - bg(x,y))D_1g(x,y) + (ag(x,y) - cx)D_2g(x,y) = bx - ay$$

Calcule  $D_1g(0,0)$  y  $D_2g(0,0)$ .

◇ **8.6.14** Compruebe que la ecuación  $z^3 + 2z + e^{z-x-y^2} + x + y^2 - \cos(x-y+z) = 0$  define, en un entorno de  $(0,0,0)$ , a la variable  $z$  como función implícita de las variables  $x, y$ . Demuestre que la función implícita  $z = f(x,y)$ , definida en un entorno de  $(0,0)$ , presenta un máximo relativo en  $(0,0)$ .

◇ **8.6.15** Compruebe que la ecuación  $xyz + \sin(z-6) - 2(x+y+x^2y^2) = 0$  define, en un entorno de  $(1,1,6)$ , una función implícita  $z = f(x,y)$  que presenta un mínimo relativo en  $(1,1)$ .

◇ **8.6.16** Compruebe que la ecuación  $\sin z + (1+x^2)^y + z + y^2 - 2y = 0$  define, en un entorno de  $(0,1,0)$ , una función implícita  $z = f(x,y)$  que verifica  $f(0,1) = 1$ . Estudie si  $f$  presenta un extremo local en  $(0,1)$ .

◇ **8.6.17** Compruebe que la ecuación  $x^3 + y^3 + z^3 + x + y - z = 4$  define, en un entorno de  $\mathbf{p} = (1, 1, 1)$ , una función implícita  $z = f(x, y)$  con la propiedad de que en un entorno de  $\mathbf{a} = (1, 1)$  la gráfica de  $f$  queda por debajo de su plano tangente en  $\mathbf{p}$ .

◇ **8.6.18** Justifique que la ecuación  $x^3 + 3z^2 + 8xy^2 - 3y^3z = 2$  define en un entorno de  $\mathbf{p} = (-1, 0, 1)$  una función implícita  $z = f(x, y)$ . Demuestre que en un entorno de  $\mathbf{a} = (-1, 0)$  la gráfica de  $f$  queda por encima de su plano tangente en  $\mathbf{p}$ .

◇ **8.6.19** Sea  $g(x, y, z) = x^3 + y^3 + z^3 + x^2 + y^2 + z^2 + x + y - z$ . Compruebe que en un entorno de  $\mathbf{p} = (1, 1, 1)$ , la ecuación  $g(x, y, z) = 7$  define, una función implícita  $z = f(x, y)$ . Obtenga el polinomio de Taylor de grado 2, de  $f$  en el punto  $(1, 1)$ .

Si  $M = \{(x, y, z) \in \mathbb{R}^3 : g(x, y, z) = 7\}$  demuestre que  $\mathbf{p}$  posee un entorno  $V_{\mathbf{p}}$  tal que  $M \cap V_{\mathbf{p}}$  queda a un lado del plano tangente a  $M$  en  $\mathbf{p}$ .

◇ **8.6.20** Compruebe que las ecuaciones  $x \cos y + y \cos z + z \cos x = \pi$ ;  $x^2 + y^2 + z^2 - xy = \pi^2$  definen, en un entorno de  $(x_0, y_0, z_0) = (0, 0, \pi)$ , a las variables  $y, z$  como funciones implícitas de  $x$ :  $(y, z) = (f_1(x), f_2(x))$ . Calcule las derivadas  $f'_i(0)$ ,  $f''_i(0)$ ,  $i = 1, 2$ .

◇ **8.6.21** Compruebe que las ecuaciones

$$yt + e^{xt} + e^{xy} = 3; \quad y^x - y + t = 1$$

definen en un entorno de  $(x_0, y_0, t_0) = (0, 1, 1)$  a las variables  $x, y$  como funciones implícitas de la variable  $t$ . Si  $x = f(t)$  y  $y = g(t)$  son las funciones implícitas, calcule las derivadas  $f'(1), g'(1), f''(1), g''(1)$ .

◇ **8.6.22** Compruebe que en un entorno de  $(x_0, y_0, u_0, v_0) = (\sqrt{2}, 0, 1, -1)$ , el sistema de ecuaciones

$$2uv + x^2 - y^2 = 0, \quad u^2 - v^2 + 2xy = 0$$

define a las variables  $(u, v)$  como funciones implícitas de las variables  $(x, y)$ .

Demuestre que la función implícita  $(u, v) = \mathbf{f}(x, y)$  es invertible en un entorno  $U$  de  $(\sqrt{2}, 0)$ . Si  $\mathbf{g}$  es la inversa de  $\mathbf{f}|_U$ , obtenga  $d\mathbf{g}((1, -1))$ .

◇ **8.6.23** Compruebe que en un entorno de  $(1, 1, 1, 1)$  el sistema de ecuaciones

$$x^2 + y^2 - u - v^2 = 0, \quad x^2 - y^2 + u^3 - v^3 = 0$$

define funciones implícitas  $(u, v) = \mathbf{f}(x, y)$ ,  $(x, y) = \mathbf{g}(u, v)$ . Explique la relación que hay entre las matrices  $\mathbf{f}'(1, 1)$  y  $\mathbf{g}'(1, 1)$  y calcule una de ellas.

◇ **8.6.24** Determine los valores de  $a \in \mathbb{R}$  para los que el sistema

$$xz^3 + yu + ax = 1, \quad 2xy^3 + u^2z + ay = a$$

define, en un entorno de  $(0, 1)$  una función implícita  $(x, y) = \varphi(z, u)$  que cumple  $\varphi(0, 1) = (0, 1)$  y estudie si  $\varphi$  es localmente invertible en  $(0, 1)$ .

◇ **8.6.25** Compruebe que las ecuaciones

$$xz^3 + y^2u^3 = 1; \quad 2xy^3 + u^2z = 0$$

definen, en un entorno de  $(x_0, y_0, z_0, u_0) = (0, 1, 0, 1)$ , a las variables  $x, y$  como funciones implícitas de las variables  $z, u$ . Si  $x = g_1(z, u)$ ,  $y = g_2(z, u)$  son las funciones implícitas, definidas en un entorno de  $(0, 1)$ , demuestre que existe un entorno abierto de  $(0, 1)$ , donde  $g = (g_1, g_2)$  es invertible con inversa indefinidamente derivable.

◇ **8.6.26** Obtenga el gradiente de una función diferenciable  $f : \mathbb{R}^3 \rightarrow \mathbb{R}$  en términos de las coordenadas cilíndricas,  $x = r \cos \theta$ ,  $y = r \sin \theta$ ,  $z = t$ , y de las coordenadas esféricas,  $x = \rho \cos \varphi \cos \theta$ ,  $y = \rho \cos \varphi \sin \theta$ ,  $z = \rho \sin \varphi$ .

◇ **8.6.27** Efectuando el cambio de variable  $x = (u+v)/2$ ,  $y = (u-v)/2$  obtenga la forma general de las funciones de clase  $C^2$ ,  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ , que satisfacen la ecuación en derivadas parciales

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} - 2 \frac{\partial^2 f}{\partial x \partial y} = 0$$

◇ **8.6.28** Utilizando el cambio de variable  $x = v$ ,  $y = uv$  obtenga las funciones  $f$  de clase  $C^2$  en  $\Omega = \{(x, y) : x > 0\}$  que verifican

$$x^2 D_{11}f(x, y) + 2xy D_{12}f(x, y) + y^2 D_{22}f(x, y) = 0, \quad \text{para todo } (x, y) \in \Omega$$

◇ **8.6.29** Utilizando el cambio de variable  $x = (u^2 + v^2)/2$ ,  $y = u/v$  obtenga las funciones  $f : \Omega \rightarrow \mathbb{R}$  de clase  $C^2$  en  $\Omega = \{(x, y) : x > 0\}$  que verifican

$$2xy D_1f(x, y) + (1 + y^2) D_2f(x, y) = 0 \quad \text{para todo } (x, y) \in \Omega$$

◇ **8.6.30** Usando el cambio de variable  $x = u^2 + v^2$ ,  $y = u + v$  obtenga las funciones  $f : \Omega \rightarrow \mathbb{R}$  de clase  $C^2$  en  $\Omega = \{(x, y) : 2x > y^2\}$  que verifican

$$2(y^2 - x^2) D_{11}f(x, y) + 2y D_{12}f(x, y) + D_{22}f(x, y) = y^2 - x^2 \quad \text{para todo } (x, y) \in \Omega$$

◇ **8.6.31** Dado el abierto  $A = \{(x, y) : 0 < x < y\}$ , justifique la existencia de un abierto  $B \subset \mathbb{R}^2$  tal que para cada  $f \in C^1(A)$  existe una única  $F \in C^1(B)$  que verifica :  $f(x, y) = F(x + y, xy)$  para todo  $(x, y) \in A$ .

Demuestre que son equivalentes

- i)  $D_1f(x, y) - D_2f(x, y) + 3(x - y)f(x, y) = 0$  para todo  $(x, y) \in A$ .
- ii)  $D_2F(u, v) = 3F(u, v)$  para todo  $(u, v) \in B$ .

# Capítulo 9

## Extremos condicionados

*Subvariedades diferenciables de  $\mathbb{R}^n$ . Espacio tangente en un punto. Extremos condicionados: Método de los multiplicadores de Lagrange. Aplicaciones geométricas.*

El objetivo central del capítulo es la optimización de funciones reales de varias variables reales,  $f(x_1, x_2, \dots, x_n)$ , cuando las variables están sometidas a ligaduras:

$$g_1(x_1, x_2, \dots, x_n) = 0, \quad g_2(x_1, x_2, \dots, x_n) = 0, \dots, \quad g_m(x_1, x_2, \dots, x_n) = 0$$

En otras palabras, se trata de calcular los extremos absolutos o relativos (si existen) de  $f$  sobre el subconjunto  $M \subset \mathbb{R}^n$  formado por los puntos que cumplen las condiciones de ligadura. Cuando este conjunto tiene una estructura geométrica apropiada, que se formula mediante la noción de subvariedad diferenciable, el problema se puede abordar con el método de los multiplicadores de Lagrange.

Por ello el capítulo comienza con el estudio de esta noción geométrica que proporciona el marco natural para los problemas de optimización con restricciones de ligadura en forma de igualdad. Después de estudiar estos problemas se pueden abordar los de optimización con restricciones en forma de desigualdad, de los que no haremos el estudio sistemático (con las condiciones de Khun-Tucker) que se puede encontrar en textos más especializados como [10].

Cuando  $M \subset \mathbb{R}^n$  sea un subconjunto definido mediante restricciones de desigualdad, el problema de obtener los extremos absolutos de  $f|_M$  (que existirán con seguridad cuando  $f$  sea continua y  $M$  sea compacto) se puede abordar considerando por separado la restricción de  $M$  al interior de  $M$ , y a su frontera. Lo primero conduce a un problema ordinario de extremos sin ligaduras, que ya han sido considerados en el capítulo 5. La restricción de  $f$  a la frontera de  $M$  puede conducir a varios problemas de optimización con ligaduras en forma de igualdad: Generalmente la frontera de  $M$  no es será subvariedad diferenciable pero habitualmente, cuando  $M$  está definido con un número finito de desigualdades, su frontera se puede descomponer en un número finito de subvariedades diferenciables (de diferentes dimensiones). Entonces la optimización de  $f$  sobre la frontera de  $M$  se podrá atacar con el método de los multiplicadores de Lagrange sobre cada una de las subvariedades diferenciables que componen la frontera.

## 9.1. Subvariedades diferenciables

En esta sección se introduce la noción de subvariedad diferenciable de  $\mathbb{R}^n$ , de dimensión  $k$ , como un subconjunto  $M \subset \mathbb{R}^n$  que localmente tiene una estructura geométrica similar a  $\mathbb{R}^k$ . Esta estructura geométrica local se puede describir bajo tres formas equivalentes (teorema 9.4) que comenzamos definiendo de manera precisa.

Una de ellas se formula en términos de la gráfica de una función diferenciable de  $k$  variables independientes. En las gráficas de este tipo, según la costumbre habitual, las primeras  $k$  variables son las independientes que desempeñan un papel especial. Esta restricción artificial se elimina en la siguiente definición considerando cambios de orden en las variables, es decir cambios de variable lineales  $T_\sigma : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,

$$T_\sigma(x_1, x_2, \dots, x_n) = (x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(n)})$$

asociados a permutaciones  $\sigma : \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, n\}$ .

**Definición 9.1** Diremos que  $M \subset \mathbb{R}^n$  admite una representación explícita de clase  $C^m$  y dimensión  $k < n$ , o que  $M$  es una gráfica de esa clase y dimensión, si existe una permutación  $\sigma : \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, n\}$  tal que  $M = T_\sigma(G(\mathbf{f}))$  donde

$$G(\mathbf{f}) = \{(\mathbf{u}, \mathbf{f}(\mathbf{u})) : \mathbf{u} \in A\}$$

y  $\mathbf{f} : A \rightarrow \mathbb{R}^{n-k}$  es de clase  $C^m$  en un abierto  $A \subset \mathbb{R}^k$ .

Otra caracterización de las subvariedades diferenciables de  $\mathbb{R}^n$  se expresará en términos de parametrizaciones regulares, que se definen a continuación

**Definición 9.2** Si  $\varphi : U \rightarrow \mathbb{R}^n$  es una aplicación de clase  $C^m$  ( $m \geq 1$ ) definida en un abierto  $U \subset \mathbb{R}^k$ , con  $1 \leq k \leq n$ , se dice que  $\varphi$  es una parametrización de clase  $C^m$  y dimensión  $k$ . Si además se cumplen las condiciones

- i)  $\varphi$  es un homeomorfismo entre  $U$  y su imagen  $\varphi(U)$  (con la topología relativa).
- ii) Para cada  $\mathbf{u} \in U$ , los vectores  $D_1\varphi(\mathbf{u}), D_2\varphi(\mathbf{u}), \dots, D_k\varphi(\mathbf{u})$  son linealmente independientes.

se dice que  $\varphi$  es una parametrización regular. En este caso, si  $M = \varphi(U)$  diremos que  $\varphi$  es una parametrización regular de  $M$  y también que  $M \subset \mathbb{R}^n$  admite una parametrización regular (de clase  $C^m$  y dimensión  $k \leq n$ ).

Toda representación explícita (de clase  $C^m$  y dimensión  $k < n$ ) de  $M \subset \mathbb{R}^n$  lleva asociada de modo natural una parametrización regular de  $M$  de la misma clase y dimensión,  $\varphi(\mathbf{u}) = T_\sigma((\mathbf{u}, \mathbf{f}(\mathbf{u})))$ , definida en  $U = A$ , es decir:

$$\varphi(u_1, u_2, \dots, u_k) = T_\sigma(u_1, u_2, \dots, u_k, f_1(u_1, \dots, u_k), \dots, f_{n-k}(u_1, \dots, u_k))$$

Dejamos como ejercicio al cuidado del lector la comprobación de que se cumplen las condiciones requeridas en la definición 9.2.

Un ejemplo sencillo de parametrización regular lo proporciona la parametrización habitual de un trozo de esfera mediante la longitud y la latitud. Esto se puede ver en H.7 donde también se muestra un ejemplo interesante de lo que puede ocurrir cuando una parametrización no es regular aunque sea inyectiva.

**Definición 9.3** Diremos que  $M \subset \mathbb{R}^n$  admite una representación implícita de clase  $C^m$  y dimensión  $k < n$  si se puede representar en la forma

$$M = \{\mathbf{x} \in \Omega : \mathbf{g}(\mathbf{x}) = 0\}$$

donde  $\mathbf{g} : \Omega \rightarrow \mathbb{R}^{n-k}$  es de clase  $C^m$  en un abierto  $\Omega \subset \mathbb{R}^n$ , y para cada  $\mathbf{x} \in M$  los vectores  $\nabla g_1(\mathbf{x}), \nabla g_2(\mathbf{x}), \dots, \nabla g_{n-k}(\mathbf{x})$  son linealmente independientes (e.d. la matriz  $(D_i g_j(\mathbf{x}))_{1 \leq i \leq n-k, 1 \leq j \leq n-k}$  tiene rango  $n-k$ ).

Toda representación explícita de clase  $C^m$  y dimensión  $k < n$  de  $M \subset \mathbb{R}^n$  proporciona de modo natural una representación implícita de la misma clase y dimensión, dada en términos de la función:

$$\mathbf{g}(\mathbf{u}, \mathbf{v}) = \mathbf{f}(\mathbf{u}) - \mathbf{v} = (f_1(u_1, u_2, \dots, u_k) - v_1, \dots, f_{n-k}(u_1, u_2, \dots, u_k) - v_{n-k})$$

definida en  $\Omega = A \times \mathbb{R}^{n-k}$ . Dejamos al cuidado del lector la comprobación de que se cumplen las condiciones requeridas en la definición 9.3.

**Teorema 9.4** Si  $1 \leq k < n$ , las siguientes propiedades de  $M \subset \mathbb{R}^n$  son equivalentes

- a) Cada  $\mathbf{p} \in M$  posee un entorno abierto  $\Omega_{\mathbf{p}} \subset \mathbb{R}^n$  tal que  $M \cap \Omega_{\mathbf{p}}$  admite una parametrización regular de clase  $C^m$  y dimensión  $k$ .
- b) Cada  $\mathbf{p} \in M$  posee un entorno abierto  $\Omega_{\mathbf{p}} \subset \mathbb{R}^n$  tal que  $M \cap \Omega_{\mathbf{p}}$  admite una representación implícita de clase  $C^m$  y dimensión  $k$ .
- c) Cada  $\mathbf{p} \in M$  posee un entorno abierto  $\Omega_{\mathbf{p}} \subset \mathbb{R}^n$  tal que  $M \cap \Omega_{\mathbf{p}}$  admite una representación explícita de clase  $C^m$  y dimensión  $k$ .

DEM: Según lo comentado después de las definiciones 9.2 y 9.3 es claro que c)  $\Rightarrow$  a), y c)  $\Rightarrow$  b).

b)  $\Rightarrow$  c) es consecuencia directa del teorema de la función implícita: Por hipótesis, cada  $\mathbf{p} \in M$  posee un entorno abierto  $\Omega_{\mathbf{p}} \subset \mathbb{R}^n$  tal que

$$M \cap \Omega_{\mathbf{p}} = \{\mathbf{x} \in \Omega_{\mathbf{p}} : \mathbf{g}(\mathbf{x}) = 0\}$$

donde  $\mathbf{g}$  está definida en  $\Omega_{\mathbf{p}}$  y cumple las condiciones de la definición 9.3. Como la matriz

$$\begin{pmatrix} D_1 g_1(\mathbf{p}) & D_2 g_1(\mathbf{p}) & \cdots & D_n g_1(\mathbf{p}) \\ \cdots & \cdots & \cdots & \cdots \\ D_1 g_{n-k}(\mathbf{p}) & D_2 g_{n-k}(\mathbf{p}) & \cdots & D_n g_{n-k}(\mathbf{p}) \end{pmatrix}$$

tiene rango  $n-k$  podemos suponer, para simplificar la notación, que no es nulo el determinante de la matriz cuadrada formada con las últimas  $n-k$  columnas. Entonces, aplicando el teorema de la función implícita, se puede asegurar que existe un entorno abierto de  $\mathbf{p}$ , de la forma  $\Omega'_{\mathbf{p}} = A \times B \subset \Omega_{\mathbf{p}}$ , con  $A \subset \mathbb{R}^k$ ,  $B \subset \mathbb{R}^{n-k}$ , abiertos, y una función implícita de clase  $C^m$ ,  $\mathbf{f} : A \rightarrow B$  tal que

$$M \cap \Omega'_{\mathbf{p}} = \{(\mathbf{u}, \mathbf{f}(\mathbf{u})) : \mathbf{u} \in A\}$$



Si fuese nulo el determinante formado por las últimas  $n - k$  columnas, tendríamos que seleccionar otras columnas para conseguir una matriz cuadrada de determinante no nulo y el teorema de la función implícita permitiría despejar localmente las correspondientes  $n - k$  variables en función de las restantes. Con una permutación  $\sigma$  de  $\{1, 2, \dots, n\}$  se consigue que estas  $n - k$  variables pasen a ser las últimas, y esta permutación es la que hace que se cumpla la definición 9.1.

a)  $\Rightarrow$  c) Se demuestra con ayuda del teorema de la función inversa: Por hipótesis  $\mathbf{p} \in M$  posee un entorno abierto  $\Omega_{\mathbf{p}} \subset \mathbb{R}^n$  tal que  $M \cap \Omega_{\mathbf{p}} = \varphi(U)$ , donde  $\varphi : U \rightarrow \mathbb{R}^n$  es una parametrización regular de clase  $C^m$  y dimensión  $k$  definida en un abierto  $U \subset \mathbb{R}^k$ . Si  $\mathbf{p} = \varphi(\mathbf{a})$ , como los vectores  $D_1\varphi(\mathbf{a}), D_2\varphi(\mathbf{a}), \dots, D_k\varphi(\mathbf{a})$  son linealmente independientes, la matriz

$$\begin{pmatrix} D_1\varphi_1(\mathbf{a}) & D_2\varphi_1(\mathbf{a}) & \cdots & D_k\varphi_1(\mathbf{a}) \\ \cdots & \cdots & \cdots & \cdots \\ D_1\varphi_n(\mathbf{a}) & D_2\varphi_n(\mathbf{a}) & \cdots & D_k\varphi_n(\mathbf{a}) \end{pmatrix}$$

tiene rango  $k$ , y podemos suponer, para simplificar la escritura, que la matriz formada por la  $k$  primeras filas tiene determinante no nulo. En ese caso escribimos

$$\varphi(\mathbf{u}) = (\boldsymbol{\alpha}(\mathbf{u}), \boldsymbol{\beta}(\mathbf{u}))$$

donde  $\boldsymbol{\alpha} : U \rightarrow \mathbb{R}^k, \boldsymbol{\beta} : U \rightarrow \mathbb{R}^{n-k}$ .

Como  $\det \alpha'(\mathbf{a}) \neq 0$ , aplicando el teorema de la función inversa a la aplicación  $\boldsymbol{\alpha}$  en el punto  $\mathbf{a}$  obtenemos  $A \subset U$ , entorno abierto de  $\mathbf{a}$ , tal que  $\boldsymbol{\alpha}(A) \subset \mathbb{R}^k$  es abierto y  $\boldsymbol{\alpha} : A \rightarrow \boldsymbol{\alpha}(A)$  es de clase  $C^m$  con inversa de clase  $C^m$ . Entonces  $\mathbf{f} = \boldsymbol{\beta} \circ \boldsymbol{\alpha}^{-1}$ , definida en  $\boldsymbol{\alpha}(A)$ , con valores en  $\mathbb{R}^{n-k}$ , también es de clase  $C^m$ .

Como  $\varphi$  establece un homeomorfismo entre  $U$  y  $\varphi(U)$  se tiene que  $\varphi(A)$  es un entorno abierto del punto  $\mathbf{p} = \varphi(\mathbf{a})$  en el conjunto  $\varphi(U) = M \cap \Omega_{\mathbf{p}}$ , para la topología relativa de este conjunto, luego es de la forma  $\varphi(A) = M \cap \Omega'_{\mathbf{p}}$ , donde  $\Omega'_{\mathbf{p}} \subset \mathbb{R}^n$  es un abierto que se puede suponer incluido en  $\Omega_{\mathbf{p}}$ . Este abierto cumple las condiciones requeridas en la definición 9.1:

$$M \cap \Omega'_{\mathbf{p}} = \varphi(A) = \{(\boldsymbol{\alpha}(\mathbf{u}), \boldsymbol{\beta}(\mathbf{u})) : \mathbf{u} \in A\} = \{(\mathbf{v}, \mathbf{f}(\mathbf{v})) : \mathbf{v} \in \boldsymbol{\alpha}(A)\}$$

En caso de que fuese nulo el determinante de la matriz formada con las  $k$  primeras filas de la matriz  $(D_i\varphi_j(\mathbf{a}))$  tendríamos que seleccionar otras filas para obtener una matriz cuadrada de determinante no nulo. Estas filas se pueden llevar a las primeras posiciones mediante una permutación  $\sigma$  de las variables, y con esta permutación se cumplen las condiciones de la definición 9.1. ■

OBSERVACIÓN: Sea  $M = \varphi(U)$ , donde  $\varphi : U \rightarrow \mathbb{R}^n$  es una parametrización regular de clase  $C^m$  y dimensión  $k$ . Según la definición 9.2 la inversa  $\varphi^{-1} : M \rightarrow U$  es continua. Ahora podemos decir algo más: Cada  $\mathbf{p} \in M$  posee un entorno abierto  $W_{\mathbf{p}} \subset \mathbb{R}^n$ , donde hay definida una función de clase  $C^m$ ,  $\Psi : W_{\mathbf{p}} \rightarrow U$ , que verifica

$$\Psi(\mathbf{x}) = \varphi^{-1}(\mathbf{x}) \quad \text{para todo } \mathbf{x} \in M \cap W_{\mathbf{p}}.$$

Para ver esto basta continuar con el razonamiento de a)  $\Rightarrow$  c) en el teorema anterior y definir en  $W_{\mathbf{p}} = \boldsymbol{\alpha}(A) \times \mathbb{R}^{n-k}$  la aplicación  $\Psi(\mathbf{s}, \mathbf{t}) = \boldsymbol{\alpha}^{-1}(\mathbf{s})$ . Es claro que para cada  $\mathbf{x} = \varphi(\mathbf{u}) = (\boldsymbol{\alpha}(\mathbf{u}), \boldsymbol{\beta}(\mathbf{u})) \in W_{\mathbf{p}}$  se cumple  $\varphi^{-1}(\mathbf{x}) = \mathbf{u} = \Psi(\mathbf{x})$ .

**Definición 9.5** Un subconjunto  $M \subset \mathbb{R}^n$  se dice que es una subvariedad diferenciable de clase  $C^m$  y dimensión  $k$ , ( $1 \leq k < n$ ), cuando posee las propiedades equivalentes del teorema 9.4

**Ejemplos 9.6** En cada uno de los siguientes casos  $M \subset \mathbb{R}^n$  es una subvariedad de clase  $C^m$  y dimensión  $k$ :

- a)  $M = \varphi(U)$  donde  $\varphi : U \rightarrow \mathbb{R}^n$  es una parametrización regular de clase  $C^m$  definida en un abierto  $U \subset \mathbb{R}^k$ .
- b)  $M = \{\mathbf{x} \in \Omega : \mathbf{g}(\mathbf{x}) = 0\}$  donde  $\mathbf{g} : \Omega \rightarrow \mathbb{R}^{n-k}$  es una aplicación de clase  $C^m$ , definida en un abierto  $\Omega \subset \mathbb{R}^n$ , con la propiedad de que, para cada  $\mathbf{x} \in M$  los vectores  $\nabla g_1(\mathbf{x}), \nabla g_2(\mathbf{x}), \dots, \nabla g_{n-k}(\mathbf{x})$  son linealmente independientes.
- c)  $M = T_\sigma(G(\mathbf{f}))$  donde  $G(\mathbf{f}) = \{(\mathbf{u}, \mathbf{f}(\mathbf{u})) : \mathbf{u} \in A\}$  es la gráfica de una aplicación de clase  $C^m$ ,  $\mathbf{f} : A \rightarrow \mathbb{R}^{n-k}$ , definida en un abierto  $A \subset \mathbb{R}^k$ , y  $\sigma$  es una permutación de  $\{1, 2, \dots, n\}$ .

**Espacio tangente a una subvariedad diferenciable.** Según la definición 5.3 un vector  $\mathbf{w} \in \mathbb{R}^n$  es tangente a  $M \subset \mathbb{R}^n$  en  $\mathbf{p} \in M$  si existe una aplicación  $\gamma : (-\alpha, \alpha) \rightarrow M$ , derivable en  $t = 0$  tal que  $\gamma(0) = \mathbf{p}$  y  $\gamma'(0) = \mathbf{w}$ . Si  $T_{\mathbf{p}}(M)$  es el conjunto de los vectores tangentes a  $M$  en  $\mathbf{p}$  y  $\Omega_{\mathbf{p}}$  es un entorno abierto de  $\mathbf{p}$ , es inmediato que  $T_{\mathbf{p}}(M) = T_{\mathbf{p}}(M \cap \Omega_{\mathbf{p}})$ .

Si  $A \subset \mathbb{R}^k$  es abierto y  $\mathbf{f} : A \rightarrow \mathbb{R}^{n-k}$  es diferenciable en  $\mathbf{a} \in A$ , según la proposición 5.28, el conjunto de vectores tangentes en  $\mathbf{p} = (\mathbf{a}, \mathbf{f}(\mathbf{a}))$  a la gráfica

$$M = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^k \times \mathbb{R}^{n-k} : \mathbf{x} \in A, \mathbf{y} = \mathbf{f}(\mathbf{x})\}$$

es la gráfica de la diferencial  $d\mathbf{f}(\mathbf{a})$ , es decir el subespacio vectorial, de dimensión  $k$

$$T_{\mathbf{p}}(M) = \{(\mathbf{u}, \mathbf{v}) \in \mathbb{R}^k \times \mathbb{R}^{n-k} : \mathbf{v} = d\mathbf{f}(\mathbf{a})\mathbf{u}\}$$

Cuando  $M = \{\mathbf{x} \in \Omega : \mathbf{g}(\mathbf{x}) = 0\}$  con  $\mathbf{g} : \Omega \rightarrow \mathbb{R}^{n-k}$  definida en un abierto  $\Omega \subset \mathbb{R}^n$ , y diferenciable en  $\mathbf{p} \in M$ , según la proposición 5.29 se verifica

$$T_{\mathbf{p}}(M) \subset \text{Ker } d\mathbf{g}(\mathbf{p}) = \bigcap_{j=1}^{n-k} \{\mathbf{u} \in \mathbb{R}^n : \langle \nabla g_j(\mathbf{p}) \mid \mathbf{u} \rangle = 0\}$$

Por consiguiente, cuando los vectores  $\nabla g_1(\mathbf{p}), \nabla g_2(\mathbf{p}), \dots, \nabla g_{n-k}(\mathbf{p})$  son linealmente independientes  $\text{Ker } d\mathbf{g}(\mathbf{p})$  es un subespacio de dimensión  $k$  que contiene a  $T_{\mathbf{p}}(M)$ .

**Proposición 9.7** Sea  $M = \{\mathbf{x} \in \Omega : g_1(\mathbf{x}) = 0, \dots, g_{n-k}(\mathbf{x}) = 0\}$ , donde las funciones  $g_1, g_2, \dots, g_{n-k} : \Omega \rightarrow \mathbb{R}$  son de clase  $C^1(\Omega)$ . Si  $\mathbf{p} \in M$  y los vectores

$$\nabla g_1(\mathbf{p}), \nabla g_2(\mathbf{p}), \dots, \nabla g_{n-k}(\mathbf{p})$$

son linealmente independientes entonces

$$T_{\mathbf{p}}(M) = \bigcap_{j=1}^{n-k} \{\mathbf{u} \in \mathbb{R}^n : \langle \nabla g_j(\mathbf{p}) \mid \mathbf{u} \rangle = 0\}$$

Es decir,  $T_{\mathbf{p}}(M)$  es el subespacio vectorial de  $\mathbb{R}^n$ , de dimensión  $k$ , ortogonal a los vectores  $\nabla g_1(\mathbf{p}), \nabla g_2(\mathbf{p}), \dots, \nabla g_{n-k}(\mathbf{p})$ .

DEM: Como  $T_{\mathbf{p}}(M) \subset \text{Ker } d\mathbf{g}(\mathbf{p})$ , y  $\text{Ker } d\mathbf{g}(\mathbf{p})$  es un subespacio vectorial de dimensión  $k$ , para obtener la igualdad  $T_{\mathbf{p}}(M) = \text{Ker } d\mathbf{g}(\mathbf{p})$  basta ver  $T_{\mathbf{p}}(M)$  es un subespacio vectorial de la misma dimensión.

Razonando como en la demostración de b)  $\Rightarrow$  c) en el teorema 9.4, el teorema de la función implícita permite obtener  $\Omega_{\mathbf{p}} \subset \mathbb{R}^n$ , entorno abierto de  $\mathbf{p}$ , con tal que  $M \cap \Omega_{\mathbf{p}}$ , es (salvo una permutación de las variables) la gráfica de una aplicación  $\mathbf{f} : A \rightarrow \mathbb{R}^{n-k}$ , diferenciable en  $\mathbf{a} \in A$ , donde  $\mathbf{p} = (\mathbf{a}, \mathbf{f}(\mathbf{a}))$ . Basta tener en cuenta ahora que  $T_{\mathbf{p}}(\Omega_{\mathbf{p}} \cap M)$  es un espacio vectorial de dimensión  $k$  y que  $T_{\mathbf{p}}(M) = T_{\mathbf{p}}(\Omega_{\mathbf{p}} \cap M)$ . ■

Finalmente, consideramos el caso de un conjunto de la forma  $M = \varphi(U)$ , donde  $\varphi : U \rightarrow \mathbb{R}^n$  está definida en un abierto  $U \subset \mathbb{R}^k$  y es diferenciable en  $\mathbf{a} \in U$ . Recordemos que el conjunto de vectores tangentes a  $M$  en  $\mathbf{p} = \varphi(\mathbf{a})$  contiene al subespacio  $d\varphi(\mathbf{a})(\mathbb{R}^k)$ , es decir,  $T_{\mathbf{p}}(M) \supset \{D_{\mathbf{u}}\varphi(\mathbf{a}) : \mathbf{u} \in \mathbb{R}^k\}$ . La dimensión de este subespacio es  $\leq k$  y será igual a  $k$ , cuando las derivadas parciales  $D_j\varphi(\mathbf{a}) = d\varphi(\mathbf{a})\mathbf{e}_j$ ,  $1 \leq j \leq k$ , sean linealmente independientes. En este caso, si se sabe que  $T_{\mathbf{p}}(M)$  es un subespacio vectorial de dimensión  $k$  se obtendrá que

$$T_{\mathbf{p}}(M) = \{D_{\mathbf{u}}\varphi(\mathbf{a}) : \mathbf{u} \in \mathbb{R}^k\} = d\varphi(\mathbf{a})(\mathbb{R}^k)$$

Esto es lo que ocurre cuando  $\varphi$  es una parametrización regular de clase  $C^m$  y dimensión  $k$ .

**Proposición 9.8** *Si  $M \subset \mathbb{R}^n$  es una subvariedad diferenciable de clase  $C^m$  y dimensión  $k$  entonces en cada punto  $\mathbf{p} \in M$  el espacio tangente  $T_{\mathbf{p}}(M)$  es un espacio vectorial de dimensión  $k$ .*

DEM: Basta considerar que  $M$  tiene localmente la propiedad 9.4 c): Existe  $\Omega_{\mathbf{p}} \subset \mathbb{R}^n$ , entorno abierto de  $\mathbf{p}$ , tal que  $M \cap \Omega_{\mathbf{p}} = \{(\mathbf{x}, \mathbf{f}(\mathbf{x})) : \mathbf{x} \in A\}$  donde  $\mathbf{f} : A \rightarrow \mathbb{R}^{n-k}$  es diferenciable en el abierto  $A \subset \mathbb{R}^{n-k}$ . Se sigue de esto que

$$T_{\mathbf{p}}(M) = T_{\mathbf{p}}(M \cap \Omega_{\mathbf{p}}) = \{(\mathbf{u}, \mathbf{v}) \in \mathbb{R}^k \times \mathbb{R}^{n-k} : \mathbf{v} = d\mathbf{f}(\mathbf{a})\mathbf{u}\}$$

luego  $T_{\mathbf{p}}(M)$  es un espacio vectorial de dimensión  $k$ . (Para simplificar la notación, al aplicar la condición 9.4 c) hemos supuesto  $\sigma = \text{identidad}$ ) ■

Cuando  $M \subset \mathbb{R}^n$  es una subvariedad diferenciable de clase  $C^m$  y dimensión  $k$ , podemos aplicar en cada punto  $\mathbf{p} \in M$  lo indicado anteriormente. Así, para escribir las ecuaciones en forma implícita del espacio tangente  $T_{\mathbf{p}}(M)$  podemos utilizar, en un entorno abierto  $\Omega_{\mathbf{p}}$  de  $\mathbf{p}$ , una representación implícita de  $M \cap \Omega_{\mathbf{p}}$ . Análogamente, para escribir las ecuaciones paramétricas de  $T_{\mathbf{p}}(M)$ , podemos utilizar una parametrización regular de un entorno relativo de  $\mathbf{p}$  (véase la sección 5.5).

## 9.2. Extremos condicionados

Dada una función  $f : \Omega \rightarrow \mathbb{R}$  definida en un abierto  $\Omega \subset \mathbb{R}^n$  se trata de determinar, si existen, los extremos (relativos o absolutos) de la restricción de  $f$  a un conjunto  $M \subset \Omega$  de la forma  $M = \{\mathbf{x} \in \Omega : \mathbf{g}(\mathbf{x}) = 0\}$  donde  $\mathbf{g} : \Omega \rightarrow \mathbb{R}^m$ . Es decir, se trata de obtener los extremos (relativos o absolutos) de  $f(x_1, x_2, \dots, x_n)$  cuando las variables  $(x_1, x_2, \dots, x_n)$  están sometidas a las condiciones de ligadura

$$g_1(x_1, x_2, \dots, x_n) = 0, \quad g_2(x_1, x_2, \dots, x_n) = 0, \dots, \quad g_m(x_1, x_2, \dots, x_n) = 0.$$

Las definiciones referentes a extremos condicionados son obvias: Si  $\mathbf{p} \in M$  posee un entorno abierto  $V_{\mathbf{p}} \subset \Omega$  tal que  $f(\mathbf{x}) \leq f(\mathbf{p})$  (resp.  $f(\mathbf{x}) \geq f(\mathbf{p})$ ) para todo  $\mathbf{x} \in M \cap V_{\mathbf{p}}$ , se dice que  $f|_M$  presenta en  $\mathbf{p}$  un máximo (resp. mínimo) relativo o bien que  $f$  presenta en  $\mathbf{p}$  un máximo (resp. mínimo) condicionado por las ligaduras  $g_k(x_1, x_2, \dots, x_n) = 0$ ,  $1 \leq k \leq m$ . Cuando se reemplaza el signo de desigualdad  $\leq$  por el de desigualdad estricta  $<$ , se obtiene la definición de máximo (resp. mínimo) relativo estricto condicionado.

Los principales resultados sobre extremos condicionados se obtienen cuando  $f$  es diferenciable y el conjunto  $M$  definido por las condiciones de ligadura, tiene una estructura geométrica sencilla desde el punto de vista del cálculo diferencial: Las propiedades que se exigirán a las condiciones de ligadura  $g_k$ ,  $1 \leq k \leq m$ , serán las adecuadas para garantizar que en cada  $\mathbf{p} \in M$  el conjunto  $T_{\mathbf{p}}(M)$  es un espacio vectorial de dimensión  $n - m$  (véase 9.7).

Para motivar los resultados generales consideramos primero el caso particular para  $n = 2$ ,  $m = 1$ . En este caso  $M = \{(x, y) \in \Omega : g(x, y) = 0\}$  es una "curva" definida implícitamente por una función  $g$  de clase  $C^1$  en un abierto  $\Omega \subset \mathbb{R}^2$ .

Si  $\mathbf{p} = (a, b) \in M$  es un punto donde  $\nabla g(\mathbf{p}) = (D_1g(a, b), D_2g(a, b)) \neq (0, 0)$ , en virtud de la proposición 9.7 el espacio tangente  $T_{\mathbf{p}}(M)$  es el subespacio de  $\mathbb{R}^2$ , de dimensión 1, ortogonal al gradiente  $\nabla g(\mathbf{p})$ . Es decir, existe la recta tangente a la "curva"  $M$  en el punto  $\mathbf{p}$ , y su ecuación es

$$D_1g(a, b)(x - a) + D_2g(a, b)(y - b) = 0$$

Representemos la función  $f : \Omega \rightarrow \mathbb{R}$  dibujando en el plano sus curvas de nivel  $N_c = \{(x, y) \in \Omega : f(x, y) = c\}$ , para diferentes valores de  $c$ . Si  $\mathbf{p} \in \Omega$  y  $c = f(\mathbf{p})$  entonces  $N_c$  es la curva de nivel que pasa por  $\mathbf{p} \in \Omega$  y sabemos que  $\nabla f(\mathbf{p})$  es ortogonal todos los vectores tangentes a  $N_c$  en  $\mathbf{p}$ . Supongamos además que existe un entorno  $B(\mathbf{p}, r) \subset \Omega$  que queda descompuesto por la curva  $N_c$  en dos regiones separadas:

$$\{(x, y) \in B(\mathbf{p}, r) : f(x, y) < c\}, \quad \{(x, y) \in B(\mathbf{p}, r) : f(x, y) > c\}$$

que llamaremos, respectivamente, lado izquierdo y lado derecho de  $N_c$ . Si la curva  $M$  pasa por  $\mathbf{p}$  atravesando la curva de nivel  $N_c$  entonces  $f(x, y) - c$  cambia de signo cuando  $(x, y) \in M$  pasa de un lado al otro de  $N_c$ , luego  $f|_M$  no puede presentar en  $\mathbf{p}$  ni un máximo ni un mínimo relativo. Por lo tanto, para que  $f|_M$  presente en  $\mathbf{p} \in M$  un extremo relativo, es necesario que la curva  $M$  no atraviese a la línea de nivel  $N_c$ ,

es decir  $M$  y  $N_c$  deben tener en  $\mathbf{p}$  la misma recta tangente. Esto ocurre si y sólo si  $\nabla f(\mathbf{p}) = \mu \nabla g(\mathbf{p})$  para algún  $\mu \in \mathbb{R}$ . Lo que acabamos de ver es una justificación heurística, y una interpretación geométrica, de la siguiente proposición

**Proposición 9.9** *Sea  $M = \{(x, y) \in \Omega : g(x, y) = 0\}$ , donde  $g : \Omega \rightarrow \mathbb{R}$  es de clase  $C^1(\Omega)$  en un abierto  $\Omega \subset \mathbb{R}^2$ . Se supone que  $f : \Omega \rightarrow \mathbb{R}$  es diferenciable en  $\mathbf{p} \in M$  y que  $\nabla g(\mathbf{p}) \neq (0, 0)$ . Si  $f|_M$  presenta un extremo relativo en  $\mathbf{p}$  entonces existe  $\mu \in \mathbb{R}$  tal que  $\nabla f(\mathbf{p}) = \mu \nabla g(\mathbf{p})$ .*

DEM: Según la proposición 9.7 el subespacio ortogonal a  $T_{\mathbf{p}}(M)$  es el generado por  $\nabla g(\mathbf{p}) \neq (0, 0)$ . Por lo tanto basta demostrar que si  $f|_M$  presenta en  $\mathbf{p} \in M$  un extremo relativo entonces  $\nabla f(\mathbf{p})$  es ortogonal a  $T_{\mathbf{p}}(M)$ .

Supongamos que el extremo relativo es un máximo y sea  $V_{\mathbf{p}} \subset \Omega$  un entorno abierto de  $\mathbf{p}$  tal que  $f(\mathbf{x}) \leq f(\mathbf{p})$  para todo  $\mathbf{x} \in V_{\mathbf{p}}$ .

Dado  $\mathbf{u} \in T_{\mathbf{p}}(M)$ , según la definición de vector tangente, existe una aplicación  $\gamma : (-\delta, \delta) \rightarrow M$  derivable en  $t = 0$ , tal que  $\gamma(0) = \mathbf{p}$  y  $\gamma'(0) = \mathbf{u}$ . Como  $\gamma$  es continua en  $t = 0$  existe  $0 < r < \delta$  tal que  $|t| < r \Rightarrow \gamma(t) \in V_{\mathbf{p}} \cap M$ , luego

$$|t| < r \Rightarrow f(\gamma(t)) \leq f(\gamma(0))$$

La función real de variable real  $\varphi(t) = f(\gamma(t))$  presenta un máximo relativo en  $t = 0$  donde es derivable (en virtud de la regla de la cadena) luego

$$0 = \varphi'(0) = df(\gamma(0))\gamma'(0) = df(\mathbf{p})\mathbf{u} = \langle \nabla f(\mathbf{p}) | \mathbf{u} \rangle$$

Es decir,  $\nabla f(\mathbf{p})$  es ortogonal a  $T_{\mathbf{p}}(M)$ . ■

En las condiciones de la proposición anterior, los puntos  $\mathbf{p} \in M$  que cumplen  $\nabla g(\mathbf{p}) \neq (0, 0)$  diremos que son puntos regulares de  $M$ . Los puntos regulares  $\mathbf{p}$  donde  $f$  es diferenciable y  $\nabla f(\mathbf{p}) = \mu \nabla g(\mathbf{p})$  para algún  $\mu \in \mathbb{R}$  se dice que son *puntos estacionarios* de  $f|_M$ . Denotaremos por  $E(f, M)$  el conjunto de todos los puntos estacionarios de  $f|_M$ . Si  $\mathbf{p} \in E(f, M)$ , el coeficiente  $\mu$  que hace que se cumplan las igualdades

$$D_1 f(\mathbf{p}) = \mu D_1 g(\mathbf{p}), \quad D_2 f(\mathbf{p}) = \mu D_2 g(\mathbf{p})$$

se llama multiplicador de Lagrange asociado al punto estacionario  $\mathbf{p}$ . Cada punto estacionario  $\mathbf{p} \in E(f, M)$  lleva asociado su multiplicador de Lagrange.

Con esta terminología la proposición 9.9 se enuncia diciendo que los puntos regulares  $\mathbf{p} \in M$  donde una función diferenciable  $f$  presenta extremos relativos condicionados por  $M$  son puntos estacionarios de  $f|_M$ .

En las aplicaciones habituales  $f$  es diferenciable en todos los puntos de  $M$  y todos los puntos de  $M$  son regulares. En este caso los extremos relativos de  $f|_M$  se alcanzan en puntos estacionario y para determinar los extremos relativos de  $f|_M$  se comienza calculando el conjunto de los puntos estacionarios  $E(f, M)$ . Esto se consigue resolviendo (cuando sea posible) el sistema de tres ecuaciones con tres incógnitas,  $x, y, \mu$ .

$$D_1 f(x, y) - \mu D_1 g(x, y) = 0; \quad D_2 f(x, y) - \mu D_2 g(x, y) = 0; \quad g(x, y) = 0.$$

Si el sistema no tiene solución se puede asegurar que  $f|_M$  no presenta extremos relativos. Si el sistema tiene un número finito de soluciones

$$(x_1, y_1, \mu_1), (x_2, y_2, \mu_2), \dots (x_k, y_k, \mu_k)$$

tendremos un conjunto finito de puntos estacionarios

$$\mathbf{p}_1 = (x_1, y_1), \quad \mathbf{p}_2 = (x_2, y_2), \quad \dots \quad \mathbf{p}_k = (x_k, y_k)$$

y sus correspondientes multiplicadores  $\mu_1, \mu_2, \dots, \mu_k$ .

Si además  $M$  es compacto y el problema que nos ocupa consiste en determinar el máximo y el mínimo absoluto de  $f$  sobre  $M$ , bastará evaluar  $f$  en cada uno de los puntos estacionarios para obtener, por inspección directa, los extremos absolutos:

$$\max_{\mathbf{x} \in M} f(\mathbf{x}) = \max\{f(\mathbf{p}_i) : 1 \leq i \leq k\}$$

$$\min_{\mathbf{x} \in M} f(\mathbf{x}) = \min\{f(\mathbf{p}_i) : 1 \leq i \leq k\}$$

Lo mismo se podrá hacer cuando  $M$  no sea compacto pero por la naturaleza del problema sepamos que  $f$  debe alcanzar en  $M$  un máximo o un mínimo absoluto, o ambos. Así por ejemplo, como  $M$  es cerrado, la función  $f(x, y) = (x - a)^2 + (y - b)^2$  siempre alcanza un mínimo absoluto sobre  $M$  (recuérdese que, según lo indicado a continuación de proposición 3.17, fijado un punto  $\mathbf{a} \in \mathbb{R}^n$ , la función distancia  $\mathbf{x} \rightarrow \|\mathbf{x} - \mathbf{a}\|_2$  siempre alcanza un mínimo absoluto sobre cada cerrado  $M \subset \mathbb{R}^n$ ).

A la hora de las aplicaciones concretas, cuando  $f$  no sea diferenciable en todos los puntos de  $M$  o algún punto de  $M$  no sea regular, para determinar los extremos absolutos de  $f$  se puede proceder como antes, pero considerando también los puntos de  $M$  donde  $f$  no es diferenciable o  $g$  no es regular. Si el conjunto de estos puntos  $S(f, M)$  es finito y sabemos que existe un máximo, o un mínimo absolutos, lo podremos calcular como antes evaluando la función en un conjunto finito de puntos:

$$\max_{\mathbf{x} \in M} f(\mathbf{x}) = \max\{f(\mathbf{x}) : \mathbf{x} \in E(f, M) \cup S(f, M)\}$$

$$\min_{\mathbf{x} \in M} f(\mathbf{x}) = \min\{f(\mathbf{x}) : \mathbf{x} \in E(f, M) \cup S(f, M)\}$$

Un problema más delicado es el de determinar, entre los puntos estacionarios los que proporcionan un máximo relativo, o un mínimo relativo. Más adelante daremos condiciones de segundo orden suficientes para que un punto estacionario sea un punto de máximo relativo o un punto de mínimo relativo. Pero este criterio no será aplicable a los puntos donde  $f$  o  $g$  no sean dos veces diferenciables.

En el caso que nos ocupa, para funciones de dos variables reales, dado un punto estacionario  $\mathbf{p} \in M$  se puede proceder de la siguiente forma: En un entorno  $V_{\mathbf{p}}$  de  $\mathbf{p}$  se busca una parametrización del trozo de curva  $V_{\mathbf{p}} \cap M$ , es decir un homeomorfismo diferenciable  $\varphi : (\alpha, \beta) \rightarrow M \cap V_{\mathbf{p}}$ . Si  $\varphi(t_0) = \mathbf{p}$  es claro que  $\mathbf{p}$  será un punto de máximo (resp. mínimo) relativo para  $f|_M$  si y sólo si  $t_0$  es un punto de máximo (resp. mínimo) de la función real de variable real  $f(\varphi(t))$ . Este último estudio se podrá abordar con las técnicas usuales de las funciones de una variable.

La parametrización local  $\varphi$  a veces se puede conseguir considerando la ecuación de la curva  $g(x, y) = 0$  y despejando, en un entorno de  $\mathbf{p}$  suficientemente pequeño, una variable en función de la otra, es decir obteniendo un entorno rectangular de  $\mathbf{p}$ ,  $V_{\mathbf{p}} = (a, b) \times (c, d)$ , tal que  $V_{\mathbf{p}} \cap M$  se pueda expresar de una de las dos formas siguientes

$$\{(x, \alpha(x)) : a < x < b\}; \quad \{(\beta(y), y) : c < y < d\}$$

(El teorema de la función implícita nos asegura que esto siempre se puede hacer) En el primer caso se consigue la parametrización local  $\varphi(t) = (t, \alpha(t))$ ,  $a < t < b$  y en el segundo caso  $\varphi(t) = (\beta(t), t)$ ,  $c < t < d$ .

Después de haber visto con detalle el caso particularmente simple de las funciones de dos variables sometidas a una ligadura consideramos el caso general de los extremos relativos de una función de  $n$  variables sometida a  $m$  condiciones de ligadura ( $m < n$ ).

**Teorema 9.10** *Sea  $\Omega \subset \mathbb{R}^n$  abierto y  $M = \{\mathbf{x} \in \Omega : g_k(\mathbf{x}) = 0, 1 \leq k \leq m\}$  donde las funciones  $g_k : \Omega \rightarrow \mathbb{R}$  son de clase  $C^1(\Omega)$ . Se supone que  $f : \Omega \rightarrow \mathbb{R}$  es diferenciable en  $\mathbf{p} \in M$  y que los vectores  $\nabla g_1(\mathbf{p}), \nabla g_2(\mathbf{p}), \dots, \nabla g_m(\mathbf{p})$  son linealmente independientes.*

*Si  $f|_M$  presenta en  $\mathbf{p} \in M$  un extremo relativo entonces  $df(\mathbf{p})$  se anula sobre el espacio tangente  $T_{\mathbf{p}}(M)$ , y por lo tanto existen coeficientes  $\mu_1, \mu_2, \dots, \mu_m \in \mathbb{R}$  tales que*

$$\nabla f(\mathbf{p}) = \mu_1 \nabla g_1(\mathbf{p}) + \mu_2 \nabla g_2(\mathbf{p}) + \dots + \mu_m \nabla g_m(\mathbf{p})$$

DEM: Supongamos que  $f|_M$  presenta en  $\mathbf{p}$  un máximo relativo y sea  $V_{\mathbf{p}} \subset \Omega$  un entorno abierto de  $\mathbf{p}$  tal que  $f(\mathbf{x}) \leq f(\mathbf{p})$  para todo  $\mathbf{x} \in V_{\mathbf{p}}$ .

Dado  $\mathbf{u} \in T_{\mathbf{p}}(M)$ , según la definición de vector tangente, existe una aplicación  $\gamma : (-\delta, \delta) \rightarrow M$  derivable en  $t = 0$ , tal que  $\gamma(0) = \mathbf{p}$ , y  $\gamma'(0) = \mathbf{u}$ . Como  $\gamma$  es continua en  $t = 0$  existe  $0 < r < \delta$  tal que  $|t| < r \Rightarrow \gamma(t) \in V_{\mathbf{p}} \cap M$ , luego

$$|t| < r \Rightarrow f(\gamma(t)) \leq f(\gamma(0))$$

La función real de variable real  $\varphi(t) = f(\gamma(t))$  presenta un máximo relativo en  $t = 0$  donde (en virtud de la regla de la cadena) es derivable luego

$$0 = \varphi'(0) = df(\gamma(0))\gamma'(0) = df(\mathbf{p})\mathbf{u} = \langle \nabla f(\mathbf{p}) | \mathbf{u} \rangle$$

Si  $E \subset \mathbb{R}^n$  es el subespacio  $m$ -dimensional generado por los vectores linealmente independientes  $\nabla g_k(\mathbf{p})$ ,  $1 \leq k \leq m$ , la proposición 9.7 asegura que  $T_{\mathbf{p}}(M)$  es el subespacio ortogonal a  $E$ . Hemos demostrado que  $\nabla f(\mathbf{p})$  es ortogonal a  $T_{\mathbf{p}}(M)$ , luego  $\nabla f(\mathbf{p}) \in E$  se puede expresar como una combinación lineal de los vectores  $\nabla g_k(\mathbf{p})$ ,  $1 \leq k \leq m$ , que forman una base de  $E$ . ■

De forma similar al caso  $n = 2, m = 1$ , los puntos  $\mathbf{p} \in M$  donde los vectores  $\nabla g_1(\mathbf{p}), \nabla g_2(\mathbf{p}), \dots, \nabla g_m(\mathbf{p})$ , son linealmente independientes diremos que son *puntos regulares* de  $\mathbf{g}$ . Los puntos regulares  $\mathbf{p}$  donde  $f$  es diferenciable y  $\nabla f(\mathbf{p})$  es combinación lineal de estos vectores se dice que son *puntos estacionarios* de  $f|_M$ .

Denotaremos por  $E(f, M)$  el conjunto de tales puntos. Para cada  $\mathbf{p} \in E(f, M)$ , los coeficientes  $\mu_k$ ,  $1 \leq k \leq m$ , que hacen que se cumpla la igualdad

$$\nabla f(\mathbf{p}) = \mu_1 \nabla g_1(\mathbf{p}) + \mu_2 \nabla g_2(\mathbf{p}) + \cdots + \mu_m \nabla g_m(\mathbf{p})$$

se llaman *multiplicadores de Lagrange* asociados al punto estacionario  $\mathbf{p}$ . Cada punto estacionario  $\mathbf{p} \in E(f, M)$  lleva asociados sus correspondientes multiplicadores.

Con esta terminología el teorema 9.13 se enuncia diciendo que los puntos regulares  $\mathbf{p} \in M$  en los que la función diferenciable  $f$  presenta extremos relativos, condicionados por  $M$ , son necesariamente puntos estacionarios de  $f|_M$ .

En las aplicaciones más habituales todos los puntos de  $M$  son regulares y  $f$  es diferenciable en todos los puntos de  $M$ . En este caso cada extremo relativo de  $f|_M$  se alcanza en un punto estacionario y la determinación de los extremos relativos de  $f|_M$  comienza calculando los puntos estacionarios. Para ello se resuelve (cuando sea posible) el sistema de  $n + m$  ecuaciones con  $n + m$  incógnitas,  $x_1, x_2, \dots, x_n, \mu_1, \mu_2, \dots, \mu_m$

$$D_i f(x_1, x_2, \dots, x_n) - \sum_{k=1}^m \mu_k D_i g_k(x_1, x_2, \dots, x_n) = 0, \quad i = 1, 2, \dots, n$$

$$g_k(x_1, x_2, \dots, x_n) = 0, \quad k = 1, 2, \dots, m.$$

Cuando el sistema no tiene solución se puede asegurar que  $f|_M$  no presenta extremos relativos. Si el sistema tiene un número finito de soluciones tendremos un conjunto finito de puntos estacionarios  $E(f, M)$ , acompañados sus correspondientes multiplicadores.

Si además  $M$  es compacto y el problema que nos ocupa consiste en determinar el máximo y el mínimo absoluto de  $f$  sobre  $M$ , bastará evaluar  $f$  en cada uno de los puntos estacionarios para obtener, por inspección directa, los extremos absolutos:

$$\max_{\mathbf{x} \in M} f(\mathbf{x}) = \max\{f(\mathbf{p}) : \mathbf{p} \in E(f, M)\}$$

$$\min_{\mathbf{x} \in M} f(\mathbf{x}) = \min\{f(\mathbf{p}) : \mathbf{p} \in E(f, M)\}$$

Lo mismo se puede hacer cuando  $M$  no sea compacto pero, por la naturaleza del problema, estemos seguros de que  $f$  alcanza en  $M$  un máximo, o un mínimo relativo, o ambos. Así por ejemplo la función

$$f(\mathbf{x}) = \|\mathbf{x} - \mathbf{a}\|^2 = (x_1 - a_1)^2 + (x_2 - a_2)^2 + \cdots + (x_n - a_n)^2$$

siempre alcanza un mínimo absoluto sobre  $M$  cuando  $M \subset \mathbb{R}^n$  es cerrado y  $\mathbf{a} \in \mathbb{R}^n$  (Recuérdese que la distancia de un punto  $\mathbf{a} \in \mathbb{R}^n$  a un cerrado  $M$ , siempre se alcanza en algún  $\mathbf{p} \in M$ ).

A la hora de las aplicaciones concretas, cuando  $f$  no es diferenciable en todos los puntos de  $M$  o existe algún punto de  $M$  donde  $\mathbf{g}$  no es regular, para determinar los extremos absolutos de  $f$  se puede proceder como antes, pero añadiendo a los puntos estacionarios  $E(f, M)$  el conjunto  $S(f, M)$  formado por los puntos de  $M$  donde  $f$  no es diferenciable o  $\mathbf{g}$  no es regular.



Un problema más delicado es el de determinar, entre los puntos estacionarios los que proporcionan un máximo relativo, o un mínimo relativo. En el siguiente teorema veremos condiciones de segundo orden suficientes para que un punto estacionario sea un punto de máximo relativo o un punto de mínimo relativo. Pero este criterio no es aplicable a los puntos de  $S(f, M)$  ni a los puntos estacionarios donde  $f$  o  $\mathbf{g}$  no sean diferenciables dos veces.

Dado un punto estacionario  $\mathbf{p} \in E(f, M)$ , para decidir si es punto de máximo o de mínimo relativo se puede aplicar el siguiente método: Se busca una parametrización de un entorno relativo  $V_{\mathbf{p}} \cap M$ , (donde  $V_{\mathbf{p}} \subset \Omega$  es entorno abierto de  $\mathbf{p}$ ), es decir un homeomorfismo de clase  $C^1$ ,  $\varphi : U \rightarrow M \cap V_{\mathbf{p}}$ , definido en un abierto  $U \subset \mathbb{R}^k$ ,  $k = m - n$ , tal que  $\mathbf{0} \in U$  y  $\varphi(\mathbf{0}) = \mathbf{p}$ . Entonces  $\mathbf{p}$  será punto de máximo (resp. mínimo) relativo para  $f|_M$  si y sólo si  $\mathbf{0}$  es un punto de máximo (resp. mínimo) para la función real de  $k$  variables  $F = f \circ \varphi$ , definida en  $U$ . A esta función  $F$  se le aplicarán con los métodos expuestos anteriormente para el caso de extremos ordinarios (sin ligaduras) y para ello será preciso calcular las derivadas segundas  $D_{ij}F(\mathbf{p})$ . El teorema de la función implícita asegura que existe esta parametrización local la cual, después de una permutación en las variables, se puede suponer de la forma  $\varphi(\mathbf{x}) = (\mathbf{x}, \mathbf{f}(\mathbf{x}))$ . Es decir, en las ecuaciones de ligadura

$$g_i(x_1, x_2, \dots, x_n), \quad 1 \leq i \leq m$$

es posible despejar localmente  $m$  variables en función de las restantes. En general no es posible, o es bastante engorroso, obtener explícitamente las ecuaciones de las funciones implícitas. Sin embargo esto no supone una dificultad seria porque, aunque no conozcamos estas ecuaciones, es posible calcular, en el punto  $\mathbf{p}$ , las derivadas parciales primeras y segundas de las funciones implícitas. Usando la regla de la cadena, estas derivadas permiten calcular las derivadas segundas  $D_{ij}F(\mathbf{p})$ , que son las que intervienen en las condiciones de segundo orden expuestas en 6.17.

**Teorema 9.11** *Sea  $f : \Omega \rightarrow \mathbb{R}$  una función de clase  $C^2(\Omega)$  definida en un abierto  $\Omega \subset \mathbb{R}^n$  y  $M \subset \Omega$  un conjunto de la forma*

$$M = \{\mathbf{x} \in \Omega : g_1(\mathbf{x}) = 0, g_2(\mathbf{x}) = 0, \dots, g_m(\mathbf{x}) = 0\}$$

donde las funciones  $g_k$ ,  $1 \leq k \leq m$ , son de clase  $C^2(\Omega)$ . Sea  $\mathbf{p} \in M$  un punto estacionario de  $f$  sobre  $M$ , es decir, un punto donde los vectores  $\nabla g_1(\mathbf{p}), \nabla g_2(\mathbf{p}), \dots, \nabla g_m(\mathbf{p})$  son linealmente independientes y existen coeficientes  $\mu_1, \mu_2, \dots, \mu_m \in \mathbb{R}$  tales que

$$\nabla f(\mathbf{p}) = \mu_1 \nabla g_1(\mathbf{p}) + \mu_2 \nabla g_2(\mathbf{p}) + \dots + \mu_m \nabla g_m(\mathbf{p})$$

Consideremos la función auxiliar  $H_{\mathbf{p}}(\mathbf{x}) = f(\mathbf{x}) - \sum_{k=1}^m \mu_k g_k(\mathbf{x})$ , y sea

$$Q_{\mathbf{p}}(\mathbf{u}) = d^2 H(\mathbf{p})\mathbf{u}^2 = \sum_{i,j=1}^n D_{ij}H_{\mathbf{p}}(\mathbf{p})u_i u_j$$

la forma cuadrática asociada a esta función en el punto  $\mathbf{p}$ . Se verifica:

- a) Si  $f|_M$  presenta en  $\mathbf{p}$  un mínimo (resp. máximo) relativo entonces  $Q_{\mathbf{p}}(\mathbf{u}) \geq 0$  (resp.  $\leq 0$ ) para todo  $\mathbf{u} \in T_{\mathbf{p}}(M)$ .
- b) Si  $Q_{\mathbf{p}}(\mathbf{u}) > 0$  (resp.  $< 0$ ) para todo  $\mathbf{u} \in T_{\mathbf{p}}(M) \setminus \{\mathbf{0}\}$  entonces  $f|_M$  presenta en  $\mathbf{p}$  un mínimo (resp. máximo) relativo estricto.

DEM: Según el teorema de la función implícita la condición de que los vectores  $\nabla g_1(\mathbf{p}), \nabla g_2(\mathbf{p}), \dots, \nabla g_m(\mathbf{p})$  sean linealmente independientes garantiza la existencia de un entorno abierto de  $\mathbf{p}$ ,  $\Omega_{\mathbf{p}} \subset \Omega$ , tal que  $M \cap \Omega_{\mathbf{p}} = \varphi(U)$ , donde  $\varphi : U \rightarrow \mathbb{R}^n$  es una función de clase  $C^2$  en un abierto  $U \subset \mathbb{R}^k$ , que establece un homeomorfismo entre  $U$  y  $M \cap \Omega_{\mathbf{p}}$ , con la propiedad de que los vectores

$$D_1\varphi(\mathbf{s}), D_2\varphi(\mathbf{s}), \dots, D_k\varphi(\mathbf{s}) \in \mathbb{R}^n$$

son linealmente independientes para todo  $\mathbf{s} \in U$ . Podemos suponer que  $\mathbf{0} \in U$  y que  $\varphi(\mathbf{0}) = \mathbf{p}$ , con lo cual  $T_{\mathbf{p}}(M) = d\varphi(\mathbf{0})(\mathbb{R}^k)$ . Como  $\varphi : U \rightarrow M \cap \Omega_{\mathbf{p}}$  es un homeomorfismo se sigue que  $f|_M$  presenta en  $\mathbf{p}$  un extremo relativo si y sólo si  $F(\mathbf{s}) = f(\varphi(\mathbf{s}))$  presenta en  $\mathbf{s} = \mathbf{0}$  un extremo relativo del mismo tipo.

Cada vector  $\mathbf{u} \in T_{\mathbf{p}}(M)$  es de la forma  $\mathbf{u} = d\varphi(\mathbf{0})\mathbf{h}$  con  $\mathbf{h} \in \mathbb{R}^k$ , y demostraremos que

$$d^2F(\mathbf{0})\mathbf{h}^2 = d^2H_{\mathbf{p}}(\mathbf{p})\mathbf{u}^2 \text{ para todo } \mathbf{u} = d\varphi(\mathbf{0})\mathbf{h} \in T_{\mathbf{p}}(M)$$

Entonces aplicando los teoremas 6.11 y 6.15 a la función  $F$ , se obtendrá a) y b):

- a) Si  $f|_M$  presenta en  $\mathbf{p}$  un mínimo relativo, entonces  $d^2F(\mathbf{0})\mathbf{h}^2 \geq 0$  para todo  $\mathbf{h} \in \mathbb{R}^k$ , lo que equivale a que  $d^2H_{\mathbf{p}}(\mathbf{p})\mathbf{u}^2 \geq 0$  para todo  $\mathbf{u} \in T_{\mathbf{p}}(M)$ .
- b) Si  $d^2H_{\mathbf{p}}(\mathbf{p})\mathbf{u}^2 > 0$  para todo  $\mathbf{u} \in T_{\mathbf{p}}(M) \setminus \{\mathbf{0}\}$ , entonces  $d^2F(\mathbf{0})\mathbf{h}^2 > 0$  para todo  $\mathbf{h} \in \mathbb{R}^k \setminus \{\mathbf{0}\}$ , luego  $F$  presenta en  $\mathbf{s} = \mathbf{0}$  un mínimo relativo estricto y por lo tanto  $f|_M$  también presenta en  $\mathbf{p}$  un mínimo relativo estricto.

Si  $(\varphi_1, \varphi_2, \dots, \varphi_n)$  son las componentes de  $\varphi$ , las componentes de  $\mathbf{u} = d\varphi(\mathbf{0})\mathbf{h}$  vienen dadas por

$$u_k = d\varphi_k(\mathbf{0})\mathbf{h} = D_{\mathbf{h}}\varphi_k(\mathbf{0}), \quad k = 1, \dots, n$$

y la igualdad que queremos establecer,  $d^2F(\mathbf{0})\mathbf{h}^2 = d^2H_{\mathbf{p}}(\mathbf{p})\mathbf{u}^2$ , se escribe en la forma

$$D_{\mathbf{h}\mathbf{h}}F(\mathbf{0}) = \sum_{k,j=1}^n D_{kj}H_{\mathbf{p}}(\mathbf{p})u_ku_j$$

Observemos en primer lugar que para todo  $\mathbf{s} \in U$  y cada  $1 \leq j \leq m$  se cumple  $g_j(\varphi(\mathbf{s})) = 0$ , luego  $F(\mathbf{s}) = f(\varphi(\mathbf{s})) = H_{\mathbf{p}}(\varphi(\mathbf{s}))$ . Por consiguiente, según 7.3 tanto  $F$  como  $H_{\mathbf{p}}$  son de clase  $C^2$  en los abiertos donde están definidas.

Emprendemos el cálculo de  $D_{\mathbf{h}\mathbf{h}}F(\mathbf{0})$  a partir de la igualdad

$$dF(\mathbf{s}) = dH_{\mathbf{p}}(\varphi(\mathbf{s})) \circ d\varphi(\mathbf{s})$$

que se cumple para todo  $\mathbf{s} \in U$ . Si  $\mathbf{h} \in \mathbb{R}^k$  resulta

$$D_{\mathbf{h}}F(\mathbf{s}) = dF(\mathbf{s})\mathbf{h} = dH_{\mathbf{p}}(\varphi(\mathbf{s}))[d\varphi(\mathbf{s})\mathbf{h}] = dH_{\mathbf{p}}(\varphi(\mathbf{s}))D_{\mathbf{h}}\varphi(\mathbf{s})$$

Existe  $\delta > 0$  tal que si  $t \in \mathbb{R}$  y  $|t| < \delta$  entonces  $t\mathbf{h} \in U$ . Como  $D_{\mathbf{h}}\varphi_i(\mathbf{s})$  son las componentes de  $D_{\mathbf{h}}\varphi(\mathbf{s})$ , aplicando la igualdad anterior con  $\mathbf{s} = t\mathbf{h}$  resulta:

$$D_{\mathbf{h}}F(t\mathbf{h}) = \sum_{j=1}^n D_j H_{\mathbf{p}}(\varphi(t\mathbf{h})) D_{\mathbf{h}}\varphi_j(t\mathbf{h})$$

La derivada en  $t = 0$  de la función en el lado izquierdo de la igualdad anterior es  $D_{\mathbf{h}\mathbf{h}}F(\mathbf{0})$ . Por otra parte, la derivada de la función en el lado derecho de la igualdad se puede calcular usando la regla para derivar un producto y teniendo presente que la derivada en  $t = 0$  de  $D_j H_{\mathbf{p}}(\varphi(t\mathbf{h}))$  viene dada por  $\sum_{k=1}^n D_k D_j H_{\mathbf{p}}(\varphi(0)) D_{\mathbf{h}}\varphi_k(0)$ . Se obtiene así

$$D_{\mathbf{h}\mathbf{h}}F(\mathbf{0}) = \sum_{j=1}^n \left[ \sum_{k=1}^n D_{kj} H_{\mathbf{p}}(\mathbf{p}) D_{\mathbf{h}}\varphi_k(\mathbf{0}) \right] D_{\mathbf{h}}\varphi_j(\mathbf{0}) + \sum_{j=1}^n D_j H_{\mathbf{p}}(\mathbf{p}) D_{\mathbf{h}\mathbf{h}}\varphi_j(\mathbf{0})$$

y sustituyendo  $D_{\mathbf{h}}\varphi_k(\mathbf{0}) = u_k$ , resulta

$$D_{\mathbf{h}\mathbf{h}}F(\mathbf{0}) = \sum_{k,j=1}^n D_{kj} H_{\mathbf{p}}(\mathbf{p}) u_k u_j + \sum_{j=1}^n D_j H_{\mathbf{p}}(\mathbf{p}) D_{\mathbf{h}\mathbf{h}}\varphi_j(\mathbf{0})$$

En virtud de su definición la función  $H_{\mathbf{p}}$  cumple  $D_j H_{\mathbf{p}}(\mathbf{p}) = 0$ , si  $1 \leq j \leq n$ , luego

$$D_{\mathbf{h}\mathbf{h}}F(\mathbf{0}) = \sum_{k,j=1}^n D_{k,j} H_{\mathbf{p}}(\mathbf{p}) u_k u_j = d^2 H_{\mathbf{p}}(\mathbf{p}) \mathbf{u}^2$$

■

### 9.3. Ejercicios resueltos

**Ejercicio 9.12** *Compruebe que la curva*

$$C = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1; \quad 2z = x^2 + 2y^2\}$$

*es una subvariedad de  $\mathbb{R}^3$ , de clase  $C^\infty$  y dimensión 1. Obtenga los puntos de  $C$  donde la recta tangente es horizontal (paralela al plano  $z = 0$ ).*

SOLUCIÓN

Como las funciones  $g_1(x, y, z) = x^2 + y^2 + z^2 - 1$ ,  $g_2(x, y, z) = x^2 + 2y^2 - 2z$ , son de clase  $C^\infty$ , basta ver que para todo  $(x, y, z) \in C$  son linealmente independientes los vectores  $\nabla g_1(x, y, z) = (2x, 2y, 2z)$ ,  $\nabla g_2(x, y, z) = (2x, 4y, -2)$ , o lo que es lo mismo, que la matriz

$$\begin{pmatrix} x & y & z \\ x & 2y & -1 \end{pmatrix}$$

tiene rango 2 para cada punto  $(x, y, z) \in C$ .

El determinante formado con las dos primeras columnas no es nulo si  $xy \neq 0$ . Cuando  $x = 0$ , en los puntos de la forma  $(0, y, z) \in C$  debe ser  $y \neq 0$  y entonces no es nulo el determinante formado con las dos últimas columnas, que vale  $-y(1 + 2z)$  (obsérvese que  $z \geq 0$  para todo  $(x, y, z) \in C$ ). Análogamente, cuando  $y = 0$ , en los puntos  $(x, 0, z) \in C$  debe ser  $x \neq 0$ , y entonces no es nulo el determinante formado con las columnas primera y tercera.

Obtengamos los puntos de  $C$  donde la recta tangente es horizontal:

En un punto genérico  $\mathbf{p} = (x_0, y_0, z_0) \in C$ , los vectores  $\nabla g_1(\mathbf{p})$ ,  $\nabla g_2(\mathbf{p})$  son ortogonales, respectivamente, a los planos tangentes en ese punto a las superficies que determinan la curva,  $g_1(x, y, z) = 0$ ,  $g_2(x, y, z) = 0$ . La recta tangente a  $C$  en  $\mathbf{p}$  es la intersección de estos dos planos, y así, con el producto vectorial  $\nabla g_1(\mathbf{p}) \times \nabla g_2(\mathbf{p})$  se consigue un vector de dirección de la recta:

$$(-y_0(1 + 2z_0), x_0(1 + z_0), x_0y_0)$$

La recta tangente a  $C$  en  $\mathbf{p}$  es horizontal cuando la tercera componente de este vector es nula, es decir, cuando  $x_0y_0 = 0$ , y esto ocurre en los puntos donde la curva corta a los planos  $x = 0$ , e  $y = 0$ . Estos puntos se calculan resolviendo primero el sistema de dos ecuaciones con incógnitas  $y, z$ , que resulta al sustituir  $x = 0$  en las ecuaciones de las superficies, y luego el sistema de dos ecuaciones con las incógnitas  $x, z$  obtenido al sustituir  $y = 0$  en dichas ecuaciones. ■

**Ejercicio 9.13** Calcule los puntos de la hipérbola  $xy = 1$  que están más cerca del punto  $(2, 1)$  (con la distancia usual).

SOLUCIÓN

La distancia del punto  $(2, 1)$  a la hipérbola  $M = \{(x, y) \in \mathbb{R}^2 : xy - 1 = 0\}$  se alcanza en algún punto  $(a, b) \in M$ , ya que  $M$  es un subconjunto cerrado de  $\mathbb{R}^2$ .

El punto  $(a, b)$  que minimiza la distancia de  $(2, 1)$  a la hipérbola también minimiza su cuadrado  $f(x, y) = (x - 2)^2 + (y - 1)^2$ , que es una función más cómoda de manejar. Observemos en primer lugar que en cada  $(x, y) \in M$  la condición de ligadura  $g(x, y) = xy - 1$  tiene gradiente no nulo  $\nabla g(x, y) = (y, x)$ . luego el punto  $(a, b)$  que andamos buscando se encontrará entre los puntos estacionarios de  $f$  sobre  $M$ . Estos puntos son soluciones del sistema de tres ecuaciones con tres incógnitas

$$D_1f(x, y) - \mu D_1g(x, y) = 0, \quad D_2f(x, y) - \mu D_2g(x, y) = 0, \quad xy - 1 = 0$$

que se concretan en

$$x - \mu y = 2, \quad -\mu x + y = 1, \quad xy = 1$$

Multiplicando las dos primeras ecuaciones por  $x$ , y usando la tercera ecuación se elimina la variable  $y$  y se obtienen las dos ecuaciones

$$\mu = x^2 - 2x, \quad \mu x^2 = 1 - x$$

Eliminando  $\mu$ , se llega a la ecuación  $x^4 - 2x^3 + x - 1 = 0$  con dos soluciones reales, que se pueden calcular con un programa de cálculo simbólico (p.e DERIVE)

$$x_1 = \frac{1}{2} - \sqrt{\sqrt{5}/2 + 3/4}; \quad x_2 = \frac{1}{2} + \sqrt{\sqrt{5}/2 + 3/4}$$

Sus valores aproximados,  $x_1 = -0,866760$ ,  $x_2 = 1,86676$ , también se pueden calcular con los métodos habituales de cálculo numérico. Son las abscisas de dos puntos, uno en cada rama de la hipérbola, donde la distancia alcanza mínimos relativos. Es claro que el mínimo absoluto se alcanza en el punto  $(a, b) = (x_2, 1/x_2)$ , situado en la rama del primer cuadrante. ■

**Ejercicio 9.14** *Utilice el método de los multiplicadores de Lagrange para calcular las dimensiones de la caja de superficie mínima que encierra un volumen de 1 litro.*

SOLUCIÓN

Si  $x > 0$ ,  $y > 0$ ,  $z > 0$  son las dimensiones de la caja, expresadas en cm., se trata de minimizar el área  $S(x, y, z) = 2xy + 2xz + 2yz$  cuando las variables  $x, y, z$  están sometidas a la condición de que el volumen encerrado sea  $xyz = 1000 \text{ cm}^3$ .

Este problema ya fue resuelto en 5.39 convirtiéndolo en un problema de extremos ordinarios para la función de dos variables  $f(x, y) = S(x, y, 1000/(xy))$ , sobre el abierto  $U = \{(x, y) \in \mathbb{R}^2 : x > 0, y > 0\}$ . Ahora se trata de resolverlo como un problema de extremos condicionados, para la función  $S(x, y, z) = 2xy + 2xz + 2yz$  sobre la superficie

$$M = \{(x, y, z) \in \Omega : xyz = 1000\}, \quad \text{donde } \Omega = \{(x, y, z) : x > 0, y > 0, z > 0\}$$

Un razonamiento similar al efectuado en 5.39 permite justificar que  $S|_M$  alcanza un mínimo absoluto: El trozo de superficie

$$K = \{(x, y, z) \in M : x \geq 1, y \geq 1, z \geq 1\}$$

es cerrado y acotado (pues  $M \subset [1, 1000]^3$ ) y por lo tanto compacto.

Cuando  $(x, y, z) \in M \setminus K$  alguna de sus componentes es menor que 1, y si suponemos que  $x < 1$  se tendrá  $2yz = 2000/x > 2000$ , luego  $S(x, y, z) > 2000$ .

Como existe  $\mathbf{p} \in K$  con  $S(\mathbf{p}) < 2000$ , podemos asegurar que el mínimo absoluto de  $S$  sobre el compacto  $K$  también es el mínimo absoluto de  $S|_M$ .

Para cada  $(x, y, z) \in M$ , la función  $g(x, y, z) = xyz - 1000$  cumple

$$\nabla g(x, y, z) = (yz, xz, xy) \neq (0, 0, 0)$$

luego, en virtud de 9.10, el mínimo absoluto de  $S|_M$  se alcanza en uno de los puntos estacionarios de  $S|_M$ , es decir, en una de las soluciones del sistema de ecuaciones

$$2y + 2z - \mu yz = 0; \quad 2x + 2z - \mu xz = 0; \quad 2y + 2x - \mu xy = 0; \quad xyz = 1000.$$

Multiplicando la primera ecuación por  $x > 0$ , la segunda por  $y > 0$  y la tercera por  $z > 0$  resulta

$$2xy + 2xz = 2000\mu; \quad 2xy + 2zy = 2000\mu; \quad 2yz + 2xz = 2000\mu;$$

cuya única solución en  $\Omega$  es  $x = y = z = 10$ ,  $\mu = 2/5$ . La unicidad de la solución permite afirmar que el mínimo absoluto de  $S|_M$  se alcanza cuando  $x = y = z = 10$ .  
NOTA: Aunque ya sabemos que en  $\mathbf{p} = (10, 10, 10)$  hay un mínimo absoluto, a título ilustrativo comprobaremos que se cumple la condición suficiente de mínimo relativo vista en el teorema 9.11: Para ello formamos la función auxiliar

$$H(x, y, z) = S(x, y, z) - \frac{2}{5}(xyz - 1000)$$

Su matriz hessiana en  $(10, 10, 10)$  es

$$\begin{pmatrix} 0 & -2 & -2 \\ -2 & 0 & -2 \\ -2 & -2 & 0 \end{pmatrix}$$

y la forma cuadrática asociada a esta matriz es

$$Q_{\mathbf{p}}(\mathbf{u}) = \sum_{i,j=1}^n D_{ij}H(\mathbf{p})u_iu_j = -4(u_1u_2 + u_1u_3 + u_2u_3)$$

Tenemos que comprobar que  $Q_{\mathbf{p}}(\mathbf{u}) > 0$  para todo  $\mathbf{u} \in T_{\mathbf{p}}(M) \setminus \{\mathbf{0}\}$ .

Efectivamente, la ecuación implícita del plano tangente  $T_{\mathbf{p}}(M)$  es  $u_1 + u_2 + u_3 = 0$ , y para restringir la forma cuadrática a los vectores de este plano sustituimos  $u_3 = -u_1 - u_2$  y así obtenemos una forma cuadrática en dos variables

$$q(u_1, u_2) = -4(u_1u_2 - (u_1 + u_2)^2) = 4(u_1^2 + u_2^2 + u_1u_2) = (u_1 + \frac{1}{2}u_2)^2 + \frac{3}{4}u_2^2$$

luego  $Q(u_1, u_2, u_3) = q(u_1, u_2) > 0$  cuando  $u_1 + u_2 + u_3 = 0$ , y  $(u_1, u_2, u_3) \neq (0, 0, 0)$ .

■

**Ejercicio 9.15** Compruebe que  $\mathbf{p} = (1, 1, 2)$  es un punto estacionario de la función  $f(x, y, z) = (2 - x)yz$  sobre la superficie  $M = \{(x, y, z) : 8x - 4y^2 - z^2 = 0\}$ , y que en este punto  $f|_M$  presenta un máximo relativo.

SOLUCIÓN

Si  $g(x, y, z) = 8x - 4y^2 - z^2$ , se tiene  $\nabla g(\mathbf{p}) = (8, -8, -4)$ , y  $\nabla f(\mathbf{p}) = (-2, 2, 1)$ , luego  $\nabla f(\mathbf{p}) = -\frac{1}{4}\nabla g(\mathbf{p})$ . Por lo tanto  $\mathbf{p}$  es un punto estacionario de  $f|_M$  con multiplicador  $\mu = -1/4$ . Para estudiar la naturaleza de este punto estacionario consideramos la función  $H(x, y, z) = f(x, y, z) - \frac{1}{4}g(x, y, z)$ , cuya matriz Hessiana en el punto  $\mathbf{p}$  es

$$\begin{pmatrix} 0 & -2 & -1 \\ -2 & -2 & 1 \\ -1 & 1 & -1/2 \end{pmatrix}$$

La forma cuadrática asociada a esta matriz es

$$Q(u_1, u_2, u_3) = -2u_2^2 - \frac{1}{2}u_3^2 - 4u_1u_2 - 2u_1u_3 + 2u_2u_3$$

Esta forma cuadrática la tenemos que restringir al plano tangente  $T_{\mathbf{p}}(M)$  cuya ecuación es  $2u_1 - 2u_2 - u_3 = 0$ . Sustituyendo  $u_3 = 2(u_1 - u_2)$  obtenemos

$$q(u_1, u_2) = Q(u_1, u_2, 2(u_1 - u_2)) = -6u_1^2 - 8u_2^2 + 8u_1u_2 = Au_1^2 + 2Bu_1u_2 + Cu_2^2$$

donde  $A = -6$ ,  $B = 4$  y  $C = -8$ . Como  $AC - B^2 > 0$  y  $A < 0$ , la forma cuadrática  $q(u_1, u_2)$  es definida negativa, luego  $Q(\mathbf{u}) = q(u_1, u_2) < 0$  para cada vector tangente no nulo  $\mathbf{u} \in T_{\mathbf{p}}(M)$ . Entonces, según el teorema 9.11, podemos afirmar que  $f|_M$  presenta un máximo relativo en  $\mathbf{p} = (1, 1, 2)$ .

*Segunda solución.* También se puede justificar la última afirmación usando la técnica de la función implícita: Si  $z = z(x, y)$  es la función implícita definida por  $8x - 4y^2 - z^2 = 0$  en un entorno de  $(1, 1, 2)$ , se considera la función compuesta  $F(x, y) = f(x, y, z(x, y))$  y se comprueba que  $F$  presenta un máximo relativo en el punto  $(x_0, y_0) = (1, 1)$ . Para calcular las derivadas parciales primeras y segundas de  $F$  en  $(1, 1)$ , debemos comenzar calculando

$$z_x(1, 1), \quad z_y(1, 1), \quad z_{xy}(1, 1) = z_{yx}(1, 1), \quad z_{yy}(1, 1)$$

Aunque en el caso que nos ocupa tenemos una fórmula concreta para la función implícita,  $z = 2\sqrt{2x - y^2}$ , preferimos efectuar el cálculo de estas derivadas parciales sin usar la fórmula, mediante la técnica usual de derivación de funciones implícitas: Se calculan las derivadas parciales sucesivas, respecto a  $x$  y respecto a  $y$ , en la identidad  $8x - 4y^2 - z(x, y)^2 = 0$ , y se sustituyen los valores concretos  $x = 1$ ,  $y = 1$ ,  $z = 2$ . (Por comodidad escribimos  $z, z_x, z_y, z_{xx}, \dots$ , omitiendo el punto  $(x, y)$  en el que se evalúan estas funciones). En lo que sigue la notación  $\rightarrow$  indica que el resultado escrito a la derecha de  $\rightarrow$  se ha obtenido con la sustitución mencionada en el término de la izquierda.

$$\begin{aligned} \text{i)} \quad & 8 - 2zz_x = 0; \quad \rightarrow \quad z_x(1, 1) = 2. \\ \text{ii)} \quad & -8y - 2zz_y = 0; \quad \rightarrow \quad z_y(1, 1) = -2. \end{aligned}$$

Derivando respecto a  $x$  y respecto a  $y$  en i) y derivando respecto a  $y$  en ii) se obtiene

$$\begin{aligned} \text{iii)} \quad & (z_x)^2 + zz_{xx} = 0; \quad \rightarrow \quad z_{xx}(1, 1) = -2. \\ \text{iv)} \quad & z_y z_x + zz_{xy} = 0; \quad \rightarrow \quad z_{xy}(1, 1) = z_{yx}(1, 1) = -2. \\ \text{v)} \quad & -8 - 2(z_y)^2 - 2zz_{yy} = 0; \quad \rightarrow \quad z_{yy}(1, 1) = -4. \end{aligned}$$

Utilizando la regla de la cadena

$$D_1F(x, y) = -yz + (2 - x)yz_x; \quad D_2F(x, y) = (2 - x)z + (2 - x)yz_y$$

y sustituyendo  $x = 1$ ,  $y = 1$ ,  $z = 2$ , se llega a los valores

$$D_1F(1, 1) = -2 + 2 = 0; \quad D_2F(1, 1) = 2 - 2 = 0,$$

luego  $(1, 1)$  es un punto estacionario de  $F$ . Para ver que  $F$  presenta un máximo relativo en este punto necesitamos calcular

$$A = D_{11}F(1, 1); \quad B = D_{12}F(1, 1) = D_{21}F(1, 1); \quad C = D_{22}F(1, 1).$$

$$\begin{aligned} D_{11}F &= -2yz_x + (2-x)yz_{xx}; & \rightarrow & A = D_{11}F(1, 1) = -6; \\ D_{21}F &= -z - yz_y + (2-x)(z_x + yz_{xy}); & \rightarrow & B = D_{21}F(1, 1) = 4; \\ D_{22}F &= (2-x)(2z_y + yz_{yy}); & \rightarrow & C = D_{22}F(1, 1) = -8; \end{aligned}$$

Como  $AC - B^2 > 0$  y  $A < 0$  se sigue que  $F$  presenta un máximo relativo en  $(1, 1)$ .

*Tercera solución.* Se puede ver directamente que  $f|_M$  presenta en  $\mathbf{p}$  un máximo relativo: Lo haremos viendo que hay un entorno relativo  $M_{\mathbf{p}}$  de  $\mathbf{p}$  en  $M$  que cumple

a)  $f|_{M_{\mathbf{p}}}$  alcanza un máximo absoluto en algún  $\mathbf{q} \in M_{\mathbf{p}}$

b)  $\mathbf{p}$  es el único punto estacionario de  $f|_{M_{\mathbf{p}}}$

ya que entonces, al ser  $\mathbf{p}$  y  $\mathbf{q}$  puntos estacionarios de  $f|_{M_{\mathbf{p}}}$ , deberá ser  $\mathbf{p} = \mathbf{q}$ .

Veamos que  $M_{\mathbf{p}} = \{(x, y, z) \in M : 0 < x < 2, 0 < y, 0 < z\}$  cumple estas condiciones.

a) El conjunto  $K = \{(x, y, z) : 8x - 4y^2 - z^2 = 0, 0 \leq x \leq 2, 0 \leq y, 0 \leq z\}$  es compacto (por ser cerrado y acotado), luego existe  $\mathbf{q} \in K$  tal que

$$f(\mathbf{q}) = \max\{f(x, y, z) : (x, y, z) \in K\}$$

Es claro que  $f(\mathbf{q}) > 0$  (porque hay puntos en  $K$ , como  $(5/8, 1, 1) \in K$ , con  $f(5/8, 1, 1) > 0$ ) y también es evidente que  $f(x, y, z) = 0$  cuando  $(x, y, z) \in K$  cumple alguna de las igualdades  $y = 0$ ,  $z = 0$ ,  $x = 2$ ,  $x = 0$  (obsérvese que  $x = 0 \Rightarrow 4y^2 + z^2 = 0$ , luego  $y = 0$  y  $z = 0$ ). Se sigue de esto que  $\mathbf{q} \in M_{\mathbf{p}} \subset K$ , luego  $f(\mathbf{q})$  es el máximo absoluto de  $f|_{M_{\mathbf{p}}}$ .

Según el método de los multiplicadores de Lagrange los puntos estacionarios de  $f$  sobre  $M_{\mathbf{p}}$  se obtienen calculando las soluciones  $(x, y, z, \mu)$  del sistema

$$\begin{aligned} -yz - 8\mu &= 0; \\ (2-x)z + 8\mu y &= 0; \\ (2-x)y + 2\mu z &= 0; \\ 8x - 4y^2 - z^2 &= 0; \end{aligned}$$

que pertenecen a  $M_{\mathbf{p}}$  y es fácil ver que  $x = 1, y = 1, z = 2, \mu = -1/4$  es la única solución del sistema que cumple esta condición.

Efectivamente, multiplicando la segunda ecuación por  $y$  y la tercera por  $z$  se obtiene  $8\mu y^2 = 2\mu z^2$ . Como buscamos soluciones con  $y > 0, z > 0$  podemos asegurar, en virtud de la primera ecuación, que  $\mu \neq 0$ , luego  $8y^2 = 2z^2$  y por lo tanto  $2y = z$ . Sustituyendo  $z = 2y$  en la primera y en la última ecuación obtenemos  $8\mu = -2y^2$ ,  $x = y^2$ . Llevando estos valores a la tercera ecuación se llega a la ecuación  $0 = (2 - y^2)2y - 2y^3 = 4(y - y^3)$ , cuya única solución con  $y > 0$  es  $y = 1$ , luego  $\mathbf{p} = (1, 1, 2)$  es el único punto estacionario de  $F|_{M_{\mathbf{p}}}$ . ■



**Ejercicio 9.16** Determine los valores de los parámetros  $a, b$  para los que la función

$$f(x, y, z) = x^2 + y^2 + 2axy + 2bz$$

presenta en  $\mathbf{p} = (1, 1, 1)$  un máximo relativo sobre la esfera  $x^2 + y^2 + z^2 = 3$ .

SOLUCIÓN

La función  $g(x, y, z) = x^2 + y^2 + z^2 - 3 = 0$  cumple que el vector  $\nabla g(\mathbf{p}) = (2, 2, 2)$  no es nulo. Si  $f$  presenta un extremo relativo sobre la esfera en  $\mathbf{p}$ , según la condición necesaria de extremo condicionado, debe existir  $\mu \in \mathbb{R}$  verificando:

$$0 = D_1 f(1, 1, 1) - \mu D_1 g(1, 1, 1) = 2 + 2a - 2\mu,$$

$$0 = D_2 f(1, 1, 1) - \mu D_2 g(1, 1, 1) = 2 + 2a - 2\mu b,$$

$$0 = D_3 f(1, 1, 1) - \mu D_3 g(1, 1, 1) = 2b - 2\mu b$$

de donde se obtiene que  $\mu = b = a + 1$ . Por lo tanto, en lo que sigue suponemos que  $b = a + 1$ , y así tenemos garantizado que  $\mathbf{p}$  es un punto estacionario de  $f$  sobre la esfera. Para discutir cuando este punto estacionario es un punto de máximo relativo condicionado, consideramos la función

$$H(x, y, z) = f(x, y, z) - b(x^2 + y^2 + z^2 - 3)$$

cuya matriz Hessiana en el punto  $\mathbf{p}$  es

$$\begin{pmatrix} 2(1-b) & 2(b-1) & 0 \\ 2(b-1) & 2(1-b) & 0 \\ 0 & 0 & -2b \end{pmatrix}$$

luego, la forma cuadrática asociada es

$$Q(u_1, u_2, u_3) = 2(b-1)(u_1^2 + u_2^2 - 2u_1u_2) - 2bu_3^2$$

Una condición suficiente para que  $f$  presente en  $\mathbf{p}$  un máximo relativo condicionado es que esta forma cuadrática sea definida negativa sobre el plano tangente a la esfera en  $\mathbf{p}$ , cuya ecuación es  $u_1 + u_2 + u_3 = 0$ . Sustituyendo  $u_3 = -(u_1 + u_2)$  se obtiene una forma cuadrática en dos variables

$$q(u_1, u_2) = Q(u_1, u_2, -(u_1 + u_2)) = Au_1^2 + 2Bu_1u_2 + Cu_2^2$$

donde  $A = C = 2 - 4b$ ,  $B = 2$ , luego  $AC - B^2 = 16b(b - 1)$ .

Si  $b > 1$  es  $AC - B^2 > 0$  y  $A < 0$ , luego la forma cuadrática  $q(u_1, u_2)$  es definida negativa, lo que significa que  $Q(\mathbf{u}) < 0$  para cada cada vector no nulo  $\mathbf{u}$  tangente a la esfera en  $\mathbf{p}$ , luego  $\mathbf{p}$  es un punto de máximo relativo condicionado.

Cuando  $0 < b < 1$  se cumple  $AC - B^2 < 0$  y por lo tanto la forma cuadrática  $q(u_1, u_2)$  es indefinida, lo que significa que existen  $\mathbf{u}, \mathbf{v}$ , vectores tangentes en  $\mathbf{p}$  a la esfera tales que  $Q(\mathbf{u}) < 0 < Q(\mathbf{v})$ , luego, en este caso en el punto  $\mathbf{p}$  no hay extremo condicionado.

Cuando  $b < 0$  se cumple  $AC - B^2 > 0$ , y  $A > 0$  luego la forma cuadrática  $q(u_1, u_2)$  es definida positiva, luego  $Q(\mathbf{u}) > 0$  para cada cada vector no nulo  $\mathbf{u}$  tangente a la esfera en  $\mathbf{p}$ , y por lo tanto  $\mathbf{p}$  es un punto de mínimo relativo condicionado.

Falta discutir lo que ocurre cuando  $b = 0$  y cuando  $b = 1$ .

Si  $b = 0$  es  $a = -1$  y es claro que  $f(x, y, z) = (x - y)^2$  no presenta en  $\mathbf{p}$  un máximo relativo condicionado.

Si  $b = 1$  es  $a = 0$ , y ahora la función  $f(x, y, z) = x^2 + y^2 + 2z$ , al restringirla a la esfera sólo depende de  $z$ :  $f(x, y, z) = 3 - z^2 + 2z$ . Como  $3 - z^2 + 2z$  presenta en  $z = 1$  un máximo relativo, se sigue que, en este caso,  $f(x, y, z) = x^2 + y^2 + 2z$  presenta en el punto  $(1, 1, 1)$  un máximo relativo condicionado.

*Segunda solución:* La discusión anterior también se puede realizar con la técnica de la función implícita, estudiando cuando  $(1, 1)$  es un punto de máximo relativo ordinario para la función de dos variables reales

$$F(x, y) = f(x, y, z(x, y)) = x^2 + y^2 + 2axy + 2bz(x, y)$$

donde  $z(x, y)$  es la función implícita que define la ecuación  $g(x, y, z) = 0$  en un entorno del punto  $(1, 1, 1)$ . (Aunque en este caso tenemos una fórmula explícita  $z(x, y) = \sqrt{3 - x^2 - y^2}$ , haremos las cuentas que siguen, sin usarla).

Para este estudio necesitamos calcular la matriz Hessiana de  $F$  en  $(1, 1)$ .

$$F_x(x, y) = 2x + 2ay + 2bz_x(x, y); \quad F_y(x, y) = 2y + 2ax + 2bz_x(x, y);$$

$$F_{xx}(x, y) = 2 + 2bz_{xx}(x, y); \quad F_{yy}(x, y) = 2 + 2bz_{yy}(x, y);$$

$$F_{xy}(x, y) = 2a + 2bz_{xy}(x, y); \quad F_{yx}(x, y) = 2a + 2bz_{xy}(x, y);$$

Calculamos las derivadas parciales  $z_{xx}(1, 1)$ ,  $z_{yy}(1, 1)$ ,  $z_{xy}(1, 1)$ , con la técnica de derivación implícita: Derivando respecto a  $x$  en la identidad  $x^2 + y^2 + z(x, y)^2 = 3$  resulta  $2x + 2z(x, y)z_x(x, y) = 0$  y cuando  $x = y = z = 1$  se obtiene  $z_x(1, 1) = -1$ . Análogamente se calcula  $z_y(1, 1) = -1$ . Derivando respecto a  $x$  y respecto a  $y$  en la identidad  $2x + 2z(x, y)z_x(x, y) = 0$ , y sustituyendo los valores particulares  $x = y = z = 1$  se obtiene  $z_{xx}(1, 1) = -2$ . Análogamente  $z_{xy}(1, 1) = z_{yx}(1, 1) = -1$ , y  $z_{yy}(1, 1) = -2$ . Sustituyendo arriba estos valores resulta:

$$F_{xx}(1, 1) = F_{yy}(1, 1) = 2 - 4b, \quad F_{xy}(1, 1) = F_{yx}(1, 1) = -2$$

luego el determinante Hessiano de  $F$  en  $(1, 1)$  vale

$$\Delta(b) = \begin{vmatrix} 2 - 4b & -2 \\ -2 & 2 - 4b \end{vmatrix} = 16b(b - 1)$$

y se acaba la discusión como antes. ■

**Ejercicio 9.17** Sea  $Q(x, y) = Ax^2 + 2Bxy + Cy^2$  una forma cuadrática no idénticamente nula. Demuestre que

$$m_1 = \max\{Q(x, y) : x^2 + y^2 = 1\} \quad y \quad m_2 = \min\{Q(x, y) : x^2 + y^2 = 1\}$$

son las soluciones de la ecuación de segundo grado  $\begin{vmatrix} A - \mu & B \\ B & C - \mu \end{vmatrix} = 0$ .

Deduzca de ello:

- i)  $Q$  es indefinida si y sólo si  $AC - B^2 < 0$ .
- ii)  $Q$  es definida positiva si y sólo si  $AC - B^2 > 0$  y  $A > 0$ .
- iii)  $Q$  es definida negativa si y sólo si  $AC - B^2 > 0$  y  $A < 0$ .

Si  $\mathbf{u}_1 = (x_1, y_1)$ , y  $\mathbf{u}_2 = (x_2, y_2)$  son puntos de la circunferencia  $x^2 + y^2 = 1$  donde  $Q$  alcanza el máximo y el mínimo absoluto,  $m_1 = Q(x_1, y_1)$ ,  $m_2 = Q(x_2, y_2)$ , demuestre que  $\mathbf{u}_1$  y  $\mathbf{u}_2$  son ortogonales. ¿Cuál es la interpretación geométrica de estos resultados.?

(Indicación : La aplicación lineal  $L(x, y) = (Ax + By, Bx + Cy)$  es simétrica,  $\langle L(\mathbf{w}) \mid \mathbf{w} \rangle = \langle \mathbf{w} \mid L(\mathbf{w}) \rangle$ , y verifica  $Q(\mathbf{w}) = \langle \mathbf{w} \mid L(\mathbf{w}) \rangle$ .)

## SOLUCIÓN

$M = \{(x, y) : x^2 + y^2 - 1 = 0\}$  es compacto, luego existen  $(x_1, y_1) \in M$ ,  $(x_2, y_2) \in M$  donde la función continua  $Q$  alcanza sus extremos absolutos:

$$m_1 = Q(x_1, y_1), \quad m_2 = Q(x_2, y_2).$$

En todo  $(x, y) \in M$  la función  $g(x, y) = x^2 + y^2 - 1$  cumple  $\nabla g(x, y) \neq (0, 0)$ , luego  $(x_1, y_1)$ ,  $(x_2, y_2)$  son puntos estacionarios de  $Q|_M$ , es decir, existen  $\mu_1, \mu_2 \in \mathbb{R}$  tales que  $(x_1, y_1, \mu_1)$ ,  $(x_2, y_2, \mu_2)$  son soluciones del sistema de tres ecuaciones

$$D_1 Q(x, y) - \mu D_1 g(x, y) = 0; \quad D_2 Q(x, y) - \mu D_2 g(x, y) = 0; \quad g(x, y) = 0;$$

que se concretan en

$$Ax + By = \mu x; \quad Bx + Cy = \mu y; \quad x^2 + y^2 = 1.$$

luego toda solución  $(x, y, \mu)$  de este sistema cumple  $L(x, y) = \mu(x, y)$ .

Multiplicando la primera ecuación por  $x$ , la segunda por  $y$ , y utilizando la tercera ecuación se obtiene que también se cumple  $Q(x, y) = \mu$ . En particular

$$\text{a) } L(x_1, y_1) = \mu_1(x_1, y_1); \quad L(x_2, y_2) = \mu_2(x_2, y_2);$$

$$\text{b) } m_1 = Q(x_1, y_1) = \mu_1; \quad m_2 = Q(x_2, y_2) = \mu_2.$$

Cuando  $\mu = \mu_1$  y cuando  $\mu = \mu_2$ , el sistema lineal

$$(A - \mu)x + By = 0; \quad Bx + (C - \mu)y = 0;$$

tiene, respectivamente, las soluciones no triviales  $(x_1, y_1)$ ,  $(x_2, y_2)$ , luego su determinante es nulo, luego  $\mu_1, \mu_2$  son las soluciones de la ecuación del enunciado:

$$\mu^2 - (A + C)\mu + (AC - B^2) = 0.$$

De las dos soluciones reales de esta ecuación la mayor es el máximo  $\mu_1 = m_1$ , y la menor el mínimo  $\mu_2 = m_2$ . Es claro que

- i)  $Q$  es indefinida si y sólo si  $m_2 < 0 < m_1$ .
- ii)  $Q$  es definida positiva si y sólo si  $m_2 > 0$ .
- iii)  $Q$  es definida negativa si y sólo si  $m_1 < 0$ .

Teniendo en cuenta que  $m_1 + m_2 = A + C$ ,  $m_1 m_2 = AC - B^2$ , se obtiene:

a)  $AC - B^2 < 0$  si y sólo si  $m_1$  y  $m_2$  tienen distinto signo, lo que ocurre si y sólo si  $Q$  es indefinida.

b)  $AC - B^2 > 0$  si y sólo si  $m_1$  y  $m_2$  tienen el mismo signo, lo que ocurre si y sólo si  $Q$  es definida positiva o definida negativa.  $Q$  será definida positiva (resp. negativa) cuando  $m_1 + m_2 = A + C > 0$  (resp.  $< 0$ ) lo que ocurre si y sólo si  $A > 0$  (resp.  $A < 0$ ). Obsérvese que  $A$  y  $C$  tienen el mismo signo porque  $AC > B^2 \geq 0$ .

Para lo que sigue, consideramos una base ortonormal de  $\mathbb{R}^2$ ,  $(\mathbf{v}_1, \mathbf{v}_2)$  con  $\mathbf{v}_1 = \mathbf{u}_1$ . Para demostrar que  $\langle \mathbf{u}_1 | \mathbf{u}_2 \rangle = 0$  basta ver que cuando se expresa  $\mathbf{u}_2$  respecto a esta base,  $\mathbf{u}_2 = \alpha \mathbf{v}_1 + \beta \mathbf{v}_2$ , se cumple que  $\alpha = 0$ . Obsérvese que  $\alpha^2 + \beta^2 = \langle \mathbf{u}_2 | \mathbf{u}_2 \rangle = 1$ .

Utilizando que  $Q(\mathbf{w}) = \langle \mathbf{w} | L(\mathbf{w}) \rangle$  y la bilinealidad del producto escalar

$$\begin{aligned} Q(\mathbf{u}_2) &= \langle \mathbf{u}_2 | L(\mathbf{u}_2) \rangle = \langle \alpha \mathbf{v}_1 + \beta \mathbf{v}_2 | \alpha L(\mathbf{v}_1) + \beta L(\mathbf{v}_2) \rangle = \\ &= \alpha^2 \langle \mathbf{v}_1 | L(\mathbf{v}_1) \rangle + \alpha \beta \langle \mathbf{v}_1 | L(\mathbf{v}_2) \rangle + \alpha \beta \langle \mathbf{v}_2 | L(\mathbf{v}_1) \rangle + \beta^2 \langle \mathbf{v}_2 | L(\mathbf{v}_2) \rangle \end{aligned}$$

Utilizando la simetría de  $L$  y que  $L(\mathbf{v}_1) = \mu_1 \mathbf{v}_1$ , (véase a)) resulta

$$\langle \mathbf{v}_1 | L(\mathbf{v}_2) \rangle = \langle L(\mathbf{v}_1) | \mathbf{v}_2 \rangle = \langle \mathbf{v}_2 | L(\mathbf{v}_1) \rangle = \mu_1 \langle \mathbf{v}_2 | \mathbf{v}_1 \rangle = 0$$

luego

$$Q(\mathbf{u}_2) = \alpha^2 Q(\mathbf{v}_1) + \beta^2 Q(\mathbf{v}_2)$$

De aquí se deduce que  $Q(\mathbf{v}_2) < Q(\mathbf{v}_1)$ : Sabemos que  $Q(\mathbf{v}_2) \leq m_1 = Q(\mathbf{v}_1)$ , pero no se puede dar la igualdad porque en ese caso se tendría

$$m_2 = Q(\mathbf{u}_2) = \alpha^2 Q(\mathbf{v}_1) + \beta^2 Q(\mathbf{v}_1) = Q(\mathbf{v}_1) = m_1$$

y  $Q$  sería constante.

Con la desigualdad  $Q(\mathbf{v}_2) < Q(\mathbf{v}_1)$  se obtiene que  $\alpha = 0$ : Si fuese  $\alpha \neq 0$  se tendría  $\alpha^2 Q(\mathbf{v}_2) < \alpha^2 Q(\mathbf{v}_1)$  es decir  $(1 - \beta^2)Q(\mathbf{v}_2) < \alpha^2 Q(\mathbf{v}_1)$ , luego

$$Q(\mathbf{v}_2) < \alpha^2 Q(\mathbf{v}_1) + \beta^2 Q(\mathbf{v}_2) = Q(\mathbf{u}_2)$$

y esta desigualdad es falsa porque  $Q(\mathbf{u}_2)$  es el mínimo absoluto de  $Q$  sobre  $M$ .

*Interpretación geométrica:*

Para cada  $c > 0$  las curvas de nivel  $N_c = \{(x, y) : Q(x, y) = c\}$  son simétricas respecto al origen. Obtengamos la ecuación de  $N_c$  respecto a la base ortonormal  $\{\mathbf{u}_1, \mathbf{u}_2\}$ . Si  $(s, t)$  son las coordenadas de  $(x, y)$  respecto a esta base, usando que  $L(\mathbf{u}_1) = \mu_1 \mathbf{u}_1$  y  $L(\mathbf{u}_2) = \mu_2 \mathbf{u}_2$  podemos escribir

$$Q(x, y) = Q(s\mathbf{u}_1 + t\mathbf{u}_2) = \langle s\mathbf{u}_1 + t\mathbf{u}_2 | sL(\mathbf{u}_1) + tL(\mathbf{u}_2) \rangle =$$

$$= \langle s\mathbf{u}_1 + t\mathbf{u}_2 \mid \mu_1 s\mathbf{u}_1 + \mu_2 t\mathbf{u}_2 \rangle = \mu_1 s^2 + \mu_2 t^2.$$

luego  $(x, y) = s\mathbf{u}_1 + t\mathbf{u}_2$  pertenece a  $N_c$  si y sólo si  $\mu_1 s^2 + \mu_2 t^2 = c$ .

Si  $AC - B^2 > 0$  y  $A > 0$  sabemos que  $0 < \mu_1 < \mu_2$ , y la ecuación de  $N_c$  se puede escribir en la forma

$$a^2 s^2 + b^2 t^2 = c \quad \text{con } a = \sqrt{\mu_1}, \quad b = \sqrt{\mu_2}$$

luego sólo hay curvas de nivel para  $c > 0$  y estas son elipses que van aumentando de tamaño conforme  $c > 0$  crece. La función  $Q$  toma el valor  $c$  en algún punto de  $M$  si y sólo si  $M \cap N_c \neq \emptyset$ , luego el máximo de  $Q$  sobre  $M$  será el mayor valor de  $c$  para el cual podamos asegurar que  $M \cap N_c \neq \emptyset$ . Esto ocurre precisamente cuando la elipse  $N_c$  es tangente, por fuera, a la circunferencia  $M$ ; los dos puntos de tangencia diametralmente opuestos,  $\mathbf{u}_1, -\mathbf{u}_1 \in M$ , donde se alcanza el máximo  $m_1 = \max Q(M)$  determinan la dirección común del eje mayor de las elipses. Análogamente el mínimo valor de  $Q$  sobre  $M$  es el valor de  $c$  para el cual la elipse  $N_c$  es tangente, por dentro, a la circunferencia  $M$ . El mínimo  $m_2 = \min Q(M)$  se alcanza en dos puntos de tangencia diametralmente opuestos  $\mathbf{u}_2, -\mathbf{u}_2$ , que determinan la dirección común del eje menor de las elipses. Evidentemente  $\mathbf{u}_1$  y  $\mathbf{u}_2$  son ortogonales.

Cuando  $AC - B^2 > 0$  y  $A < 0$  la interpretación geométrica se reduce a la acabamos de hacer, considerando la función  $-Q$ .

Si  $AC - B^2 < 0$  sabemos que  $\mu_1 < 0 < \mu_2$ , y la ecuación de  $N_c$  se puede escribir en la forma

$$a^2 s^2 - b^2 t^2 = c \quad \text{con } a = \sqrt{-\mu_1}, \quad b = \sqrt{\mu_2}$$

Ahora las curvas de nivel  $N_c$  son hipérbolas que llenan el abierto

$$A = \{(s, t) : (as + bt)(as - bt) \neq 0\}.$$

Si  $c > 0$  las hipérbolas ocupan  $A^+ = \{(s, t) : (as + bt)(as - bt) > 0\}$ , formado por dos de las regiones angulares opuestas que determinan las rectas  $as + bt = 0$ ,  $as - bt = 0$ . Cuando  $c < 0$  las hipérbolas ocupan,  $A^- = \{(s, t) : (as + bt)(as - bt) < 0\}$ , formado por las otras dos regiones angulares limitadas por las mismas rectas. Estas rectas son las asíntotas comunes de todas las hipérbolas, y conforme  $c$  se aproxima a 0 las hipérbolas van quedando más cerca de las asíntotas.

Ahora el máximo de  $Q$  sobre  $M$  es positivo. Es el mayor valor de  $c$  para el cual  $M \cap N_c \neq \emptyset$ . Esto ocurre precisamente para la única hipérbola  $N_c$ , con  $c > 0$ , que es tangente a la circunferencia  $M$ . Hay dos puntos de tangencia, diametralmente opuestos,  $\mathbf{u}_1, -\mathbf{u}_1 \in M$ , donde  $Q$  alcanza el máximo  $m_1 = \max Q(M)$ . Análogamente el mínimo de  $Q$  sobre  $M$  es negativo. Es el único  $c < 0$  que hace que la hipérbola  $N_c$  sea tangente a la circunferencia  $M$ . Hay dos puntos de tangencia diametralmente opuestos,  $\mathbf{u}_2, -\mathbf{u}_2 \in M$ , que están en una recta perpendicular a la determinada por  $\mathbf{u}_1$  y  $-\mathbf{u}_1$ . ■

**Ejercicio 9.18** Calcule los extremos absolutos de la forma cuadrática  $Q(x, y, z) = x^2 + y^2 + z^2 + 4xy$  sobre la circunferencia

$$C = \{(x, y, z) : x^2 + y^2 + z^2 = 1, \quad x + y + 2z = 0\}$$

Estudie los planos de la forma  $z = ax - ay$  sobre los que la forma cuadrática es definida positiva.

SOLUCIÓN

Los extremos absolutos de  $Q$  sobre  $C$  se pueden obtener de varias formas

a) En la forma usual, como un problema de extremos con dos ligaduras:  $C$  es compacto (cerrado y acotado) luego la función continua  $Q$  alcanza sobre  $C$  un mínimo absoluto  $\mu_1$ , y un máximo absoluto  $\mu_2$ , es decir, existen  $\mathbf{p}_1, \mathbf{p}_2 \in C$  tales que

$$\mu_1 = \min\{Q(\mathbf{p}) : \mathbf{p} \in C\} = Q(\mathbf{p}_1); \quad \mu_2 = \max\{Q(\mathbf{p}) : \mathbf{p} \in C\} = Q(\mathbf{p}_2);$$

Se comprueba fácilmente que en todo punto  $(x, y, z) \in C$  los vectores  $(2x, 2y, 2z)$ ,  $(1, 1, 2)$  son linealmente independientes, luego  $C$  es una subvariedad de dimensión 1 y clase  $C^\infty$ . Se sigue que  $\mathbf{p}_1$  y  $\mathbf{p}_2$  son puntos estacionarios de  $Q$  sobre  $C$  y según el método de los multiplicadores de Lagrange, estos puntos son soluciones del sistema

$$\begin{aligned} i) \quad & 2x + 4y - 2\mu x - \lambda = 0 \\ ii) \quad & 2y + 4x - 2\mu y - \lambda = 0 \\ iii) \quad & 2z - 2\mu z - 2\lambda = 0 \\ iv) \quad & x^2 + y^2 + z^2 = 1 \\ v) \quad & x + y - 2z = 0 \end{aligned}$$

Multiplicando la primera ecuación por  $y$ , la segunda por  $x$  y sumando miembro a miembro resulta  $4(y^2 - x^2) = \lambda(y - x)$  de donde se obtiene que, o bien  $(y - x) = 0$  o bien  $4(y + x) = \lambda$ . Con la primera alternativa, y utilizando las ecuaciones iv) y v) se obtienen los puntos estacionarios

$$\mathbf{a} = (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3}), \quad -\mathbf{a} = (-1/\sqrt{3}, -1/\sqrt{3}, -1/\sqrt{3})$$

Con la segunda alternativa, sustituyendo  $\lambda = 4(x + y)$  en i) y ii), se obtiene  $x = -y$ , y utilizando otra vez iv) y v) salen los puntos estacionarios

$$\mathbf{b} = (1/\sqrt{3}, -1/\sqrt{3}, 0), \quad -\mathbf{b} = (-1/\sqrt{3}, 1/\sqrt{3}, 0)$$

Se concluye así que el máximo absoluto es  $4/3 = Q(\mathbf{a}) = Q(-\mathbf{a})$  y que el mínimo absoluto es  $-1 = Q(\mathbf{b}) = Q(-\mathbf{b})$ .

b) También podemos empezar como en a) y terminar como en la demostración de la proposición I.2 (estamos considerando un caso particular de la situación considerada allí). Ahora, en vez de resolver completamente el sistema que proporciona los puntos estacionarios, podemos razonar como en I.2 y obtener que si  $\mu, \lambda$  son los multiplicadores asociados a un punto estacionario  $\mathbf{p}$  entonces  $\mu = Q(\mathbf{p})$ . Se sigue de esto que el mínimo y el máximo de  $Q$  sobre  $C$  vienen dados por las soluciones  $\mu_1 = -1, \mu_2 = 7/5$  de la ecuación de segundo grado

$$\begin{vmatrix} (1 - \mu) & 2 & 0 & 1 \\ 2 & (1 - \mu) & 0 & 1 \\ 0 & 0 & (1 - \mu) & 2 \\ 1 & 1 & 2 & 0 \end{vmatrix} = 0$$

c) Otra alternativa para resolver la primera parte del problema consiste en reducirlo a uno de extremos condicionados de una función de dos variables con una sola ligadura: Para  $(x, y, z) \in C$  el valor  $Q(x, y, z) = 1 + 4xy$  no depende de  $z$ , luego el problema es equivalente al de obtener los extremos absolutos de la función de dos variables  $1 + 4xy$  sobre la elipse  $x^2 + y^2 + (x + y)^2/4 = 1$  (la proyección de  $C$  sobre el plano  $(x, y)$ ). Probablemente este es el camino más breve.

Para discutir cuando la forma cuadrática es definida positiva sobre el plano  $z = ax - ay$  proponemos dos alternativas:

- i) Restringir  $Q(x, y, z)$  al plano  $z = ax - ay$ . Se obtiene así una forma cuadrática en dos variables  $q(x, y) = Ax^2 + Cy^2 + 2Bxy$  con  $A = C = 1 + a^2$ ,  $B = 2 - a^2$ , que es definida positiva cuando  $A > 0$  y  $AC - B^2 > 0$ , lo que ocurre si y sólo si  $a^2 > 1/2$ .
- ii) Acudir otra vez a la proposición I.2, que nos dice que la forma cuadrática es  $Q$  es definida positiva sobre el plano  $z = ax - ay$  cuando son positivas las dos soluciones de la ecuación de segundo grado

$$\begin{vmatrix} (1 - \mu) & 2 & 0 & a \\ 2 & (1 - \mu) & 0 & -a \\ 0 & 0 & (1 - \mu) & -1 \\ a & -a & -1 & 0 \end{vmatrix} = 0$$

Resolviendo la ecuación se obtiene fácilmente que ambas soluciones son positivas si y sólo si  $a^2 > 1/2$ . ■

## 9.4. Ejercicios propuestos

◇ **9.4.1** Utilice el método de los multiplicadores de Lagrange para obtener, en cada caso, los extremos absolutos y relativos de la función  $f(x, y)$  sobre la elipse  $E$ :

- a)  $f(x, y) = x + y^2$ ;  $E = \{(x, y) : 2x^2 + y^2 = 1\}$ .  
 b)  $f(x, y) = x^2 + y^2 - 4xy + 20x + 20y$ ;  $E = \{(x, y) : x^2 + y^2 + xy = 12\}$ .  
 c)  $f(x, y) = xy - 4x - 4y$ ;  $E = \{(x, y) : x^2 + y^2 + xy = 12\}$ .

Interprete los resultados considerando las curvas de nivel  $N_c = \{(x, y) : f(x, y) = c\}$ .

◇ **9.4.2** Utilice el método de los multiplicadores de Lagrange para obtener, en cada caso, los extremos absolutos de  $f(x, y)$  sobre el compacto  $K \subset \mathbb{R}^2$ :

- a)  $f(x, y) = x - x^2 - y^2$ ;  $K = \{(x, y) : x \geq 0, x^2 + y^2 \leq 1\}$ .  
 b)  $f(x, y) = \text{sen}(xy)$ ;  $K = \{(x, y) : x^2 + y^2 \leq 1\}$ .  
 c)  $f(x, y) = x^2 + y^2 + xy + x$   $K = \{(x, y) : x^2 + y^2 \leq 1\}$ .

◇ **9.4.3** Utilice el método de los multiplicadores de Lagrange para obtener:

- a) La mínima distancia entre la recta  $y - x + 5 = 0$  y la parábola  $y = x^2$ .  
 b) La máxima y mínima distancia del origen a la elipse  $5x^2 + 6xy + 5y^2 = 8$ .

◇ **9.4.4** Si  $g : \mathbb{R}^3 \rightarrow \mathbb{R}$  es continua y  $S = \{(x, y, z) : g(x, y, z) = 0\}$  no es vacío demuestre para cada  $\mathbf{p} \in \mathbb{R}^3 \setminus S$  hay un punto  $\mathbf{q} \in S$  cuya distancia a  $\mathbf{p}$  es mínima:

$$\|\mathbf{p} - \mathbf{q}\|_2 = \min\{\|\mathbf{p} - \mathbf{x}\|_2 : \mathbf{x} \in S\}$$

Si  $g$  es de clase  $C^1$  y  $\nabla g(x, y, z) \neq 0$  para todo  $(x, y, z) \in S$  justifique que el vector  $\mathbf{p} - \mathbf{q}$  es normal a la superficie  $S$  en el punto  $\mathbf{q}$ . Calcule la mínima distancia de  $\mathbf{p} = (1, 1, 1)$  a la superficie  $x^2 + y^2 - z^2 - 2x + 2 = 0$ .

◇ **9.4.5** Utilice el método de los multiplicadores de Lagrange para obtener los puntos de la superficie  $z^2 - xy - 1 = 0$  más próximos al origen.

◇ **9.4.6** Obtenga los puntos de la curva  $x^2 - xy + y^2 - z^2 = 1$ ,  $x^2 + y^2 = 1$ , que están más cerca del origen.

◇ **9.4.7** Obtenga los puntos de  $C = \{(x, y, z) : 2z + x^2 + y^2 = 16; x + y = 4\}$  que están más cerca del origen. Compruebe que  $C_1 = \{(x, y, z) \in C : x \geq 0, y \geq 0, z \geq 0\}$  es compacto y obtenga los puntos de  $C_1$  que están más lejos del origen.

◇ **9.4.8** Sean  $S, T \subset \mathbb{R}^3$  subvariedades diferenciables disjuntas de clase  $C^1$ , tales que existen  $\mathbf{p} \in S$ ,  $\mathbf{q} \in T$  verificando  $\|\mathbf{p} - \mathbf{q}\|_2 = \min\{\|\mathbf{x} - \mathbf{y}\|_2 : \mathbf{x} \in S, \mathbf{y} \in T\}$ . Demuestre que el vector  $\mathbf{p} - \mathbf{q}$  es ortogonal a  $S$  y  $T$  en  $\mathbf{p}$  y  $\mathbf{q}$  respectivamente.

◇ **9.4.9** Utilice el método de los multiplicadores de Lagrange para obtener la mínima distancia entre la superficie  $z = x^2 + y^2$  y el plano  $x + 2y - z = 4$ .



◇ **9.4.10** Se supone que el plano  $P = \{(x, y, z) : Ax + By + Cz = D\}$  no corta a la superficie  $S = \{(x, y, z) : x^2 + y^2 - z = 0\}$ .

Justifique que la distancia entre un punto genérico  $\mathbf{p} \in P$  y un punto genérico  $\mathbf{q} \in S$  alcanza un mínimo absoluto en un único par de puntos  $\mathbf{a} \in P, \mathbf{b} \in S$ . Calcule  $\mathbf{a}$  y un vector normal a  $S$  en  $\mathbf{b}$ . Obtenga la mínima distancia entre  $S$  y  $P = \{(x, y, z) : x + 2y - z = 4\}$ .

◇ **9.4.11** Calcule la máxima y la mínima distancia del origen a la elipse

$$E_a = \{(x, y, z) : x^2 + y^2 \leq 1, x + y + az = 1\}$$

◇ **9.4.12** De todos los paralelepípedos inscritos en el elipsoide  $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ , obtenga el de mayor volumen. Calcule también el mínimo volumen encerrado por un plano tangente al elipsoide y los planos  $x = 0, y = 0, z = 0$ .

◇ **9.4.13** De todos los planos tangentes al trozo de elipsoide

$$S = \{(x, y, z) : x^2 + 2y^2 + 3z^2 = 1, x > 0, y > 0, z > 0\}$$

determine el que forma con los planos  $x = 0, y = 0, z = 0$  un tetraedro de volumen mínimo.

◇ **9.4.14** Determine el elipsoide  $E(a, b, c) = \{(x, y, z) : \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1\}$  que pasa por  $\mathbf{p} = (1, 2, 3)$  y encierra mayor volumen.

◇ **9.4.15** Utilice el método de los multiplicadores de Lagrange para determinar las dimensiones de una caja rectangular sin tapa, con superficie de  $16 \text{ m}^2$ , que encierra un volumen máximo. Justifique la existencia de la caja de volumen máximo.

◇ **9.4.16** Determine las dimensiones del bote cilíndrico (con tapa) de mayor superficie que se puede inscribir en una esfera de radio  $R$ .

◇ **9.4.17** Obtenga los extremos relativos de la función  $3x - 4y + 2z$  sobre la superficie  $x^3 - 2y^2 + z^2 = 0$ .

◇ **9.4.18** Obtenga los extremos absolutos y relativos de las siguientes funciones sobre las superficies que se indican:

- |    |                         |                       |                               |
|----|-------------------------|-----------------------|-------------------------------|
| a) | $x - 2y + 2z,$          | sobre la esfera       | $x^2 + y^2 + z^2 = 1.$        |
| b) | $x + y + z,$            | sobre el elipsoide    | $x^2 + 2y^2 + 3z^2 = 1.$      |
| c) | $x + y + z,$            | sobre el hiperboloide | $x^2 + y^2 - 2z^2 = 6.$       |
| d) | $x^2 + y^2 + z^2,$      | sobre la superficie   | $x^2 + y^2 - z^2 + 2xy = 16.$ |
| e) | $x^2 + y^2 + 8xy + 10z$ | sobre la esfera       | $x^2 + y^2 + z^2 = 3.$        |

◇ **9.4.19** Estudie si la función  $x^2 + y^2 + z^2$  posee en el punto  $(0, 1, 0)$  un extremo condicionado por la ligadura  $z^2 + 2xyz + y^2 + x^3 = 1$

◇ **9.4.20** Calcule el máximo y el mínimo absoluto de la función  $x^2 + y^2 + z$  sobre la curva  $C = \{(x, y, z) : x^2 + y^2 + z^2 = 4, x^2 + y^2 = 2x\}$ . Compruebe que existe  $\mathbf{p} \in C$  tal que  $C_0 = C \setminus \{\mathbf{p}\}$  es subvariedad diferenciable de  $\mathbb{R}^3$ .

◇ **9.4.21** a) Compruebe que  $C = \{(x, y, z) \in \mathbb{R}^3 : 3z^2 - 2xy - 1 = 0, x^2 + y^2 - 2z^2 = 0\}$  es una subvariedad diferenciable de clase  $C^\infty$  y dimensión 1.

b) Justifique que el punto  $\mathbf{p} = (1, 1, 1)$  posee un entorno abierto  $U_p \subset \mathbb{R}^3$  tal que el trozo de curva  $C \cap U_p$  admite una parametrización de clase  $C^\infty$  la forma  $\gamma(t) = (t, y(t), z(t))$ . Calcule los vectores  $\mathbf{u} = \gamma'(1)$  y  $\mathbf{v} = \gamma''(1)$ .

c) Sea  $f(x, y, z) = x + y^3 - 2y + z^2$ . Calcule  $D_{\mathbf{u}}f(\mathbf{p})$ . Compruebe que  $f|_C$  presenta un extremo relativo en  $\mathbf{p}$  y determine su naturaleza.

◇ **9.4.22** Compruebe que  $C = \{(x, y, z) \in \mathbb{R}^3 : x + y = 1, x^2 + y^2 + z^2 = 1\}$  es una subvariedad diferenciable de  $\mathbb{R}^3$  de clase  $C^\infty$  y dimensión 1, y que  $\mathbf{p} = (0, 1, 0)$  es un punto estacionario de  $f(x, y, z) = x(y + z)$  sobre la curva  $C$ . Estudie si  $f|_C$  presenta en  $\mathbf{p}$  un extremo relativo.

◇ **9.4.23** Compruebe que la curva  $C = \{(x, y, z) : x^2 - y^2 - 1 = 0, 2x + z - 1 = 0\}$  es una subvariedad diferenciable de clase  $C^\infty$  y dimensión 1, y escriba la ecuación de la recta tangente a  $C$  en un punto genérico  $(x_0, y_0, z_0) \in C$ .

Calcule los puntos estacionarios de  $f(x, y, z) = x + y^2 + z$  sobre  $C$  y compruebe que  $f$  presenta mínimos relativos en todos ellos.

Estudie si alguno de los mínimos relativos es un mínimo absoluto.

◇ **9.4.24** Compruebe que la curva  $C = \{(x, y, z) : x^2 + y^2 = z^2, y = \sqrt{3}x + 2\}$  es una subvariedad diferenciable de clase  $C^\infty$  y dimensión 1. Determine los extremos relativos y absolutos de la función  $f(x, y, z) = x^2 + y^2 + (z - 1)^2$  sobre la curva. Interpretación geométrica del resultado.

◇ **9.4.25** Compruebe que  $\mathbf{p} = (\sqrt{2}/4, -1/2, 0)$ ,  $\mathbf{q} = (\sqrt{2}/4, 1/2, 0)$  son puntos estacionarios de la función  $f(x, y, z) = 2x^2 + y^2 + z^2 - xy$  sobre el elipsoide  $S = \{(x, y, z) : 4x^2 + 2y^2 + z^2 = 1\}$ , y estudie su naturaleza.

¿Hay más puntos estacionarios?. Justifique la respuesta.

◇ **9.4.26** Compruebe que  $\mathbf{p} = (1/3, -1/3, -5/3)$  y  $\mathbf{q} = (1, 1, 1)$  son puntos estacionarios de la función  $x^2 + y^2 + 8xy + 10z$  sobre la esfera  $x^2 + y^2 + z^2 = 3$ . Determine su naturaleza.

◇ **9.4.27** Calcule los extremos relativos y absolutos de  $f(x, y, z) = z - (1 + x^2 + y^2)e^z$  sobre la elipse  $E = \{(x, y, z) : x^2 + y^2 = 1, x + z = 1\}$ . Si  $\mathbf{p} \in E$  es el punto donde  $f|_E$  alcanza el mínimo absoluto, obtenga la ecuación de la recta tangente a  $E$  en  $\mathbf{p}$ .

◇ **9.4.28** Compruebe que  $C = \{(x, y, z) : x^2 + y^2 - 2 = 0, x - yz + 1 = 0\}$  es una subvariedad de dimensión 1 y clase  $C^\infty$ .

a) Obtenga un vector  $\mathbf{w} = (w_1, w_2, w_3)$  tangente a  $C$  en  $\mathbf{p} = (1, 1, 2)$ .

b) Para la función  $f(x, y, z) = xy - ayz + z$  calcule la derivada  $D_{\mathbf{w}}f(\mathbf{p})$ .

c) Determine  $a \in \mathbb{R}$  para que  $\mathbf{p}$  sea punto estacionario de  $f|_C$  y en ese caso estudie su naturaleza (e.d. si es punto máximo o mínimo relativo).

◇ **9.4.29** En cada caso calcule los extremos absolutos sobre  $K$ , de la función dada:

- a)  $x^2 + y^2 + z^2 + x + y + z$ ;  $K = \{x^2 + y^2 + z^2 \leq 4, z \leq 1\}$ .  
 b)  $x + y + z$ ;  $K = \{0 \leq z \leq 2, x^2 + y^2 - 2z^2 \leq 6\}$ .  
 c)  $x + y - \sqrt{2}z + x^2 + y^2 + z^2$ ;  $K = \{0 \leq z \leq 1, x^2 + y^2 = 2z^2\}$ .  
 d)  $x^2 + y^2 + z^2 + xy + x - 2z$ ;  $K = \{x^2 + y^2 \leq 1, |z| \leq 2\}$ .  
 e)  $x^2 - y^2 + z^2 - z$ ;  $K = \{x^2 + y^2 + z^2 \leq 1, x^2 + y^2 \geq z^2, z \geq 0\}$ .  
 f)  $x + y + z$ ;  $K = \{x^2 + y^2 \leq z \leq 2\}$ .  
 g)  $x^2 + y^2 + z^2 + x + y - \sqrt{2}z$ ;  $K = \{x^2 + y^2 - z^2 = 0; 0 \leq z \leq 1\}$

◇ **9.4.30** Sea  $(a_{ij})_{1 \leq i, j \leq n}$  una matriz simétrica, y  $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , la aplicación lineal asociada  $A(\mathbf{x})_i = \sum_{j=1}^n a_{ij}x_j$ .

- i) Demuestre que el máximo valor  $\mu$  del polinomio  $p(\mathbf{x}) = \sum_{i,j=1}^n a_{ij}x_i x_j$  sobre la esfera  $\{(x_1, x_2, \dots, x_n) : x_1^2 + x_2^2 + \dots + x_n^2 = 1\}$  es el máximo autovalor de  $A$ .  
 ii) En el caso  $n = 2$  se considera un polinomio  $p(x, y) = ax^2 + 2bxy + cy^2$  tal que  $C = \{(x, y) \in \mathbb{R}^2 : p(x, y) = 1\}$  no es vacío. Demuestre que  $\mu > 0$  y que la mínima distancia de la cónica  $C$  al origen es  $1/\sqrt{\mu}$ . Si  $C$  es una elipse, ¿Cuál es la máxima distancia de  $C$  al origen?

◇ **9.4.31** Sea  $(a_{ij})_{1 \leq i, j \leq 3}$  una matriz simétrica. Demuestre que los extremos de  $f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2$  sujetos a las ligaduras

$$\sum_{1 \leq i, j \leq 3} a_{ij}x_i x_j = 1; \quad a_1x_1 + a_2x_2 + a_3x_3 = 0$$

son  $1/\mu_1$  y  $1/\mu_2$  donde  $\mu_1, \mu_2$  son las soluciones de la ecuación de segundo grado

$$\begin{vmatrix} a_1 & a_2 & a_3 & 0 \\ a_{11} - \mu & a_{12} & a_{13} & a_1 \\ a_{21} & a_{22} - \mu & a_{23} & a_2 \\ a_{31} & a_{32} & a_{33} - \mu & a_3 \end{vmatrix} = 0$$

Mediante una interpretación geométrica del resultado justifique que esta ecuación ha de tener dos raíces reales positivas.

# Capítulo 10

## Integral de Riemann

*Funciones integrables Riemann en un intervalo y propiedades de la integral. Integración sobre conjuntos medibles Jordan. Conjuntos de contenido nulo de medida nula. Caracterización de las funciones integrables.*

En este capítulo y en el siguiente se desarrollan las técnicas básicas de cálculo integral para funciones de varias variables, y se muestran las aplicaciones clásicas de la integral al cálculo de áreas y volúmenes.

La integral de Riemann proporciona una introducción rápida y elemental al problema del cálculo efectivo de integrales, áreas y volúmenes, y permite definir de forma rigurosa una amplia clase de conjuntos del plano o del espacio ordinario que tienen asignada un área o un volumen. Esta clase de conjuntos, llamados medibles Jordan, incluye a los recintos acotados que se pueden describir geoméricamente como intersección de figuras geométricas elementales (conos, cilindros, esferas, etc).

Además de las aplicaciones geométricas usuales al cálculo de áreas y volúmenes se menciona en este capítulo la noción de función de densidad de un sólido y se describen las aplicaciones de las integrales al cálculo de masas, centros de masas y momentos de inercia de sólidos cuya distribución de masa no uniforme está descrita mediante una función de densidad.

En este capítulo, dedicado esencialmente a los fundamentos teóricos de la integral de Riemann, se demuestra el clásico teorema de Lebesgue que caracteriza las funciones integrables Riemann mediante el conjunto de sus discontinuidades.

Hay que advertir que la integral de Riemann es una noción insuficiente como instrumento teórico para el Análisis Matemático avanzado (Análisis de Fourier, Análisis Funcional) donde se requiere la noción más general de integral de Lebesgue.

### 10.1. Funciones integrables Riemann

**Notaciones y terminología.** Una subdivisión  $p$  de  $[a, b] \subset \mathbb{R}$ , es una sucesión finita creciente  $t_0 < t_1 < \cdots < t_m$ , con  $t_0 = a$ ,  $t_m = b$ . Usaremos la notación  $\Delta(p)$  para la familia de los intervalos cerrados determinados por  $p$ , es decir

$$\Delta(p) = \{[t_0, t_1], [t_1, t_2], [t_2, t_3], \cdots [t_{m-1}, t_m]\}$$

En lo que sigue  $\mathcal{P}([a, b])$  designará la colección de todas las subdivisiones de  $[a, b]$ . Si  $p, p' \in \mathcal{P}([a, b])$ , y  $p'$  contiene todos los puntos de  $p$ , se dice que  $p'$  es más fina que  $p$ , y se escribe  $p' \geq p$ . Así se tiene definida en  $\mathcal{P}([a, b])$  una relación de orden parcial con la cual  $\mathcal{P}([a, b])$  es un conjunto dirigido:

Dadas  $p', p'' \in \mathcal{P}([a, b])$  existe  $p \in \mathcal{P}([a, b])$ , tal que  $p \geq p'$ , y  $p \geq p''$

En lo que sigue llamaremos rectángulo o intervalo cerrado  $n$ -dimensional a un conjunto  $A \subset \mathbb{R}^n$ , de la forma  $A = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$ , que salvo mención expresa de lo contrario, se supondrá no degenerado, es decir con  $a_i < b_i$ , para  $1 \leq i \leq n$ . Análogamente se definen los rectángulos o intervalos abiertos  $n$ -dimensionales:  $U = (a_1, b_1) \times (a_2, b_2) \times \cdots \times (a_n, b_n)$ .

El volumen  $n$ -dimensional de los rectángulos abiertos o cerrados se define como el producto de las longitudes de sus lados:

$$v(A) = v(U) = (b_1 - a_1) \times (b_2 - a_2) \times \cdots \times (b_n - a_n).$$

Cuando sea conveniente hacer explícita la dimensión  $n$  que se está considerando, en lugar de  $v$ , escribiremos  $v_n$  para designar el volumen  $n$ -dimensional.

Una subdivisión del rectángulo cerrado  $A \subset \mathbb{R}^n$  es una  $n$ -pla,  $p = (p_1, p_2, \cdots, p_n)$ , donde  $p_i \in \mathcal{P}([a_i, b_i])$ ,  $1 \leq i \leq n$ . El conjunto de las subdivisiones de  $A$

$$\mathcal{P}(A) := \mathcal{P}([a_1, b_1]) \times \mathcal{P}([a_2, b_2]) \times \cdots \times \mathcal{P}([a_n, b_n])$$

también se puede dirigir por refinamiento: Se dice que  $p' = (p'_1, p'_2, \cdots, p'_n) \in \mathcal{P}(A)$  es más fina que  $p = (p_1, p_2, \cdots, p_n) \in \mathcal{P}(A)$ , y se escribe  $p' \geq p$ , cuando  $p'_i \geq p_i$ , para  $1 \leq i \leq n$ .

Cada  $p = (p_1, p_2, \cdots, p_n) \in \mathcal{P}(A)$  determina una colección finita de rectángulos cerrados que no se solapan (esto significa que sus interiores son disjuntos):

$$\Delta(p) = \{J_1 \times J_2 \times \cdots \times J_n : J_k \in \Delta(p_k), 1 \leq k \leq n\}$$

Si  $p, p' \in \mathcal{P}(A)$ , y  $p'$  es más fina que  $p$ , cada  $S \in \Delta(p)$  es unión de los rectángulos  $S' \in \Delta(p')$  que están contenidos en  $S$ , y es fácil ver que

$$v(S) = \sum_{S' \in \Delta(p'), S' \subset S} v(S').$$

**Integral inferior e integral superior. Funciones integrables.** Sea  $f : A \rightarrow \mathbb{R}$  una función acotada definida en el rectángulo cerrado  $A \subset \mathbb{R}^n$ ,  $|f(\mathbf{x})| \leq C$  para todo  $\mathbf{x} \in A$ . Entonces, para cada  $S \subset A$  se pueden definir

$$M(f, S) = \sup\{f(\mathbf{x}) : \mathbf{x} \in S\} \leq C, \quad m(f, S) = \inf\{f(\mathbf{x}) : \mathbf{x} \in S\} \geq -C.$$

A cada subdivisión  $p \in \mathcal{P}(A)$  se le asocian las sumas superior e inferior de Darboux:

$$S(f, p) = \sum_{S \in \Delta(p)} M(f, S)v(S); \quad s(f, p) = \sum_{S \in \Delta(p)} m(f, S)v(S).$$

Es inmediato que  $s(f, p) \leq S(f, p)$  para cada  $p \in \mathcal{P}(A)$ .

**Lema 10.1** Si  $p' \in \mathcal{P}(A)$  es más fina que  $p \in \mathcal{P}(A)$ , se cumple

$$S(f, p') \leq S(f, p), \quad s(f, p') \geq s(f, p),$$

y si  $p, q \in \mathcal{P}(A)$  son subdivisiones arbitrarias,  $s(f, p) \leq S(f, q)$ .

DEM: Fijado  $S \in \Delta(p)$ , si  $S' \in \Delta(p')$ , y  $S' \subset S$ , se verifica  $M(f, S') \leq M(f, S)$ , luego,  $M(f, S')v(S') \leq M(f, S)v(S')$ . Sumando las desigualdades que corresponden a los  $S' \in \Delta(p')$  contenidos en  $S$ , resulta

$$\sum_{S' \subset S} M(f, S')v(S') \leq M(f, S) \sum_{S' \subset S} v(S') = M(f, S)v(S)$$

Volviendo a sumar cuando  $S$  recorre  $\Delta(p)$ , se obtiene

$$S(f, p') = \sum_{S \in \Delta(p)} \left[ \sum_{S' \subset S} M(f, S')v(S') \right] \leq \sum_{S \in \Delta(p)} M(f, S)v(S) = S(f, p)$$

Análogamente se demuestra que  $s(f, p') \geq s(f, p)$ .

Finalmente, si  $p, q \in \mathcal{P}(A)$  son subdivisiones arbitrarias, considerando una subdivisión  $p' \in \mathcal{P}(A)$ , más fina que  $p$  y que  $q$ , aplicando las desigualdades que acabamos de establecer resulta  $s(f, p) \leq s(f, p') \leq S(f, p') \leq S(f, q)$ . ■

Fijado  $q \in \mathcal{P}(A)$ , el número  $S(f, q)$  es cota superior del conjunto de números reales  $\{s(f, p) : p \in \mathcal{P}(A)\}$ , y se puede definir la integral inferior

$$\underline{\int}_A f = \sup\{s(f, p) : p \in \mathcal{P}(A)\} \leq S(f, q)$$

Para cada  $q \in \mathcal{P}(A)$  se cumple  $\underline{\int}_A f \leq S(f, q)$ , y se puede definir la integral superior

$$\overline{\int}_A f = \inf\{S(f, q) : q \in \mathcal{P}(A)\} \geq \underline{\int}_A f$$

**Definición 10.2** Una función acotada  $f : A \rightarrow \mathbb{R}$ , definida en un rectángulo cerrado  $A \subset \mathbb{R}^n$ , es integrable Riemann sobre  $A$  cuando  $\underline{\int}_A f = \overline{\int}_A f$ . En este caso se define su integral como el valor común

$$\int_A f = \underline{\int}_A f = \overline{\int}_A f$$

Obsérvese que la integrabilidad de  $f$  significa que existe un único número real  $I = \int_A f(\mathbf{x}) \, d\mathbf{x}$ , tal que  $s(f, p) \leq I \leq S(f, p)$ , para todo  $p \in \mathcal{P}(A)$ .

En lo que sigue  $\mathcal{R}(A)$  designará el conjunto de las funciones integrables Riemann  $f : A \rightarrow \mathbb{R}$ . Si  $B \subset A$  es un subrectángulo cerrado y  $f|_B \in \mathcal{R}(B)$ , se dice que  $f$  es integrable sobre  $B$ , y en lugar de  $f|_B \in \mathcal{R}(B)$ ,  $\int_B f|_B$ , se escribe  $f \in \mathcal{R}(B)$ ,

$\int_B f$ , respectivamente. Si  $f$  es integrable sobre  $A$ , a veces es conveniente utilizar la notación habitual que hace explícitas las variables:

$$\int_A f = \int_A f(\mathbf{x}) \, d\mathbf{x} = \int_A f(x_1, x_2, \dots, x_n) \, dx_1 \, dx_2 \cdots dx_n.$$

Cuando  $n = 1$ , y  $a < b$  se define  $\int_a^b f(x) \, dx = -\int_b^a f(x) \, dx = \int_{[a,b]} f(x) \, dx$ .

La siguiente proposición, que es consecuencia directa de las definiciones, expresa en una forma bastante útil la condición de que  $f$  sea integrable, sin mencionar explícitamente las integrales superior e inferior.

**Proposición 10.3** *Una condición necesaria y suficiente para que una función acotada  $f : A \rightarrow \mathbb{R}$  sea integrable sobre el rectángulo cerrado  $A \subset \mathbb{R}^n$ , es que se verifique: Para cada  $\epsilon > 0$  existe  $p_\epsilon \in \mathcal{P}(A)$  tal que  $S(f, p_\epsilon) - s(f, p_\epsilon) < \epsilon$ .*

DEM: Si se cumple la condición del enunciado, para cada  $\epsilon > 0$ , se verifica

$$0 \leq \overline{\int_A f} - \underline{\int_A f} \leq S(f, p_\epsilon) - s(f, p_\epsilon) < \epsilon, \quad \text{luego} \quad \overline{\int_A f} = \underline{\int_A f}.$$

Recíprocamente, si  $f$  es integrable, su integral  $\int_A f$  es el extremo superior de las sumas  $s(f, p)$ , y el extremo inferior de las sumas  $S(f, p)$  luego, dado  $\epsilon > 0$ , existen  $p', p'' \in \mathcal{P}(A)$  verificando

$$s(f, p') \geq \int_A f - \epsilon/2, \quad S(f, p'') \leq \int_A f + \epsilon/2.$$

Si  $p_\epsilon \in \mathcal{P}(A)$  es una subdivisión más fina que  $p'$  y que  $p''$ , se cumple

$$S(f, p_\epsilon) - s(f, p_\epsilon) \leq S(f, p'') - s(f, p') \leq \int_A f + \epsilon/2 - \left( \int_A f - \epsilon/2 \right) = \epsilon$$

■

**Teorema 10.4** *Toda función continua  $f : A \rightarrow \mathbb{R}$  en un rectángulo cerrado  $A \subset \mathbb{R}^n$ , es integrable Riemann.*

DEM: Como  $A$  es compacto,  $f$  es uniformemente continua (3.24), luego para cada  $\epsilon > 0$  existe  $\delta > 0$ , tal que

$$\mathbf{x}, \mathbf{y} \in A, \quad \|\mathbf{x} - \mathbf{y}\| < \delta \Rightarrow |f(\mathbf{x}) - f(\mathbf{y})| < \epsilon/v(A)$$

Sea  $p \in \mathcal{P}(A)$  tal que  $\text{diam}(S) < \delta$  para todo  $S \in \Delta(p)$ . La función continua  $f$  alcanza en cada rectángulo compacto  $S \in \Delta(p)$  un máximo y un mínimo absolutos (3.16), es decir, existen  $\mathbf{x}_S, \mathbf{y}_S \in S$ , tales que

$$M(f, S) = f(\mathbf{x}_S), \quad m(f, S) = f(\mathbf{y}_S)$$

Como  $\|\mathbf{x}_S - \mathbf{y}_S\| < \delta$ , se verifica  $0 \leq M(f, S) - m(f, S) = f(\mathbf{x}_S) - f(\mathbf{y}_S) \leq \epsilon/v(A)$ , luego

$$S(f, p) - s(f, p) = \sum_{S \in \Delta(p)} [M(f, S) - m(f, S)]v(S) \leq \frac{\epsilon}{v(A)} \sum_{S \in \Delta(p)} v(S) = \epsilon$$

Aplicando la proposición 10.3 se concluye que  $f$  es integrable. ■

NOTA: Hay otras formas alternativas de definir la integral de Riemann que comentamos a continuación: Si  $A \subset \mathbb{R}^n$  es un rectángulo cerrado, sea  $\Pi(A)$  la familia de los pares  $\pi = (p, \tau(p))$ , donde  $p \in \mathcal{P}(A)$  y  $\tau(p) = \{\tau_S : S \in \Delta(p)\}$  es una colección finita de puntos de  $A$  tal que  $\tau_S \in S$  para cada  $S \in \Delta(p)$ . Diremos que  $\pi = (p, \tau(p)) \in \Pi(A)$  es más fina que  $\pi' = (p', \tau(p'))$  cuando  $p$  es más fina que  $p'$ , y en ese caso escribiremos  $\pi \geq \pi'$ .

Si  $f : A \rightarrow \mathbb{R}$  es acotada, a cada  $\pi = (p, \tau(p))$  se le asocia la suma de Riemann

$$\Sigma(f, \pi) = \sum_{S \in \Delta(p)} f(\tau_S)v(S)$$

Es fácil ver que  $f$  es integrable Riemann sobre  $A$ , con integral  $I$ , si y sólo si para cada  $\epsilon > 0$  existe  $\pi_\epsilon \in \Pi(A)$ , tal que toda  $\pi \in \Pi(A)$  más fina que  $\pi_\epsilon$  cumple  $|\Sigma(f, \pi) - I| < \epsilon$ . Esto significa que la red  $(\Sigma(f, \pi))_{\pi \in \Pi(A)}$  converge hacia  $I$ , cuando  $\Pi(A)$  está dirigido por refinamiento (es decir, por la relación de orden  $\geq$  definida anteriormente).

Para  $\pi = (p, \tau(p)) \in \Pi(A)$  se define  $\text{diam}(\pi) = \text{máx}\{\text{diam}(S) : S \in \Delta(p)\}$ . Se puede demostrar que  $f$  es integrable Riemann sobre  $A$ , con integral  $I$ , si y sólo si para cada  $\epsilon > 0$  existe  $\delta > 0$ , tal que toda  $\pi \in \Pi(A)$ , con  $\text{diam}(\pi) < \delta$ , cumple  $|\Sigma(f, \pi) - I| < \epsilon$ . (Cuando  $f$  es continua el cumplimiento de esta condición está implícito en la demostración del teorema 10.4). En el lenguaje de la teoría de redes esta caracterización se expresa diciendo que  $f$  es integrable Riemann con integral  $I$  si y sólo si la red  $(\Sigma(f, \pi))_{\pi \in \Pi(A)}$  converge hacia  $I$  cuando  $\Pi(A)$  está dirigido mediante la relación de orden:  $\pi \succeq \pi'$  si  $\text{diam}(\pi) \leq \text{diam}(\pi')$ .

En términos de sucesiones, esto significa que para cada sucesión  $\pi_n \in \Pi(A)$ , con  $\lim_n \text{diam}(\pi_n) = 0$ , la sucesión  $\Sigma(f, \pi_n)$  converge hacia  $I$  (Véase [12], vol. III, pág 29, y [5], pág 34, para el caso  $n = 1$ ).

**Propiedades de la integral.** Dada una función  $f : A \rightarrow \mathbb{R}$ , su parte positiva y su parte negativa se definen por  $f^+(\mathbf{x}) = \text{máx}\{f(\mathbf{x}), 0\}$ ,  $f^-(\mathbf{x}) = -\text{mín}\{f(\mathbf{x}), 0\}$ , Obsérvese que  $f = f^+ - f^-$ , y que  $|f| = f^+ + f^-$ .

**Proposición 10.5** *Sea  $A \subset \mathbb{R}^n$  un rectángulo cerrado y  $\mathcal{R}(A)$  el conjunto de las funciones  $f : A \rightarrow \mathbb{R}$  que son integrables Riemann.*

- a)  $\mathcal{R}(A)$  es un espacio vectorial sobre  $\mathbb{R}$ , y la integral es lineal: Si  $f, g \in \mathcal{R}(A)$ , y  $\alpha, \beta \in \mathbb{R}$ , entonces  $\alpha f + \beta g \in \mathcal{R}(A)$  y  $\int_A (\alpha f + \beta g) = \alpha \int_A f + \beta \int_A g$
- b) Si  $f, g \in \mathcal{R}(A)$ , y  $f \leq g$  entonces  $\int_A f \leq \int_A g$ .



- c) Si  $f \in \mathcal{R}(A)$ , entonces  $f^+, f^-, |f| \in \mathcal{R}(A)$ , y  $|\int_A f| \leq \int_A |f|$ .
- d) Si  $f, g \in \mathcal{R}(A)$ , entonces  $fg \in \mathcal{R}(A)$ .
- e) Si  $f : A \rightarrow \mathbb{R}$  es acotada y  $p \in \mathcal{P}(A)$  entonces  $f$  es integrable sobre  $A$  si y sólo si  $f|_S$  es integrable sobre cada  $S \in \Delta(p)$ , y en este caso

$$\int_A f(\mathbf{x}) \, d\mathbf{x} = \sum_{S \in \Delta(p)} \int_S f(\mathbf{x}) \, d\mathbf{x}$$

- f) Si  $f \in \mathcal{R}(A)$ , y  $\mathbf{u} \in \mathbb{R}^n$ , sea  $A_{\mathbf{u}} = \mathbf{u} + A$ . Entonces  $f_{\mathbf{u}}(\mathbf{x}) = f(\mathbf{x} - \mathbf{u})$  es integrable sobre  $A_{\mathbf{u}}$  y se verifica  $\int_A f = \int_{A_{\mathbf{u}}} f_{\mathbf{u}}$ .

DEM: a) Sean  $p', p'' \in \mathcal{P}(A)$ , y  $p \in \mathcal{P}(A)$  más fina que  $p'$  y  $p''$ . Para cada  $S \in \Delta(p)$ , es  $M(f+g, S) \leq M(f, S) + M(g, S)$ , luego  $S(f+g, p) \leq S(f, p) + S(g, p)$  y se sigue que  $\overline{\int_A}(f+g) \leq S(f+g, p) \leq S(f, p) + S(g, p) \leq S(f, p') + S(g, p'')$ . Considerando el extremo inferior de las sumas  $S(f, p')$ , y luego el extremo inferior de las sumas  $S(g, p'')$  se obtiene:  $\overline{\int_A}(f+g) \leq \overline{\int_A}f + \overline{\int_A}g$ . Un razonamiento análogo conduce a  $\underline{\int_A}f + \underline{\int_A}g \geq \underline{\int_A}(f+g)$ .

Si  $f, g$  son integrables sobre  $A$ , en virtud de las desigualdades anteriores,

$$\underline{\int_A}(f+g) = \overline{\int_A}(f+g) = \int_A f + \int_A g$$

es decir,  $f+g$  es integrable sobre  $A$ , y  $\int_A(f+g) = \int_A f + \int_A g$ .

Si  $\alpha \geq 0$ , se cumple  $S(\alpha f, p) = \alpha S(f, p)$ ,  $s(\alpha f, p) = \alpha s(f, p)$ , mientras que para  $\alpha < 0$ , se tiene  $S(\alpha f, p) = \alpha s(f, p)$ ,  $s(\alpha f, p) = \alpha S(f, p)$ . De aquí se deduce

$$\overline{\int_A}(\alpha f) = \alpha \overline{\int_A} f; \quad \underline{\int_A}(\alpha f) = \alpha \underline{\int_A} f \quad \text{si } \alpha \geq 0.$$

$$\overline{\int_A}(\alpha f) = \alpha \underline{\int_A} f; \quad \underline{\int_A}(\alpha f) = \alpha \overline{\int_A} f \quad \text{si } \alpha < 0.$$

En ambos casos se concluye que  $f \in \mathcal{R}(A) \Rightarrow \alpha f \in \mathcal{R}(A)$ , con  $\int_A(\alpha f) = \alpha \int_A f$ .

- b) Es consecuencia de  $S(f, p) \leq S(g, p)$ , que se cumple para todo  $p \in \mathcal{P}(A)$ .

c) Observemos que  $M(f^+, S) - m(f^+, S) \leq M(f, S) - m(f, S)$  (es evidente cuando  $f$  no cambia de signo en  $S$ , y cuando  $f$  cambia de signo en  $S$  basta tener en cuenta que  $M(f^+, S) = M(f, S)$ ,  $m(f^+, S) = 0$ , y  $m(f, S) < 0$ ). Esta desigualdad conduce a  $S(f^+, p) - s(f^+, p) \leq S(f, p) - s(f, p)$  de donde se sigue, usando 10.3, que  $f \in \mathcal{R}(A) \Rightarrow f^+ \in \mathcal{R}(A)$ . Utilizando la propiedad a) se concluye que  $f^- = f^+ - f$  y  $|f| = f^+ + f^-$  son integrables. Finalmente, en virtud de b) y de la desigualdad  $-|f| \leq f \leq |f|$ , resulta  $-\int_A |f| \leq \int_A f \leq \int_A |f|$ .

d) Empezamos con el caso  $0 \leq f = g \in \mathcal{R}(A)$ . Como  $M(f^2, S) = M(f, S)^2$ , y  $m(f^2, S) = m(f, S)^2$ , se tiene

$$\begin{aligned} M(f^2, S) - m(f^2, S) &= [M(f, S) + m(f, S)][M(f, S) - m(f, S)] \leq \\ &\leq 2C[M(f, S) - m(f, S)] \end{aligned}$$

donde  $C = M(f, A)$ . Multiplicando esta desigualdad por  $v(S)$ , y sumando cuando  $S$  recorre  $\Delta(p)$ , resulta

$$S(f^2, p) - s(f^2, p) \leq 2C[S(f, p) - s(f, p)]$$

Como esta desigualdad es válida para cada  $p \in \mathcal{P}(A)$ , con 10.3 se obtiene que  $f \in \mathcal{R}(A) \Rightarrow f^2 \in \mathcal{R}(A)$ .

Cuando  $0 \leq f, g \in \mathcal{R}(A)$ , como  $fg = \frac{1}{2}[(f+g)^2 - f^2 - g^2]$ , según lo que se acaba de demostrar y la propiedad a) resulta  $fg \in \mathcal{R}(A)$ . Finalmente, cuando  $f, g \in \mathcal{R}(A)$  son arbitrarias, usando las propiedades a) y c) y lo que ya se ha demostrado resulta  $fg = f^+g^+ + f^-g^- - f^+g^- - f^-g^+ \in \mathcal{R}(A)$ .

e) La demostración se reduce al caso en que  $\Delta(p) = \{A_1, A_2\}$ , consta de dos rectángulos. Dadas sendas subdivisiones  $p' \in \mathcal{P}(A_1)$ ,  $p'' \in \mathcal{P}(A_2)$ , podemos obtener una subdivisión  $q \in \mathcal{P}(A)$ , que induce en  $A_1$ , y en  $A_2$ , subdivisiones  $q'$ , y  $q''$ , más finas que  $p'$ , y  $p''$ , respectivamente. Entonces

$$\begin{aligned} S(f, p') + S(f, p'') &\geq S(f, q') + S(f, q'') = S(f, q) \geq \overline{\int_A} f. \\ s(f, p') + s(f, p'') &\leq s(f, q') + s(f, q'') = s(f, q) \leq \underline{\int_A} f. \end{aligned}$$

Si  $f \in \mathcal{R}(A_1)$ , y  $f \in \mathcal{R}(A_2)$ , considerando el extremo inferior de las sumas  $S(f, p')$ , y luego el de las sumas  $S(f, p'')$ , se obtiene

$$\int_{A_1} f + \int_{A_2} f \geq \overline{\int_A} f$$

Análogamente, considerando el extremo superior de las sumas  $s(f, p')$ , y luego el de las sumas  $s(f, p'')$ , resulta

$$\int_{A_1} f + \int_{A_2} f \leq \underline{\int_A} f$$

Se concluye así que  $f \in \mathcal{R}(A)$ , y que  $\int_A f = \int_{A_1} f + \int_{A_2} f$ .

Recíprocamente, si  $f \in \mathcal{R}(A)$ , en virtud de 10.3, dado  $\epsilon > 0$ , existe  $q \in \mathcal{R}(A)$ , tal que  $S(f, q) - s(f, q) < \epsilon$ . No hay inconveniente en suponer que  $q$  es más fina que  $p$ . Entonces, podemos considerar las subdivisiones  $q' \in \mathcal{P}(A_1)$ , y  $q'' \in \mathcal{P}(A_2)$ , que  $q$  induce en  $A_1$ , y en  $A_2$ , respectivamente, para las que se cumple

$$S(f, q') - s(f, q') < \epsilon, \quad S(f, q'') - s(f, q'') < \epsilon,$$

luego, en virtud de 10.3,  $f$  es integrable sobre  $A_1$ , y sobre  $A_2$ .

f) La traslación  $\mathbf{x} \rightarrow \mathbf{x} + \mathbf{u}$ , establece una biyección natural  $p \rightarrow p'$ , entre las subdivisiones  $p \in \mathcal{P}(A)$ , y las subdivisiones  $p' \in \mathcal{P}(A_{\mathbf{u}})$ , en la que a cada rectángulo  $S \in \Delta(p)$ , le corresponde el trasladado  $S' = \mathbf{u} + S \in \Delta(p')$ . Es evidente que

$$M(f, S) = M(f_{\mathbf{u}}, S'), \quad m(f, S) = m(f_{\mathbf{u}}, S'), \quad v(S) = v(S')$$

luego  $S(f, p) = S(f_{\mathbf{u}}, p')$ , y  $s(f, p) = s(f_{\mathbf{u}}, p')$ . De estas igualdades se desprende el resultado. ■

El siguiente resultado, que ha quedado establecido en la demostración del apartado a) de la proposición anterior, conviene hacerlo explícito para que sirva de referencia en algunas de las demostraciones que siguen.

**Proposición 10.6** Si  $f, g : A \rightarrow \mathbb{R}$  son funciones acotadas sobre el rectángulo cerrado  $A \subset \mathbb{R}^n$ , se verifica,  $\int_A (f + g) \leq \int_A f + \int_A g$ .

El objetivo del siguiente lema es establecer, con los recursos disponibles en este momento, un resultado parcial que sirve para justificar una afirmación que se hace al comienzo de la siguiente sección.

**Lema 10.7** Si  $f : A \rightarrow \mathbb{R}$  es acotada en un intervalo cerrado  $A \subset \mathbb{R}^n$ , y  $f(\mathbf{x}) = 0$  para todo  $\mathbf{x} \in A^\circ$  entonces  $f$  es integrable y  $\int_A f = 0$ .

DEM: Basta demostrarlo cuando  $f \geq 0$  (ya que el caso general se reduce a este considerando la descomposición  $f = f^+ - f^-$ ).

Para cada  $\epsilon > 0$  sea  $A_\epsilon \subset A^\circ$  un rectángulo cerrado tal que  $v(A) - v(A_\epsilon) < \epsilon/C$ , donde  $C > M(f, A)$  es una cota superior de  $f$ . Sea  $q \in \mathcal{P}(A)$  tal que  $A_\epsilon \in \Delta(q)$ . Como  $M(f, A_\epsilon) = 0$ , se tiene

$$S(f, q) = \sum_{S \in \Delta(q), S \neq A_\epsilon} M(f, S)v(S) \leq \sum_{S \in \Delta(q), S \neq A_\epsilon} Cv(S) = C(v(A) - v(A_\epsilon)) < \epsilon$$

luego,  $0 \leq \int_A f \leq \overline{\int_A f} \leq S(f, q) \leq \epsilon$ . Como  $\epsilon > 0$  es arbitrario se obtiene que  $f$  es integrable sobre  $A$  con  $\int_A f = 0$ . ■

## 10.2. Conjuntos medibles Jordan

Un conjunto acotado  $E \subset \mathbb{R}^n$  se dice que es *medible Jordan* si está contenido en un rectángulo cerrado  $A \subset \mathbb{R}^n$  tal que la función característica  $\chi_E$  es integrable Riemann sobre  $A$ . En este caso, si  $A' \subset \mathbb{R}^n$  es otro rectángulo cerrado que contiene a  $E$ , es fácil ver que  $\chi_E$  también es integrable Riemann sobre  $A'$ , con la misma integral  $\int_{A'} \chi_E = \int_A \chi_E$ .

Efectivamente,  $B = A \cap A'$ , es un rectángulo cerrado y existen subdivisiones  $p \in \mathcal{P}(A)$ ,  $p' \in \mathcal{P}(A')$  tales que  $B \in \Delta(p)$ , y  $B \in \Delta(p')$ . Como  $E$  está contenido

en  $B$ , para cada  $S \in \Delta(p)$  con  $S \neq B$  la función  $\chi_E$  es nula sobre  $S^\circ$ , y según el lema 10.7 es integrable sobre  $S$ , con integral nula. Usando 10.5 e), se obtiene

$$\chi_E \in \mathcal{R}(A) \Leftrightarrow \chi_E \in \mathcal{R}(B), \quad \text{y} \quad \int_A \chi_E = \int_B \chi_E$$

Análogamente se razona con  $A'$  y se concluye que

$$\chi_E \in \mathcal{R}(A) \Leftrightarrow \chi_E \in \mathcal{R}(B) \Leftrightarrow \chi_E \in \mathcal{R}(A') \quad \text{y} \quad \int_A \chi_E = \int_B \chi_E = \int_{A'} \chi_E.$$

En lo que sigue denotaremos por  $\mathcal{M}_n$  la familia de los subconjuntos medibles Jordan de  $\mathbb{R}^n$ . Las consideraciones anteriores ponen de manifiesto que para cada  $E \in \mathcal{M}_n$ , el valor de la integral  $\int_A \chi_E$  es el mismo para todo rectángulo cerrado  $A$  que contenga a  $E$ , por lo que, en lo que sigue, podemos omitir el rectángulo  $A \supset E$ , y escribir  $\int \chi_E = \int_A \chi_E$ .

El número  $c(E) = \int \chi_E$  se llama *contenido de Jordan* ( $n$ -dimensional) de  $E$ . Cuando sea preciso especificar la dimensión  $n$  que se está considerando escribiremos  $c_n(E)$ , en lugar de  $c(E)$ . Así  $c_3$  mide volúmenes de sólidos en  $\mathbb{R}^3$ ,  $c_2$  mide áreas de regiones planas en  $\mathbb{R}^2$ , etc.

La siguiente proposición es consecuencia inmediata de las definiciones y de las propiedades básicas de la integral establecidas en 10.5.

### Proposición 10.8

- a) Si  $E, F \in \mathcal{M}_n$ , y  $E \subset F$ , entonces  $c(E) \leq c(F)$ .
- b) Si  $E, F \in \mathcal{M}_n$ , entonces  $E \cap F \in \mathcal{M}_n$ ,  $E \cup F \in \mathcal{M}_n$ ,  $E \setminus F \in \mathcal{M}_n$ , y  $c(E \cup F) \leq c(E) + c(F)$ . Si  $c(E \cap F) = 0$ , se cumple  $c(E \cup F) = c(E) + c(F)$ .
- c) Si  $E_1, E_2, \dots, E_m \in \mathcal{M}_n$ , entonces  $E = \bigcup_{k=1}^m E_k \in \mathcal{M}_n$ , y  $c(E) \leq \sum_{k=1}^m c(E_k)$ , y si los conjuntos son disjuntos se cumple la igualdad.
- d) Si  $A \subset \mathbb{R}^n$  es un rectángulo cerrado, y  $A^\circ \subset R \subset A$ , entonces  $R$  es medible Jordan y  $c(R) = v(A)$ . En particular,  $A^\circ$  y  $A$  son medibles Jordan y  $c(A^\circ) = c(A) = v(A)$ .
- e) Si  $E \in \mathcal{M}_n$  y  $\mathbf{u} \in \mathbb{R}$ , entonces  $\mathbf{u} + E \in \mathcal{M}_n$ , y  $c(E) = c(\mathbf{u} + E)$ .

DEM:

- a) es consecuencia directa de 10.5 b), ya que  $\chi_E \leq \chi_F$ .
- b) se obtiene aplicando 10.5 d) y 10.5 a), ya que

$$\chi_{E \cap F} = \chi_E \chi_F, \quad \chi_{E \cup F} = \chi_E + \chi_F - \chi_E \chi_F, \quad \chi_{E \setminus F} = \chi_E - \chi_E \chi_F$$

c) Según acabamos de ver, el resultado es cierto para  $m = 2$ , y la demostración se completa por inducción sobre  $m$ .

d) La función acotada  $\chi_R - 1$  se anula sobre  $A^\circ$  luego, según el lema 10.7, es

integrable sobre  $A$  con  $\int_A(\chi_R - 1) = 0$ . Como la función constante 1 es integrable sobre  $A$  se sigue que  $\chi_R$  también lo es, y  $\int_A \chi_R = \int_A 1 = v(A)$ .

e) Basta aplicar 10.5 f) a la función  $\chi_E$ , teniendo en cuenta que  $\chi_E(\mathbf{x} - \mathbf{u})$  es la función característica de  $\mathbf{u} + E$ . ■

NOTA: Se puede demostrar que toda función de conjunto  $\mu : \mathcal{M}_n \rightarrow [0, +\infty)$  que verifique

$$\text{i) } \mu([0, 1]^n) = 1; \quad \mu(\mathbf{x} + E) = \mu(E) \quad \text{para todo } E \in \mathcal{M}_n \text{ y todo } \mathbf{x} \in \mathbb{R}^n.$$

$$\text{ii) } \mu(E_1 \cup E_2) = \mu(E_1) + \mu(E_2) \quad \text{si } E_1, E_2 \in \mathcal{M}_n \text{ y } E_1 \cap E_2 = \emptyset$$

coincide con el contenido de Jordan (véase ([5], prob. 38, pág 159).

**Contenido interior y contenido exterior.** Dado un conjunto acotado  $E \subset \mathbb{R}^n$ , su *contenido exterior* y su *contenido interior* de Jordan se pueden definir en términos de las integrales superior e inferior

$$c^*(E) = \overline{\int}_A \chi_E, \quad c_*(E) = \underline{\int}_A \chi_E$$

donde  $A$  es cualquier rectángulo cerrado que contiene a  $E$  (utilizando lo que se establece en el ejercicio 10.28 es fácil justificar que los valores  $\overline{\int}_A \chi_E$ ,  $\underline{\int}_A \chi_E$ , no dependen del rectángulo cerrado  $A \supset E$ ).

Es obvio que  $c_*(E) \leq c^*(E)$ , y que se cumple la igualdad si y sólo si  $E$  es medible Jordan. También es inmediato que las funciones de conjunto  $c^*$  y  $c_*$  son monótonas: Si  $E \subset F \subset \mathbb{R}^n$  son acotados entonces  $c_*(E) \leq c_*(F)$ , y  $c^*(E) \leq c^*(F)$ .

Con el fin de dar una interpretación geométrica del contenido interior y del contenido exterior, introducimos la siguiente terminología: Llamamos *figura elemental* a un conjunto acotado  $Z \subset \mathbb{R}^n$  que admite una representación de la forma  $Z = \cup\{S : S \in \Gamma\}$ , donde  $\Gamma \subset \Delta(p)$  y  $p \in \mathcal{P}(A)$  es una subdivisión de algún rectángulo cerrado  $A \supset Z$ . En ese caso diremos que  $Z = \cup\{S : S \in \Gamma\}$  es una representación asociada a la partición  $p$ . Es claro que esta representación no es única, pues si  $q \in \mathcal{P}(A)$  es más fina que  $p$ , entonces  $Z$  también admite una representación asociada a  $q$ . Toda figura elemental  $Z \subset \mathbb{R}^n$  es medible Jordan y si  $Z = \cup\{S : S \in \Gamma\}$  es una representación asociada a  $p \in \mathcal{P}(A)$  entonces  $c(Z) = \sum_{S \in \Gamma} v(S)$  (la justificación detallada de esta afirmación se puede ver en el ejercicio 10.29).

Si  $\mathcal{E}_n \subset \mathcal{M}_n$  es la familia de las figuras elementales se puede demostrar (véase el ejercicio 10.30) para un conjunto acotado  $E \subset \mathbb{R}^n$  se verifica

$$c^*(E) = \inf\{c(Z) : E \subset Z \in \mathcal{E}_n\}$$

$$c_*(E) = \sup\{c(Z') : E \supset Z' \in \mathcal{E}_n\}$$

**Conjuntos de contenido nulo.** Un conjunto medible Jordan  $E \subset \mathbb{R}^n$  que cumple  $c_n(E) = 0$ , se dice que tiene *contenido nulo* ( $n$ -dimensional).

**Proposición 10.9** *Las siguientes propiedades de un conjunto acotado  $E \subset \mathbb{R}^n$  son equivalentes:*

- i)  $E$  tiene contenido nulo.
- ii) Para cada  $\epsilon > 0$  existe una familia finita de rectángulos cerrados  $\{S_1, S_2, \dots, S_m\}$ , tal que  $E \subset \bigcup_{k=1}^m S_k$ , y  $\sum_{k=1}^m v(S_k) < \epsilon$ .
- iii) Para cada  $\epsilon > 0$  existe una familia finita de rectángulos abiertos,  $\{U_1, U_2, \dots, U_m\}$ , tal que  $E \subset \bigcup_{k=1}^m U_k$ , y  $\sum_{k=1}^m v(U_k) < \epsilon$ .

DEM: i)  $\Rightarrow$  ii): Si se cumple i) existe un rectángulo cerrado  $A \supset E$  tal que  $\chi_E$  es integrable sobre  $A$ , con  $\int_A \chi_E = 0$ . Para cada  $\epsilon > 0$  existe  $p \in \mathcal{P}(A)$  tal que  $S(\chi_E, p) < \epsilon$ . El recubrimiento finito de  $E$  formado por los rectángulos (cerrados)  $S_1, S_2, \dots, S_m \in \Delta(p)$  que tienen intersección no vacía con  $E$  verifica

$$\sum_{j=1}^m v(S_j) = S(\chi_E, p) < \epsilon$$

ii)  $\Rightarrow$  iii): Dado  $\epsilon > 0$ , sean  $\{S_j : 1 \leq j \leq m\}$ , los rectángulos cerrados suministrados por la hipótesis ii). Como  $r = \epsilon - \sum_{j=1}^m v(S_j) > 0$ , para cada  $j \in \{1, 2, \dots, m\}$  existe un rectángulo abierto  $U_j \supset S_j$  con  $v(U_j) < v(S_j) + r/m$ . Estos rectángulos abiertos cubren  $E$ , y verifican

$$\sum_{j=1}^m v(U_j) < \sum_{j=1}^m v(S_j) + r = \epsilon$$

iii)  $\Rightarrow$  i): Para cada  $\epsilon > 0$ , la hipótesis iii) proporciona un abierto  $G = \bigcup_{k=1}^m U_k$  que contiene a  $E$ . En virtud de los apartados c) y d) de la proposición 10.8 este abierto es medible Jordan y verifica

$$c(G) \leq \sum_{k=1}^m c(U_k) = \sum_{k=1}^m v(U_k) < \epsilon$$

luego  $0 \leq c_*(E) \leq c^*(E) \leq c(G) < \epsilon$ . Como  $\epsilon > 0$  es arbitrario se concluye que  $E$  es medible Jordan con  $c(E) = 0$ . ■

Con la caracterización de 10.9 ii) se obtiene que si  $H \subset \mathbb{R}^n$  tiene contenido nulo entonces  $\overline{H}$  también lo tiene. Es inmediato que todo subconjunto de un conjunto de contenido nulo tiene contenido nulo y que la unión de una familia finita de conjuntos de contenido nulo tiene contenido nulo. Los conjuntos finitos tienen contenido nulo, pero hay conjuntos numerables que no tienen contenido nulo:  $H = [0, 1] \cap \mathbb{Q}$  no tiene contenido nulo en  $\mathbb{R}$  porque  $\overline{H} = [0, 1]$  no tiene contenido nulo. En  $\mathbb{R}$  existen conjuntos no numerables de contenido nulo (véase [5] pág 316).

La siguiente proposición proporciona una clase amplia de conjuntos de contenido nulo.

**Proposición 10.10** Si  $f : A \rightarrow \mathbb{R}$  es integrable Riemann en un rectángulo cerrado  $A \subset \mathbb{R}^n$ , su gráfica  $G(f) = \{(\mathbf{x}, f(\mathbf{x})) : \mathbf{x} \in A\}$  tiene contenido nulo en  $\mathbb{R}^{n+1}$ .

DEM: Según la proposición 10.3, para cada  $\epsilon > 0$  existe  $p \in \mathcal{P}(A)$ , tal que

$$S(f, p) - s(f, p) = \sum_{S \in \Delta(p)} [M(f, S) - m(f, S)]v(S) < \epsilon$$

Los sumandos que intervienen en esta suma se pueden interpretar como volúmenes de los rectángulos cerrados  $R_S = S \times [m(f, S), M(f, S)] \subset \mathbb{R}^{n+1}$ , que recubren  $G(f)$ , luego, en virtud de la proposición 10.9,  $G(f)$  tiene contenido nulo. ■

**Integración sobre conjuntos medibles Jordan.** Si  $f : D \rightarrow \mathbb{R}$  es una función con dominio  $D \subset \mathbb{R}^n$ , dado  $E \subset D$  se define  $f_E$  como la función que coincide con  $f$  en  $E$  y vale 0 en  $\mathbb{R}^n \setminus E$ .

**Definición 10.11** Si  $E$  es acotado y  $f_E$  es integrable Riemann sobre algún rectángulo cerrado  $n$ -dimensional  $A \supset E$ , se dice que  $f$  es integrable Riemann sobre  $E$ . En ese caso se define

$$\int_E f(\mathbf{x}) \, d\mathbf{x} = \int_A f_E(\mathbf{x}) \, d\mathbf{x}$$

Razonando como se hizo al definir el contenido  $c(E)$ , (donde se hizo lo mismo con  $f = 1$ ) es fácil comprobar que la definición anterior no depende del rectángulo cerrado  $A$  en el que se considere incluido  $E$ .

Obsérvese que una condición necesaria y suficiente para que las funciones constantes sean integrables Riemann sobre el conjunto acotado  $E \subset \mathbb{R}^n$  es que  $E$  sea medible Jordan. Por ello, en lo que sigue, sólo consideramos integrales sobre conjuntos medibles Jordan. Denotaremos por  $\mathcal{R}(E)$  el conjunto de las funciones integrables Riemann sobre  $E$ .

Utilizando la proposición 10.5, es inmediato comprobar que  $\mathcal{R}(E)$  es un espacio vectorial sobre el cuerpo de los números reales sobre el cual la integral  $f \rightarrow \int_E f$  es una forma lineal monótona. Además,  $f \in \mathcal{R}(E) \Rightarrow |f| \in \mathcal{R}(E)$  y  $|\int_E f| \leq \int_E |f|$ .

### Proposición 10.12

- a) Si  $H \subset \mathbb{R}^n$  tiene contenido nulo y  $f : H \rightarrow \mathbb{R}$  es acotada entonces  $f$  es integrable sobre  $H$ , con integral nula,  $\int_H f = 0$ .
- b) Sean  $E_1, E_2 \subset \mathbb{R}^n$  medibles Jordan y  $E = E_1 \cup E_2$ . Una función acotada  $f : E \rightarrow \mathbb{R}$ , es integrable sobre  $E$  si y sólo si es integrable sobre  $E_1$  y sobre  $E_2$ . En este caso, si  $E_1 \cap E_2$  tiene contenido nulo,

$$\int_E f(\mathbf{x}) \, d\mathbf{x} = \int_{E_1} f(\mathbf{x}) \, d\mathbf{x} + \int_{E_2} f(\mathbf{x}) \, d\mathbf{x}$$

DEM: a) Basta demostrarlo cuando  $f \geq 0$ , pues el caso general se reduce a este considerando la descomposición  $f = f^+ - f^-$ .

Sea  $A \subset \mathbb{R}^n$  un rectángulo cerrado que contiene a  $H$ . Si  $\alpha > 0$  es una cota de  $f$ , se cumple  $0 \leq f_H \leq \alpha \chi_H$ , luego

$$0 \leq \int_A f_H \leq \overline{\int_A f_H} \leq \alpha \int_A \chi_H = \alpha c(H) = 0$$

b) Sea  $A \subset \mathbb{R}^n$  un rectángulo cerrado que contiene a  $E$ .

Si  $f$  es integrable sobre  $E$  se verifica  $f_E \in \mathcal{R}(A)$ , y en virtud de 10.5 d), podemos afirmar que  $f_{E_i} = \chi_{E_i} f_E \in \mathcal{R}(A)$ . Análogamente,  $f_{E_1 \cap E_2} \in \mathcal{R}(A)$ .

Recíprocamente, si las funciones  $f_{E_1}$ ,  $f_{E_2}$ , son integrables sobre  $A$ , por lo que acabamos de demostrar, también lo es  $f_{E_1 \cap E_2}$ , luego  $f_E = f_{E_1} + f_{E_2} - f_{E_1 \cap E_2}$  es integrable sobre  $A$ . Si  $c(E_1 \cap E_2) = 0$ , según a),  $\int_A f_{E_1 \cap E_2} = \int_{E_1 \cap E_2} f = 0$ , luego

$$\int_E f = \int_{E_1} f + \int_{E_2} f.$$

■

El siguiente corolario pone de manifiesto que los conjuntos de contenido nulo son despreciables frente a la integral de Riemann: La modificación de una función integrable en un conjunto de puntos de contenido nulo no perturba ni la integrabilidad de la función ni el valor de su integral.

**Corolario 10.13** Sean  $f, g : E \rightarrow \mathbb{R}$ , funciones acotadas en un conjunto medible Jordan  $E \subset \mathbb{R}^n$ , tales que  $H = \{\mathbf{x} \in E : f(\mathbf{x}) \neq g(\mathbf{x})\}$  tiene contenido nulo. Entonces  $f$  es integrable Riemann sobre  $E$  si y sólo si lo es  $g$ , y en ese caso

$$\int_E f(\mathbf{x}) \, d\mathbf{x} = \int_E g(\mathbf{x}) \, d\mathbf{x}$$

DEM: Según 10.12 a), la diferencia  $\varphi = f - g$  es integrable Riemann sobre  $H$ , y  $\int_H \varphi = 0$ . Como  $\varphi$  es idénticamente nula sobre  $E \setminus H$ , aplicando 10.12 b) con  $E_1 = H$ ,  $E_2 = E \setminus H$ , se obtiene que  $\int_E \varphi = 0$ . ■

Aunque el siguiente resultado quedará incluido en uno posterior (10.27) que depende del teorema de Lebesgue 10.24, merece la pena dar ver una demostración directa directa del mismo basada en el teorema 10.4.

**Proposición 10.14** Toda función continua acotada  $f : E \rightarrow \mathbb{R}$  en un conjunto medible Jordan  $E \subset \mathbb{R}^n$ , es integrable Riemann sobre  $E$ .

DEM: Como  $f = f^+ - f^-$ , donde  $f^+, f^- \geq 0$  son continuas sobre  $E$ , basta hacer la demostración en el caso particular  $f \geq 0$ .

Sea  $\alpha > 0$  una cota de  $f$  sobre  $E$ , y  $A \subset \mathbb{R}^n$  un rectángulo cerrado que contiene a  $E$ . Como  $\chi_E$  es integrable sobre  $A$ , dado  $\epsilon > 0$  existe  $p \in \mathcal{P}(A)$  tal que  $S(\chi_E, p) - s(\chi_E, p) < \epsilon/\alpha$ , es decir

$$\sum_{S \in \Delta'(p)} v(S) - \sum_{S \in \Delta''(p)} v(S) < \epsilon/\alpha$$



donde  $\Delta'(p) = \{S \in \Delta(p) : S \cap E \neq \emptyset\}$ ,  $\Delta''(p) = \{S \in \Delta(p) : S \subset E\}$ .

Sea  $\Gamma = \Delta'(p) \setminus \Delta''(p)$ , y  $Z = \bigcup_{S \in \Gamma} S$ . Como  $\chi_Z \leq \sum_{S \in \Gamma} \chi_S$ , resulta

$$\int_A \chi_Z \leq \sum_{S \in \Gamma} \int_A \chi_S = \sum_{S \in \Gamma} c(S) = \sum_{S \in \Gamma} v(S) < \epsilon/\alpha$$

y teniendo en cuenta que  $f_Z \leq \alpha \chi_Z$ , se obtiene que  $\overline{\int}_A f_Z \leq \int_A \alpha \chi_Z < \epsilon$ .

Consideremos ahora las figuras elementales

$$Z_1 = \bigcup_{S \in \Delta'(p)} S, \quad Z_2 = \bigcup_{S \in \Delta''(p)} S$$

Cada rectángulo cerrado  $S \in \Delta''(p)$  está contenido en  $E$ , luego  $f$  es continua sobre  $S$ , y por lo tanto integrable (10.4). Entonces, en virtud de 10.12 b),  $f$  es integrable sobre  $Z_2$ .

Como  $Z_2 \subset E \subset Z_1 = Z_2 \cup Z$ , se cumple  $f_{Z_2} \leq f_E \leq f_{Z_1} \leq f_{Z_2} + f_Z$ , y utilizando 10.6 se obtiene

$$\int_A f_{Z_2} \leq \underline{\int}_A f_E \leq \overline{\int}_A f_E \leq \overline{\int}_A f_{Z_1} \leq \overline{\int}_A f_{Z_2} + \overline{\int}_A f_Z = \int_A f_{Z_2} + \overline{\int}_A f_Z \leq \int_A f_{Z_2} + \epsilon$$

Como  $\epsilon > 0$  es arbitrario se concluye que  $\underline{\int}_A f_E = \overline{\int}_A f_E$ , lo que significa que  $f$  es integrable sobre  $E$ . ■.

El siguiente objetivo es establecer un criterio útil para justificar que cierto tipo de conjuntos, que surgen habitualmente en el cálculo integral, son medibles Jordan.

Si  $f, g : A \rightarrow \mathbb{R}$  son funciones definidas en un rectángulo cerrado  $A \subset \mathbb{R}^n$ , y  $g \leq f$ , denotaremos por  $R(g, f, A)$  el subconjunto de  $\mathbb{R}^{n+1}$  definido así

$$R(g, f, A) = \{(\mathbf{x}, y) \in \mathbb{R}^n \times \mathbb{R} : \mathbf{x} \in A, g(\mathbf{x}) \leq y \leq f(\mathbf{x})\}$$

A veces también es conveniente considerar

$$R_0(g, f, A) = \{(\mathbf{x}, y) \in \mathbb{R}^n \times \mathbb{R} : \mathbf{x} \in A, g(\mathbf{x}) < y < f(\mathbf{x})\}$$

En el caso particular  $g = 0 \leq f$ , escribiremos más brevemente  $R(f, A)$ ,  $R_0(f, A)$ . La diferencia  $R(g, f, A) \setminus R_0(g, f, A)$  es la unión de las gráficas  $G(f) \cup G(g)$  y por lo tanto tendrá contenido nulo cuando  $f$  y  $g$  sean integrables Riemann (véase la proposición 10.10).

**Proposición 10.15** Sean  $f, g : A \rightarrow [0 + \infty)$  funciones integrables Riemann en un rectángulo cerrado  $A \subset \mathbb{R}^n$ , tales que  $g \leq f$ . Entonces cualquier conjunto  $W$  que verifique  $R_0(g, f, A) \subset W \subset R(g, f, A)$  es medible Jordan y

$$c_{n+1}(W) = \int_A (f(\mathbf{x}) - g(\mathbf{x})) \, d\mathbf{x}$$

DEM: i) Empecemos con el caso particular  $g = 0 \leq f$ , considerando un conjunto  $W$  tal que  $R_0(f, A) \subset W \subset R(f, A)$ . Para cada  $p \in \mathcal{P}(A)$  podemos interpretar la suma inferior

$$s(f, p) = \sum_{S \in \Delta(p)} m(f, S)v(S)$$

como suma de volúmenes de rectángulos abiertos disjuntos  $U_S = S^\circ \times (0, m(f, S))$  contenidos en  $W$ . En virtud de 10.8 c) la unión  $U$  de estos rectángulos abiertos es un abierto medible Jordan contenido en  $W$ , luego

$$s(f, p) = \sum_{S \in \Delta(p)} v_{n+1}(U_S) = c_{n+1}(U) \leq c_*(W)$$

Considerando el supremo de las sumas  $s(f, p)$ , resulta  $\int_A f \leq c_*(W)$ .

Análogamente podemos interpretar la suma superior

$$S(f, p) = \sum_{S \in \Delta(p)} M(f, S)v(S)$$

como una suma de volúmenes de rectángulos cerrados  $B_S = S \times [0, M(f, S)]$ , cuya unión  $B$  contiene a  $W$ . En virtud de la proposición 10.8  $B$  es un conjunto medible Jordan que cumple

$$c^*(W) \leq c_{n+1}(B) \leq \sum_{S \in \Delta(p)} c_{n+1}(B_S) = \sum_{S \in \Delta(p)} v_{n+1}(B_S) = S(f, p)$$

y considerando el extremo inferior de las sumas  $S(f, p)$ , resulta  $c^*(W) \leq \int_A f$ .

Juntando las dos desigualdades que hemos establecido queda demostrado que  $c_*(W) = c^*(W) = \int_A f$ , luego  $W$  es medible Jordan en  $\mathbb{R}^{n+1}$ , y

$$c_{n+1}(W) = \int_A f$$

ii) Para demostrar el caso general,  $g \leq f$ , considerando una cota inferior  $\alpha$  de las funciones  $f, g$ , podemos escribir  $R(g, f, A) = F_\alpha - E_\alpha$ , donde

$$E_\alpha = \{(\mathbf{x}, y) \in \mathbb{R}^n \times \mathbb{R} : \mathbf{x} \in A, \alpha \leq y < g(\mathbf{x})\}$$

$$F_\alpha = \{(\mathbf{x}, y) \in \mathbb{R}^n \times \mathbb{R} : \mathbf{x} \in A, \alpha \leq y \leq f(\mathbf{x})\}$$

Estos conjuntos son trasladados de los conjuntos

$$E = \{(\mathbf{x}, y) \in \mathbb{R}^n \times \mathbb{R} : \mathbf{x} \in A, 0 \leq y < g(\mathbf{x}) - \alpha\}$$

$$F = \{(\mathbf{x}, y) \in \mathbb{R}^n \times \mathbb{R} : \mathbf{x} \in A, 0 \leq y \leq f(\mathbf{x}) - \alpha\}$$

que son medibles Jordan por lo demostrado en el caso preliminar i). Se sigue que  $E_\alpha$  y  $F_\alpha$  son medibles, y según 10.8, e) se verifica

$$c_{n+1}(E_\alpha) = \int_A (g - \alpha); \quad c_{n+1}(F_\alpha) = \int_A (f - \alpha).$$

luego  $R(g, f, A) = F_\alpha - E_\alpha$ , es medible Jordan y

$$c_{n+1}(R(g, f, A)) = c_{n+1}(F_\alpha) - c_{n+1}(E_\alpha) = \int_A (f - g)$$

Finalmente, como  $R(g, f, A) \setminus W \subset G(f) \cup G(g)$  tiene contenido nulo, se sigue que  $W$  es medible Jordan en  $\mathbb{R}^{n+1}$ , y además  $c_{n+1}(W) = c_{n+1}(R(g, f, A))$ . ■

**Corolario 10.16** *Los resultados de la proposición 10.15 siguen siendo ciertos para funciones integrables Riemann  $g \leq f$ , sobre un conjunto medible Jordan  $E \subset \mathbb{R}^{n+1}$ .*

DEM: Sean  $g \leq f$  integrables Riemann sobre un conjunto medible Jordan  $E \subset \mathbb{R}^n$  y  $A \subset \mathbb{R}^n$  un rectángulo cerrado que contiene a  $E$ . Las funciones  $f_E, g_E$  son integrables Riemann sobre  $A$ , y es claro que  $R_0(g, f, E) = R_0(g_E, f_E, A)$ , luego, en virtud de 10.15,  $R_0(g, f, E)$  es medible Jordan y

$$c_{n+1}(R_0(g, f, E)) = \int_A (f_E - g_E) = \int_E (f - g)$$

Según 10.10, los trozos de gráficas  $G(f_E), G(g_E)$ , que determina  $A$ , tienen contenido nulo, luego también lo tienen  $G(f) \subset G(f_E)$ , y  $G(g) \subset G(g_E)$ . Entonces  $W \setminus R_0(g, f, E)$  tiene contenido nulo y se sigue que  $W$  es medible Jordan con  $c_{n+1}(W) = c_{n+1}(R_0(g, f, E))$ . ■

**Algunas aplicaciones del cálculo integral.** Los resultados establecidos 10.15 y 10.16, combinados con la proposición 10.8 permiten establecer que las figuras geométricas elementales de  $\mathbb{R}^2$  (triángulos, polígonos, círculos, elipses, etc.) son medibles Jordan y que su contenido es el área que la geometría elemental asigna a tales figuras. Lo mismo se puede decir de las figuras geométricas elementales de  $\mathbb{R}^3$ , como pirámides, poliedros, esferas, elipsoides, etc.

Si  $f$  es integrable Riemann sobre el conjunto medible Jordan  $E \subset \mathbb{R}$ , teniendo en cuenta la descomposición  $f = f^+ - f^-$ , y el corolario 10.16 resulta que el valor de la integral

$$\int_E f(\mathbf{x}) \, d\mathbf{x} = \int_E f^+(\mathbf{x}) \, d\mathbf{x} - \int_E f^-(\mathbf{x}) \, d\mathbf{x}$$

se puede interpretar como la diferencia de los volúmenes en  $\mathbb{R}^{n+1}$ , de los recintos  $R(f^+, E), R(f^-, E)$ . Hablando de manera informal,  $\int_E f(\mathbf{x}) \, d\mathbf{x}$  es la suma de los volúmenes determinados por la gráfica de  $f$ , contando con volumen positivo el que queda por encima de  $E$ , y con volumen negativo el que queda por debajo.

Además del cálculo de áreas y volúmenes la integral tiene diversas aplicaciones. Con la integral triple de una función no negativa se puede describir la distribución de la masa en un cuerpo no homogéneo. Con el fin de motivar las definiciones que siguen comenzamos con algunas consideraciones heurísticas procedentes de la física: Supongamos que el cuerpo es un bloque de un material no homogéneo que ocupa un intervalo cerrado  $A \subset \mathbb{R}^3$  y que la densidad del material en cada punto  $\mathbf{x} \in A$  viene

dada por una función  $\rho(\mathbf{x}) \geq 0$ . Esto significa que la masa  $\mu(S)$  de un bloque muy pequeño  $S \subset A$ , con  $\mathbf{x} \in S$ , es aproximadamente  $\rho(\mathbf{x})v(S)$ , y que la aproximación mejora conforme disminuye el tamaño del bloque, lo que se puede expresar así

$$\lim_{\mathbf{x} \in S, \text{diam}(S) \rightarrow 0} \frac{\mu(S)}{v(S)} = \rho(\mathbf{x})$$

Según esto, un valor aproximado de la masa total  $\mu(A)$  del bloque se consigue con una suma de Riemann  $\sum_{S \in \Delta(p)} \rho(\mathbf{x}_S)v(S)$ , donde  $\mathbf{x}_S \in S$  para cada  $S \in \Delta(p)$ , y  $p \in \mathcal{P}(A)$  es una partición suficientemente fina de  $A$ . Refinando la partición  $p$ , de modo que tienda hacia 0 su diámetro,

$$\text{diam}(p) := \text{máx}\{\text{diam}(S) : S \in \Delta(p)\}$$

lograremos aproximaciones, cada vez más precisas, de la masa total  $\mu(A)$ , y suponiendo que la densidad puntual  $\rho(\mathbf{x})$ , es una función integrable sobre  $A$  es natural definir la masa total del bloque mediante la integral triple

$$\mu(A) = \int_A \rho(\mathbf{x}) \, d\mathbf{x}$$

Consideraciones análogas se pueden hacer para un cuerpo no homogéneo que ocupa un recinto medible Jordan  $M \subset \mathbb{R}^3$ , con función de densidad puntual  $\rho(\mathbf{x}) \geq 0$ , que permite obtener la masa total de cada trozo medible  $E \subset M$  mediante la integral

$$\mu(E) = \int_E \rho(\mathbf{x}) \, d\mathbf{x}$$

Estas consideraciones preliminares son la motivación de la siguiente definición

**Definición 10.17** *La masa de un sólido, que ocupa un recinto medible Jordan  $M \subset \mathbb{R}^3$ , se dice que está distribuida según la función de densidad  $\rho : M \rightarrow [0, +\infty)$ , cuando  $\rho$  es integrable Riemann sobre  $M$  y la masa de cada porción medible  $E \subset M$  viene dada por  $\mu(E) = \int_E \rho(\mathbf{x}) \, d\mathbf{x}$ .*

La siguiente proposición pone de manifiesto que, en las condiciones de la definición anterior, si  $\mathbf{x}$  es interior a  $M$ , y la función de densidad es continua en  $\mathbf{x}$ , entonces  $\rho(\mathbf{x})$  es realmente el límite del cociente entre la masa y el volumen de las intervalos cerrados  $S \subset M$ , que se contraen hacia  $\mathbf{x}$  (e.d. tales que  $\mathbf{x} \in S$ , y  $\text{diam}(S) \rightarrow 0$ ).

**Proposición 10.18** *Sea  $f : M \rightarrow \mathbb{R}$  una función integrable Riemann sobre un recinto medible Jordan  $M \subset \mathbb{R}^n$  y  $\mathbf{a}$  un punto interior de  $M$  donde  $f$  es continua: Entonces la función de conjunto  $\mu(E) = \int_E f$ , definida sobre todos los conjuntos medibles Jordan  $E \subset M$ , verifica*

$$\lim_{\mathbf{a} \in S, \text{diam}(S) \rightarrow 0} \frac{\mu(S)}{v(S)} = f(\mathbf{a})$$

DEM: Como  $\mathbf{a}$  es interior a  $M$ , y  $f$  es continua en  $\mathbf{a}$ , existe  $\delta > 0$  tal que  $B_\infty(\mathbf{a}, \delta) \subset M$  y se cumple

$$\mathbf{x} \in B_\infty(\mathbf{a}, \delta) \Rightarrow |f(\mathbf{x}) - f(\mathbf{a})| < \epsilon$$

Si  $S$  es un intervalo cerrado con  $\mathbf{a} \in S$ , y  $\text{diam}(S) < \delta$ , se cumple  $S \subset B_\infty(\mathbf{a}, \delta)$ , luego, para todo  $\mathbf{x} \in S$  se verifica  $f(\mathbf{a}) - \epsilon < f(\mathbf{x}) < f(\mathbf{a}) + \epsilon$ , y se sigue que  $\int_S [f(\mathbf{a}) - \epsilon] \leq \int_S f \leq \int_S [f(\mathbf{a}) + \epsilon]$ , es decir,  $v(S)[f(\mathbf{a}) - \epsilon] \leq \int_S f \leq v(S)[f(\mathbf{a}) + \epsilon]$ . Dividiendo por  $v(S) > 0$ , queda establecido que

$$\mathbf{a} \in S, \text{diam}(S) < \delta \Rightarrow \left| \frac{\mu(S)}{v(S)} - f(\mathbf{a}) \right| \leq \epsilon$$

y con ello lo que se deseaba demostrar. ■

NOTA: Cuando  $n = 1$  la proposición anterior no es otra cosa que el teorema fundamental del cálculo para funciones de una variable. Si consideramos la función  $F(x) = \int_a^x f(t)dt$ , y utilizamos intervalos de la forma  $S = [a, a+h]$ , donde  $h > 0$ , se tiene  $\mu(S) = F(a+h) - F(a)$ , y la conclusión se escribe ahora en la forma

$$\lim_{h \rightarrow 0^+} \frac{F(a+h) - F(a)}{h} = f(a)$$

es decir,  $F$  es derivable por la derecha en  $a$ , con derivada lateral  $F'_d(a) = f(a)$ . Análogamente, con intervalos de la forma  $[a-h, a]$ , donde  $h > 0$ , se obtiene que  $F$  es derivable por la izquierda en  $a$ , con derivada lateral  $F'_i(a) = f(a)$ .

**Definición 10.19** Si la masa un sólido que ocupa un recinto medible Jordan  $M \subset \mathbb{R}^3$ , se distribuye según la función de densidad  $\rho : M \rightarrow [0, +\infty)$ , se llama centro de masa del sólido al punto  $\mathbf{b} = (b_1, b_2, b_3) \in \mathbb{R}^3$ , de coordenadas

$$b_j = \frac{1}{\mu(M)} \int_M x_j \rho(x_1, x_2, x_3) dx_1 dx_2 dx_3, \text{ donde } \mu(M) = \int_M \rho(\mathbf{x}) d\mathbf{x}$$

(es decir,  $b_j$  es el valor medio de la función  $x_j$ , ponderado mediante la función de densidad del sólido).

NOTA: Cuando la función de densidad es constante,  $\rho(\mathbf{x}) = k$ , (caso de una distribución de masa homogénea) el centro de masa recibe el nombre de *baricentro*. Obsérvese que, en este caso,  $\mu(M) = kc_3(M)$ , luego

$$b_j = \frac{1}{c_3(M)} \int_M x_j dx_1 dx_2 dx_3$$

Dejamos al cuidado del lector la formulación de las definiciones anteriores, 10.17 y 10.19, para el caso de cuerpos en el espacio euclídeo  $\mathbb{R}^n$ . Estas nociones, que carecen de interpretación física para  $n > 3$ , la siguen teniendo en los casos  $n = 1$  y  $n = 2$ . Cuando  $n = 1$ , las correspondientes versiones de estas definiciones intervienen al

considerar una varilla muy fina que ocupa un segmento  $[a, b] \subset \mathbb{R}$ , con una masa se distribuye según una función de densidad  $\rho : [a, b] \rightarrow [0, +\infty)$ . En este caso, el centro de masa de la varilla es el punto de abscisa

$$x_0 = \frac{1}{\mu([a, b])} \int_a^b x\rho(x)dx, \quad \text{donde } \mu([a, b]) = \int_a^b \rho(x)dx$$

Análogamente, el caso  $n = 2$ , que interviene al considerar la distribución de masa en una placa plana muy delgada, se modeliza suponiendo que la placa ocupa un recinto medible  $M \subset \mathbb{R}^2$  donde está definida su función de densidad  $\rho : M \rightarrow [0, +\infty)$ . Ahora la masa de una porción medible  $E \subset M$  la proporciona la integral doble

$$\mu(E) = \int_E \rho(x, y) dx dy$$

y el centro de masa de la placa  $(x_0, y_0)$ , viene dado por las integrales dobles

$$x_0 = \frac{1}{\mu(M)} \int_M x\rho(x, y) dx dy; \quad y_0 = \frac{1}{\mu(M)} \int_M y\rho(x, y) dx dy$$

Otro concepto importante de la Mecánica que interviene al estudiar el movimiento de un cuerpo rígido que gira alrededor de un eje es el de momento de inercia. Si el sólido no es homogéneo y su masa se distribuye según la función de densidad  $\rho(x, y, z) \geq 0$ , los momentos de inercia  $I_x, I_y, I_z$  respecto a los ejes  $Ox, Oy, Oz$  se definen, respectivamente, mediante las integrales triples

$$I_x = \int_M (y^2 + z^2)\rho(x, y, z) dx dy dz; \quad I_y = \int_M (x^2 + z^2)\rho(x, y, z) dx dy dz$$

$$I_z = \int_M (x^2 + y^2)\rho(x, y, z) dx dy dz$$

De la misma forma que la noción de masa mide la respuesta de un cuerpo a las fuerzas que le imprimen una traslación, la noción de momento de inercia de un sólido respecto a un eje de giro mide su respuesta a las fuerzas que lo someten a rotación.

### 10.3. Caracterización de las funciones integrables

Los conjuntos de medida nula que se definen a continuación intervienen en la caracterización de las funciones integrables Riemann (teorema 10.24)

**Definición 10.20** *Se dice que  $H \subset \mathbb{R}^n$  tiene medida nula si para cada  $\epsilon > 0$  existe una sucesión de rectángulos cerrados  $\{R_k : k \in \mathbb{N}\}$ , tal que*

$$H \subset \bigcup_{k=1}^{\infty} R_k \quad \text{y} \quad \sum_{k=1}^{\infty} v(R_k) < \epsilon$$

Los conjuntos de contenido nulo tienen medida nula. Razonando como en la demostración de 10.9 es fácil ver que la definición 10.20 es equivalente a la que resulta considerando rectángulos abiertos. Utilizando este hecho y 10.9 se obtiene que todo conjunto compacto de medida nula tiene contenido nulo. Los conjuntos numerables tienen medida nula como consecuencia de la siguiente proposición:

**Proposición 10.21** *La unión de una familia numerable de conjuntos de medida nula tiene medida nula.*

DEM: Sea  $\{H_j : j \in \mathbb{N}\}$  una familia numerable de conjuntos de medida nula. Para cada  $\epsilon > 0$  hay una sucesión de rectángulos cerrados  $\{R_{j,k} : k \in \mathbb{N}\}$ , que recubre  $H_j$ , y verifica  $\sum_{k=1}^{\infty} v(R_{j,k}) < \epsilon/2^j$ .

La familia numerable  $\{R_{j,k} : (j,k) \in \mathbb{N} \times \mathbb{N}\}$  se puede ordenar formando una sucesión de rectángulos cerrados  $\{R'_m : m \in \mathbb{N}\}$ , que recubre  $H$  y verifica

$$\sum_{m=1}^{\infty} v(R'_m) = \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} v(R_{j,k}) \leq \sum_{j=1}^{\infty} \epsilon/2^j = \epsilon$$

■

Aunque la clausura de un conjunto de contenido nulo sigue teniendo contenido nulo, no es cierto un resultado análogo para los conjuntos de medida nula: El conjunto numerable  $\mathbb{Q} \cap [0, 1]$  es de medida nula pero su clausura  $[0, 1]$  no tiene medida nula porque es compacto y no tiene contenido nulo. La proposición 10.13 no se verifica cuando se sustituye la noción de contenido nulo por la de medida nula: La función  $\psi = \chi_{[0,1] \cap \mathbb{Q}}$  no es integrable Riemann en  $[0, 1]$ , aunque  $\{t \in [0, 1] : |\psi(t)| > 0\}$  tiene medida nula.

Antes de emprender la demostración del teorema de Lebesgue 10.24 que caracteriza las funciones integrables Riemann mediante el conjunto de sus discontinuidades conviene describir este conjunto usando la noción de oscilación.

Dada una función acotada  $f : A \rightarrow \mathbb{R}$  definida en un rectángulo cerrado  $A \subset \mathbb{R}^n$ , el conjunto de sus puntos de discontinuidad lo denotaremos

$$D(f) = \{\mathbf{x} \in A : f \text{ es discontinua en } \mathbf{x}\}$$

La oscilación de  $f$  en  $U \subset A$  es el número  $O(f, U) = \sup f(U) - \inf f(U)$ .

Sea  $A(\mathbf{x}, r) = A \cap B(\mathbf{x}, r)$ , la bola relativa en  $A$ , de centro  $\mathbf{x}$  y radio  $r > 0$ . La oscilación de  $f$  en  $A(\mathbf{x}, r)$  decrece con  $r > 0$ , luego existe el límite

$$o(f, \mathbf{x}) = \lim_{r \rightarrow 0} O(f, A(\mathbf{x}, r))$$

que recibe el nombre de *oscilación* de  $f$  en  $\mathbf{x}$ . Conviene observar que si  $\mathbf{x}$  es interior a  $U \subset A$ , en la topología relativa de  $A$ , entonces  $O(f, U) \geq o(f, \mathbf{x})$ .

Utilizando la definición de continuidad en un punto es inmediato comprobar que  $f$  es continua en  $\mathbf{x}$  si y sólo si  $o(f, \mathbf{x}) = 0$ . Se sigue de esto que las discontinuidades de  $f$  se pueden clasificar usando el concepto de oscilación:

$$D(f) = \bigcup_{\epsilon > 0} D_{\epsilon}(f) \quad \text{con} \quad D_{\epsilon}(f) = \{\mathbf{x} \in A : o(f, \mathbf{x}) \geq \epsilon\}$$

**Lema 10.22**  $D_\epsilon(f)$  es un conjunto compacto.

DEM:  $G_\epsilon(f) = \{\mathbf{x} \in A : o(f, \mathbf{x}) < \epsilon\}$  es abierto en la topología relativa de  $A$ :

Dado  $\mathbf{x} \in G_\epsilon$ , existe  $r > 0$  tal que  $O(f, A(\mathbf{x}, r)) < \epsilon$ . Como  $A(\mathbf{x}, r)$  es abierto relativo en  $A$ , para todo  $\mathbf{y} \in A(\mathbf{x}, r)$  se cumple  $o(f, \mathbf{y}) \leq O(f, A(\mathbf{x}, r)) < \epsilon$ , es decir  $A(\mathbf{x}, r) \subset G_\epsilon(f)$ .

Como  $A$  es cerrado y  $D_\epsilon(f) = A \setminus G_\epsilon(f)$  es cerrado en la topología relativa de  $A$ , se sigue que el conjunto acotado  $D_\epsilon(f)$  es cerrado en  $\mathbb{R}^n$ , y por lo tanto compacto. ■

**Lema 10.23** Sea  $f : S \rightarrow \mathbb{R}$  una función acotada definida en un rectángulo cerrado  $S$ , tal que  $o(f, \mathbf{x}) < \epsilon$  para todo  $\mathbf{x} \in S$ . Entonces existe  $p \in \mathcal{P}(S)$  verificando

$$\sum_{S' \in \Delta(p)} [M(f, S') - m(f, S')]v(S') < \epsilon v(S)$$

DEM: Según la definición de oscilación, cada  $\mathbf{x} \in S$  tiene un entorno relativo  $S \cap B(\mathbf{x}, r)$ , tal que  $O(f, S \cap B(\mathbf{x}, r)) < \epsilon$ . Sea  $V_{\mathbf{x}}$  un rectángulo abierto tal que  $\mathbf{x} \in V_{\mathbf{x}} \subset \overline{V_{\mathbf{x}}} \subset B(\mathbf{x}, r)$ . Entonces  $R_{\mathbf{x}} = \overline{V_{\mathbf{x}}} \cap S$  es un rectángulo cerrado que verifica  $O(f, R_{\mathbf{x}}) < \epsilon$ . Una cantidad finita de estos rectángulos,  $R_{\mathbf{x}_1}, R_{\mathbf{x}_2}, \dots, R_{\mathbf{x}_m}$ , recubre el compacto  $S$ , y existe una subdivisión  $p \in \mathcal{P}(S)$  tal que cada  $S' \in \Delta(p)$  está contenido en algún  $R_{\mathbf{x}_j}$ , luego

$$M(f, S') - m(f, S') = O(f, S') \leq O(f, R_{\mathbf{x}_j}) < \epsilon$$

Multiplicando por  $v(S')$  y sumando se obtiene la desigualdad del enunciado. ■

**Teorema 10.24** (Lebesgue) Sea  $f : A \rightarrow \mathbb{R}$  una función acotada definida en un rectángulo cerrado  $A \subset \mathbb{R}^n$ . Una condición necesaria y suficiente para que  $f$  sea integrable Riemann sobre  $A$  es que el conjunto de sus puntos de discontinuidad  $D(f)$  tenga medida nula.

DEM: La condición es necesaria:

Con las notaciones anteriores demostraremos que cada  $D_{1/m}(f)$  tiene contenido nulo y se seguirá que  $D(f) = \bigcup_{m \in \mathbb{N}} D_{1/m}(f)$  tiene medida nula.

Para demostrar que  $D_{1/m}(f)$  tiene contenido nulo veremos que para cada  $\epsilon > 0$  hay una descomposición  $D_{1/m}(f) = E_\epsilon \cup F_\epsilon$  donde  $E_\epsilon$  tiene contenido nulo y  $F_\epsilon$  se puede recubrir con una cantidad finita de rectángulos cerrados cuya suma de volúmenes es menor que  $\epsilon$ .

Como  $f$  es integrable, existe  $p \in \mathcal{P}(A)$  tal que  $S(f, p) - s(f, p) < \epsilon/m$ . Sea  $F_\epsilon$  la parte de  $D_{1/m}(f)$  cubierta por los rectángulos de

$$\Delta^m = \{S \in \Delta(p) : S^\circ \cap D_{1/m}(f) \neq \emptyset\}$$

y  $E_\epsilon$  la parte de  $D_{1/m}(f)$  no cubierta por estos rectángulos.



El conjunto  $E_\epsilon$  tiene contenido nulo porque está contenido en la unión de las caras de los rectángulos  $S \in \Delta(p)$ . Por otra parte, cuando  $S \in \Delta^m$  existe  $\mathbf{x} \in S^\circ$ , con  $o(f, \mathbf{x}) \geq 1/m$ , luego  $M(f, S) - m(f, S) = O(f, S) \geq 1/m$ , y así

$$\frac{1}{m} \sum_{S \in \Delta^m} v(S) \leq \sum_{S \in \Delta^m} [M(f, S) - m(f, S)]v(S) \leq S(f, p) - s(f, p) \leq \epsilon/m$$

Es decir, la familia  $\Delta^m$  que recubre  $F_\epsilon$ , verifica  $\sum_{S \in \Delta^m} v(S) < \epsilon$ .

La condición es suficiente:

Si  $D(f)$  tiene medida nula, para cada  $\epsilon > 0$ , el conjunto  $D_\epsilon(f)$  tiene medida nula y es compacto. Según el lema 10.22,  $D_\epsilon(f)$  tiene contenido nulo, luego existe una familia finita de rectángulos abiertos  $U_1, U_2, \dots, U_m$ , que recubren  $D_\epsilon(f)$ , y verifica  $\sum_{j=1}^m v(U_j) < \epsilon$  (véase la proposición 10.9).

Es fácil ver que existe  $p \in \mathcal{P}(A)$  con la siguiente propiedad: Si  $S \in \Delta(p)$  corta a  $D_\epsilon(f)$ , entonces  $S \subset \overline{U_j}$  para algún  $j \in \{1, 2, \dots, m\}$ . Clasificamos los rectángulos de  $\Delta(p)$  en dos familias

$$\Delta_1 = \{S \in \Delta(p) : S \cap D_\epsilon(f) \neq \emptyset\}, \quad \Delta_2 = \{S \in \Delta(p) : S \cap D_\epsilon(f) = \emptyset\}.$$

Cada  $S \in \Delta_1$  está contenido en algún  $\overline{U_j}$ , luego

$$\sum_{S \in \Delta_1} v(S) \leq \sum_{j=1}^m v(U_j) \leq \epsilon$$

Si  $S \in \Delta_2$ , para todo  $\mathbf{x} \in S$  se cumple  $o(f, \mathbf{x}) < \epsilon$ , y según el lema 10.23 existe  $p_S \in \mathcal{P}(S)$  verificando

$$\sum_{S' \in \Delta(p_S)} [M(f, S') - m(f, S')]v(S') < \epsilon v(S)$$

Consideremos una subdivisión  $p' \in \mathcal{P}(A)$ , más fina que  $p$ , que induzca en cada  $S \in \Delta_2$  una subdivisión más fina que  $p_S$ . En los siguientes sumatorios  $S'$  denota un elemento genérico de  $\Delta(p')$ , y  $S$  un elemento genérico de  $\Delta(p)$ :

$$\begin{aligned} S(f, p') - s(f, p') &= \sum_{S'} [M(f, S') - m(f, S')]v(S') = \\ &= \sum_{S \in \Delta_1} \sum_{S' \subset S} [M(f, S') - m(f, S')]v(S') + \sum_{S \in \Delta_2} \sum_{S' \subset S} [M(f, S') - m(f, S')]v(S') \end{aligned}$$

Para cada  $S' \subset S \in \Delta_1$  se utiliza la acotación  $M(f, S') - m(f, S') \leq 2C$ , donde  $C = \sup_{\mathbf{x} \in A} |f(\mathbf{x})|$ , y se obtiene:

$$\sum_{S \in \Delta_1} \sum_{S' \subset S} [M(f, S') - m(f, S')]v(S') \leq \sum_{S \in \Delta_1} \sum_{S' \subset S} 2Cv(S') = 2C \sum_{S \in \Delta_1} v(S) \leq 2C\epsilon$$

Por otra parte, cuando  $S \in \Delta_2$ , los rectángulos  $S' \subset S$  forman una subdivisión de  $S$ , más fina que  $p_S$ , y por ello se sigue verificando

$$\sum_{S' \subset S} [M(f, S') - m(f, S')]v(S') < \epsilon v(S)$$

luego

$$\sum_{S \in \Delta_2} \sum_{S' \subset S} [M(f, S') - m(f, S')]v(S') \leq \sum_{S \in \Delta_2} \epsilon v(S) \leq \epsilon v(A)$$

Entonces  $S(f, p') - s(f, p') \leq (2C + v(A))\epsilon$ , y utilizando la proposición 10.3 se concluye que  $f$  es integrable Riemann sobre  $A$ . ■

**Corolario 10.25** Si  $f : A \rightarrow \mathbb{R}$  es acotada en el rectángulo cerrado  $A \subset \mathbb{R}^n$  y  $D(f)$  es numerable entonces  $f$  es integrable Riemann sobre  $A$ .

DEM: Es consecuencia inmediata del teorema 10.24 ya que todo conjunto numerable tiene medida nula. ■

NOTA: Es bien conocido que si  $f : [a, b] \rightarrow \mathbb{R}$  es monótona entonces  $D(f)$  es numerable, de modo que el corolario anterior incluye, como caso particular, el resultado elemental que afirma que toda función monótona es integrable Riemann.

Una consecuencia directa del teorema 10.24 es la siguiente caracterización de los conjuntos medibles Jordan:

**Teorema 10.26** Una condición necesaria y suficiente para que un conjunto acotado  $E \subset \mathbb{R}^n$  sea medible Jordan es que su frontera  $\partial E$  tenga contenido nulo.

DEM: Sea  $f$  la restricción de  $\chi_E$  a un rectángulo cerrado  $A \subset \mathbb{R}^n$  tal que  $E \subset A^\circ$ . Es claro que el conjunto de puntos de discontinuidad de  $f : A \rightarrow \mathbb{R}$ , es  $\partial E$ . Como este conjunto es compacto, tendrá contenido nulo si y sólo si tiene medida nula, y aplicando el teorema de Lebesgue 10.24 se obtiene el resultado. ■

Este resultado queda englobado en la siguiente caracterización de las funciones integrables sobre un conjunto medible Jordan

**Teorema 10.27** Una función acotada  $f : E \rightarrow \mathbb{R}$ , sobre un conjunto medible Jordan  $E \subset \mathbb{R}^n$ , es integrable Riemann sobre  $E$  si y sólo si

$$D(f) = \{\mathbf{x} \in E : f \text{ es discontinua en } \mathbf{x}\}$$

tiene medida nula.

DEM: Sea  $A \subset \mathbb{R}^n$  un rectángulo cerrado tal que  $E \subset A^\circ$ . Debemos considerar

$$D(f_E) = \{x \in A : f_E \text{ es discontinua en } x\}$$

Se comprueba fácilmente que  $D(f) \subset D(f_E) \subset D(f) \cup \partial E$ , donde  $\partial E$  tiene contenido nulo. Se sigue que  $D(f)$  tiene medida nula si y sólo si  $D(f_E)$  tiene medida nula. Aplicando el teorema 10.24 se concluye que  $D(f)$  tiene medida nula si y sólo si  $f$  es integrable sobre  $E$ . ■

En el ejercicio resuelto 10.33 se muestra la patología que puede presentar una función integrable Riemann.

## 10.4. Ejercicios resueltos

**Ejercicio 10.28** Sea  $f : A \rightarrow [0, +\infty)$  una función acotada y  $B \subset A$  un intervalo cerrado tal que  $\{\mathbf{x} \in A : f(\mathbf{x}) \neq 0\} \subset B$ . Demuestre que  $\overline{\int_A f} = \overline{\int_B f}$ ,  $\underline{\int_A f} = \underline{\int_B f}$ .

SOLUCIÓN

Sea  $q \in \mathcal{P}(A)$  tal que  $B \in \Delta(q)$ . Razonando como en la demostración de 10.5 e)

$$\overline{\int_A f(\mathbf{x}) \, d\mathbf{x}} \leq \sum_{S \in \Delta(q)} \overline{\int_S f(\mathbf{x}) \, d\mathbf{x}}$$

Obsérvese que todos los sumandos son nulos, excepto el que corresponde a  $S = B$  (si  $S \neq B$ , entonces  $f(\mathbf{x}) = 0$  para todo  $\mathbf{x} \in S^\circ$ , y según el lema 10.7,  $f|_S$  es integrable y  $\int_S f = 0$ ). Se obtiene así la desigualdad  $\overline{\int_A f} \leq \overline{\int_B f}$ .

Por otra parte, si  $p \in \mathcal{P}(A)$ , es más fina que  $q \in \mathcal{P}(A)$ , y  $p_B \in \mathcal{P}(B)$  es la subdivisión que  $p$  induce en  $B$ , como  $f \geq 0$ , se cumple

$$\overline{\int_B f} \leq S(f, p_B) \leq S(f, p)$$

Es claro que  $\overline{\int_A f}$  es el extremo inferior de las sumas  $S(f, p)$ , cuando  $p$  recorre las subdivisiones de  $A$  que son más finas que  $q$ , luego  $\overline{\int_B f} \leq \overline{\int_A f}$ , y queda demostrada la igualdad  $\overline{\int_A f} = \overline{\int_B f}$ . Con un razonamiento análogo se demuestra que  $\underline{\int_A f} = \underline{\int_B f}$ . ■

**Ejercicio 10.29** Utilice la proposición 10.8 para demostrar que toda figura elemental  $Z \subset \mathbb{R}^n$  es medible Jordan y que si  $Z = \cup\{S : S \in \Gamma\}$  es una representación asociada a  $p \in \mathcal{P}(A)$  entonces  $c(Z) = \sum_{S \in \Gamma} v(S)$ .

SOLUCIÓN

Según las propiedades c) y d) en la proposición 10.8  $Z$  y  $G = \cup\{S^\circ : S \in \Gamma\} \subset Z$ , son medibles Jordan y se verifica

$$c(Z) \leq \sum_{S \in \Gamma} c(S) = \sum_{S \in \Gamma} v(S)$$

Como los rectángulos abiertos que intervienen en la unión  $G = \cup\{S^\circ : S \in \Gamma\}$ , son disjuntos, podemos escribir

$$c(G) = \sum_{S \in \Gamma} c(S^\circ) = \sum_{S \in \Gamma} v(S)$$

Se obtiene así que  $\sum_{S \in \Gamma} v(S) = c(G) \leq c(Z) \leq \sum_{S \in \Gamma} c(S) = \sum_{S \in \Gamma} v(S)$ . ■

**Ejercicio 10.30** Si  $E \subset \mathbb{R}^n$  es acotado demuestre que

$$c^*(E) = \inf\{c(Z) : E \subset Z \in \mathcal{E}_n\}$$

$$c_*(E) = \sup\{c(Z') : E \supset Z' \in \mathcal{E}_n\}$$

donde  $\mathcal{E}_n \subset \mathcal{M}_n$  es la familia de las figuras elementales.

SOLUCIÓN

Sea  $\alpha = \inf\{c(Z) : E \subset Z \in \mathcal{E}_n\}$ ,  $\beta = \sup\{c(Z') : E \supset Z' \in \mathcal{E}_n\}$ . Si  $Z, Z'$  son figuras elementales y  $Z' \subset E \subset Z$ , es claro que  $c(Z') \leq c_*(E) \leq c^*(E) \leq c(Z)$ , luego  $\beta \leq c_*(E) \leq c^*(E) \leq \alpha$ . Por otra parte, fijando un rectángulo cerrado  $A \supset E$ , para cada  $\epsilon > 0$  existen  $p, p' \in \mathcal{P}(A)$  tales que

$$S(\chi_E, p) \leq \overline{\int_A \chi_E} + \epsilon = c^*(E) + \epsilon$$

$$s(\chi_E, p') \geq \underline{\int_A \chi_E} - \epsilon = c_*(E) - \epsilon$$

Es claro que  $S(\chi_E, p) = \sum_{S \in \Gamma} v(S)$ , donde  $\Gamma = \{S : S \in \Delta(p) : S \cap E \neq \emptyset\}$ , luego  $Z = \cup\{S : S \in \Gamma\}$  es una figura elemental que contiene a  $E$  y, en virtud del ejercicio, 10.29 cumple que  $c(Z) = S(\chi_E, p)$  luego

$$c^*(E) \leq \alpha \leq c(Z) = S(\chi_E, p) \leq c^*(E) + \epsilon$$

Como esto es cierto para cada  $\epsilon > 0$  se concluye que  $\alpha = c^*(E)$ .

Por otra parte, si  $\Gamma' = \{S \in \Delta(p) : S \subset E\}$ , entonces  $Z' = \cup\{S : S \in \Gamma'\}$ , es una figura elemental contenida en  $E$  que cumple  $c(Z') = s(\chi_E, p)$ , luego

$$c_*(E) \geq \beta \geq c(Z') = s(\chi_E, p) \geq c_*(E) - \epsilon$$

y se concluye que  $\beta = c^*(E)$  ■

**Ejercicio 10.31** Si  $f : A \rightarrow [0, +\infty)$  es una función continua en un rectángulo cerrado  $A \subset \mathbb{R}^n$ , y  $\int_A f(\mathbf{x})d\mathbf{x} = 0$ , demuestre que  $f$  es idénticamente nula.

SOLUCIÓN

Si se supone que para algún  $\mathbf{a} \in A$  es  $f(\mathbf{a}) > 0$ , por la continuidad de  $f$  debe existir un rectángulo cerrado no degenerado  $S \subset A$ , tal que  $\mathbf{a} \in S$ , y  $f(\mathbf{a})/2 \leq f(\mathbf{x})$ , para todo  $\mathbf{x} \in S$ . Así se llega a la desigualdad contradictoria

$$0 < v(S)f(\mathbf{a})/2 \leq \int_S f(\mathbf{x})d\mathbf{x} \leq \int_A f(\mathbf{x})d\mathbf{x} = 0$$

■

**Ejercicio 10.32** Sea  $f : A \rightarrow [0, +\infty)$  una función integrable Riemann sobre un rectángulo cerrado  $A \subset \mathbb{R}^n$  tal que  $\int_A f(\mathbf{x})d\mathbf{x} = 0$ . Demuestre que  $H = \{\mathbf{x} \in A : f(\mathbf{x}) > 0\}$  tiene medida nula y muestre un ejemplo donde  $H$  no tenga contenido nulo.

SOLUCIÓN

Demostraremos que, para cada  $m \in \mathbb{N}$ , el conjunto  $H_m = \{\mathbf{x} \in A : f(\mathbf{x}) > 1/m\}$  tiene contenido nulo y por lo tanto medida nula. Aplicando la proposición 10.21 se obtendrá que  $H = \cup_{m=1}^{\infty} H_m$  tiene medida nula.

Según la definición de integral superior, para cada  $\epsilon > 0$  existe  $p \in \mathcal{P}(A)$  tal que  $S(f, p) < \epsilon/m$ . Cuando  $S \in \Delta(p)$  y  $S \cap H_m \neq \emptyset$ , es claro que se cumple  $1/m \leq M(f, S)$  luego

$$\sum_{S \cap H_m \neq \emptyset} \frac{v(S)}{m} \leq \sum_{S \cap H_m \neq \emptyset} M(f, S)v(S) \leq S(f, p) \leq \frac{\epsilon}{m}$$

Como la familia finita de rectángulos cerrados  $\{S \in \Delta(p) : S \cap H_m \neq \emptyset\}$ , recubre  $H_m$  y la suma de sus volúmenes es menor que  $\epsilon$ , queda demostrado que  $H_m$  tiene contenido nulo, y por lo tanto medida nula.

Sea  $f : [0, 1] \rightarrow [0, +\infty)$  definida por  $f(x) = 1/q$  si  $x = p/q \in \mathbb{Q} \cap (0, 1]$  (fracción irreducible)  $f(x) = 0$  en los restantes puntos. Es fácil ver que esta función es integrable Riemann con integral nula. En este caso

$$H = \{x \in [0, 1] : f(x) > 0\} = (0, 1] \cap \mathbb{Q}$$

tiene medida nula pero no tiene contenido nulo (ya que  $\overline{H}$  no tiene contenido nulo).

■

**Ejercicio 10.33** Sea  $\varphi : [0, 1] \rightarrow \{0, 1\}$  la función característica de  $[0, 1] \cap \mathbb{Q}$ ,  $A = [0, 1] \times [0, 1]$  y  $f : A \rightarrow \mathbb{R}$  la función definida por

$$f(x, y) = 0 \text{ si } x \in [0, 1] \text{ es irracional ó si } x = 0.$$

$$f(x, y) = \frac{1}{q}\varphi(y) \text{ si } x \in (0, 1] \text{ es un número racional que se expresa en la forma irreducible } x = p/q.$$

Justifique, sin utilizar el teorema de Lebesgue, las siguientes afirmaciones: La función  $f$  es integrable Riemann sobre  $A$ , el conjunto de sus puntos de discontinuidad,  $D(f) = \{(a, b) : 0 < a \leq 1, a \in \mathbb{Q}\}$ , es denso en  $A$ , y las funciones parciales,  $y \rightarrow f(a, y)$ , no son integrables cuando  $a \in \mathbb{Q} \cap (0, 1]$ .

SOLUCIÓN

Dado  $\epsilon > 0$ , sea  $H$  el conjunto finito formado por los puntos  $x = p/q \in (0, 1]$  tales que  $1/q > \epsilon$ .

Sea  $p_1 \in \mathcal{P}([0, 1])$ ,  $p_1 = (0 = x_0 < x_1 < \dots < x_m = 1)$ , verificando

$$\sum_{j \in J} (x_j - x_{j-1}) < \epsilon, \text{ donde } J = \{j \in \{1, 2, \dots, m\} : [x_{j-1}, x_j] \cap H \neq \emptyset\}$$

Consideremos una subdivisión  $p_\epsilon = (p_1, p_2) \in \mathcal{P}(A)$ , tal que  $p_2 = \{0, 1\}$ , de modo que  $\Delta(p_\epsilon) = \{S_j : 1 \leq j \leq m\}$ , con  $S_j = [x_{j-1}, x_j] \times [0, 1]$ .

Como  $M(f, S_j) \leq 1$  cuando  $j \in J$ , y  $M(f, S_j) \leq \epsilon$  cuando  $j \notin J$ , se cumple

$$\begin{aligned} S(f, p_\epsilon) &= \sum_{j \in J} M(f, S_j)(x_j - x_{j-1}) + \sum_{j \notin J} M(f, S_j)(x_j - x_{j-1}) \leq \\ &\leq \sum_{j \in J} (x_j - x_{j-1}) + \epsilon \sum_{j \notin J} (x_j - x_{j-1}) \leq 2\epsilon \end{aligned}$$

Se sigue que para todo  $\epsilon > 0$  se cumple

$$0 \leq \int_A f \leq \overline{\int_A f} \leq S(f, p_\epsilon) \leq 2\epsilon$$

y por lo tanto  $f$  es integrable sobre  $A$ , con  $\int_A f = 0$ .

Si  $0 < a \leq 1$ , y  $a \in \mathbb{Q}$ , la función parcial  $y \rightarrow f(a, y)$  es discontinua en todo  $b \in [0, 1]$ , luego  $\{(a, b) \in A : a \in (0, 1] \cap \mathbb{Q}\} \subset D(f)$ . Por otra parte, si  $a \notin (0, 1] \cap \mathbb{Q}$ , dado  $\epsilon > 0$ , existe  $\delta > 0$  tal que  $|x - a| < \delta \Rightarrow x \notin H$ , luego  $0 \leq f(x, y) \leq \epsilon$  para todo  $y \in [0, 1]$ . Como  $f(a, b) = 0$  resulta

$$\max\{|x - a|, |y - b|\} < \delta \Rightarrow |f(x, y) - f(a, b)| < \epsilon$$

Queda demostrado que  $f$  es continua en  $(a, b)$ , y con ello la igualdad

$$D(f) = \{(a, b) \in A : a \in (0, 1] \cap \mathbb{Q}\}$$

Finalmente, si  $a = p/q \in (0, 1]$ , (fracción irreducible), como  $\varphi$  no es integrable sobre  $[0, 1]$ , tampoco lo es la función parcial  $y \rightarrow f(x, y) = \frac{1}{q}\varphi(y)$ . ■

**Ejercicio 10.34** Si  $K \subset \mathbb{R}^k$ ,  $M \subset \mathbb{R}^n$  son conjuntos medibles Jordan, demuestre que  $K \times M$  es medible Jordan en  $\mathbb{R}^k \times \mathbb{R}^n$ .

SOLUCIÓN

Según el teorema 10.26 basta ver que si  $\partial K$  tiene contenido nulo en  $\mathbb{R}^k$ , y  $\partial M$  tiene contenido nulo en  $\mathbb{R}^n$ , entonces  $\partial(K \times M)$ , tiene contenido nulo en  $\mathbb{R}^{k+n}$ . Si  $\partial K$  tiene contenido nulo en  $\mathbb{R}^k$  es fácil ver que  $\partial K \times M$  tiene contenido nulo en  $\mathbb{R}^{k+n}$ . Análogamente  $K \times \partial M$ , tiene contenido nulo en  $\mathbb{R}^{k+n}$  y usando la inclusión

$$\partial(K \times M) \subset (\overline{K} \times \overline{M}) \setminus K^\circ \times M^\circ \subset (\partial K \times \overline{M}) \cup (\overline{K} \times \partial M)$$

se concluye que  $\partial(K \times M)$  tiene contenido nulo en  $\mathbb{R}^{k+n}$ . ■

## 10.5. Ejercicios propuestos

◇ **10.5.1** Si  $f : A \rightarrow \mathbb{R}$  es integrable Riemann sobre el rectángulo cerrado  $A \subset \mathbb{R}^n$  y  $f(\mathbf{x}) \geq \alpha > 0$  para todo  $\mathbf{x} \in A$ , demuestre directamente, (sin usar el teorema de Lebesgue) que la función  $1/f$  también es integrable Riemann sobre  $A$ .

◇ **10.5.2** Si  $A \subset \mathbb{R}^n$  es un intervalo cerrado, una función acotada  $f : A \rightarrow \mathbb{R}$  se dice que es escalonada cuando existe  $p \in \mathcal{P}(A)$  tal que en el interior de cada  $S \in \Delta(p)$   $f$  toma un valor constante  $\alpha(S)$ . Demuestre que  $f$  es integrable Riemann sobre  $A$  y  $\int_A f(\mathbf{x})d\mathbf{x} = \sum_{S \in \Delta(p)} \alpha(S)v(S)$ .

◇ **10.5.3** Si  $f, g : [0, 1] \rightarrow \mathbb{R}$  son integrables Riemann y  $f(x) \geq m > 0$  para todo  $x \in [0, 1]$ , demuestre que  $F(x, y) = f(x)g(y)$  es integrable Riemann sobre  $A = [0, 1] \times [0, 1]$ .

◇ **10.5.4** Justifique que la función  $f(x, y) = \frac{1}{x-y}$  si  $x \neq y$ ,  $f(x, x) = 1$  es integrable Riemann sobre  $E = \{(x, y) \in [0, 1] \times [0, 1] : x \leq y\}$

◇ **10.5.5** Demuestre las desigualdades:

$$a) \quad \frac{1}{e} \leq \frac{1}{4\pi^2} \int_A e^{\sin(x+y)} dx dy \leq e \quad \text{donde } A = [-\pi, \pi] \times [-\pi, \pi]$$

$$b) \quad \frac{1}{6} \leq \int_B \frac{dx dy}{y-x+3} \leq \frac{1}{4} \quad \text{donde } B = \{(x, y) : 0 \leq x \leq 1, 0 \leq y \leq x\}.$$

◇ **10.5.6** Se supone que la función  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es integrable Riemann sobre cada cubo  $Q(r) = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_\infty \leq r\}$ . Demuestre:

i) Si  $\lim_{\|\mathbf{x}\| \rightarrow +\infty} f(\mathbf{x}) = \alpha$  entonces  $\lim_{r \rightarrow +\infty} \frac{1}{2^n r^n} \int_{Q(r)} f(\mathbf{x})d\mathbf{x} = \alpha$ .

ii) Si  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es continua en  $\mathbf{a}$  y  $Q(\mathbf{a}, r) = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{a}\|_\infty \leq r\}$  entonces

$$\lim_{r \rightarrow 0} \frac{1}{2^n r^n} \int_{Q(\mathbf{a}, r)} f(\mathbf{x})d\mathbf{x} = f(\mathbf{a})$$

◇ **10.5.7** Demuestre que el conjunto  $\{(x, \sin(1/x)) : 0 < x \leq 1\}$  tiene contenido nulo en  $\mathbb{R}^2$ .

◇ **10.5.8** Si  $M \subset \mathbb{R}^n$  es acotado sea  $M'$  el conjunto de sus puntos de acumulación. De modo recurrente se define  $M_1 = M'$ ,  $M_{n+1} = M'_n$ . Demuestre que si  $M_n = \emptyset$  para algún  $n \in \mathbb{N}$  entonces  $M$  tiene contenido nulo.

◇ **10.5.9** Si  $H \subset \mathbb{R}^n$  es de medida nula demuestre que  $H \times \mathbb{R}^k \subset \mathbb{R}^{n+k}$  es de medida nula en  $\mathbb{R}^{n+k}$ .

◇ **10.5.10** Demuestre que las siguientes afirmaciones

i)  $H \subset \mathbb{R}^n$  tiene medida nula si y sólo si para cada  $\epsilon > 0$  existe una sucesión de cubos abiertos (o cerrados)  $\{Q_k : k \in \mathbb{N}\}$ , que cubre  $H$  y verifica  $\sum_{k=1}^{\infty} v(Q_k) < \epsilon$ .

ii) Un conjunto acotado  $H \subset \mathbb{R}^n$  tiene contenido nulo si y sólo si para cada  $\epsilon > 0$  existe una sucesión finita de cubos abiertos (o cerrados)  $\{Q_k : 1 \leq k \leq m\}$  que cubre  $H$  y verifica  $\sum_{k=1}^m v(Q_k) < \epsilon$ .

◇ **10.5.11** Sea  $H \subset \mathbb{R}^n$  y  $\mathbf{g} : H \rightarrow \mathbb{R}^m$  una aplicación lipschitziana. Utilice el ejercicio 10.5.10 para demostrar las siguientes afirmaciones:

- Si  $n < m$ ,  $\mathbf{g}(H)$  tiene medida nula.
- Si  $n < m$  y  $H$  es acotado,  $\mathbf{g}(H)$  tiene contenido nulo.
- Si  $n = m$  y  $H \subset \mathbb{R}^n$  es de medida nula (resp. contenido nulo) entonces  $\mathbf{g}(H)$  tiene medida nula (resp. contenido nulo).

◇ **10.5.12** Sea  $\Omega \subset \mathbb{R}^n$  abierto y  $\mathbf{g} : \Omega \rightarrow \mathbb{R}^m$  una aplicación de clase  $C^1(\Omega)$ . Justifique las siguientes afirmaciones:

- Si  $n < m$  entonces  $\mathbf{g}(\Omega)$  tiene medida nula.
- Si  $n < m$ ,  $H$  es acotado, y  $\overline{H} \subset \Omega$ , entonces  $\mathbf{g}(H)$  tiene contenido nulo.
- Si  $n = m$  y  $H \subset \Omega$  tiene medida nula, entonces  $\mathbf{g}(H)$  tiene medida nula.
- Si  $n = m$ ,  $H$  tiene contenido nulo, y  $\overline{H} \subset \Omega$ , entonces  $\mathbf{g}(H)$  tiene contenido nulo.

Muestre un ejemplo que ponga de manifiesto que no se cumple d) cuando la condición  $\overline{H} \subset \Omega$  se sustituye por  $H \subset \Omega$ .

◇ **10.5.13** Sea  $M = \{(x, y, z) \in \mathbb{R}^3 : f(x, y, z) = 0\}$  donde  $f : \mathbb{R}^3 \rightarrow \mathbb{R}$  es de clase  $C^1$  y  $\nabla f(x, y, z) \neq (0, 0, 0)$  para cada  $(x, y, z) \in M$ . Demuestre que si  $M$  es acotado entonces tiene contenido nulo. ¿Qué se puede decir si  $M$  no es acotado?

◇ **10.5.14** Demuestre que  $H \subset \mathbb{R}$  es de medida nula si y sólo si existe una sucesión de intervalos acotados  $\{(a_n, b_n) : n \in \mathbb{N}\}$ , tal que  $\sum_{n=1}^{+\infty} (b_n - a_n) < +\infty$ , y para cada  $x \in H$  el conjunto  $\{n \in \mathbb{N} : x \in (a_n, b_n)\}$  es infinito. Si  $H \subset \mathbb{R}$  es de medida nula, a una sucesión de intervalos con las propiedades anteriores se le asocia la sucesión de funciones continuas  $f_n : \mathbb{R} \rightarrow \mathbb{R}$  definida por

$$f_n(x) = 0 \text{ si } x < a_n; \quad f_n(x) = x - a_n \text{ si } a_n \leq x \leq b_n; \quad f_n(x) = b_n - a_n \text{ si } b_n < x$$

Demuestre que la serie  $f_n(x) = \sum_{n=1}^{+\infty} f_n(x)$  converge para todo  $x \in \mathbb{R}$  y su suma es una función creciente continua no derivable en los puntos de  $H$ .

◇ **10.5.15** Sea  $f : A \rightarrow \mathbb{R}$  una función continua en un intervalo cerrado y acotado  $A \subset \mathbb{R}^n$ . Para cada  $p \geq 1$  se define

$$\|f\|_p = \left( \int_A |f|^p \right)^{1/p}, \quad \|f\|_\infty = \max\{|f(\mathbf{x})| : \mathbf{x} \in A\}$$

Demuestre que  $\lim_{p \rightarrow +\infty} \|f\|_p = \|f\|_\infty$ .



# Capítulo 11

## Técnicas de cálculo integral

*Integración iterada y cambio de variable. Cálculo de integrales dobles y triples. Aplicaciones geométricas y físicas del cálculo integral.*

En este capítulo se exponen técnicas para el cálculo efectivo de una integral múltiple o del contenido (área, volumen...) de un conjunto medible Jordan, y se exponen algunas de las aplicaciones clásicas del cálculo integral.

En el caso de las funciones continuas, el teorema de Fubini sobre integración iterada, que se estudia en primer lugar, resuelve teóricamente el problema, pero no conduce siempre a buenos resultados desde el punto de vista práctico porque las integrales iteradas que aparecen, o no se pueden calcular con los métodos usuales del cálculo de una variable (regla de Barrow), o requieren cálculos muy engorrosos. En estos casos la técnica del cambio de variable puede resolver el problema transformando la integral en otra cuyo cálculo sea posible, o más sencillo de realizar.

Una de las primeras aplicaciones del teorema de Fubini es la justificación del clásico principio de Cavalieri que afirma que si en dos conjuntos medibles las secciones producidas por planos perpendiculares a una recta arbitraria tienen siempre la misma área entonces los dos conjuntos tienen el mismo volumen. En este principio se basa el método de las secciones para el cálculo de volúmenes, muy útil cuando las secciones son figuras geométricas sencillas de área conocida, como ocurre en el caso de los sólidos de revolución. En relación con el teorema de Fubini también se explica con detalle el procedimiento para expresar, mediante integrales iteradas, las integrales dobles o triples sobre recintos simples. En este caso suelen haber varias alternativas para representar el dominio de integración en forma de recinto simple (o de unión sin solapamiento de tales recintos) a las que corresponden diferentes integrales iteradas. Por ello uno de los aspectos prácticos de este capítulo está encaminado a adquirir experiencia en la elección atinada de un orden de integración que conduzca a cálculos factibles.

En lo que se refiere al teorema del cambio de variable para la integral de Riemann, como su demostración es demasiado larga y técnica para abordarla ahora hemos optado por relegarla al apéndice J donde se ofrece una demostración detallada y completa que utiliza fuertes recursos de álgebra lineal y de cálculo diferencial.

El teorema del cambio de variable, aunque no se demuestra en este capítulo, se toma como base para justificar teóricamente diversas fórmulas y reglas clásicas del cálculo integral: Cálculo de áreas limitadas por curvas dadas en coordenadas polares, utilización de las simetrías en el cálculo de integrales, propiedades geométricas de los baricentros y justificación de las reglas de cálculo para volúmenes de sólidos de revolución. Otro aspecto práctico en el que se insiste en este capítulo es el de adquirir experiencia en elegir y aplicar el cambio de variable apropiado que facilita el cálculo de una integral múltiple dada.

Finaliza el capítulo con un amplio repertorio de ejercicios resueltos que ilustran los aspectos prácticos de la teoría. Entre otras cosas se muestra como se pueden usar los cambios de variable usuales (polares, cilíndricas, y esféricas) para calcular algunas integrales dobles o triples.

## 11.1. Integración iterada

En lo que sigue,  $n = k + m$ , con  $k, m \in \mathbb{N}$ , y se supone  $\mathbb{R}^n$  identificado con  $\mathbb{R}^k \times \mathbb{R}^m$ . Cada  $\mathbf{z} \in \mathbb{R}^n$  se representa como un par

$$\mathbf{z} = (\mathbf{x}, \mathbf{y}) \quad \text{con } \mathbf{x} \in \mathbb{R}^k, \mathbf{y} \in \mathbb{R}^m$$

y el rectángulo cerrado  $n$ -dimensional  $A = [a_1, b_1] \times \cdots \times [a_n, b_n]$ , lo expresaremos como el producto cartesiano  $A = X \times Y$ , de los rectángulos cerrados

$$X = [a_1, b_1] \times \cdots \times [a_k, b_k] \subset \mathbb{R}^k, \quad Y = [a_{k+1}, b_{k+1}] \times \cdots \times [a_n, b_n] \subset \mathbb{R}^m.$$

Entonces cada función acotada  $f : A \rightarrow \mathbb{R}$  se puede considerar como función de dos variables vectoriales  $(\mathbf{x}, \mathbf{y}) \in X \times Y$ . Fijado  $\mathbf{x} \in X$  (resp.  $\mathbf{y} \in Y$ ), queda definida la función parcial  $f_{\mathbf{x}}(\mathbf{y}) = f(\mathbf{x}, \mathbf{y})$  (resp.  $f^{\mathbf{y}}(\mathbf{x}) = f(\mathbf{x}, \mathbf{y})$ ), definida en  $Y$  (resp. en  $X$ ). Si  $f^{\mathbf{y}}$  es integrable sobre  $X$  su integral se denotará

$$\int_X f^{\mathbf{y}} = \int_X f^{\mathbf{y}}(\mathbf{x}) d\mathbf{x} = \int_X f(\mathbf{x}, \mathbf{y}) d\mathbf{x}$$

y en caso de no ser integrable, las integrales inferior y superior de  $f^{\mathbf{y}}$  se denotarán

$$\underline{\int}_X f(\mathbf{x}, \mathbf{y}) d\mathbf{x}, \quad \overline{\int}_X f(\mathbf{x}, \mathbf{y}) d\mathbf{x}.$$

Análogamente se definen  $\int_Y f(\mathbf{x}, \mathbf{y}) d\mathbf{y}$ ,  $\underline{\int}_Y f(\mathbf{x}, \mathbf{y}) d\mathbf{y}$ ,  $\overline{\int}_Y f(\mathbf{x}, \mathbf{y}) d\mathbf{y}$ .

**Teorema 11.1** (Fubini) *Sea  $A = X \times Y \subset \mathbb{R}^n$ , donde  $X \subset \mathbb{R}^k$ ,  $Y \subset \mathbb{R}^m$  son rectángulos cerrados. Si  $f : A \rightarrow \mathbb{R}$  es integrable Riemann entonces las funciones*

$$\underline{J}(\mathbf{y}) = \underline{\int}_X f(\mathbf{x}, \mathbf{y}) d\mathbf{x}, \quad \overline{J}(\mathbf{y}) = \overline{\int}_X f(\mathbf{x}, \mathbf{y}) d\mathbf{x},$$

son integrables sobre  $Y$ , y se cumple  $\int_A f = \int_Y \underline{J}(\mathbf{y}) d\mathbf{y} = \int_Y \overline{J}(\mathbf{y}) d\mathbf{y}$ , es decir

$$\int_A f(\mathbf{x}, \mathbf{y}) d\mathbf{x} d\mathbf{y} = \int_Y \left[ \int_{\underline{X}} f(\mathbf{x}, \mathbf{y}) d\mathbf{x} \right] d\mathbf{y} = \int_Y \left[ \int_{\overline{X}} f(\mathbf{x}, \mathbf{y}) d\mathbf{x} \right] d\mathbf{y}.$$

Análogamente, las funciones

$$\underline{I}(\mathbf{x}) = \int_{\underline{Y}} f(\mathbf{x}, \mathbf{y}) d\mathbf{y}, \quad \overline{I}(\mathbf{x}) = \int_{\overline{Y}} f(\mathbf{x}, \mathbf{y}) d\mathbf{y},$$

son integrables sobre  $X$ , y se verifica  $\int_A f = \int_X \underline{I}(\mathbf{x}) d\mathbf{x} = \int_X \overline{I}(\mathbf{x}) d\mathbf{x}$ , es decir

$$\int_A f(\mathbf{x}, \mathbf{y}) d\mathbf{x} d\mathbf{y} = \int_X \left[ \int_{\underline{Y}} f(\mathbf{x}, \mathbf{y}) d\mathbf{y} \right] d\mathbf{x} = \int_X \left[ \int_{\overline{Y}} f(\mathbf{x}, \mathbf{y}) d\mathbf{y} \right] d\mathbf{x}.$$

DEM: Cada  $p \in \mathcal{P}(A)$ ,  $p = (p_1, p_2, \dots, p_k, p_{k+1}, \dots, p_n)$ , se identifica con el par  $(p', p'')$ , donde  $p' \in \mathcal{P}(X)$ ,  $p'' \in \mathcal{P}(Y)$ , vienen dadas por  $p' = (p_1, p_2, \dots, p_k)$ ,  $p'' = (p_{k+1}, \dots, p_n)$ , y es claro que  $\Delta(p) = \{S' \times S'' : S' \in \Delta(p'), S'' \in \Delta(p'')\}$ .

Fijado  $S = S' \times S'' \in \Delta(p)$ , con  $S' \in \Delta(p')$ ,  $S'' \in \Delta(p'')$ , para cada  $\mathbf{y} \in S''$  se cumple  $m(f, S' \times S'') \leq m(f^{\mathbf{y}}, S')$ , luego

$$\sum_{S' \in \Delta(p')} m(f, S' \times S'') v(S') \leq \sum_{S' \in \Delta(p')} m(f^{\mathbf{y}}, S') v(S') = s(f^{\mathbf{y}}, p') \leq \int_{\underline{X}} f^{\mathbf{y}}(\mathbf{x}) d\mathbf{x} = \underline{J}(\mathbf{y})$$

Considerando el extremo inferior de  $\underline{J}(\mathbf{y})$  cuando  $\mathbf{y}$  recorre  $S''$ , se obtiene

$$\sum_{S' \in \Delta(p')} m(f, S' \times S'') v(S') \leq m(\underline{J}, S'').$$

Multiplicando miembro a miembro por  $v(S'')$ , y sumando cuando  $S''$  recorre  $\Delta(p'')$  se llega a la desigualdad

$$s(f, p) \leq s(\underline{J}, p'')$$

Análogamente se demuestra que

$$S(f, p) \geq S(\overline{J}, p'')$$

Quedan establecidas así las dos desigualdades no triviales de la cadena

$$s(f, p) \leq s(\underline{J}, p'') \leq s(\overline{J}, p'') \leq \int_{\underline{Y}} \overline{J}(\mathbf{y}) d\mathbf{y} \leq \int_{\overline{Y}} \overline{J}(\mathbf{y}) d\mathbf{y} \leq S(\overline{J}, p'') \leq S(f, p)$$

de las que se deduce que  $\overline{J}$  es integrable sobre  $Y$ , con  $\int_Y \overline{J}(\mathbf{y}) d\mathbf{y} = \int_A f(\mathbf{x}, \mathbf{y}) d\mathbf{x} d\mathbf{y}$ .  
Análogamente, utilizando las desigualdades

$$s(f, p) \leq s(\underline{J}, p'') \leq \int_{\underline{Y}} \underline{J}(\mathbf{y}) d\mathbf{y} \leq \int_{\overline{Y}} \underline{J}(\mathbf{y}) d\mathbf{y} \leq S(\underline{J}, p'') \leq S(\overline{J}, p'') \leq S(f, p)$$

se concluye que  $\underline{J}$  es integrable sobre  $Y$ , con  $\int_Y \underline{J}(\mathbf{y}) \, d\mathbf{y} = \int_A f(\mathbf{x}, \mathbf{y}) \, d\mathbf{x} \, d\mathbf{y}$ .

De forma similar se demuestran las afirmaciones que conciernen a las integrales sobre  $X$  de las funciones  $\underline{I}(\mathbf{x})$ ,  $\bar{I}(\mathbf{x})$ . ■

OBSERVACIONES:

i) En las condiciones del teorema de Fubini, si  $f$  es continua, todas las funciones parciales  $f_{\mathbf{x}}$ ,  $f_{\mathbf{y}}$  son continuas (y por lo tanto integrables) y se puede escribir

$$\int_A f(\mathbf{x}, \mathbf{y}) \, d\mathbf{x} \, d\mathbf{y} = \int_X \left( \int_Y f(\mathbf{x}, \mathbf{y}) \, d\mathbf{y} \right) d\mathbf{x} = \int_Y \left( \int_X f(\mathbf{x}, \mathbf{y}) \, d\mathbf{x} \right) d\mathbf{y}$$

ii) Es instructivo comprobar directamente la tesis del teorema de Fubini para la función considerada en el ejercicio resuelto 10.33. Es claro que  $\underline{J}(y) = 0 = \bar{J}(y)$  para todo  $y \in [0, 1]$ , mientras que para todo  $x \in (0, 1] \cap \mathbb{Q}$ , es  $\underline{I}(x) < \bar{I}(x)$ .

Si  $x = p/q$  en forma irreducible entonces  $\underline{I}(x) = 0$ , y  $\bar{I}(x) = 1/q$ . Obsérvese que, de acuerdo con el teorema de Fubini, 11.1, las funciones  $\bar{I}$ ,  $\underline{I}$  son integrables sobre  $[0, 1]$ , con el mismo valor de la integral.

iii) La función considerada en el ejercicio resuelto 10.33 también muestra que, en las condiciones del teorema de Fubini, pueden existir puntos  $\mathbf{x} \in X$  (resp.  $\mathbf{y} \in Y$ ), tales que  $f_{\mathbf{x}}$  (resp.  $f_{\mathbf{y}}$ ) no es integrable sobre  $Y$  (resp. sobre  $X$ ). Sin embargo la función  $\varphi(\mathbf{x}) = \bar{I}(\mathbf{x}) - \underline{I}(\mathbf{x}) \geq 0$  es integrable, con integral nula, y según el ejercicio resuelto 10.32, podemos asegurar que el conjunto  $\{\mathbf{x} \in X : \varphi(\mathbf{x}) > 0\}$  es de medida nula en  $\mathbb{R}^k$ . Es decir, el conjunto de puntos  $\mathbf{x} \in X$  tales que  $f_{\mathbf{x}}$  no es integrable, es de medida nula. Análogamente, el conjunto de puntos  $\mathbf{y} \in Y$  tales que  $f_{\mathbf{y}}$  no es integrable, es de medida nula.

iv) La existencia de una integral iterada  $\int_X [\int_Y f(\mathbf{y}) \, d\mathbf{y}] \, d\mathbf{x}$  no implica que exista la integral  $\int_A f(\mathbf{x}, \mathbf{y}) \, d\mathbf{x} \, d\mathbf{y}$ . En el ejercicio propuesto 11.4.3 se muestra una función acotada no integrable Riemann  $f : [0, 1] \times [0, 1] \rightarrow \mathbb{R}$  tal que existe la integral iterada.

Para el cálculo de una integral múltiple sobre un conjunto medible Jordan también se puede aplicar el teorema de Fubini. Suponemos, como en el enunciado de este teorema,  $\mathbb{R}^n$  identificado con  $\mathbb{R}^k \times \mathbb{R}^m$ , ( $n = k + m$ ). Si  $f : E \rightarrow \mathbb{R}$  es integrable Riemann sobre un conjunto medible Jordan  $E \subset \mathbb{R}^n$ , para calcular la integral  $\int_E f$  debemos considerar un rectángulo cerrado  $A = X \times Y$ , con  $X \subset \mathbb{R}^k$ ,  $Y \subset \mathbb{R}^m$ , que contenga a  $E$ . Para simplificar la escritura no es restrictivo suponer que  $f$  está definida en todo  $\mathbb{R}^n$ , (en caso contrario podemos extender  $f : E \rightarrow \mathbb{R}^n$  a todo  $\mathbb{R}^n$ , definiendo  $\mathbf{f}(\mathbf{x}, \mathbf{y}) = 0$  cuando  $(\mathbf{x}, \mathbf{y}) \notin E$ ), de modo que  $f_E(\mathbf{x}, \mathbf{y}) = \chi_E(\mathbf{x}, \mathbf{y})f(\mathbf{x}, \mathbf{y})$  para todo  $(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^k \times \mathbb{R}^m$ . Según la definición de integral sobre un conjunto medible Jordan, se tiene

$$\int_E f(\mathbf{x}, \mathbf{y}) \, d\mathbf{x} \, d\mathbf{y} = \int_A \chi_E(\mathbf{x}, \mathbf{y})f(\mathbf{x}, \mathbf{y}) \, d\mathbf{x} \, d\mathbf{y}$$

Para cada  $\mathbf{x} \in X$ , sea  $E_{\mathbf{x}} = \{\mathbf{y} \in Y : (\mathbf{x}, \mathbf{y}) \in E\}$  la sección de  $E$  por  $\mathbf{x}$ . Con el fin de simplificar la exposición supondremos en lo que sigue que cada  $E_{\mathbf{x}}$  es medible Jordan y que  $f|_E$  es continua (el caso más habitual en las aplicaciones). En este caso cada  $f_{\mathbf{x}}$  es continua, y por lo tanto integrable, sobre  $E_{\mathbf{x}}$ , y en virtud del teorema de Fubini se tiene

$$\begin{aligned} \int_E f(\mathbf{x}, \mathbf{y}) \, d\mathbf{x} \, d\mathbf{y} &= \int_X \left( \int_Y \chi_E(\mathbf{x}, \mathbf{y}) f(\mathbf{x}, \mathbf{y}) \, d\mathbf{y} \right) d\mathbf{x} = \\ &= \int_X \left( \int_Y \chi_{E_{\mathbf{x}}}(\mathbf{y}) f_{\mathbf{x}}(\mathbf{y}) \, d\mathbf{y} \right) d\mathbf{x} = \int_X \left( \int_{E_{\mathbf{x}}} f_{\mathbf{x}}(\mathbf{y}) \, d\mathbf{y} \right) d\mathbf{x} \end{aligned}$$

Naturalmente que, en esta discusión, el papel de las variables  $(x_1, x_2, \dots, x_k) = \mathbf{x}$  lo puede desempeñar cualquier subconjunto de las variables  $(x_1, x_2, \dots, x_n)$ , pero hemos considerado las  $k$  primeras para simplificar la notación.

Para el cálculo de integrales múltiples mediante integración iterada conviene examinar atentamente la función y el conjunto  $E$ , con el fin de plantear la integración iterada que conduzca a los cálculos más sencillos (véanse los ejercicios resueltos 11.13 y 11.14).

**Cálculo de volúmenes por el método de las secciones.** En la discusión anterior, si  $f$  es la función constante 1, resulta

$$c_n(E) = \int_X c_k(E_{\mathbf{x}}) \, d\mathbf{x}$$

Cuando  $n = 3$ , y  $k = 1$ , como  $E$  es acotado, existe  $[a, b] \in \mathbb{R}$ , tal que

$$E \subset \{(x, y, z) : a \leq x \leq b\}$$

y podemos tomar  $X = [a, b]$ . Entonces, el contenido (volumen) de  $E \subset \mathbb{R}^3$ , se expresa mediante la integral simple

$$c_3(E) = \int_a^b S(x) \, dx$$

donde  $S(x) = c_2(E_x)$  es el área de la sección  $E_x$

Este método de cálculo de volúmenes es especialmente útil cuando las secciones  $E_x$  son figuras geométricas sencillas de área conocida, como ocurre en el caso de los sólidos de revolución (véanse los ejercicios resueltos 11.12 y 11.11).

Con él queda justificado el clásico principio de Cavalieri que afirma que si la secciones  $E_x, F_x$  de dos conjuntos medibles  $E, F \subset \mathbb{R}^3$  tienen siempre la misma área entonces los dos conjuntos tienen el mismo volumen.

NOTA: En el método de las secciones el papel que desempeña la variable  $x$  lo puede desempeñar cualquiera de las otras dos variables,  $y$  ó  $z$ . Más generalmente, el método también es válido cuando las secciones se realizan mediante planos perpendiculares a una recta arbitraria (distinta de los ejes de coordenadas). Para justificar

esta afirmación basta tener en cuenta que con una traslación y un giro la recta se puede llevar a la posición del eje  $Ox$ , y que los conjuntos medibles Jordan y su contenido son invariantes por estas transformaciones (véase el corolario 11.3).

**Integrales dobles iteradas.** Consideremos primero el caso de una función continua  $f : A \rightarrow \mathbb{R}$ , sobre un rectángulo  $A = [a, b] \times [c, d]$ . En este caso, tomando  $X = [a, b]$ ,  $Y = [c, d]$ , el cálculo de la primera integral iterada  $I(x) = \int_a^b f(x, y) dy$  se podrá emprender buscando una primitiva de  $f_x$  y aplicando la regla de Barrow. La función resultante  $I(x)$  también es continua (esto se deja como ejercicio) y el valor de su integral  $\int_c^d I(x) dx$  se podrá obtener, cuando sea posible, mediante el cálculo de una primitiva. Así queda resuelto, al menos teóricamente, el cálculo de la integral doble de una función continua de dos variables reales sobre un rectángulo cerrado  $A \subset \mathbb{R}^2$ .

Consideremos ahora el caso general de integrales dobles sobre conjuntos medibles Jordan  $E \subset \mathbb{R}^2$ . Empezamos considerando el caso particular de los que admiten una representación de la forma

$$E = \{(x, y) : a \leq x \leq b, \alpha_1(x) \leq y \leq \alpha_2(x)\}$$

donde  $\alpha_1 \leq \alpha_2$  son funciones integrables Riemann sobre  $[a, b]$ . A estos conjuntos los llamaremos recintos simples de tipo  $(1, 2)$ , (también llamaremos recintos simples de tipo  $(1, 2)$  a los que admiten una representación análoga con algunas de las desigualdades  $\leq$  reemplazadas por  $<$ ). Según la proposición 10.15 los recintos de este tipo son medibles Jordan, y dada una función integrable  $f : E \rightarrow \mathbb{R}$ , vamos a exponer con detalle la aplicación del teorema de Fubini al cálculo de  $\int_E f$ .

Si  $c$  (resp.  $d$ ) es una cota inferior (resp. superior) de  $\alpha_1(x)$  (resp.  $\alpha_2(x)$ ) en  $[a, b]$ , podemos fijar el rectángulo  $A = [a, b] \times [c, d]$  para aplicar el teorema de Fubini, con  $X = [a, b]$ ,  $Y = [c, d]$ . Como estamos suponiendo que  $E$  es un recinto simple tipo  $(1, 2)$ , es natural considerar primero la integral respecto a la variable  $y$ , ya que cada sección  $E_x = [\alpha_1(x), \alpha_2(x)]$  es un intervalo. Si cada  $f_x$  es integrable sobre  $[\alpha_1(x), \alpha_2(x)]$ , (lo que ocurre cuando  $f$  es continua) podemos escribir

$$I(x) = \int_{\alpha_1(x)}^{\alpha_2(x)} f(x, y) dy$$

y la integral  $\int_E f(x, y) dx dy$  se expresa mediante la integral iterada

$$\int_E f(x, y) dx dy = \int_a^b \left[ \int_{\alpha_1(x)}^{\alpha_2(x)} f(x, y) dy \right] dx$$

El cálculo de esta integral iterada se puede abordar con los métodos del cálculo de una variable, basados en el cálculo de primitivas y la regla de Barrow.

Análogamente se pueden considerar los recintos simples de tipo  $(2, 1)$ , que son los de la forma

$$E = \{(x, y) : c \leq y \leq d, \beta_1(y) \leq x \leq \beta_2(y)\}$$

donde  $\beta_1 \leq \beta_2$  son integrables Riemann sobre  $[c, d]$ . En este caso la integral se expresa como integral iterada según el otro orden de integración

$$\int_E f(x, y) dx dy = \int_c^d \left[ \int_{\beta_1(y)}^{\beta_2(y)} f(x, y) dx \right] dy$$

NOTA: Si  $E \subset \mathbb{R}^2$  se puede expresar simultáneamente como recinto simple de tipo (1, 2), y como recinto simple de tipo (2, 1), hay dos alternativas para emprender el cálculo de la integral doble y conviene elegir, entre las dos integrales iteradas

$$\int_E f(x, y) dx dy = \int_a^b \left[ \int_{\alpha_1(x)}^{\alpha_2(x)} f(x, y) dy \right] dx = \int_c^d \left[ \int_{\beta_1(y)}^{\beta_2(y)} f(x, y) dx \right] dy$$

la que se pueda evaluar usando el cálculo de primitivas o la que conduzca a un cálculo más simple (véanse los ejercicios resueltos 11.13 y 11.14).

Finalmente, si  $E \subset \mathbb{R}^2$  no es simple, pero se puede expresar como unión disjunta de recintos simples  $E = E_1 \cup E_2 \cup \dots \cup E_r$ , el cálculo de la integral doble  $\int_E f(x, y) dx dy$  se puede abordar con los métodos anteriores utilizando la propiedad aditiva de la integral:  $\int_E f = \sum_{i=1}^r \int_{E_i} f$ .

**Integrales triples iteradas.** Consideremos en primer lugar integrales de funciones de tres variables sobre conjuntos  $E \subset \mathbb{R}^3$  que se puedan expresar en la forma

$$E = \{(x, y, z) : a \leq x \leq b, \alpha_1(x) \leq y \leq \alpha_2(x), \beta_1(x, y) \leq z \leq \beta_2(x, y)\}$$

donde  $\alpha_1 \leq \alpha_2$  son integrables sobre  $[a, b]$ , y  $\beta_1 \leq \beta_2$  son integrables sobre

$$F = \{(x, y) : a \leq x \leq b, \alpha_1(x) \leq y \leq \alpha_2(x)\}$$

A los conjuntos de este tipo los llamaremos *recintos simples* de tipo (1, 2, 3) (también incluimos en este tipo a los descritos en forma similar reemplazando algunas desigualdades  $\leq$  por desigualdades estrictas  $<$ ).

En virtud de la proposición 10.16,  $F$  es medible Jordan en  $\mathbb{R}^2$  y  $E$  medible Jordan en  $\mathbb{R}^3$ . Cada sección

$$E_x = \{(y, z) \in \mathbb{R}^2 : \alpha_1(x) \leq y \leq \alpha_2(x), \beta_1(x, y) \leq z \leq \beta_2(x, y)\}$$

es medible Jordan en  $\mathbb{R}^2$ , por ser un recinto simple de tipo (1, 2) en el plano  $yz$ . En lo que sigue se supone que  $f : E \rightarrow \mathbb{R}$  es integrable sobre  $E$ . De acuerdo con lo que se expuso en el apartado anterior, si  $f_x$  es integrable sobre  $E_x$ , su integral viene dada por

$$I(x) = \int_{E_x} f(x, y, z) dy dz = \int_{\alpha_1(x)}^{\alpha_2(x)} \left( \int_{\beta_1(x, y)}^{\beta_2(x, y)} f(x, y, z) dz \right) dy$$

luego

$$\int_E f(x, y, z) dx dy dz = \int_a^b \left( \int_{\alpha_1(x)}^{\alpha_2(x)} \left( \int_{\beta_1(x, y)}^{\beta_2(x, y)} f(x, y, z) dz \right) dy \right) dx$$

expresión que también se suele escribir, omitiendo los paréntesis, en la forma

$$\int_a^b \int_{\alpha_1(x)}^{\alpha_2(x)} \int_{\beta_1(x,y)}^{\beta_2(x,y)} f(x, y, z) dz dy dx = \int_a^b dx \int_{\alpha_1(x)}^{\alpha_2(x)} dy \int_{\beta_1(x,y)}^{\beta_2(x,y)} f(x, y, z) dz$$

Análogamente se pueden considerar los recintos simples de tipo  $\sigma = (i, j, k)$ , donde  $\sigma$  es una permutación de  $(1, 2, 3)$ . En este caso, si  $E \subset \mathbb{R}^3$  está descrito como recinto simple de tipo  $(i, j, k)$ , la integral triple

$$\int_M f(x_1, x_2, x_3) dx_1 dx_2 dx_3$$

se podrá expresar directamente como una integral en la que primero se integra respecto a la variable  $x_k$ , luego respecto a la variable  $x_j$ , y finalmente respecto a la variable  $x_i$ . A veces ocurre que un conjunto  $M \subset \mathbb{R}^3$  se puede describir como recinto simple de tipo  $\sigma$  para diferentes permutaciones  $\sigma$  de  $(1, 2, 3)$ . En este caso la integral triple se puede escribir, de varias formas, como integral iterada según los ordenes de integración que corresponden a estas permutaciones y convendrá elegir aquella integral iterada que conduzca a los cálculos más sencillos.

Finalmente, cuando  $E \subset \mathbb{R}^3$  no es simple, pero se puede expresar como unión disjunta de recintos simples  $E = E_1 \cup E_2 \cup \dots \cup E_r$ , el cálculo de la integral triple  $\int_E f(x, y, z) dx dy dz$ , se puede abordar con los métodos anteriores utilizando la propiedad aditiva de la integral:  $\int_E f = \sum_{i=1}^r \int_{E_i} f$ .

La definición de recinto simple se extiende a  $\mathbb{R}^n$  de manera obvia. Estos conjuntos son medibles Jordan en  $\mathbb{R}^n$ , y permiten expresar directamente la integral múltiple como una integral iterada.

## 11.2. Utilización del cambio de variable

Una técnica muy útil para el cálculo de integrales múltiples es la de cambio de variable. Frecuentemente el cálculo de una integral múltiple se puede simplificar bastante efectuando un cambio de variable que transforme la integral en otra cuyo cálculo sea accesible o más fácil.

Comenzamos recordando el significado de los conceptos que intervienen en el enunciado del teorema del cambio de variable:

Dada una aplicación  $g : \Omega \rightarrow \mathbb{R}$ , definida en un abierto  $\Omega \subset \mathbb{R}^n$ , cuando se fija un punto  $\mathbf{a} = (a_1, a_2, \dots, a_n) \in \Omega$ , en un entorno de  $a_1$  está definida la función parcial  $x_1 \rightarrow g(x_1, a_2, \dots, a_n)$ . Si esta función es derivable en  $a_1$ , su derivada,

$$\frac{\partial g}{\partial x_1}(\mathbf{a}) = \lim_{h \rightarrow 0} \frac{g(a_1 + h, a_2, \dots, a_n) - g(a_1, a_2, \dots, a_n)}{h}$$

también denotada  $D_1 g(\mathbf{a})$ , se llama derivada parcial de la función  $g$ , en el punto  $\mathbf{a}$ , respecto a la primera variable  $x_1$ . Análogamente se definen las derivadas parciales respecto a las restantes variables, denotadas

$$\frac{\partial g}{\partial x_j}(\mathbf{a}) = D_j g(\mathbf{a}), \quad 1 \leq j \leq n$$



Si en cada punto  $\mathbf{x} \in \Omega$  existen las derivadas parciales  $D_j g(\mathbf{x})$ ,  $1 \leq j \leq n$ , y todas las funciones  $\mathbf{x} \rightarrow D_j g(\mathbf{x})$  son continuas en  $\Omega$ , se dice que la función  $g$  es de clase  $C^1(\Omega)$ .

Una aplicación  $\mathbf{g} : \Omega \rightarrow \mathbb{R}^n$ , de componentes  $\mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), g_2(\mathbf{x}), \dots, g_n(\mathbf{x}))$ , definida en un abierto  $\Omega \subset \mathbb{R}^n$ , se dice que es de clase  $C^1(\Omega)$  cuando cada componente  $g_j$  ( $1 \leq j \leq n$ ), es de clase  $C^1(\Omega)$ . En este caso, la matriz  $(D_i g_j(\mathbf{x}))_{1 \leq i, j \leq n}$ , llamada matriz Jacobiana de  $\mathbf{g}$  en  $\mathbf{x}$ , la denotaremos por  $\mathbf{g}'(\mathbf{x})$ .

Los conceptos que acabamos de recordar son los que intervienen en el enunciado del teorema del cambio de variable para la integral de Riemann. Este teorema es un punto de encuentro del cálculo integral con el cálculo diferencial y en su demostración intervienen ambas teorías. Su demostración que es larga e involucra bastantes detalles técnicos, se expone con detalle en el apéndice J. Se suele comenzar demostrando un resultado preliminar (el caso de un cambio de variable lineal) que aquí figura como corolario 11.3.

**Teorema 11.2** [Cambio de variable] *Sea  $\mathbf{g} : \Omega \rightarrow \mathbb{R}^n$  una aplicación de clase  $C^1(\Omega)$  en un abierto  $\Omega \subset \mathbb{R}^n$ , y  $E \subset \mathbb{R}^n$  un conjunto medible Jordan que verifica*

- a)  $\overline{E} \subset \Omega$ .
- b)  $\det \mathbf{g}'(\mathbf{x}) \neq 0$  para cada  $\mathbf{x} \in E^\circ$ .
- c)  $\mathbf{g}$  es inyectiva sobre  $E^\circ$ .

Entonces  $M = \mathbf{g}(E)$  es medible Jordan, y para cada función integrable Riemann  $f : M \rightarrow \mathbb{R}$ , la función  $f \circ \mathbf{g} |\det \mathbf{g}'|$  es integrable sobre  $E$ , y se verifica

$$\int_M f(\mathbf{y}) \, d\mathbf{y} = \int_E f(\mathbf{g}(\mathbf{x})) |\det \mathbf{g}'(\mathbf{x})| \, d\mathbf{x}$$

En particular, cuando  $f = 1$ , se obtiene,

$$c_n(\mathbf{g}(E)) = \int_E |\det \mathbf{g}'(\mathbf{x})| \, d\mathbf{x}$$

OBSERVACIONES:

ii) Como consecuencia del teorema de la función inversa 8.13, cuando se cumplen las hipótesis del teorema anterior, la imagen  $V = \mathbf{g}(E^\circ)$  es abierta y la inversa de la biyección  $\mathbf{g} : E^\circ \rightarrow V$  también es de clase  $C^1(V)$ .

i) En las condiciones del teorema anterior, cuando  $n = 1$ , y  $E = [a, b]$  las hipótesis b) y c) significan que la restricción de  $g$  al intervalo  $(a, b)$  es una función estrictamente monótona con derivada continua no nula. Si  $g$  es estrictamente decreciente se cumplirá que  $g'(t) < 0$ , para todo  $t \in (a, b)$ , y la imagen  $g([a, b])$  será un intervalo compacto  $M = [c, d]$ , con  $c = g(b)$ , y  $d = g(a)$  de modo que la fórmula del cambio de variable se escribe en la forma

$$\int_c^d f(y) \, dy = - \int_a^b f(g(x)) g'(x) \, dx$$

Utilizando el convenio  $\int_d^c = -\int_c^d$  se obtiene la fórmula del cambio de variable en la forma habitual

$$\int_{g(a)}^{g(b)} f(y) dy = \int_a^b f(g(x))g'(x) dx$$

Se deja al cuidado del lector la comprobación de esta fórmula también es válida en el caso de un cambio de variable estrictamente creciente.

**Corolario 11.3** Si  $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$  es una aplicación lineal y  $E \subset \mathbb{R}^n$  es medible Jordan, entonces  $T(E)$  también lo es y  $c_n(T(E)) = |\det T|c_n(E)$ .

Si además  $T$  es una isometría para la distancia euclídea,  $c_n(T(E)) = c_n(E)$ .

DEM: En el teorema J.8 y también en [5], pág 151, se puede ver una demostración directa de este resultado preliminar. Para obtenerlo como corolario del teorema 11.2 basta aplicar este teorema cuando  $f$  es la función constante 1 y  $\mathbf{g} = T$  es una aplicación lineal pues, en ese caso, la matriz jacobiana  $\mathbf{g}'(\mathbf{x})$  coincide, en todo punto  $\mathbf{x} \in \mathbb{R}^n$ , con la matriz de la aplicación lineal  $T$ , luego

$$c_n(T(E)) = \int_{T(E)} 1 d\mathbf{y} = \int_E |\det(T)| d\mathbf{x} = |\det(T)|c_n(E)$$

Finalmente, si  $T$  es una isometría, y  $B = \{\mathbf{x} : \|\mathbf{x}\|_2 \leq 1\}$  se cumple  $T(B) = B$ , luego  $c_n(B) = c_n(T(B)) = |\det T|c_n(B)$ , luego,  $|\det T| = 1$ . (Se deja al cuidado del lector la comprobación de que  $B$  es medible Jordan y  $c_n(B) > 0$ ). ■

Los resultados en la siguiente proposición, que son consecuencia del teorema del cambio de variable, pueden resultar útiles en la práctica para simplificar el cálculo de algunas integrales (véanse los ejercicios resueltos 11.15, 11.16, y 11.19).

**Proposición 11.4** Sea  $M \subset \mathbb{R}^n$  medible Jordan,  $f : M \rightarrow \mathbb{R}$  integrable Riemann sobre  $M$ , y  $S_j(x_1, \dots, x_j, \dots, x_n) = (x_1, \dots, x_{j-1}, -x_j, x_{j+1}, \dots, x_n)$  la simetría respecto al hiperplano  $x_j = 0$ .

- Si  $M$  es simétrico respecto al origen y  $f$  es impar, ( $\mathbf{x} \in M \Rightarrow -\mathbf{x} \in M$ , y  $f(\mathbf{x}) + f(-\mathbf{x}) = 0$ ) entonces  $\int_M f(\mathbf{x}) d\mathbf{x} = 0$ .
- Si  $M$  simétrico respecto al hiperplano  $x_j = 0$ , y  $f$  es impar respecto a la variable  $x_j$ , ( $\mathbf{x} \in M \Rightarrow S_j(\mathbf{x}) \in M$ , y  $f(S_j(\mathbf{x})) = -f(\mathbf{x})$ ) entonces  $\int_M f(\mathbf{x}) d\mathbf{x} = 0$ .
- Si  $M$  simétrico respecto al hiperplano  $x_j = 0$ , y  $f$  es par respecto a la variable  $x_j$ , ( $\mathbf{x} \in M \Rightarrow S_j(\mathbf{x}) \in M$ , y  $f(S_j(\mathbf{x})) = f(\mathbf{x})$ ) entonces  $\int_M f(\mathbf{x}) d\mathbf{x} = 2 \int_{M^+} f(\mathbf{x}) d\mathbf{x} = 0$ , donde  $M^+ = \{\mathbf{x} \in M : x_j \geq 0\}$ .

DEM: a) Hacemos el cambio de variable lineal  $\mathbf{y} = T(\mathbf{x})$ , donde  $T(\mathbf{x}) = -\mathbf{x}$  es la simetría respecto al origen. Como  $M = T(M)$ ,  $f(T(\mathbf{x})) = -f(\mathbf{x})$ , y  $|\det(T)| = 1$ , resulta

$$\int_M f(\mathbf{y}) d\mathbf{y} = \int_{T(M)} f(\mathbf{y}) d\mathbf{y} = \int_M f(T(\mathbf{x})) d\mathbf{x} = - \int_M f(\mathbf{x}) d\mathbf{x}$$

luego  $\int_M f(\mathbf{x}) \, d\mathbf{x} = 0$ .

b) En el caso de una función impar se razona como en a), usando el cambio de variable lineal  $\mathbf{y} = S_j(\mathbf{x})$ , que cumple  $|\det(S_j)| = 1$ .

c) En el caso de una función par, en virtud de la proposición 10.12

$$\int_M f(\mathbf{x}) \, d\mathbf{x} = \int_{M^+} f(\mathbf{x}) \, d\mathbf{x} + \int_{M^-} f(\mathbf{x}) \, d\mathbf{x}, \quad \text{donde } M^- = \{\mathbf{x} \in M : x_j \leq 0\}.$$

Obsérvese que  $M^+, M^-$  son conjuntos medibles (porque se obtienen como intersección de  $M$  con intervalos cerrados apropiados) y que  $M^+ \cap M^-$  tiene contenido nulo (por ser un conjunto acotado contenido en el hiperplano  $x_j = 0$ ). Como  $M^- = S_j(M^+)$ ,  $f(S_j(\mathbf{x})) = f(\mathbf{x})$ , con el cambio de variable  $\mathbf{y} = S_j(\mathbf{x})$ , se obtiene

$$\int_{M^-} f(\mathbf{y}) \, d\mathbf{y} = \int_{S_j(M^+)} f(\mathbf{y}) \, d\mathbf{y} = \int_{M^+} f(S_j(\mathbf{x})) \, d\mathbf{x} = \int_{M^+} f(\mathbf{x}) \, d\mathbf{x}$$

■

NOTA: También se cumplen resultados análogos a los establecidos en los apartados b) y c) de la proposición 11.4, reemplazando el hiperplano  $x_j = 0$  por un hiperplano afín arbitrario. El enunciado preciso de los resultados y su demostración se deja como ejercicio al cuidado del lector.

Nuestro siguiente objetivo es mostrar algunas aplicaciones interesantes del teorema de cambio de variable. Comenzamos con su aplicación al cálculo efectivo de integrales dobles y triples mediante los siguiente cambios de variable que se usan con bastante frecuencia, y tienen interés especial:

a) *Coordenadas polares* (para integrales dobles):  $(x, y) = \mathbf{g}(r, \theta)$ , con

$$x = r \cos \theta, \quad y = r \sin \theta, \quad \det \mathbf{g}'(r, \theta) = r.$$

b) *Coordenadas cilíndricas* (para integrales triples):  $(x, y, t) = \mathbf{g}(r, \theta, t)$ , con

$$x = r \cos \theta, \quad y = r \sin \theta, \quad z = t, \quad \det \mathbf{g}'(r, \theta, t) = r.$$

c) *Coordenadas esféricas* (para integrales triples):  $(x, y, z) = \mathbf{g}(\rho, \theta, \varphi)$ , con

$$x = \rho \cos \varphi \cos \theta, \quad y = \rho \cos \varphi \sin \theta, \quad z = \rho \sin \varphi, \quad |\det \mathbf{g}'(\rho, \varphi, \theta)| = \rho^2 \cos \varphi$$

**Cambio de variable a coordenadas polares.** Para calcular una integral doble  $\int_M f(x, y) \, dx \, dy$ , mediante un cambio de variable a coordenadas polares, lo primero que hay que hacer es describir el recinto  $M$  en términos de las nuevas coordenadas, es decir, hay que expresarlo como imagen  $M = \mathbf{g}(E)$ , mediante el cambio de variable  $\mathbf{g}(r, \theta) = (r \cos \theta, r \sin \theta)$ , de un recinto  $E$ , que podemos suponerlo incluido

en un conjunto de la forma  $B = \{(r, \theta) : r \geq 0, \alpha \leq \theta \leq \beta\}$ , donde  $\beta - \alpha \leq 2\pi$ . Obsérvese que, al ser  $\mathbf{g}$  inyectiva sobre el interior de  $B$ , también lo es sobre el interior de  $E$ . También es claro que cada punto  $(r, \theta)$  interior a  $E$  es interior a  $B$ , y por lo tanto se cumple la condición  $\det \mathbf{g}'(r, \theta) = r > 0$ .

Con el cambio de variable a coordenadas polares se obtiene fácilmente la fórmula clásica para el cálculo de áreas de recintos planos limitados por una curva dada en coordenadas polares.

**Proposición 11.5** *Sea  $f : [\alpha, \beta] \rightarrow \mathbb{R}$  integrable Riemann,  $\beta - \alpha \leq 2\pi$ , y*

$$M = \{(r \cos \theta, r \sin \theta) : \alpha \leq \theta \leq \beta, 0 \leq r \leq f(\theta)\}$$

*el recinto limitado por la curva (en polares)  $r = f(\theta)$ , y los radios vectores en los extremos. Entonces  $M$  es medible Jordan en  $\mathbb{R}^2$ , y su área viene dada por*

$$c_2(M) = \frac{1}{2} \int_{\alpha}^{\beta} f(\theta)^2 d\theta$$

DEM: El recinto simple  $E = \{(r, \theta) : \alpha \leq \theta \leq \beta, 0 \leq r \leq f(\theta)\}$  es medible Jordan en  $\mathbb{R}^2$ , y el cambio de variable a coordenadas polares,  $\mathbf{g}(r, \theta) = (r \cos \theta, r \sin \theta)$ , cumple sobre  $E$  las hipótesis del teorema 11.2: La condición  $\beta - \alpha \leq 2\pi$  garantiza que  $\mathbf{g}$  es inyectiva en el interior de  $E$ , y es claro que en cada  $(r, \theta)$  interior a  $E$ , se cumple  $\det \mathbf{g}'(r, \theta) = r > 0$ . En virtud del teorema de cambio de variable  $M = \mathbf{g}(E)$  es medible Jordan en  $\mathbb{R}^2$ , con área

$$c_2(M) = \int_E \det |\mathbf{g}'(r, \theta)| dr d\theta$$

Para  $\theta \in [\alpha, \beta]$ , la sección  $E^\theta$  es el intervalo  $\{r : 0 \leq r \leq f(\theta)\}$ , luego

$$c_2(M) = \int_{\alpha}^{\beta} \left( \int_0^{f(\theta)} r dr \right) d\theta = \frac{1}{2} \int_{\alpha}^{\beta} f(\theta)^2 d\theta$$

■

**Cambio de variable a coordenadas esféricas y cilíndricas.** Si en una integral triple  $\int_M f(x, y, z) dx dy dz$  deseamos efectuar un cambio de variable a coordenadas esféricas o cilíndricas debemos empezar describiendo  $M \subset \mathbb{R}^3$  en términos de las nuevas variables. En el cambio de variable a coordenadas esféricas

$$\mathbf{g}(\rho, \theta, \varphi) = (\rho \cos \varphi \cos \theta, \rho \cos \varphi \sin \theta, \rho \sin \varphi)$$

debemos describir  $M$  en términos de las nuevas variables  $(\rho, \theta, \varphi)$ , es decir, en la forma  $M = \mathbf{g}(E)$ , donde

$$E \subset B = \{(\rho, \theta, \varphi) : \rho \geq 0, \alpha \leq \theta \leq \beta, -\pi/2 \leq \varphi \leq \pi/2\}, \text{ donde } \beta - \alpha \leq 2\pi$$

Es fácil ver que  $\mathbf{g}$  es inyectiva sobre el interior de  $B$ , luego también lo es sobre el interior de  $E$ . Además, en cada punto  $(\rho, \theta, \varphi)$  interior a  $E$ , se cumple

$$|\det \mathbf{g}'(\rho, \theta, \varphi)| = \rho^2 \cos \varphi > 0$$

Dejamos al cuidado del lector una comprobación análoga para el caso del cambio de variable a coordenadas cilíndricas.

**Volúmenes de cuerpos de revolución.** Con el teorema del cambio de variable también se pueden obtener fórmulas útiles para el cálculo de volúmenes de cuerpos de revolución. Dado un conjunto medible Jordan  $E \subset \mathbb{R}^2$ , lo consideramos inmerso en  $\mathbb{R}^3$ , mediante la aplicación natural  $j(x, y) = (x, y, 0)$  y denotamos por  $R_x(E) \subset \mathbb{R}^3$  el conjunto engendrado por la rotación de  $j(E)$  alrededor del eje  $Ox$ .

$$R_x(E) = \{(x, y \cos \varphi, y \sin \varphi) : (x, y) \in E, 0 \leq \varphi \leq 2\pi\}$$

(Análogamente se define  $R_y(E)$ ).

**Teorema 11.6** *Si  $E \subset \{(x, y) \in \mathbb{R}^2 : y \geq 0\}$  es medible Jordan en  $\mathbb{R}^2$ , entonces  $R_x(E) \subset \mathbb{R}^3$  es medible Jordan en  $\mathbb{R}^3$ , y su volumen viene dado por la fórmula*

$$c_3(R_x(E)) = 2\pi \int_E y dx dy$$

DEM: Según la proposición 10.15,  $M = \{(x, y, \varphi) : (x, y) \in E, 0 \leq \varphi \leq 2\pi\}$  es medible Jordan en  $\mathbb{R}^3$ , y con el cambio de variable  $\mathbf{g}(x, y, \varphi) = (x, y \cos \varphi, y \sin \varphi)$ , el conjunto  $R_x(E)$  se expresa en la forma  $R_x(E) = \mathbf{g}(M)$ . Con el teorema del cambio de variable se obtiene que  $R_x(E)$  es medible Jordan en  $\mathbb{R}^3$ , y

$$c_3(\mathbf{g}(M)) = \int_M \det |\mathbf{g}'(x, y, \varphi)| dx dy d\varphi$$

(Obsérvese que  $\mathbf{g}$  es inyectiva sobre el interior de  $M$ , y que  $\det \mathbf{g}'(x, y, \varphi) = y > 0$  en cada punto  $(x, y, \varphi)$  del interior de  $M$ ). Para todo  $\varphi \in [0, 2\pi]$ , la sección  $M^\varphi = \{(x, y) : (x, y, \varphi) \in M\}$  coincide con  $E$ , luego es medible, y en virtud del teorema de Fubini podemos escribir

$$c_3(\mathbf{g}(M)) = \int_0^{2\pi} \left( \int_E y dx dy \right) d\varphi = 2\pi \int_E y dx dy$$

En particular, cuando  $E = \{(x, y) \in \mathbb{R}^2 : a \leq x \leq b, 0 \leq y \leq f(x)\} \subset \mathbb{R}^2$  es el recinto de ordenadas asociado a una función integrable no negativa  $f : [a, b] \rightarrow \mathbb{R}$ , se cumplen las condiciones del teorema anterior, y efectuando una integración iterada

$$\int_E y \, dx \, dy = \int_a^b \left( \int_0^{f(x)} y \, dy \right) dx = \frac{1}{2} \int_a^b f(x)^2 \, dx$$

que proporciona la fórmula  $c_3(R_x(E)) = \pi \int_a^b f(x)^2 \, dx$ , que también se podría haber obtenido con el método de las secciones.

Si además  $a \geq 0$ , para el volumen del sólido que genera  $E$  al girar alrededor del eje  $Oy$  se obtiene

$$c_3(R_y(E)) = 2\pi \int_E x \, dx \, dy = 2\pi \int_a^b x f(x) \, dx$$

**Baricentros.** Si  $E \subset \mathbb{R}^n$  es medible Jordan, y  $c_n(E) > 0$ , se define el *baricentro* de  $E$  como el punto  $\mathbf{b}(E) \in \mathbb{R}^n$  de coordenadas

$$b_j(E) = \frac{1}{c_n(E)} \int_E x_j \, dx_1 \, dx_2 \cdots dx_n$$

Cuando  $E$  tiene alguna simetría es fácil obtener el baricentro, o algunas de sus coordenadas, sin necesidad de calcular integrales: Si  $E$  tiene centro de simetría entonces el centro de simetría es el baricentro y si  $E$  es simétrico respecto a un hiperplano entonces el baricentro está en el hiperplano. Para obtener estos resultados conviene establecer primero el siguiente

**Lema 11.7** *En las condiciones anteriores, si  $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$  es una aplicación lineal no singular y  $M = T(E)$ , se cumple que  $\mathbf{b}(M) = T(\mathbf{b}(E))$ , es decir, las aplicaciones lineales no singulares conservan los baricentros. Análogamente, las traslaciones conservan los baricentros.*

DEM: Observemos en primer lugar que al ser  $T$  una aplicación lineal no singular, según el corolario 11.3 se cumple  $c_n(M) = |\det(T)|c_n(E) > 0$ , y por lo tanto está definido el baricentro  $\mathbf{b}(M) = (b_1(M), b_2(M), \dots, b_n(M))$ , donde

$$c_n(M)b_k(M) = \int_M y_k \, dy_1 \, dy_2 \cdots dy_n$$

Si  $y_k = \sum_{j=1}^n a_{kj}x_j$ ,  $1 \leq k \leq n$ , son las ecuaciones de la transformación lineal  $\mathbf{y} = T(\mathbf{x})$ , con este cambio de variable la última integral se convierte en

$$\begin{aligned} c_n(M) b_k(M) &= \int_E \left( \sum_{j=1}^n a_{kj} x_j \right) |\det(T)| \, dx_1 \, dx_2 \cdots dx_n = \\ &= |\det(T)| \sum_{j=1}^n a_{kj} \int_E x_j \, dx_1 \, dx_2 \cdots dx_n = \\ &= |\det(T)| c_n(E) \sum_{j=1}^n a_{kj} b_j(E) = c_n(M) T(\mathbf{b}(E))_k \end{aligned}$$

y dividiendo por  $c_n(M) > 0$  se obtiene el resultado  $\mathbf{b}(M) = T(\mathbf{b}(E))$ . Dejamos al cuidado del lector la afirmación, más sencilla, referente a las traslaciones. ■

**Proposición 11.8** *Sea  $E \subset \mathbb{R}^n$  un conjunto medible Jordan tal que  $c_n(E) > 0$ .*

- i) *Si  $E$  es simétrico respecto a un hiperplano, su baricentro está en el hiperplano.*  
 ii) *Si  $E$  es simétrico respecto a un punto, su baricentro es ese punto.*

DEM: i) Según el lema 11.7, los baricentros se conservan mediante traslaciones y aplicaciones lineales luego no es restrictiva hacer la demostración cuando el hiperplano de simetría es  $x_1 = 0$ . En este caso, la simetría viene dada por la aplicación lineal  $T(x_1, x_2, \dots, x_n) = (-x_1, x_2, \dots, x_n)$ . Por hipótesis  $T(E) = E$ , y el cambio de variable lineal  $\mathbf{y} = T(\mathbf{x})$ , cumple  $|\det(T)| = 1$ , luego

$$\int_E y_1 dy_1 dy_2 \cdots dy_n = \int_{T(E)} y_1 dy_1 dy_2 \cdots dy_n = - \int_E x_1 dx_1 dx_2 \cdots dx_n$$

luego la primera coordenada de  $\mathbf{b}(E)_1$  es nula, lo que significa que el baricentro  $\mathbf{b}(E)$  está en el hiperplano de simetría.

ii) En virtud del lema 11.7, no es restrictivo suponer que  $E$  es simétrico respecto al origen  $\mathbf{0}$ , y la demostración es análoga a la anterior considerando ahora el cambio de variable  $T(\mathbf{x}) = -\mathbf{x}$ , para el que se cumple  $T(E) = E$ . ■

El interés de estos resultados sobre a baricentros se pondrá de manifiesto después del siguiente teorema, que es una reformulación de 11.6, en la que se hace intervenir la noción de baricentro.

**Teorema 11.9** [Guldin] *Sea  $E \subset \{(x, y) \in \mathbb{R}^2 : y \geq 0\}$  un recinto medible Jordan de área  $c_2(E) > 0$ . Si  $\mathbf{b} = (x_0, y_0)$  es su baricentro, se verifica*

$$c_3(R_x(E)) = 2\pi y_0 c_2(E)$$

*(El volumen del cuerpo generado por  $E$  al girar alrededor del eje  $Ox$  es igual al producto de su área por la longitud de la circunferencia que describe su baricentro).*

DEM: Es consecuencia directa del teorema 11.6 y de la definición de baricentro. ■

### Ejemplo 11.10

Se considera el cuerpo de revolución generado al girar alrededor del eje  $Ox$  el disco

$$\{(x, y) : x^2 + (y - b)^2 \leq r^2\}, \text{ donde } r > b > 0$$

Según la proposición 11.8 el baricentro del disco es su centro, y aplicando el teorema de Guldin se obtiene que el volumen del cuerpo es  $2\pi^2 br^2$ . ■

NOTA: Para la validez de teorema de Guldin no hace falta suponer que  $E$  esta inmerso en el plano  $z = 0$ , ni que el eje de giro sea  $Ox$ . El resultado es válido

para un subconjunto  $E$  de un plano afín arbitrario  $H$  que gira alrededor de un eje  $\mathbf{e}$  situado en ese plano. Naturalmente que hay que suponer que  $E$  es medible Jordan dentro de  $H$ , con  $c_H(E) > 0$ , y que  $E$  está contenido en uno de los dos semiplanos que la recta  $\mathbf{e}$  determina en  $H$  (véase [5], pág. 165-66).

La definición, en un subespacio afín  $H \subset \mathbb{R}^n$ , de la noción intrínseca de subconjunto medible Jordan, y del contenido de Jordan asociado  $c_H$ , se puede ver con detalle en la página 149 de [5].

### 11.3. Ejercicios resueltos

**Ejercicio 11.11** *Arquímedes, en una carta a su amigo Eratóstenes (matemático y bibliotecario de Alejandría) donde le expuso su Método de cálculo de áreas, volúmenes y centros de gravedad, enunció el siguiente teorema: Si en un cubo se inscribe un cilindro que tiene las bases situadas en dos cuadrados opuestos y la superficie tangente a los cuatro planos restantes, y se inscribe en el mismo cubo otro cilindro con las bases en otros dos cuadrados y la superficie tangente a los cuatro planos restantes, la figura comprendida por las superficies de los cilindros e inscrita en ambos es dos tercios del cubo entero.*

*Utilice el método de las secciones para demostrar este resultado.*

SOLUCIÓN

Supongamos que la figura del enunciado está inscrita en un cubo de lado  $2r$ , de modo que las bases de los cilindros son círculos de radio  $r$ . Fijando un sistema de referencia formado por unos ejes cartesianos rectangulares donde los ejes  $Ox$ ,  $Oy$  coinciden con los ejes de los cilindros, podemos describir analíticamente la figura en la forma siguiente:

$$E = \{(x, y, z) : x^2 + z^2 \leq r^2, y^2 + z^2 \leq r^2\}$$

Para justificar que  $E$  es medible Jordan basta observar que es un recinto del tipo considerado en la proposición 10.15:

$$E = \{(x, y, z) : (x, y) \in Q(r) : -f(x, y) \leq z \leq f(x, y)\}$$

donde  $f(x, y) = \min\{\sqrt{r^2 - y^2}, \sqrt{r^2 - x^2}\}$  es una función continua sobre el cuadrado  $Q(r) = [-r, r] \times [-r, r]$ .

La intersección del recinto  $E$  con el plano  $z = t$  es no vacía si y sólo si  $t \in [-r, r]$ , y en este caso la sección  $\{(x, y) : (x, y, t) \in E\}$  es el cuadrado

$$E^t = \{(x, y) : |x| \leq \sqrt{r^2 - t^2}, |y| \leq \sqrt{r^2 - t^2}, \}$$

de área  $S(t) = 4(r^2 - t^2)$ , luego

$$c_3(E) = \int_{-r}^r S(t) dt = \frac{16}{3} R^3 = \frac{2}{3}(8r^3)$$

donde  $8r^3$  es el volumen del cubo circunscrito. ■



**Ejercicio 11.12** Calcule el área del recinto elíptico

$$E(a, b) = \{(x, y) : x^2/a^2 + y^2/b^2 \leq 1\}$$

y volumen del cuerpo limitado por el elipsoide

$$E(a, b, c) = \{(x, y, z) : x^2/a^2 + y^2/b^2 + z^2/c^2 \leq 1\}$$

donde  $a > 0, b > 0, c > 0$ .

SOLUCIÓN

$$E(a, b) = \{(x, y) : -a \leq x \leq a, -b\sqrt{1 - x^2/a^2} \leq y \leq b\sqrt{1 - x^2/a^2}\}$$

es un recinto simple de tipo (1, 2) luego, según la proposición 10.15,

$$c_2(E(a, b)) = \int_{-a}^a 2b\sqrt{1 - x^2/a^2} dx$$

Esta integral se calcula con el cambio de variable  $x = a \operatorname{sen} t$ , y se obtiene

$$c_2(E(a, b)) = 2ab \int_{-\pi/2}^{\pi/2} \cos^2 t dt = \pi ab$$

El volumen del recinto  $E(a, b, c)$  se puede calcular por el método de las secciones: La intersección de este recinto con el plano  $z = t$  es no vacía si y sólo si  $t \in [-c, c]$ , y en este caso la sección  $\{(x, y) : (x, y, t) \in E(a, b, c)\}$  es el recinto elíptico

$$\{(x, y) : x^2/a^2 + y^2/b^2 \leq 1 - c^2/t^2\}$$

limitado por una elipse de semiejes,  $a\sqrt{1 - c^2/t^2}$ ,  $b\sqrt{1 - c^2/t^2}$ , cuya área es

$$S(t) = \pi ab(1 - t^2/c^2)$$

luego el volumen de  $E(a, b, c)$  viene dado por la integral  $\int_{-c}^c S(t) dt = \frac{4}{3}\pi abc$ .

El mismo resultado se obtiene considerando que  $E(a, b, c) = T(B(\mathbf{0}, 1))$ , donde  $B(\mathbf{0}, 1)$  es la bola euclídea de centro el origen  $\mathbf{0}$  y radio 1, y  $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  es la transformación lineal  $T(z, y, z) = (ax, by, cz)$ , de determinante  $\det(T) = abc$ .

El volumen de la bola euclídea  $B(\mathbf{0}, 1)$ , se calcula fácilmente mediante un cambio de variable a coordenadas esféricas

$$c_3(B(\mathbf{0}, 1)) = \int_0^{2\pi} d\theta \int_{-\pi/2}^{\pi/2} d\varphi \int_0^1 \rho^2 \cos \varphi d\rho = \frac{4}{3}\pi$$

y según el ejercicio 11.3, el volumen de  $E(a, b, c)$  viene dado por

$$c_3(E(a, b, c)) = |\det(T)|c_3(B(\mathbf{0}, 1)) = \frac{4}{3}\pi abc$$

■

**Ejercicio 11.13** Calcule las siguientes integrales dobles

$$a) \quad I = \int_A ye^{xy} dx dy, \quad A = [0, 1] \times [0, 1].$$

$$b) \quad J = \int_E \sqrt{a^2 - y^2} dx dy, \quad E = \{(x, y) : x^2 + y^2 \leq a^2, x \geq 0, y \geq 0\}, (a > 0).$$

$$c) \quad K = \int_E (x-1)\sqrt{1+e^{2y}} dx dy, \quad E = \{(x, y) : 1 \leq x \leq 2, 0 \leq y \leq \log x\}.$$

SOLUCIÓN

a) Según el teorema de Fubini,

$$I = \int_0^1 \left( \int_0^1 ye^{xy} dy \right) dx = \int_0^1 \left( \int_0^1 ye^{xy} dx \right) dy$$

El cálculo de la primera integral iterada comienza con la búsqueda de una primitiva de la función  $y \rightarrow ye^{xy}$  (que se podría encontrar mediante una rutinaria integración por partes). Abandonamos este camino porque el cálculo de la segunda integral iterada es mucho más rápido, ya que la función parcial  $x \rightarrow ye^{xy}$  tiene la primitiva inmediata  $x \rightarrow e^{xy}$ , con la que se obtiene  $I = \int_0^1 (e^y - 1) dy = e - 2$ .

b)  $E$  se describe fácilmente como recinto simple de tipo (1, 2) y de tipo (2, 1):

$$E = \{(x, y) : 0 \leq x \leq a, 0 \leq y \leq \sqrt{a^2 - x^2}\} = \{(x, y) : 0 \leq y \leq a, 0 \leq x \leq \sqrt{a^2 - y^2}\}$$

De las dos alternativas calcular  $J$  mediante una integral iterada

$$J = \int_0^a \left( \int_0^{\sqrt{a^2 - y^2}} \sqrt{a^2 - y^2} dx \right) dy = \int_0^a \left( \int_0^{\sqrt{a^2 - x^2}} \sqrt{a^2 - y^2} dy \right) dx$$

conviene la primera, que involucra el cálculo más fácil:  $J = \int_0^a (a^2 - y^2) dy = \frac{2}{3}a^3$ .

c) El recinto de integración, que viene descrito como recinto simple de tipo (1, 2), conviene expresarlo como recinto simple de tipo (2, 1):

$$E = \{(x, y) : 0 \leq y \leq \log 2, e^y \leq x \leq 2\}$$

ya que el cálculo de  $I$  mediante una integración iterada parece más fácil si empezamos integrando respecto a la variable  $x$

$$K = \int_0^{\log 2} \left( \int_{e^y}^2 (x-1)\sqrt{1+e^{2y}} dx \right) dy = \int_0^{\log 2} \left( e^y - \frac{1}{2}e^{2y} \right) \sqrt{1+e^{2y}} dy = K_1 - \frac{K_2}{2}$$

donde  $K_1 = \int_0^{\log 2} e^y \sqrt{1+e^{2y}} dy$ ,  $K_2 = \int_0^{\log 2} e^{2y} \sqrt{1+e^{2y}} dy$ .

Con los cambios de variable  $s = e^y$ ,  $t = e^{2y}$ , se llega a las integrales

$$K_1 = \int_1^2 \sqrt{1+s^2} ds, \quad K_2 = \frac{1}{2} \int_1^4 \sqrt{1+t} dt$$

que se pueden calcular fácilmente con las técnicas usuales del cálculo de primitivas.

■

**Ejercicio 11.14** Calcule las siguientes integrales dobles

$$a) \quad I = \int_E x \, dx \, dy, \quad E = \{(x, y) : x^2 + y^2 \leq 2x\}.$$

$$b) \quad J = \int_E x \sqrt{1 - x^2 - y^2} \, dx \, dy, \quad E = \{(x, y) : x^2 + y^2 \leq 1, x \geq 0, y \geq 0\}.$$

SOLUCIÓN

a) El disco  $E = \{(x, y) : (x - 1)^2 + y^2 \leq 1\}$  se describe fácilmente como recinto simple de tipo (1, 2):

$$E = \{(x, y) : 0 \leq x \leq 2, -\sqrt{2x - x^2} \leq y \leq \sqrt{2x - x^2}\}$$

lo que permite calcular  $I$  mediante la integral iterada

$$I = \int_0^2 \left( \int_{-\sqrt{2x-x^2}}^{+\sqrt{2x-x^2}} x \, dy \right) dx = \int_0^2 2x\sqrt{2x-x^2} \, dx$$

Como  $2x - x^2 = 1 - (1 - x)^2$ , es natural efectuar el cambio de variable  $x = 1 + \sin t$  con el que se obtiene la integral

$$I = 2 \int_{-\pi/2}^{\pi/2} (1 + \sin t) \cos^2 t \, dt = \int_{-\pi/2}^{\pi/2} (1 + \cos 2t + 2 \cos^2 t \sin t) dt = \pi$$

La integral  $I$  también se puede calcular con un cambio de variable a coordenadas polares. La descripción de  $E$  en coordenadas polares es bien sencilla

$$E = \{(r \cos \theta, r \sin \theta) : -\pi/2 \leq \theta \leq \pi/2, 0 \leq r \leq 2 \cos \theta\}$$

y se obtiene

$$I = \int_{-\pi/2}^{\pi/2} \left( \int_0^{2 \cos \theta} r^2 \cos \theta \, dr \right) d\theta = \frac{8}{3} \int_{-\pi/2}^{\pi/2} \cos^4 \theta \, d\theta = \pi$$

donde la última integral se ha calculado usando el desarrollo:

$$\cos^4 \theta = \left( \frac{1 + \cos 2\theta}{2} \right)^2 = \frac{1}{4} \left( 1 + 2 \cos 2\theta + \frac{1 + \cos 4\theta}{2} \right)$$

b)  $E$  es un recinto simple de tipo (1, 2) y de tipo (2, 1):

$$E = \{(x, y) : 0 \leq x \leq 1, 0 \leq y \leq \sqrt{1 - x^2}\} = \{(x, y) : 0 \leq y \leq 1, 0 \leq x \leq \sqrt{1 - y^2}\}$$

y de las dos alternativas para calcular  $J$  mediante una integral iterada

$$J = \int_0^1 \left( \int_0^{\sqrt{1-y^2}} x \sqrt{1-x^2-y^2} dx \right) dy = \int_0^1 \left( \int_0^{\sqrt{1-x^2}} x \sqrt{1-x^2-y^2} dy \right) dx$$

elegimos la primera, que involucra el cálculo de una primitiva inmediata (respecto a la variable  $x$ ). Se obtiene así

$$J = \frac{1}{3} \int_0^1 (1-y^2)^{3/2} dy$$

Al mismo resultado se llega con un cambio de variable a coordenadas polares en la integral doble:

$$J = \int_0^1 \left( \int_0^{\pi/2} r^2 \sqrt{1-r^2} \cos \theta d\theta \right) dr = \int_0^1 r^2 \sqrt{1-r^2} dr = \frac{1}{3} \int_0^1 (1-r^2)^{3/2} dr$$

donde la última igualdad es consecuencia de una integración por partes.

Con el cambio de variable  $y = \sin t$  se obtiene

$$J = \frac{1}{3} \int_0^1 (1-y^2)^{3/2} dy = \frac{1}{3} \int_0^{\pi/2} \cos^4 t dt$$

La última integral se calcula mediante el desarrollo indicado en el apartado a), y se obtiene que  $J = \pi/16$ . ■

**Ejercicio 11.15** *Se supone que  $M \subset \mathbb{R}^2$  es medible Jordan y simétrico respecto a la recta  $x = y$ . Demuestre que  $E = \{(x, y) \in M : y \leq x\}$  también es medible Jordan y que si  $f : M \rightarrow \mathbb{R}$  es una función integrable Riemann que verifica  $f(x, y) = f(y, x)$  para todo  $(x, y) \in M$  entonces*

$$\int_M f(x, y) dx dy = 2 \int_E f(x, y) dy$$

*Utilice esta propiedad y un cambio de variable apropiado, para calcular*

$$\int_0^1 \left( \int_0^1 \sqrt{x^2 + y^2} dx \right) dy$$

SOLUCIÓN

Como  $M$  es acotado, estará contenido en un cuadrado  $Q = [-R, R] \times [-R, R]$ . Como el triángulo  $T = \{(x, y) : -R \leq x \leq R, -R \leq y \leq x\}$  es medible Jordan, también lo será la intersección  $E = T \cap M$ .

Análogamente  $F = \{(x, y) \in M : y \geq x\}$  es medible Jordan. Evidentemente  $M = E \cup F$ , y aunque  $E \cap F$  no es vacío, como tiene contenido nulo, podemos asegurar que

$$\int_M f(x, y) dx dy = \int_E f(x, y) dx dy + \int_F f(x, y) dx dy$$

En la última integral se efectúa el cambio de variable dado por  $g(u, v) = (v, u)$ . Como  $F = g(E)$ ,  $|g'(u, v)| = 1$  y  $f = f \circ g$  resulta:

$$\int_F f(x, y) dx dy = \int_{g(E)} f(x, y) dx dy = \int_E f(g(u, v)) du dv = \int_E f(u, v) du dv$$

luego

$$\int_M f(x, y) dx dy = 2 \int_E f(x, y) dx dy$$

Aplicando la fórmula anterior cuando  $f(x, y) = \sqrt{x^2 + y^2}$ ,  $M = [0, 1] \times [0, 1]$ , y  $E = \{(x, y) : 0 \leq x \leq 1, 0 \leq y \leq x\}$ , con un cambio a coordenadas polares, resulta

$$\begin{aligned} I &= \int_0^1 \left( \int_0^1 \sqrt{x^2 + y^2} dx \right) dy = 2 \int_E \sqrt{x^2 + y^2} dx dy = \int_0^{\pi/4} \left( \int_0^{1/\cos\theta} r^2 dr \right) d\theta = \\ &= \frac{1}{3} \int_0^{\pi/4} \frac{d\theta}{\cos^3\theta} = \frac{1}{3} \int_0^{\pi/4} \frac{\cos\theta d\theta}{(1 - \sin^2\theta)^2} = \frac{1}{3} \int_0^{1/\sqrt{2}} \frac{dt}{(1 - t^2)^2} \end{aligned}$$

Hemos llegado a la integral de una función racional, fácilmente calculable por los métodos estandar y con ello se considera resuelto el problema. ■

**Ejercicio 11.16** Se considera la integral  $\alpha = \int_Q f(x, y) dx dy$ , donde

$$f(x, y) = \frac{x}{(1 + x^2)(1 + xy)}; \quad Q = [0, 1] \times [0, 1]$$

Justifique la igualdad  $2\alpha = \int_Q (f(x, y) + f(y, x)) dx dy$  y utilícela para calcular  $\alpha$  y el valor de la integral  $\int_0^1 \frac{\log(1+x)}{1+x^2} dx$

SOLUCIÓN

Con el cambio de variable  $\mathbf{g}(x, y) = (y, x)$ , teniendo en cuenta que  $\mathbf{g}(Q) = Q$ , se obtiene que  $\int_Q f(y, x) dx dy = \int_Q f(x, y) dx dy$  luego  $2\alpha = \int_Q (f(x, y) + f(y, x)) dx dy$ .

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

luego

$$2\alpha = \int_Q \frac{x dx dy}{(1+x^2)(1+y^2)} + \int_Q \frac{y dx dy}{(1+x^2)(1+y^2)}$$

Estas integrales se calculan directamente usando integración iterada:

$$\int_Q \frac{x dx dy}{(1+x^2)(1+y^2)} = \int_0^1 \frac{x}{1+x^2} \left( \int_0^1 \frac{dy}{1+y^2} \right) dx = \frac{\pi}{8} \log 2$$

Análogamente

$$\int_Q \frac{y \, dx \, dy}{(1+x^2)(1+y^2)} = \frac{\pi}{8} \log 2$$

y se obtiene así el valor  $\alpha = (\pi/8)/\log 2$ . Con el teorema de Fubini se obtiene

$$\frac{\pi}{8} \log 2 = \int_Q \frac{x \, dx \, dy}{(1+x^2)(1+xy)} = \int_0^1 \frac{dx}{1+x^2} \int_0^1 \frac{x \, dy}{1+xy} = \int_0^1 \frac{\log(1+x)}{1+x^2} dx$$

■

**Ejercicio 11.17** *En la integral*

$$\int_M \frac{dx \, dy \, dz}{\sqrt{x^2 + y^2}} \quad \text{donde} \quad M = \{(x, y, z) : x^2 + y^2 \geq 1, x^2 + y^2 + z^2 \leq 4\}$$

- Efectúe un cambio de variable a coordenadas esféricas.
- Efectúe un cambio de variable a coordenadas cilíndricas.
- Calcule su valor.

SOLUCIÓN

a) Cambio de variable a coordenadas esféricas

$$x = \rho \cos \varphi \cos \theta; \quad y = \rho \cos \varphi \sin \theta; \quad z = \rho \sin \theta$$

$$\mathbf{g}(\rho, \theta, \varphi) = (x, y, z); \quad |\det \mathbf{g}'(\rho, \theta, \varphi)| = \rho^2 \cos \varphi$$

$$M = g(B), \quad \text{con } B = \{(\rho, \theta, \varphi) : 0 \leq \theta \leq 2\pi, |\varphi| < \pi/3, 1/\cos \varphi \leq \rho \leq 2\}.$$

$$\begin{aligned} I &= \int_B \frac{\rho^2 \cos \varphi}{\sqrt{\rho^2 \cos^2 \varphi}} \, d\theta \, d\varphi \, d\rho = \int_0^{2\pi} d\theta \int_{-\pi/3}^{\pi/3} d\varphi \int_{1/\cos \varphi}^2 \rho \, d\rho = \\ &= 2\pi \int_{-\pi/3}^{\pi/3} \frac{1}{2} \left( 4 - \frac{1}{\cos^2 \varphi} \right) d\varphi = \frac{8\pi^2}{3} - 2\pi\sqrt{3} \end{aligned}$$

b) Cambio de variable a coordenadas cilíndricas

$$x = r \cos \theta; \quad y = r \sin \theta; \quad z = t$$

$$\mathbf{g}(r, \theta, t) = (x, y, z); \quad |\det \mathbf{g}'(r, \theta, t)| = r$$

$$M = g(A), \quad \text{con } A = \{(r, \theta, t) : 0 \leq \theta \leq 2\pi, 1 \leq r \leq 2, |t| \leq \sqrt{4-r^2}\}.$$

$$I = \int_0^{2\pi} d\theta \int_1^2 dr \int_{-\sqrt{4-r^2}}^{\sqrt{4-r^2}} dt = 4\pi \int_1^2 \sqrt{4-r^2} \, dr = \dots = \frac{8\pi^2}{3} - 2\pi\sqrt{3}$$

■

**Ejercicio 11.18** Sea  $M = \{(x, y, z) \in \mathbb{R}^3 : a^2 < x^2 + y^2 + z^2 < b^2\}$ . Para  $R \notin [a, b]$  calcule el valor de la integral

$$I(R) = \int_M \frac{dx \, dy \, dz}{\sqrt{x^2 + y^2 + (z - R)^2}}$$

y compruebe que no depende de  $R$  para  $0 < R < a$ .

SOLUCIÓN

Con un cambio de variable a coordenadas esféricas se llega a

$$I(R) = 2\pi \int_a^b \rho^2 J(\rho) \, d\rho \quad \text{donde} \quad J(\rho) = \int_{-\pi/2}^{\pi/2} \frac{\cos \varphi \, d\varphi}{\sqrt{\rho^2 - 2R\rho \sin \varphi + R^2}}$$

Con el cambio de variable  $u = 2R\rho \sin \varphi$  se calcula

$$J(\rho) = \frac{1}{2R\rho} \int_{-2R\rho}^{2R\rho} \frac{du}{\sqrt{R^2 + \rho^2 - u}} = \dots = \frac{\sqrt{(R+\rho)^2} - \sqrt{(R-\rho)^2}}{2R\rho} = \frac{|R+\rho| - |R-\rho|}{2R\rho}$$

Teniendo en cuenta que  $a \leq \rho \leq b$ , resulta

$$\text{a) } R < a \Rightarrow R < \rho \Rightarrow J(\rho) = 1/\rho \Rightarrow I(R) = 2\pi(b^2 - a^2)$$

$$\text{b) } R > b \Rightarrow R > \rho \Rightarrow J(\rho) = 1/R \Rightarrow I(R) = \frac{2\pi}{3R}(b^3 - a^3)$$

■

**Ejercicio 11.19** Calcule la integral  $I = \int_M (x + y + z) \, dx \, dy \, dz$ , y el baricentro de

$$M = \{(x, y, z) : x^2 + y^2 \leq 2z, \quad x^2 + y^2 + z^2 \leq 3\}$$

SOLUCIÓN

La esfera  $x^2 + y^2 + z^2 = 3$  y el paraboloido de revolución  $2z = x^2 + y^2$  se cortan según una circunferencia, situada en el plano  $z = 1$ , cuya proyección en el plano  $xy$  es la circunferencia  $x^2 + y^2 = 2$ .

[En efecto, si el punto  $(x, y, z)$  satisface las ecuaciones de las dos superficies debe cumplir  $z \geq 0$  y  $2z + z^2 = 3$ , luego  $z = 1$  y  $x^2 + y^2 = 2$ ].

Es decir,  $M$  es el recinto, situado sobre el disco  $x^2 + y^2 \leq 2$  del plano  $xy$ , entre el paraboloido (abajo) y la esfera (encima):

$$M = \{(x, y, z) : x^2 + y^2 \leq 2, \quad (x^2 + y^2)/2 \leq z \leq \sqrt{3 - (x^2 + y^2)}\}$$

Como el eje  $Oz$  es un eje de simetría de  $M$ , su baricentro  $(x_0, y_0, z_0)$ , está en este eje y debe cumplir

$$x_0 = \frac{1}{v(M)} \int_M x \, dx \, dy \, dz = 0; \quad y_0 = \frac{1}{v(M)} \int_M y \, dx \, dy \, dz = 0.$$

Se sigue que

$$I = \int_M z \, dx \, dy \, dz, \quad y \quad z_0 = I/v(M)$$

Usando coordenadas cilíndricas  $x = r \cos \theta$ ,  $y = r \sin \theta$ ,  $z = t$ , se tiene

$$M = \{(r \cos \theta, r \sin \theta, t) : 0 \leq \theta \leq 2\pi, 0 \leq r \leq \sqrt{2}, r^2/2 \leq t \leq \sqrt{3-r^2}\}$$

luego

$$I = \int_0^{2\pi} d\theta \int_0^{\sqrt{2}} r \, dr \int_{r^2/2}^{\sqrt{3-r^2}} t \, dt = 2\pi \int_0^{\sqrt{2}} \frac{r}{2} (3 - r^2 - r^4/4) \, dr = \dots$$

$$v(M) = \int_0^{2\pi} d\theta \int_0^{\sqrt{2}} r \, dr \int_{r^2/2}^{\sqrt{3-r^2}} dt = 2\pi \int_0^{\sqrt{2}} r(\sqrt{3-r^2} - r^2/2) \, dr = \dots$$

integrales inmediatas que se dejan al cuidado del lector. ■

**Ejercicio 11.20** *Se consideran los recintos planos*

$$A = \{(x, y) : (x-1)^2 + y^2 \leq \sqrt{2}\}, \quad B = \{(x, y) : (x+1)^2 + y^2 \leq \sqrt{2}\}$$

- i) Calcule el volumen del sólido generado por  $A \cap B$  al girar alrededor del eje  $Oy$ .
- ii) Sean  $E, F \subset \mathbb{R}^2$  medibles Jordan de área no nula tales que  $E \cap F$  tiene contenido nulo. Obtenga una fórmula relacione el baricentro de  $E \cup F$  con los baricentros de  $E$  y  $F$ . Utilice esta relación para calcular, sin usar integrales, el volumen del sólido generado por  $A \setminus B$  al girar alrededor del eje  $Oy$ .

SOLUCIÓN

i) Volumen del sólido generado al girar  $E = A \cap B$  alrededor del eje  $Oy$ .

Las circunferencias que limitan  $A$  y  $B$ , de radio  $R = \sqrt[4]{2} > 1$ , y centros  $(1, 0)$ ,  $(-1, 0)$  se cortan en los puntos  $(0, b)$  y  $(0, -b)$  donde  $b^2 = R^2 - 1$ .

Calcularemos el volumen  $V(b)$  en función de  $b$ , y sustituyendo luego el valor concreto  $b = \sqrt{\sqrt{2} - 1}$  se obtendrá el volumen pedido. En lo que sigue  $\alpha = \arccos(b/R) \in (0, \pi/2)$ , de modo que  $R \cos \alpha = b$ ,  $R \sin \alpha = 1$  y  $\operatorname{tg} \alpha = b$ .

a) *Solución utilizando el método de las secciones:*

Cortando por planos perpendiculares al eje de revolución, para  $-b \leq y \leq b$ , se obtienen círculos de radio  $f(y) = -1 + \sqrt{R^2 - y^2}$  luego

$$\begin{aligned} V_b &= \int_{-b}^b \pi f(y)^2 \, dy = \pi \int_{-b}^b [1 + R^2 - y^2 - 2\sqrt{R^2 - y^2}] \, dy = \\ &= \pi[(1 + R^2)2b - \frac{2}{3}b^3 - 2I] = \pi[(2 + b^2)2b - \frac{2}{3}b^3 - 2I] \end{aligned}$$



donde  $I = \int_{-b}^b \sqrt{R^2 - y^2} dy$  se calcula con el cambio de variable  $y = R \operatorname{sen} t$ :

$$I = \int_{-\alpha}^{\alpha} R^2 \cos^2 t dt = \dots = R^2 \alpha + R^2 \operatorname{sen} \alpha \cos \alpha = R^2 \alpha + b$$

y con este valor de  $I$  se obtiene  $V(b) = 2\pi[b + \frac{2}{3}b^3 - \alpha(1 + b^2)]$ .

b) *Solución utilizando la fórmula general:*

$$V_b = 2\pi \int_M x dx dy, \quad \text{donde } M = \{(x, y) \in B : x \geq 0\}$$

Esta integral se calcula con el cambio de variable  $x = -1 + r \cos \theta$ ,  $y = r \operatorname{sen} \theta$ .

$$\begin{aligned} V_b &= 2\pi \int_{-\alpha}^{\alpha} d\theta \int_{1/(\cos \theta)}^R (-1 + r \cos \theta) r dr = 2\pi \int_{-\alpha}^{\alpha} \left( \frac{1}{6 \cos^2 \theta} - \frac{R^2}{2} + \frac{R^3}{3} \cos \theta \right) d\theta = \\ &= 2\pi \left[ \frac{1}{3} \operatorname{tg} \alpha - R^2 \alpha + \frac{2}{3} R^3 \operatorname{sen} \alpha \right] = 2\pi \left[ b + \frac{2}{3} b^3 - \alpha(1 + b^2) \right] \end{aligned}$$

ii) Sean  $E, F \subset \mathbb{R}^2$  medibles Jordan de área no nula tales que  $E \cap F$  tiene contenido nulo. Si  $(x_0, y_0)$ ,  $(x_E, y_E)$ ,  $(x_F, y_F)$  son los baricentros de  $E \cup F$ ,  $E$  y  $F$ , respectivamente, como  $E \cap F$  tiene contenido nulo resulta:

$$c_2(E \cup F)x_0 = \int_{E \cup F} x dx dy = \int_E x dx dy + \int_F x dx dy = c_2(E)x_E + c_2(F)x_F$$

Los números  $\alpha = c_2(E)/c_2(E \cup F)$  y  $\beta = c_2(F)/c_2(E \cup F)$  verifican  $\alpha + \beta = 1$  y con ellos se obtiene la relación:  $x_0 = \alpha x_E + \beta x_F$ . Análogamente  $y_0 = \alpha y_E + \beta y_F$ .

iii) Utilizando la relación anterior se puede calcular, sin usar integrales, el volumen del sólido generado por  $A \setminus B$  al girar alrededor del eje  $Oy$ :

Apliquemos la relación con  $E = A \cap B$ ,  $F = A \setminus B$ , de modo que  $E \cup F = A$ . Por razones de simetría  $(x_E, y_E) = (0, 0)$ ,  $(x_A, y_A) = (1, 0)$ , luego

$$1 = \alpha x_E + \beta x_F = \beta x_F = x_F c_2(F)/c_2(A)$$

Según el teorema de Guldin el volumen del sólido que genera  $F$  al girar alrededor del eje  $Oy$  viene dado por  $2\pi x_F c_2(F) = 2\pi c_2(A) = 2\pi \cdot 2\pi = 4\pi^2$ . ■

**Ejercicio 11.21** Sea  $M \subset X \times Y$ , donde  $X \subset \mathbb{R}^k$ ,  $Y \subset \mathbb{R}^m$  son rectángulos cerrados. Si  $M$  tiene contenido nulo en  $\mathbb{R}^{k+m}$ , demuestre que  $M_{\mathbf{x}} = \{\mathbf{y} \in Y : (\mathbf{x}, \mathbf{y}) \in M\}$  tiene contenido nulo (en  $\mathbb{R}^m$ ) para casi todo  $\mathbf{x} \in X$  (es decir, excepto en un conjunto de medida nula en  $\mathbb{R}^k$ ).

SOLUCIÓN

Según el teorema de Fubini la función  $u(\mathbf{x}) = \int_Y \chi_M(\mathbf{x}, \mathbf{y}) dy \geq 0$  es integrable sobre  $X$ , con

$$\int_X u(\mathbf{x}) d\mathbf{x} = \int_A \chi_M(\mathbf{x}, \mathbf{y}) d\mathbf{x} dy = c_n(M) = 0$$

y aplicando el ejercicio 10.32 se obtiene que  $N = \{\mathbf{x} \in X : u(\mathbf{x}) > 0\}$  tiene medida nula en  $\mathbb{R}^k$ . Claramente,  $M_{\mathbf{x}}$  tiene contenido nulo si y sólo si  $\mathbf{x} \notin N$ . ■

## 11.4. Ejercicios propuestos

◇ 11.4.1 Sean  $f_j : [a_j, b_j] \rightarrow \mathbb{R}$ ,  $1 \leq j \leq n$ , integrables Riemann. Demuestre que

$$f(x_1, x_2, \dots, x_n) = f_1(x_1)f_2(x_2) \cdots f_n(x_n)$$

es integrable Riemann sobre  $A = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$ , y

$$\int_A f(\mathbf{x})d\mathbf{x} = \left( \int_{a_1}^{b_1} f_1(t)dt \right) \left( \int_{a_2}^{b_2} f_2(t)dt \right) \cdots \left( \int_{a_n}^{b_n} f_n(t)dt \right)$$

◇ 11.4.2 Sea  $f : A \rightarrow \mathbb{R}$  definida en  $A = [0, 1] \times [0, 1]$  por

$$f(x, 1/2) = \chi_{\mathbb{Q}}(x), \quad f(x, y) = 1 \quad \text{si } y \neq 1/2$$

Demuestre que  $f$  es integrable sobre  $A$  y exprese la integral doble  $\int_A f(x, y)dxdy$  como integral iterada según los dos posibles ordenes de integración.

◇ 11.4.3 Sea  $f : A \rightarrow \mathbb{R}$  definida en  $A = [0, 1] \times [0, 1]$  por  $f(x, y) = 1$  si  $x \in \mathbb{Q}$ ,  $f(x, y) = 2y$  si  $x \notin \mathbb{Q}$ . Calcule  $\overline{J}(y) = \int_0^1 f(x, y)dx$ ,  $\underline{J}(y) = \int_0^1 f(x, y)dx$  y deduzca que  $f$  no es integrable sobre  $A$ . Compruebe que aunque  $f$  no es integrable sin embargo existe la integral iterada  $\int_0^1 \left( \int_0^1 f(x, y)dy \right) dx$ . Calcule  $\int_A f(x, y)dxdy$ ,  $\overline{\int}_A f(x, y)dxdy$ .

◇ 11.4.4 Sean  $X \subset \mathbb{R}^k$ ,  $Y \subset \mathbb{R}^m$  rectángulos cerrados y  $M \subset X \times Y$  medible Jordan en  $\mathbb{R}^{k+m}$ . Dado  $\mathbf{x} \in \mathbb{R}^k$  sea  $M_{\mathbf{x}} = \{\mathbf{y} \in \mathbb{R}^m : (\mathbf{x}, \mathbf{y}) \in M\}$ . Si  $S$  es el conjunto de los  $\mathbf{x} \in \mathbb{R}^k$  tales que  $M_{\mathbf{x}}$  no es medible Jordan en  $\mathbb{R}^m$ , demuestre que  $S$  tiene medida nula  $k$ -dimensional.

◇ 11.4.5 Se supone que  $f : M \rightarrow \mathbb{R}$  es integrable sobre un conjunto medible Jordan  $M \subset \mathbb{R}^2$  y que  $\varphi : [a, b] \rightarrow \mathbb{R}$  es una función creciente continua tal que el conjunto  $M_t = \{(x, y) \in M : f(x, y) \leq \varphi(t)\}$  es medible Jordan para cada  $t \in [a, b]$ ,  $M_a = \emptyset$  y  $M_b = M$ . Si la función  $A(t) = \text{área}(M_t)$  es derivable, demuestre que  $F(t) = \int_{M_t} f(x, y)dxdy$  también lo es, con derivada  $F'(t) = \varphi(t)A'(t)$ .

◇ 11.4.6 Sea  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  una función continua. Se supone que  $E_t = \{(x, y) : f(x, y) \geq t\}$  es medible Jordan para cada  $t \geq 0$  y que la función  $\varphi(t) = \text{área}(E_t)$  es continua. Demuestre:

a)  $f$  alcanza un máximo absoluto en  $E_0$ .

b)  $M_t = \{(x, y, z) : t \leq z \leq f(x, y)\}$  es medible Jordan para cada  $t \geq 0$

y la función  $v(t) = \text{Volumen}(M_t)$  es derivable con derivada continua en  $[0, +\infty)$ .

◇ 11.4.7 Se supone que  $f : [0, +\infty) \rightarrow \mathbb{R}$  integrable Riemann sobre  $[0, a]$ , para cada  $a > 0$ . Utilice el teorema de Fubini para demostrar que la función

$$F(x) = \int_0^x (x-t)f(t)dt$$

definida para todo  $x > 0$ , es derivable con derivada continua  $F'(x) = \int_0^x f(t)dt$ .

◇ **11.4.8** Sean  $f, g : \mathbb{R} \rightarrow \mathbb{R}$ , donde  $f$  es continua y  $g$  derivable. Demuestre que

$$F(t) = \int_0^t f(x) \left( \int_x^t g(y) dy \right) dx$$

es derivable dos veces y obtenga la derivada segunda  $F''$ .

◇ **11.4.9** Si  $f : \mathbb{R} \rightarrow \mathbb{R}$  es continua, establezca la igualdad

$$2 \int_a^b \left( \int_x^b f(x)f(y) dy \right) dx = \left( \int_a^b f(x) dx \right)^2$$

◇ **11.4.10** En cada caso obtenga un conjunto medible Jordan  $E \subset \mathbb{R}^2$  tal que la integral  $\int_E f(x, y) dx dy$  se exprese mediante la integral iterada que se indica. Obtenga también la integral iterada en el otro orden de integración.

$$\begin{aligned} \text{a)} \int_1^2 \left( \int_{-\sqrt{2x-x^2}}^{\sqrt{2x-x^2}} f(x, y) dy \right) dx; & \quad \text{b)} \int_0^2 \left( \int_{y/2}^{(4-y)/2} f(x, y) dx \right) dy. \\ \text{c)} \int_0^1 \left( \int_{\sqrt{y}}^{1+\sqrt{1-y^2}} f(x, y) dx \right) dy; & \quad \text{d)} \int_{-6}^2 \left( \int_{y^2/4-1}^{2-y} f(x, y) dx \right) dy. \end{aligned}$$

◇ **11.4.11** En cada caso, justifique que  $f : A \rightarrow \mathbb{R}$  es integrable Riemann sobre el rectángulo  $A$ , y calcule la integral  $\int_A f(x, y) dx dy$ .

$$\text{a)} f(x, y) = x^{[y]} y^{[x]} \text{ si } x \neq 0, y \neq 0, \quad f(x, 0) = f(0, y) = 0, \quad A = [0, 2] \times [0, 2].$$

$$\text{b)} f(x, y) = (x-1)^{[3y]-1}, \quad A = [2, 3] \times [0, 1].$$

$$\text{c)} f(x, y) = x \text{ si } x > y, \quad f(x, y) = y^2 \text{ si } x \leq y, \quad A = [0, 1] \times [0, 1].$$

$$\text{d)} f(x, y) = |y - \text{sen } x|, \quad A = [0, \pi] \times [0, 1].$$

(Nota:  $[a]$  denota la parte entera de  $a \in \mathbb{R}$ )

◇ **11.4.12** En cada caso justifique la existencia de la integral  $\int_E f(x, y) dx dy$  y calcule su valor:

$$\text{a)} f(x, y) = [x+y], \quad E = \{(x, y) : 0 \leq x \leq 2, 0 \leq y \leq 2-x\}.$$

$$\text{b)} f(x, y) = (x+y)^{-4}, \quad E = \{(x, y) \in \mathbb{R}^2 : 1 \leq x, 1 \leq y, x+y \leq 4\}.$$

(Nota: La integral b) se puede calcular interpretándola como un volumen que se calcula fácilmente considerando las secciones que producen los planos  $x+y=c$ ).

◇ **11.4.13** En cada caso calcule la integral  $\int_E f(x, y) dx dy$  mediante un cambio de variable a coordenadas polares.

$$\text{a)} f(x, y) = (x^2 + y^2)^{-3/2}, \quad E = \{(x, y) : x \leq y, 1 \leq x+y, x^2 + y^2 \leq 1\}.$$

$$\text{b)} f(x, y) = x^2 + y^2, \quad E = \{(x, y) : 0 \leq x, x^2 + y^2 \leq 2y, x^2 + y^2 \leq 1\}.$$

$$\text{c)} f(x, y) = x^2 + y^2, \quad E = \{(x, y) : 0 \leq x, (x^2 + y^2)^2 \leq 4(x^2 - y^2)\}.$$

$$\text{d)} f(x, y) = \sqrt{x^2 + y^2}, \quad E = \{(x, y) : 2x \leq x^2 + y^2 \leq 1, x \geq 0, y \geq 0\}$$

◇ **11.4.14** En cada caso calcule  $\int_E f(x, y) dx dy$  con el cambio de variable indicado

a)  $f(x, y) = e^{(y-x)/(y+x)}$ ;  $E = \{(x, y) : 0 \leq x, 0 \leq y, x + y \leq 2\}$ .  
 $x = (v - u)/2$ ;  $y = (v + u)/2$ .

b)  $f(x, y) = \frac{x^2 + y}{\sqrt{1 + 4(x - y)}}$ ;  $E = \{(x, y) : 0 \leq x \leq 2, -x^2 \leq y \leq x\}$ ,  
 $x = u + v$ ,  $y = v - u^2$ .

c)  $f(x, y) = (y^2 - x^2)xy(x^2 + y^2)$ ;  
 $E = \{(x, y) : 0 < x, 0 < y, a \leq xy \leq b, y^2 - x^2 \leq 1, x \leq y\}$ , ( $0 < a \leq b$ );  
 $u = y^2 - x^2$ ,  $v = xy$ .

◇ **11.4.15** Considerando el cambio de variable  $u = 2x/(x^2 + y^2)$ ,  $v = 2y/(x^2 + y^2)$ , en el abierto  $\Omega = \{(x, y) : x > 0, y > 0\}$ , calcule la integral doble  $\int_M \frac{dx dy}{(x^2 + y^2)^2}$  sobre el cuadrilátero curvilíneo  $M = \{(x, y) : 4x \leq x^2 + y^2 \leq 6x, 2y \leq x^2 + y^2 \leq 8y\}$ .

◇ **11.4.16** Si  $0 < a < 1$  demuestre que  $\mathbf{g}(x, y) = (x - a \cos y, y - a \cos x)$  es una transformación inyectiva de  $\mathbb{R}^2$  que transforma abiertos en abiertos y conjuntos medibles Jordan en conjuntos medibles Jordan. Calcule el área de  $\mathbf{g}(T)$  donde

$$T = \{(x, y) : 0 \leq x \leq 1, 0 \leq y \leq x\}$$

◇ **11.4.17** Si  $f : \mathbb{R} \rightarrow \mathbb{R}$  es una función continua par, y  $A = [0, a] \times [0, a]$ , establezca la igualdad

$$\int_A f(x - y) dx dy = 2 \int_0^a (a - t) f(t) dt$$

◇ **11.4.18** Si  $D(r) = \{(x, y) : x^2 + y^2 \leq r^2\}$  y  $f : D(R) \rightarrow \mathbb{R}$  es continua demuestre que  $\varphi(r) = \int_{D(r)} f(x, y) dx dy$  es de clase  $C^1$  en  $[0, R]$  y obtenga una fórmula integral para la derivada.

◇ **11.4.19** Justifique que los siguientes conjuntos son medibles Jordan y calcule sus volúmenes:

$$A = \{(x, y, z) : 0 \leq x, 0 \leq y, 0 \leq z, \sqrt{x} + \sqrt{y} + \sqrt{z} \leq 1\}.$$

$$B = \{(x, y, z) : 0 \leq x, 0 \leq y, 0 \leq z, x + y + z \leq a, az \leq xy\}; (a > 0).$$

$$C = \{(x, y, z) : 2z^2 \leq x^2 + y^2 \leq 1 + z^2\}.$$

$$D = \{(x, y, z) : x^2 + y^2 + z^2 \leq a^2, x^2 + y^2 \leq z^2\}.$$

$$E = \{(x, y, z) : x^2 + y^2 + z^2 \leq a^2, x^2 + y^2 \leq ax; 0 \leq z\}.$$

$$F = \{(x, y, z) : x^2/a^2 + y^2/b^2 \leq 1 + z^2/c^2, 0 \leq z \leq 1\}.$$

◇ **11.4.20** Justifique que el conjunto  $M = \{(x, y, z) : x^2 + y^2 \leq 2az, z \leq x + y\}$  es medible Jordan y calcule su volumen (se supone que  $a > 0$ ).

◇ **11.4.21** Calcule el volumen de

$$E = \{(x, y, z) : x^2 + y^2 \leq 2y; x^2 + y^2 \leq 1; 0 \leq x; 0 \leq z \leq x^2 + y^2\}$$

◇ **11.4.22** Calcule el volumen de

$$A = \{(x, y, z) : \frac{x^2}{a^2} + \frac{y^2}{b^2} < 1, \quad \frac{y^2}{b^2} + \frac{z^2}{c^2} < 1\}$$

◇ **11.4.23** Una copa que tiene forma de paraboloides de revolución está llena de vino hasta el borde. Bebemos hasta que el plano de la superficie del vino pasa por el fondo de la copa. Determine la cantidad de vino que queda en la copa.

◇ **11.4.24** En un recipiente que tiene la forma del paraboloides de revolución  $2z = x^2 + y^2$  se encaja una esfera maciza de radio  $\sqrt{5}$  y centro en el eje  $Oz$ . Determine el volumen del hueco que queda entre el recipiente y la esfera.

◇ **11.4.25** En el cono  $z = a\sqrt{x^2 + y^2}$  se encaja una esfera de radio  $R$ . Calcule el volumen del recinto acotado limitado por el cono y la esfera.

◇ **11.4.26** Calcule el volumen del recinto acotado limitado por el elipsoide  $3x^2 + 3y^2 + z^2 = 3$  y el cono  $z = 2 - \sqrt{x^2 + y^2}$ .

◇ **11.4.27** Considere los recintos  $C$  y  $H$  de  $\mathbb{R}^3$  definidos por:

$$C = \{(x, y, z) : \left(\frac{Rz}{h}\right)^2 \geq x^2 + y^2, 0 \leq z \leq h\}$$

$$H = \{(x, y, z) \in C : x^2 + y^2 - Ry \geq 0\}$$

a) Justifique que  $H$  es medible Jordan.

b) Calcule, por el método de las secciones  $z = \text{constante}$ , el volumen de  $C$ . Calcule el volumen de  $H$ , y compruebe que el cociente de los volúmenes de  $H$  y  $C$  no depende de los parámetros  $h, R$ .

◇ **11.4.28** Sea  $\mathbf{g}(x, y, z) = (u, v, w)$ , donde  $u = e^{2y} - e^{2z}$ ,  $v = e^{2x} - e^{2z}$ ,  $w = x - y$ . Compruebe que  $\mathbf{g}(\mathbb{R}^3) = \{(u, v, w) : u > 0, u + v > 0, e^{2w} > v/u\}$  y que  $\mathbf{g}$  establece un  $C^\infty$ -difeomorfismo entre  $\mathbb{R}^3$  y su imagen  $\mathbf{g}(\mathbb{R}^3)$ . Calcule el volumen de  $\mathbf{g}([0, 1]^3)$ .

◇ **11.4.29** El potencial gravitacional producido por un sólido  $W \subset \mathbb{R}^3$  (medible Jordan) con densidad de masa  $p(x, y, z)$  es la suma de los potenciales producidos por los "elementos de masa infinitesimal"  $dm = p(x, y, z)dx dy dz$ , y viene dado por la integral triple

$$V(x_1, y_1, z_1) = Gm \int_W \frac{p(x, y, z)dx dy dz}{\sqrt{(x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2}}$$

Sea  $W = \{(x, y, z) : r^2 \leq x^2 + y^2 + z^2 \leq R^2\}$  con densidad es uniforme. Justifique las siguientes afirmaciones

a) El potencial  $V$  es constante en el hueco  $\Omega_1 = \{(x, y, z) : x^2 + y^2 + z^2 < r^2\}$ , (es decir, no existe fuerza gravitacional dentro de un planeta hueco).

- b) En el exterior de la esfera mayor  $\Omega_2 = \{(x, y, z) : x^2 + y^2 + z^2 > R^2\}$  el potencial es el mismo que produciría toda la masa de  $W$  concentrada en el centro de la esfera.

◇ **11.4.30** Sea  $S \subset \mathbb{R}^2$  el recinto plano limitado por el eje de abscisas y el arco de cicloide

$$x = r(t - \operatorname{sen} t), \quad y = r(1 - \operatorname{cos} t), \quad t \in [0, 2\pi].$$

(El arco descrito por un punto  $\mathbf{p}$  de una circunferencia de radio  $r$  que rueda sin deslizar sobre el eje de abscisas, entre dos pasos consecutivos de  $\mathbf{p}$  por los puntos  $(0, 0)$  y  $(2\pi r, 0)$  de dicho eje.) Calcule el área de  $S$  y el volumen de los cuerpos que engendra  $S$  al girar alrededor de las siguientes rectas: a) El eje de abscisas  $OX$ ; b) El eje de ordenadas  $OY$ ; c) La recta  $x = \pi r$ .

◇ **11.4.31** Calcule el volumen de los siguientes sólidos de revolución:

a) Cuerpo engendrado al girar alrededor del eje  $OY$  el cuadrilátero curvilíneo limitado por las parábolas  $y^2 = a^3x$ ,  $y^2 = b^3x$ ,  $x^2 = c^3y$ ,  $x^2 = d^3y$ , donde  $0 < a < b$ ,  $0 < c < d$ .

b) Cuerpo engendrado al girar alrededor de la recta  $x + y = 0$  el recinto plano limitado por la parábola  $y = x - x^2$  y el eje de abscisas.

c) Cuerpo engendrado cuando el recinto  $\{(x, y) : 0 \leq y \leq 2a, 0 \leq x \leq e^{|y-a|}\}$ , ( $a > 0$ ) gira alrededor del eje  $OX$ .

d) Cuerpo engendrado cuando el recinto  $\{(x, y) : x^2 + y^2 \leq 1, x^2 + y^2 - 2x \leq 0\}$  al girar alrededor del eje  $OX$ .

e) Cuerpos engendrados por  $\{(x, y) : x^2 + y^2 \leq 1, x^2 + y^2 - 2x \geq 0, x \geq 0, y \geq 0\}$  al girar alrededor del eje  $OX$  y del eje  $OY$ .

◇ **11.4.32** Justifique que es medible Jordan el recinto

$$B = \{(y, z) \in \mathbb{R}^2 : y = \rho \operatorname{cos} \varphi, z = \rho \operatorname{sen} \varphi : |\varphi| \leq \pi/2, 0 \leq \rho \leq 1 + \operatorname{sen} \varphi\}$$

Sea  $M \subset \mathbb{R}^3$  el cuerpo de revolución que genera  $\{(0, y, z) : (y, z) \in B\}$  al girar alrededor del eje  $Oz$ . Calcule el volumen y el baricentro de  $M$ .

◇ **11.4.33** Calcule las siguientes integrales iteradas

a)  $\int_0^{2a} \left( \int_0^{\sqrt{a^2 - (x-a)^2}} \sqrt{4a^2 - x^2 - y^2} dy \right) dx.$

b)  $\int_{-a}^a dx \int_{-\sqrt{a^2 - x^2}}^{\sqrt{a^2 - x^2}} dy \int_{\sqrt{x^2 + y^2}}^{\sqrt{1 - x^2 - y^2}} z dz, \quad 0 < a < 1.$

c)  $\int_0^1 dx \int_{1 - \sqrt{1 - x^2}}^{\sqrt{2x - x^2}} dy \int_0^{\max\{x, y\}} z^2 dz.$

◇ **11.4.34** En cada caso calcule  $\int_E f(x, y, z) dx dy dz$  con un cambio de variable apropiado:

- a)  $f(x, y, z) = x^2$ ;  $E = \{0 \leq x, x^2 + y^2 + (z-1)^2 \leq 1; 4z^2 \geq 3(x^2 + y^2)\}$ .  
 b)  $f(x, y, z) = yz\sqrt{x^2 + y^2}$ ;  $E = \{0 \leq z \leq x^2 + y^2, 0 \leq y \leq \sqrt{2x - x^2}\}$ .  
 c)  $f(x, y, z) = z$ ;  $E = \{(x^2 + y^2 + z^2 \leq 2, x^2 + y^2 \leq z)\}$ .  
 d)  $f(x, y, z) = z^2$ ;  $E = \{(x, y, z) : x^2 + y^2 + z^2 \leq R^2, x^2 + y^2 + z^2 \leq 2Rz\}$ .  
 e)  $f(x, y, z) = \sqrt{x^2 + y^2 + z^2}$ ;  $E = \{(x, y, z) : \sqrt{x^2 + y^2} \leq z \leq 3\}$ .  
 f)  $f(x, y, z) = z$ ;  $E = \{(x, y, z) : 2z^2 \leq x^2 + y^2 \leq z^2 + 1, 0 \leq z\}$ .  
 g)  $f(x, y, z) = x$ ;  $E = \{(x, y, z) : x^2 + y^2 + (z-1)^2 \leq 1, 3(x^2 + y^2) \leq z^2\}$ .  
 h)  $f(x, y, z) = 1/\sqrt{x^2 + y^2 + z^2}$ ;  $E = \{(x, y, z) : \frac{1}{2} \leq z \leq 1, x^2 + y^2 + z^2 \leq 1\}$ .  
 i)  $f(x, y, z) = ze^{-(x^2+y^2)}$ ;  $E = \{(x, y, z) : z \geq 0, z^2 - 1 \leq (x^2 + y^2) \leq z^2/2\}$ .  
 j)  $f(x, y, z) = z^2$ ;  $E = \{(x, y, z) : x^2 + y^2 \leq 2y, y \leq x, 0 \leq z \leq x\}$ .  
 k)  $f(x, y, z) = z$ ;  $E = \{(x, y, z) : x^2 + (y-1)^2 + z^2 \leq 1, \sqrt{x^2 + y^2} \leq z\}$ .  
 l)  $f(x, y, z) = 1/\sqrt{x^2 + y^2}$ ;  $E = \{(x, y, z) : x^2 + y^2 \geq 1, x^2 + y^2 + z^2 \leq 4\}$ .

◇ **11.4.35** Calcule  $\int_M \frac{dx dy dz}{z^2}$ , donde

$$M = \{(x, y, z) : 1 \leq 4z, \sqrt{x^2 + y^2} \leq z \leq \sqrt{1 - x^2 - y^2}\}$$

◇ **11.4.36** Calcule la integral  $\int_M z dx dy dz$ , donde

$$M = \{(x, y, z) : x^2 + y^2 + z^2 \leq a^2, x^2 + y^2 \geq 1 + z^2, |y| \leq x\} \quad (a > 1)$$

◇ **11.4.37** Justifique que los siguientes conjuntos son medibles Jordan y calcule sus baricentros,

- a)  $M = \{(x, y, z) : x^2 + y^2 + z^2 \leq 2z, (x^2 + y^2) \operatorname{tg}^2 \alpha \leq z^2\}$ , donde  $0 < \alpha < \pi/2$ .  
 b)  $M = \{(x, y, z) : x^2 + y^2 + z^2 \leq 2z, 3(x^2 + y^2) \leq z^2\}$ .  
 c)  $M = \{(x, y, z) : z \geq 0, x^2 + y^2 \leq 1/4, x^2 + y^2 + z^2 \leq 2z\}$ .

◇ **11.4.38** Sea  $V_n$  el contenido  $n$ -dimensional de  $B_n = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_2 \leq 1\}$ . Demuestre que  $nV_n = 2\pi V_{n-2}$ .

◇ **11.4.39** Sea  $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$  continua y  $B(t) = \{(x, y) : x^2 + y^2 \leq t^2\}$ . Demuestre que la función  $g(t) = \int_{B(t)} \varphi(x, y) dx dy$ , definida para  $t > 0$ , es de clase  $C^1$  y que la condición  $\varphi(x, y) = o(x^2 + y^2)$  implica  $g(t) = o(t^4)$ .

◇ **11.4.40** Sea  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  de clase  $C^2$ . Para cada  $\epsilon > 0$  se define:

$$M(\epsilon) = \frac{1}{\pi\epsilon^2} \int_{B(\epsilon)} f(x, y) dx dy \quad \text{donde} \quad B(\epsilon) = \{(x, y) : x^2 + y^2 \leq \epsilon^2\}$$

Utilice el problema 11.4.39 y el desarrollo de Taylor de  $f$  en  $(0, 0)$  para obtener que

$$\lim_{\epsilon \rightarrow 0} \frac{M(\epsilon) - f(0, 0)}{\epsilon^2} = \frac{1}{8} [D_{11}f(0, 0) + D_{22}f(0, 0)]$$

# Capítulo 12

## Integrales impropias. Integrales dependientes de un parámetro

*Integrales impropias. Paso al límite bajo la integral. Continuidad y derivabilidad de las integrales dependientes de un parámetro*

Un inconveniente de la integral de Riemann es que sólo se aplica a funciones acotadas sobre conjuntos acotados. Este inconveniente se resuelve parcialmente con la noción de integral impropia de Riemann absolutamente convergente con la que comienza este capítulo. Luego se estudia la validez del paso al límite bajo la integral de Riemann (y bajo una integral impropia absolutamente convergente), y se obtienen resultados sobre continuidad y derivabilidad de funciones definidas por integrales que dependen de un parámetro.

Aunque la integral de Riemann es completamente satisfactoria como herramienta para el cálculo de áreas y volúmenes de figuras geométricas sin embargo, para estos problemas de paso al límite, la integral de Riemann comienza a mostrar sus deficiencias: Para garantizar la integrabilidad de la función límite y el paso al límite bajo la integral hay que recurrir a la convergencia uniforme, una hipótesis que suele ser demasiado restrictiva en las aplicaciones.

### 12.1. Integrales impropias

El objetivo de esta sección es el de extender la definición de la integral  $\int_{\Omega} f$  para el caso de funciones, que no se suponen acotadas, definidas en conjuntos  $\Omega \subset \mathbb{R}^n$  que tampoco se suponen acotados. En lo que sigue, dado un conjunto  $\Omega \subset \mathbb{R}^n$ , denotaremos por  $\mathcal{K}_{\Omega}$  la familia de los compactos  $K \subset \Omega$  que son medibles Jordan. En este capítulo sólo consideraremos dominios  $\Omega \subset \mathbb{R}^n$  que se pueden expresar en la forma  $\Omega = \cup_{j=1}^{\infty} K_j$  donde  $K_j \in \mathcal{K}_{\Omega}$  es una sucesión creciente tal que todo  $K \in \mathcal{K}_{\Omega}$  está contenido en algún  $K_m$ . Diremos entonces que  $\Omega$  es un *recinto de integración* y que  $(K_j)$  es una *sucesión expansiva* en  $\Omega$ . Es obvio que todo compacto medible Jordan es un recinto de integración. La demostración de que todo abierto  $\Omega \subset \mathbb{R}^n$  es un recinto de integración es una consecuencia inmediata del siguiente



lema: Obsérvese que si  $\{Q_j : j \in M\}$  es la sucesión que proporciona el lema, entonces  $K_j = \cup\{\overline{Q_k} : k \in M, k < j\}$ , es una sucesión expansiva de compactos medibles Jordan cuya unión es  $\Omega$ .

**Lema 12.1** *Si  $\Omega \subset \mathbb{R}^n$  es abierto existe una familia finita o numerable de cubos semiabiertos disjuntos dos a dos  $\{Q_j : j \in M\}$ ,  $M \subset \mathbb{N}$ , con  $\overline{Q_j} \subset \Omega$ , tal que  $\Omega = \cup_{j \in M} Q_j$ , y cada compacto  $K \subset \Omega$  se puede cubrir con una cantidad finita de cubos de la familia.*

DEM: Para cada  $k \in \mathbb{N}$ , sea  $D_k = \{j2^{-k} : j \in \mathbb{Z}\}$ . Llamaremos cubos diádicos de lado  $2^{-k}$  a los cubos semiabiertos de la forma

$$Q = \prod_{j=1}^n [d_j, d_j + 2^{-k}), \text{ con } d_j \in D_k, \text{ para } 1 \leq j \leq n,$$

Sea  $\mathcal{Q}_k$  la familia, finita o numerable formada por los cubos diádicos  $Q$ , de lado  $2^{-k}$ , que cumplen  $\overline{Q} \subset \Omega$ , a los que llamaremos cubos de rango  $k$ . Sea  $\mathcal{E}_1 = \mathcal{Q}_1$  la familia finita o numerable formada por los cubos de rango 1. De manera recurrente se define  $\mathcal{E}_{k+1} \subset \mathcal{Q}_{k+1}$  como la familia finita o numerable formada por los cubos de rango  $k+1$ , que no están contenidos en cubos de rango inferior. La unión de las familias  $\mathcal{E} = \mathcal{E}_1 \cup \mathcal{E}_2 \cup \dots \cup \mathcal{E}_k \cup \dots$  proporciona una familia numerable de cubos diádicos semiabiertos y disjuntos  $\{Q_j : j \in M\}$ , con la propiedad requerida.

En efecto, el compacto  $K \subset \Omega$ , es disjunto del cerrado  $F = \mathbb{R}^n \setminus \Omega$ , y por lo tanto existe  $\delta > 0$  tal que  $d(\mathbf{x}, \mathbf{y}) \geq \delta$  para cada  $\mathbf{x} \in K$ , y cada  $\mathbf{y} \in F$ .

En  $\mathbb{R}^n$  consideramos la distancia asociada a la norma  $\|\cdot\|_\infty$ , para la que se cumple que el diámetro de cada  $Q \in \mathcal{E}_k$  es  $2^{-k}$ . Elegimos  $p \in \mathbb{N}$  tal que  $2^{-p} < \delta$ . Entonces, dado  $\mathbf{x} \in K$ , hay un único cubo diádico  $Q$  de lado  $2^{-p}$  tal que  $\mathbf{x} \in Q$ , y la condición  $\text{diam}(Q) = 2^{-p} < \delta$  implica que  $\overline{Q} \subset \Omega$ , luego  $Q \in \mathcal{Q}_p$ . Se sigue que  $Q$  está contenido en algún cubo de una familia  $\mathcal{E}_j$  con  $j \leq p$ . Esto demuestra que  $K$  está cubierto por los cubos de las familias  $\mathcal{E}_1, \mathcal{E}_2, \dots, \mathcal{E}_p$ , luego los cubos de la familia  $\mathcal{E}_{p+1}$  no intersecan a  $K$ . Teniendo en cuenta que para cada  $k \leq p \in \mathbb{N}$  la familia  $\{Q \in \mathcal{E}_k : Q \cap K \neq \emptyset\}$  es finita (porque  $K$  es acotado), se obtiene que  $\{j \in M : Q_j \cap K \neq \emptyset\}$  es finito. ■

**Definición 12.2** *Una función  $f : \Omega \rightarrow \mathbb{R}$ , definida en un recinto de integración  $\Omega \subset \mathbb{R}^n$  se dice que es localmente integrable Riemann (brevemente, localmente integrable) cuando para cada  $K \in \mathcal{K}_\Omega$  la restricción  $f|_K$  es integrable Riemann sobre  $K$ . Una función localmente integrable Riemann  $f : \Omega \rightarrow \mathbb{R}$  se dice que es absolutamente integrable Riemann sobre  $\Omega$ , cuando cumple*

$$[IA] : \quad \sup_{K \in \mathcal{K}_\Omega} \int_K |f| < +\infty$$

En virtud de 10.14, toda función continua  $f : \Omega \rightarrow \mathbb{R}$  es localmente integrable. Es claro que la condición de integrabilidad absoluta  $[IA]$ , equivale a  $\sup_{j \in \mathbb{N}} \int_{K_j} |f| < +\infty$ , donde  $K_j \in \mathcal{K}_\Omega$ , es cualquier sucesión expansiva en  $\Omega$ .

**Proposición 12.3** Si  $f : \Omega \rightarrow \mathbb{R}$  es absolutamente integrable sobre el recinto de integración  $\Omega \subset \mathbb{R}^n$ , hay un único  $\alpha \in \mathbb{R}$  que verifica: Para cada  $\epsilon > 0$  existe  $K(\epsilon) \in \mathcal{K}_\Omega$  tal que  $K \in \mathcal{K}_\Omega$ ,  $K \supset K(\epsilon) \Rightarrow \left| \int_K f - \alpha \right| < \epsilon$ . Además, para cada sucesión expansiva  $K_j \in \mathcal{K}_\Omega$  se cumple  $\alpha = \lim_j \int_{K_j} f$ .

DEM: En primer lugar, es claro que sólo puede haber un número real  $\alpha$  verificando la condición del enunciado (basta razonar de forma análoga al caso de la unicidad del límite de una sucesión de números reales). Para demostrar la existencia de este número consideramos una sucesión  $K_j \in \mathcal{K}_\Omega$  expansiva en  $\Omega$ , y comenzamos viendo que la sucesión de las integrales  $\alpha_j = \int_{K_j} f$  es de Cauchy.

Según la hipótesis, es finito el supremo  $A = \sup_{K \in \mathcal{K}_\Omega} \int_K |f|$ , luego para cada  $\epsilon > 0$  existe  $K(\epsilon) \in \mathcal{K}_\Omega$  tal que

$$\int_{K(\epsilon)} |f| \geq A - \epsilon/2$$

Como la sucesión  $(K_j)$  es expansiva,  $K(\epsilon)$  está contenido en algún  $K_m$ . Entonces, si  $p \geq q \geq m$  se cumple  $K_p \supset K_q \supset K_m \supset K(\epsilon)$ , luego

$$\begin{aligned} |\alpha_p - \alpha_q| &= \left| \int_{K_p \setminus K_q} f \right| \leq \int_{K_p \setminus K_q} |f| = \int_{K_p} |f| - \int_{K_q} |f| \leq \\ &\leq A - \int_{K(\epsilon)} |f| \leq A - (A - \epsilon/2) = \epsilon/2 \end{aligned}$$

Para terminar la demostración basta ver que el límite  $\alpha = \lim_j \alpha_j$ , de esta sucesión de Cauchy cumple la condición del enunciado.

Como la desigualdad  $|\alpha_p - \alpha_q| \leq \epsilon/2$ , es válida para  $p > q \geq m$ , aplicándola con  $q = m$ , y pasando al límite cuando  $p \rightarrow +\infty$ , se obtiene que

$$|\alpha - \alpha_m| \leq \epsilon/2$$

Es claro que cada  $K \in \mathcal{K}_\Omega$  con  $K \supset K_m$  verifica

$$\begin{aligned} \left| \int_K f - \alpha_m \right| &= \left| \int_K f - \int_{K_m} f \right| = \left| \int_{K \setminus K_m} f \right| \leq \int_{K \setminus K_m} |f| = \\ &= \int_K |f| - \int_{K_m} |f| \leq \int_K |f| - \int_{K(\epsilon)} |f| \leq A - (A - \epsilon/2) = \epsilon/2 \end{aligned}$$

luego

$$\left| \alpha - \int_K f \right| \leq |\alpha - \alpha_m| + \left| \alpha_m - \int_K f \right| \leq \epsilon/2 + \epsilon/2 = \epsilon$$

■

En lo que sigue, el número  $\alpha$  que interviene en la proposición 12.3 lo denotaremos

$$\alpha = \lim_{K \in \mathcal{K}_\Omega} \int_K f(\mathbf{x}) \, d\mathbf{x}$$

(Este es un caso particular de la noción de límite de una red que se suele estudiar en los cursos de Topología General).

**Definición 12.4** Cuando  $f : \Omega \rightarrow \mathbb{R}$  es absolutamente integrable Riemann sobre el recinto de integración  $\Omega \subset \mathbb{R}^n$ , se define

$$\int_{\Omega} f(\mathbf{x}) \, d\mathbf{x} = \alpha$$

donde  $\alpha = \lim_{K \in \mathcal{K}_{\Omega}} \int_K f(\mathbf{x}) \, d\mathbf{x}$  es el número que interviene en la proposición 12.3. También se suele decir que  $\int_{\Omega} f$  es una integral de Riemann impropia absolutamente convergente, de valor  $\alpha$ .

En las condiciones de la definición 12.3, para una función  $f \geq 0$ , es fácil ver que

$$\int_{\Omega} f = \sup_{K \in \mathcal{K}_{\Omega}} \int_K f = \sup_{j \in \mathbb{N}} \int_{K_j} f$$

donde  $K_j \in \mathcal{K}_{\Omega}$  es cualquier sucesión expansiva en  $\Omega$ .

Cuando sólo se supone que  $f \geq 0$  es localmente integrable en  $\Omega$ , también se define

$$\int_{\Omega} f = \sup_{K \in \mathcal{K}_{\Omega}} \int_K f = \sup_{j \in \mathbb{N}} \int_{K_j} f \leq +\infty$$

de modo que, en este caso, la integral  $\int_{\Omega} f$  es (absolutamente) convergente cuando su valor es finito.

La segunda parte de la conclusión de la proposición 12.3 es útil a la hora de calcular el valor de una integral impropia absolutamente convergente  $\int_{\Omega} f(\mathbf{x}) \, d\mathbf{x}$ : Se elige una sucesión expansiva  $K_j \in \mathcal{K}_{\Omega}$  tal que las integrales  $\alpha_j = \int_{K_j} f(\mathbf{x}) \, d\mathbf{x}$  sean fácilmente calculables y luego se calcula el límite  $\int_{\Omega} f(\mathbf{x}) \, d\mathbf{x} = \lim_j \alpha_j$ .

## 12.2. Paso al límite bajo la integral

Los resultados sobre continuidad y derivabilidad de integrales dependientes de un parámetro se refieren en última instancia a la posibilidad de que cierto proceso de paso al límite (límite funcional, límite de cocientes incrementales) pase bajo la integral. Comenzamos estudiando la validez de la integración término a término de sucesiones (integrales que dependen del parámetro  $k \in \mathbb{N}$ ):

**Teorema 12.5** Sea  $f_k : M \rightarrow \mathbb{R}$  una sucesión de funciones integrables Riemann sobre un conjunto medible Jordan  $M \subset \mathbb{R}^n$ , que converge uniformemente hacia  $f : M \rightarrow \mathbb{R}$ . Entonces  $f$  es integrable sobre  $M$ , y

$$\int_M f(\mathbf{x}) \, d\mathbf{x} = \lim_k \int_M f_k(\mathbf{x}) \, d\mathbf{x}$$

DEM: Consideremos primero el caso  $M = A \subset \mathbb{R}^n$ , donde  $A$  es un rectángulo cerrado  $n$ -dimensional.

Por la convergencia uniforme, la sucesión  $\rho_k = \sup_{\mathbf{x} \in A} |f_k(\mathbf{x}) - f(\mathbf{x})| \leq +\infty$ , converge hacia 0, luego existe  $k_0$  tal que  $\rho_k < +\infty$ , para todo  $k \geq k_0$ .

Como  $f_k(\mathbf{x}) - \rho_k \leq f(\mathbf{x}) \leq f_k(\mathbf{x}) + \rho_k$  para todo  $\mathbf{x} \in A$ , obtenemos que  $f$  es acotada sobre  $A$ . Además, para todo  $k \in \mathbb{N}$  se cumple

$$\int_A f_k - \rho_k v(A) \leq \int_A f \leq \overline{\int_A f} \leq \int_A f_k + \rho_k v(A)$$

luego

$$0 \leq \overline{\int_A f} - \underline{\int_A f} \leq 2\rho_k v(A)$$

y pasando al límite se obtiene  $\overline{\int_A f} = \underline{\int_A f}$ , es decir,  $f$  es integrable sobre  $A$ .

Por otra parte, usando la desigualdad  $|f(\mathbf{x}) - f_k(\mathbf{x})| \leq \rho_k$ , válida para todo  $\mathbf{x} \in A$ , y todo  $k \in \mathbb{N}$ , resulta

$$\left| \int_A f_k(\mathbf{x}) d\mathbf{x} - \int_A f(\mathbf{x}) d\mathbf{x} \right| \leq \int_A |f(\mathbf{x}) - f_k(\mathbf{x})| d\mathbf{x} \leq \rho_k v(A)$$

luego,  $\lim_k \int_A f_k(\mathbf{x}) = \int_A f(\mathbf{x}) d\mathbf{x}$ .

Para demostrar el teorema cuando  $M$  es un conjunto medible Jordan arbitrario, basta fijar un rectángulo cerrado  $n$ -dimensional  $A \supset M$  y aplicar lo que se acaba de demostrar a la sucesión  $\hat{f}_k : A \rightarrow \mathbb{R}$ , definida por

$$\hat{f}_k(\mathbf{x}) = f_k(\mathbf{x}) \text{ si } \mathbf{x} \in M, \hat{f}_k(\mathbf{x}) = 0 \text{ si } \mathbf{x} \notin M$$

Como esta sucesión converge uniformemente hacia  $\hat{f} : A \rightarrow \mathbb{R}$ , (definida en forma similar), resulta que  $\hat{f}$  es integrable sobre  $A$ , y  $\lim_k \int_A \hat{f}_k = \int_A \hat{f}$ , lo que significa que  $f$  es integrable sobre  $M$ , y  $\int_M f = \lim_k \int_M f_k$ . ■

**Corolario 12.6** *Sea  $\sum_{n=1}^{\infty} f_k(\mathbf{x})$  una serie de funciones  $f_k : M \rightarrow \mathbb{R}$ , integrables Riemann sobre un conjunto medible Jordan  $M \subset \mathbb{R}^n$ , que converge uniformemente. Entonces  $f(\mathbf{x}) = \sum_{n=1}^{\infty} f_k(\mathbf{x})$  es integrable sobre  $M$ , y se cumple*

$$\int_M f(\mathbf{x}) d\mathbf{x} = \sum_{k=1}^{\infty} \int_M f_k(\mathbf{x}) d\mathbf{x}$$

DEM: La sucesión  $S_k = \sum_{j=1}^k f_j$ , converge uniformemente sobre  $M$  hacia  $f$ , y en virtud de 12.5 la sucesión  $\int_M S_k = \sum_{j=1}^k (\int_M f_j)$  converge hacia  $\int_M f$ . ■

Cuando se sabe que la función límite es integrable Riemann los siguientes resultados (teoremas 12.7 y 12.8) garantizan el paso al límite bajo la integral con hipótesis más débiles que la convergencia uniforme.

**Teorema 12.7** Sea  $f_n : M \rightarrow \mathbb{R}$  una sucesión de funciones integrables Riemann en el conjunto medible Jordan  $M \subset \mathbb{R}^n$  que converge puntualmente hacia una función  $f : M \rightarrow \mathbb{R}$  que se supone integrable Riemann sobre  $M$ . Si la sucesión  $f_n$  es uniformemente acotada, (existe  $C > 0$  tal que  $|f_n(\mathbf{t})| \leq C$  para todo  $\mathbf{t} \in M$ , y todo  $n \in \mathbb{N}$ ) entonces

$$\int_M f(\mathbf{x}) \, d\mathbf{x} = \lim_n \int_M f_n(\mathbf{x}) \, d\mathbf{x}$$

En lo que sigue diremos que la sucesión  $f_n : \Omega \rightarrow \mathbb{R}$  está dominada por la función  $g : \Omega \rightarrow [0, +\infty)$  cuando para todo  $\mathbf{t} \in \Omega$  y todo  $n \in \mathbb{N}$  se cumple  $|f_n(\mathbf{t})| \leq g(\mathbf{t})$ .

**Teorema 12.8** Sea  $f_n : \Omega \rightarrow \mathbb{R}$  una sucesión de funciones absolutamente integrables Riemann en el recinto de integración  $\Omega \subset \mathbb{R}^n$  que converge puntualmente sobre  $\Omega$  hacia una función localmente integrable  $f : \Omega \rightarrow \mathbb{R}$ . Si la sucesión  $f_n$  está dominada por una función absolutamente integrable Riemann  $g : \Omega \rightarrow [0, +\infty)$  entonces  $f$  es absolutamente integrable Riemann y se cumple  $\int_\Omega f(\mathbf{t}) \, d\mathbf{t} = \lim_n \int_\Omega f_n(\mathbf{t}) \, d\mathbf{t}$ .

Los teoremas 12.7 y 12.8, en el contexto de la teoría de la integral de Lebesgue, son versiones particulares del teorema de la convergencia dominada.

**Funciones definidas por integrales.** Sea  $M \subset \mathbb{R}^n$  medible Jordan,  $T$  un conjunto, y  $f : T \times M \rightarrow \mathbb{R}$  una función tal que para cada  $\mathbf{t} \in T$  la función parcial  $f_{\mathbf{t}} : M \rightarrow \mathbb{R}$  es integrable Riemann. En estas condiciones la integral dependiente del parámetro  $\mathbf{t} \in T$ ,

$$F(\mathbf{t}) = \int_M f(\mathbf{t}, \mathbf{x}) \, d\mathbf{x}$$

define una función  $F : T \rightarrow \mathbb{R}$ .

Cuando  $T$  es un espacio métrico (resp. un abierto de  $\mathbb{R}^k$ ) se estudian a continuación condiciones suficientes para que  $F$  sea continua (resp. derivable).

**Teorema 12.9** Sea  $(T, d)$  un espacio métrico,  $M \subset \mathbb{R}^n$  un compacto medible Jordan. Si  $f : T \times M \rightarrow \mathbb{R}$  es continua, la función

$$F(\mathbf{t}) = \int_M f(\mathbf{t}, \mathbf{x}) \, d\mathbf{x}$$

definida en  $T$ , también es continua.

DEM: Como  $(T, d)$  es un espacio métrico podemos utilizar la caracterización de la continuidad por sucesiones: Basta demostrar que si una sucesión  $\mathbf{t}_k \in T$  converge hacia  $\mathbf{t} \in T$ , entonces  $F(\mathbf{t}_k)$  converge hacia  $F(\mathbf{t})$ . Como  $K = \{\mathbf{t}_k : n \in \mathbb{N}\} \cup \{\mathbf{t}\}$ , y  $M$  son compactos también lo es  $K \times M$ , para la distancia  $\rho((\mathbf{u}, \mathbf{x}), (\mathbf{v}, \mathbf{y})) = \max\{d(\mathbf{u}, \mathbf{v}), d_2(\mathbf{x}, \mathbf{y})\}$  luego  $f$  es uniformemente continua sobre  $K \times M$ , es decir, para cada  $\epsilon > 0$ , existe  $\delta > 0$ , tal que

$$\mathbf{u}, \mathbf{v} \in T, \quad \mathbf{x}, \mathbf{y} \in M, \quad d(\mathbf{u}, \mathbf{v}) < \delta, \quad d_2(\mathbf{x}, \mathbf{y}) < \delta \Rightarrow |f(\mathbf{u}, \mathbf{x}) - f(\mathbf{v}, \mathbf{y})| < \epsilon$$

Como  $\mathbf{t}_k$  converge hacia  $\mathbf{t}$ , existe  $n_\delta \in \mathbb{N}$  tal que  $n \geq n_\delta \Rightarrow d(\mathbf{t}_k, \mathbf{t}) < \delta$ , luego, para todo  $\mathbf{x} \in M$  se cumple  $|f(\mathbf{t}, \mathbf{x}) - f(\mathbf{t}_k, \mathbf{x})| < \epsilon$ .

Esto significa que la sucesión  $f_{\mathbf{t}_k}$  converge hacia  $f_{\mathbf{t}}$  uniformemente sobre  $M$ , y aplicando el teorema 12.5 se concluye que la sucesión  $F(\mathbf{t}_k) = \int_M f_{\mathbf{t}_k}$  converge hacia  $F(\mathbf{t}) = \int_M f_{\mathbf{t}}$ . ■

**Teorema 12.10** *Sea  $\Omega \subset \mathbb{R}^k$  abierto y  $M \subset \mathbb{R}^n$  un compacto medible Jordan y  $f : \Omega \times M \rightarrow \mathbb{R}$  una función continua, tal que en cada  $(\mathbf{t}, \mathbf{x}) \in \Omega \times K$  existe y es continua la derivada parcial  $\frac{\partial f}{\partial t_j}(\mathbf{t}, \mathbf{x})$ . Entonces la función  $F(\mathbf{t}) = \int_M f(\mathbf{t}, \mathbf{x})d\mathbf{x}$ , definida en  $\Omega$ , es derivable en cada  $\mathbf{t} \in \Omega$ , respecto a la variable  $t_j$ , y se verifica*

$$\frac{\partial F(\mathbf{t})}{\partial t_j} = \int_M \frac{\partial f(\mathbf{t}, \mathbf{x})}{\partial t_j} d\mathbf{x}$$

DEM: Suponemos, para simplificar la escritura, que  $j = 1$ . Dado  $\mathbf{t} \in \Omega$ , fijemos una bola cerrada  $\overline{B(\mathbf{t}, r)} \subset \Omega$ . Así, para cada  $h \in \mathbb{R}$ , con  $|h| < r$ , está definido

$$\begin{aligned} \Delta(\mathbf{t}, h) &= \frac{F(t_1 + h, t_2, t_3, \dots, t_k) - F(t_1, t_2, t_3, \dots, t_k)}{h} = \\ &= \frac{1}{h} \int_M (f(t_1 + h, t_2, \dots, t_k, \mathbf{x}) - f(t_1, t_2, \dots, t_k, \mathbf{x}))d\mathbf{x}. \end{aligned}$$

La función  $s \rightarrow f(s, t_2, \dots, t_k, \mathbf{x})$ , es derivable en el intervalo  $(t_1 - r, t_1 + r)$ , luego, en virtud del teorema del valor medio,

$$f(t_1 + h, t_2, \dots, t_k, \mathbf{x}) - f(t_1, t_2, \dots, t_k, \mathbf{x}) = h \frac{\partial f}{\partial t_1}(t_1 + h\theta_{\mathbf{x}}, t_2, \dots, t_k, \mathbf{x})$$

donde  $\theta_{\mathbf{x}} \in [0, 1]$  depende de  $\mathbf{x}$  y de  $h$  ( $\mathbf{t}$  está fijo todo el rato). Se obtiene así que

$$\Delta(\mathbf{t}, h) = \int_M \frac{\partial f}{\partial t_1}(t_1 + h\theta_{\mathbf{x}}, t_2, \dots, t_k, \mathbf{x})d\mathbf{x}$$

luego

$$\begin{aligned} &\left| \Delta(\mathbf{t}, h) - \int_M \frac{\partial f}{\partial t_1}(\mathbf{t}, \mathbf{x})d\mathbf{x} \right| \leq \\ &\leq \int_M \left| \frac{\partial f}{\partial t_1}(t_1 + h\theta_{\mathbf{x}}, t_2, \dots, t_k, \mathbf{x}) - \frac{\partial f}{\partial t_1}(t_1, t_2, \dots, t_k, \mathbf{x}) \right| d\mathbf{x} \end{aligned}$$

Como la derivada parcial  $\partial f / \partial t_1$  es uniformemente continua sobre el compacto  $\overline{B(\mathbf{t}, r)} \times M \subset \mathbb{R}^{k+n}$ , para cada  $\epsilon > 0$  existe  $\delta \in (0, r)$  tal que para todo par  $\mathbf{t}', \mathbf{t}'' \in \overline{B(\mathbf{t}, r)}$  con  $d_{\mathbb{R}^k}(\mathbf{t}', \mathbf{t}'') < \delta$ , y todo par,  $\mathbf{x}', \mathbf{x}'' \in M$ , con  $d_{\mathbb{R}^n}(\mathbf{x}', \mathbf{x}'') < \delta$ , se cumple

$$\left| \frac{\partial f}{\partial t_1}(\mathbf{t}', \mathbf{x}') - \frac{\partial f}{\partial t_1}(\mathbf{t}'', \mathbf{x}'') \right| < \epsilon$$

Si  $|h| < \delta$  ( $< r$ ), los dos puntos  $(t_1, t_2, \dots, t_k)$ ,  $(t_1 + h\theta_{\mathbf{x}}, t_2, \dots, t_k)$  están en la bola  $B(\mathbf{t}, r)$ , y la distancia entre ellos es menor que  $\delta$ , luego, para cada  $\mathbf{x} \in M$

$$\left| \frac{\partial f}{\partial t_1}(t_1 + h\theta_{\mathbf{x}}, t_2, \dots, t_k, \mathbf{x}) - \frac{\partial f}{\partial t_1}(t_1, t_2, \dots, t_k, \mathbf{x}) \right| < \epsilon$$

Se sigue que para  $|h| < \delta$  se verifica  $\left| \Delta(\mathbf{t}, h) - \int_M \frac{\partial f}{\partial t_j}(\mathbf{t}, \mathbf{x}) d\mathbf{x} \right| \leq \epsilon c_n(M)$  y queda demostrado que

$$\lim_{h \rightarrow 0} \Delta(\mathbf{t}, h) = \int_M \frac{\partial f}{\partial t_1}(\mathbf{t}, \mathbf{x}) d\mathbf{x}$$

■

Finalmente exponemos algunos resultados de carácter complementario referentes a integrales impropias dependientes de un parámetro. Son versiones restringidas de resultados más generales cuya formulación adecuada requiere el conocimiento de la integral de Lebesgue. Se obtienen de forma natural combinando los resultados básicos obtenidos hasta ahora.

**Proposición 12.11** *Sea  $(T, d)$  un espacio métrico,  $\Omega \subset \mathbb{R}^n$  un recinto de integración y  $f : T \times \Omega \rightarrow \mathbb{R}$  una función continua tal que existe una función absolutamente integrable  $g : \Omega \rightarrow [0, +\infty)$  verificando  $|f(\mathbf{t}, \mathbf{x})| \leq g(\mathbf{x})$  para todo  $\mathbf{t} \in T$  y todo  $\mathbf{x} \in \Omega$ . Entonces la función  $F(\mathbf{t}) = \int_{\Omega} f(\mathbf{t}, \mathbf{x}) d\mathbf{x}$ , está definida y es continua en  $T$ .*

DEM: La hipótesis implica que todas las integrales impropias  $\int_{\Omega} f(\mathbf{t}, \mathbf{x}) d\mathbf{x}$  son absolutamente convergentes y por lo tanto  $F(\mathbf{t})$  está definida para todo  $\mathbf{t} \in T$ .

Sea  $K_j \in \mathcal{K}_{\Omega}$  una sucesión expansiva de compactos medibles Jordan contenidos en  $\Omega$ . Según el teorema 12.9, las funciones  $F_j(\mathbf{t}) = \int_{K_j} f(\mathbf{t}, \mathbf{x}) d\mathbf{x}$  están definidas y son continuas en  $T$ . Es evidente que la sucesión  $F_j$  converge puntualmente hacia  $F$ , y si demostramos que la convergencia es uniforme, acudiendo al teorema 12.9 obtendremos la continuidad de  $F$ .

Usando la convergencia de la sucesión  $\int_{K_j} g(\mathbf{x}) d\mathbf{x}$ , y la siguiente desigualdad, válida para  $j \geq i$ , y todo  $\mathbf{t} \in T$ ,

$$|F_j(\mathbf{t}) - F_i(\mathbf{t})| \leq \int_{K_j \setminus K_i} |f(\mathbf{t}, \mathbf{x})| d\mathbf{x} \leq \int_{K_j \setminus K_i} g(\mathbf{x}) d\mathbf{x} = \int_{K_j} g(\mathbf{x}) d\mathbf{x} - \int_{K_i} g(\mathbf{x}) d\mathbf{x}$$

se obtiene fácilmente que la sucesión  $F_j$  cumple la condición de Cauchy para la convergencia uniforme sobre  $T$ . ■

**Proposición 12.12** Sea  $\Omega \subset \mathbb{R}^n$  un recinto de integración y  $f : (a, b) \times \Omega \rightarrow \mathbb{R}$  una función continua, tal que para cada  $\mathbf{x} \in \Omega$  la función parcial  $t \rightarrow f(t, \mathbf{x})$  es derivable, y su derivada  $D_1 f(t, \mathbf{x})$ , define una función continua en  $(a, b) \times \Omega$ .

Se supone que para algún  $t_0 \in (a, b)$  la función parcial  $\mathbf{x} \rightarrow f(t_0, \mathbf{x})$  es absolutamente integrable Riemann sobre  $\Omega$  y que las funciones  $\mathbf{x} \rightarrow D_1 f(t, \mathbf{x})$  están dominadas por una función absolutamente integrable Riemann  $g : \Omega \rightarrow \mathbb{R}$ :

$$|D_1 f(t, \mathbf{x})| \leq g(\mathbf{x}) \text{ para todo } \mathbf{x} \in \Omega, \text{ y todo } t \in (a, b)$$

Entonces, todas las funciones parciales  $\mathbf{x} \rightarrow f(t, \mathbf{x})$  son absolutamente integrables, y la función definida por sus integrales

$$F(t) = \int_{\Omega} f(t, \mathbf{x}) d\mathbf{x}$$

es derivable en  $(a, b)$ , con derivada

$$F'(t) = \int_{\Omega} D_1 f(t, \mathbf{x}) d\mathbf{x}$$

DEM: Aplicando el teorema del incremento finito a la función  $t \rightarrow f(t, \mathbf{x})$ , se deduce que para cada  $t \in (a, b)$  se cumple

$$|f(t, \mathbf{x}) - f(t_0, \mathbf{x})| = |t - t_0| |D_1 f(\xi, \mathbf{x})| \leq (b - a)g(\mathbf{x})$$

(donde  $\xi \in (a, b)$  es un punto intermedio del intervalo de extremos  $t, t_0$ ).

Se obtiene así que para todo  $t \in (a, b)$  y todo  $\mathbf{x} \in \Omega$  se cumple a desigualdad

$$|f(t, \mathbf{x})| \leq (b - a)g(\mathbf{x}) + |f(t_0, \mathbf{x})|$$

con la que se obtiene fácilmente que todas las funciones  $\mathbf{x} \rightarrow f(t, \mathbf{x})$  son absolutamente integrables Riemann sobre  $\Omega$ , luego para cada  $t \in \Omega$  está definida

$$F(t) = \int_{\Omega} f(t, \mathbf{x}) d\mathbf{x}$$

Sea  $K_j \in \mathcal{K}_{\Omega}$  una sucesión expansiva de compactos medibles Jordan en  $\Omega$ . Según el teorema 12.10 las funciones  $F_j(t) = \int_{K_j} f(t, \mathbf{x}) d\mathbf{x}$ , están definidas y son derivables en  $(a, b)$ , con derivada  $F'_j(t) = \int_{K_j} D_1 f(t, \mathbf{x}) d\mathbf{x}$ .

Obsérvese que la condición de dominación que figura como hipótesis implica que para cada  $t \in (a, b)$  la función continua  $\mathbf{x} \rightarrow D_1 f(t, \mathbf{x})$  es absolutamente integrable sobre  $\Omega$  y por lo tanto, para cada  $t \in (a, b)$ , se cumple

$$\int_{\Omega} D_1 f(t, \mathbf{x}) d\mathbf{x} = \lim_j \int_{K_j} D_1 f(t, \mathbf{x}) d\mathbf{x} = \lim_j F'_j(t)$$

Además, para  $j > i$  y cualquier  $t \in (a, b)$  se verifica la desigualdad

$$|F'_j(t) - F'_i(t)| \leq \int_{K_j \setminus K_i} |D_1 f(t, \mathbf{x})| d\mathbf{x} \leq \int_{K_j \setminus K_i} g(\mathbf{x}) d\mathbf{x} = \int_{K_j} g(\mathbf{x}) d\mathbf{x} - \int_{K_i} g(\mathbf{x}) d\mathbf{x}$$



y con la condición de Cauchy se obtiene que  $F'_j$  converge uniformemente en  $(a, b)$ .

Por otra parte, como la función  $\mathbf{x} \rightarrow f(t_0, \mathbf{x})$  es absolutamente integrable, podemos asegurar que existe el límite

$$\lim_j F_j(t_0) = \lim_j \int_{K_j} f(t_0, \mathbf{x}) d\mathbf{x} = \int_{\Omega} f(t_0, \mathbf{x}) d\mathbf{x}$$

y aplicando el teorema A.11 a las sucesión  $F_j$  se deduce que  $F(t) = \lim_j F_j(t)$ , es derivable en  $(a, b)$ , con derivada

$$F'(t) = \lim_j F'_j(t) = \int_{\Omega} D_1 f(t, \mathbf{x}) d\mathbf{x}$$

■

### 12.3. Ejercicios resueltos

**Ejercicio 12.13** Compruebe que la integral impropia

$$I_p = \int_{\Omega} \frac{dx dy dz}{(x^2 + y^2 + z^2)^p}, \quad \Omega = \{(x, y, z) : r^2 \leq x^2 + y^2 + z^2\}$$

es finita si y sólo si  $2p > 3$ , y obtenga su valor.

SOLUCIÓN

Como la integral se plantea sobre un conjunto no acotado, debemos considerarla como integral impropia. Es claro que  $\Omega$  es un recinto de integración donde la sucesión de compactos medibles Jordan  $K_j = \{(x, y, z) : r^2 \leq x^2 + y^2 + z^2 \leq (r+j)^2\}$  es expansiva. La función  $f_p(x, y, z) = (x^2 + y^2 + z^2)^{-p}$  es continua en  $\Omega$  y por lo tanto localmente integrable. Para calcular  $I_p = \int_{\Omega} f_p$  comenzamos calculando las integrales  $\alpha_j = \int_{K_j} f_p$ .

Si  $2p \neq 3$ , con un cambio de variable a coordenadas esféricas se obtiene

$$\alpha_j = \int_0^{2\pi} d\theta \int_{-\pi/2}^{\pi/2} d\varphi \int_r^{r+j} \frac{\rho^2 \cos \varphi}{\rho^{2p}} d\rho = \frac{4\pi}{3-2p} [(r+j)^{3-2p} - r^{3-2p}]$$

Si  $2p - 3 > 0$ , la sucesión  $\alpha_j$  es convergente hacia el valor

$$I_p = \frac{4\pi}{3-2p} r^{3-2p}$$

Con el mismo cálculo, para  $2p - 3 < 0$ , y con un cálculo similar para el caso  $2p - 3 = 0$ , se obtiene que  $\int_{\Omega} f_p = +\infty$  cuando  $2p \leq 3$ . ■

**Ejercicio 12.14** Para  $p > 0$  se considera la integral impropia

$$I_p = \int_{\Omega} \frac{dx dy dz}{(x^2 + y^2 + z^2)^p}, \quad \Omega = \{(x, y, z) : 0 < x^2 + y^2 + z^2 \leq r^2\}$$

Compruebe que es convergente si y sólo si  $2p < 3$ , y obtenga su valor.

## SOLUCIÓN

Aunque  $\Omega$  es medible Jordan, la función  $f_p(x, y, z) = (x^2 + y^2 + z^2)^{-p}$  no está acotada sobre  $\Omega$ , por lo que la integral se debe considerar como integral impropia. Obsérvese que  $\Omega$  es un recinto de integración y que  $f_p$  es localmente integrable en  $\Omega$ , por ser continua. La sucesión de compactos medibles Jordan

$$K_j = \{(x, y, z) : j^{-2} \leq x^2 + y^2 + z^2 \leq r^2\}$$

es expansiva en  $\Omega$ , y para calcular  $I_p$  comenzamos calculando  $\alpha_j = \int_{K_j} f_p$ . Si  $2p \neq 3$ , con un cambio de variable a coordenadas esféricas se obtiene

$$\alpha_j = \int_0^{2\pi} d\theta \int_{-\pi/2}^{\pi/2} d\varphi \int_{1/j}^r \frac{\rho^2 \cos \varphi}{\rho^{2p}} d\rho = \frac{4\pi}{3-2p} [r^{3-2p} - j^{2p-3}]$$

Si  $2p - 3 < 0$ , la sucesión  $\alpha_j$  es convergente hacia el valor

$$I_p = \frac{4\pi}{3-2p} r^{3-2p}$$

Con el mismo cálculo, para  $2p - 3 > 0$ , y con un cálculo similar para el caso  $2p - 3 = 0$ , se obtiene que  $\int_{\Omega} f_p = +\infty$  cuando  $2p \geq 3$ . ■

**Ejercicio 12.15** *Compruebe que la integral impropia*

$$I = \int_{\Omega} \frac{dx \, dy \, dz}{\sqrt{x^2 + y^2 + z^2}}, \quad \Omega = \{(x, y, z) : x^2 + y^2 + (z - R)^2 \leq R^2, z > 0\}$$

es convergente y calcule su valor.

## SOLUCIÓN

Aunque  $\Omega$  es medible Jordan, la función  $f(x, y, z) = 1/\sqrt{x^2 + y^2 + z^2}$  no es acotada en  $\Omega$  (pues  $f(0, 0, z) = 1/z$ , cuando  $(0, 0, z) \in \Omega$ ) luego la integral debe considerarse como integral impropia. Es obvio que  $\Omega$  es un recinto de integración y que  $f$  es localmente integrable por ser continua. Si  $\epsilon_j \in (0, 1)$  es una sucesión decreciente que converge hacia 0, la sucesión de compactos medibles Jordan

$$K_j = \{(x, y, z) : x^2 + y^2 + (z - R)^2 \leq R^2, z \geq \epsilon_j\}$$

es expansiva en  $\Omega$ , y para calcular  $I$  comenzamos calculando  $\alpha_j = \int_{K_j} f$ . Con un cambio de variable a coordenadas cilíndricas

$$x = r \cos \theta, \quad y = r \sin \theta, \quad z = t$$

se obtiene

$$\alpha_j = \int_{\epsilon_j}^{2R} dt \int_0^{2\pi} d\theta \int_0^{\sqrt{2Rt-t^2}} \frac{r \, dr}{\sqrt{r^2 + t^2}} = 2\pi \int_{\epsilon_j}^{2R} (\sqrt{2Rt} - t) dt$$

Como la función  $t \rightarrow \sqrt{2Rt} - t$ , es integrable en  $[0, 2R]$  (por ser continua), y  $\lim_j \epsilon_j = 0$ , en virtud del teorema fundamental del cálculo existe el límite

$$\lim_j \alpha_j = \int_0^{2R} (\sqrt{2Rt} - t) dt = \frac{2}{3} R^2$$

lo que significa que la integral impropia converge y que su valor es  $I = \frac{4}{3}\pi R^2$ . ■

**Ejercicio 12.16** Compruebe que las siguientes integrales impropias son convergentes y calcule sus valores:

$$\begin{aligned} a) \quad I &= \int_{\Omega} \frac{dx \, dy}{\sqrt{x^2 + y^2 - R^2}}, \quad \Omega = \{(x, y) : x^2 + y^2 \leq R^2\}. \\ b) \quad J &= \int_{\Omega} e^{-(x^2+y^2)} dx \, dy, \quad \Omega = \mathbb{R}^2. \end{aligned}$$

SOLUCIÓN

Como las funciones son no negativas, basta ver que las integrales tienen un valor finito que se puede calcular usando la sucesión expansiva de discos cerrados que se indica en cada caso.

a) Utilizamos la sucesión expansiva de discos  $K_j = \{(x, y) : x^2 + y^2 \leq R_j^2\}$ , donde  $R_j \in (0, R)$  es una sucesión creciente convergente hacia  $R$ . Con un cambio de variable a coordenadas polares se obtiene que la integral sobre  $K_j$  vale,

$$\alpha_j = 2\pi(R_j - \sqrt{R^2 - R_j^2})$$

luego  $I = \lim_j \alpha_j = 2\pi R$ .

b) Consideramos la sucesión expansiva de discos  $D_j = \{(x, y) : x^2 + y^2 \leq R_j^2\}$ , donde  $R_j > 0$  es una sucesión creciente convergente hacia  $+\infty$ . Con un cambio de variable a coordenadas polares se obtiene que la integral sobre  $D_j$  vale,

$$\beta_j = 2\pi \int_0^{R_j} r e^{-r^2} dr = \pi(1 - e^{-R_j^2})$$

luego  $J = \lim_j \beta_j = \pi$ . ■

**Ejercicio 12.17** Se supone que  $g, h : (a, b) \rightarrow \mathbb{R}$  son continuas, y que  $\Omega \subset \mathbb{R}^2$  es un abierto que contiene al conjunto  $\{(t, x) \in \mathbb{R}^2 : a < t < b, x \in [g(t), f(t)]\}$ , donde  $[g(t), h(t)] = \{\alpha g(t) + (1 - \alpha)h(t) : 0 \leq \alpha \leq 1\}$ . Si  $f : \Omega \rightarrow \mathbb{R}$  es continua, demuestre que la siguiente integral

$$F(t) = \int_{g(t)}^{h(t)} f(t, x) dx$$

define una función continua en  $(a, b)$ .

## SOLUCIÓN

Por hipótesis, para cada  $t \in (a, b)$  el punto  $(g(t), h(t), t)$  pertenece al conjunto

$$A = \{(u, v, t) \in \mathbb{R}^3 : \{t\} \times [u, v] \subset \Omega\}$$

donde  $[u, v] = \{\alpha u + (1 - \alpha)v : 0 \leq \alpha \leq 1\} \subset \mathbb{R}$  es el segmento de extremos  $u, v$  (¡atención: no se supone  $u \leq v$ !). La función  $F$  se puede expresar como la composición  $F(t) = \varphi(g(t), h(t), t)$ , donde  $\varphi(u, v, t) = \int_u^v f(t, x)dx$ , está definida en  $A$ , luego basta demostrar que  $\varphi$  es continua en  $A$ .

Para ello comenzamos viendo que  $A$  es un subconjunto abierto de  $\mathbb{R}^3$ , y esto lo haremos comprobando, con la técnica de las sucesiones, que su complemento  $\mathbb{R}^3 \setminus A$ , es cerrado: Si  $(u_n, v_n, t_n)$  es una sucesión en  $\mathbb{R}^3 \setminus A$ , que converge hacia  $(u, v, t)$ , según la definición de  $A$ , para cada  $n \in \mathbb{N}$  existe  $\alpha_n \in [0, 1]$  tal que  $(t_n, \alpha_n u_n + (1 - \alpha_n)v_n) \notin \Omega$ . La sucesión  $\alpha_n$  posee una subsucesión  $\alpha_{n_k}$  convergente hacia un punto  $\alpha \in [0, 1]$ . Como  $(t, \alpha u + (1 - \alpha)v)$  es el límite de la sucesión  $(t_{n_k}, \alpha_{n_k} u_{n_k} + (1 - \alpha_{n_k})v_{n_k})$ , que está contenida en el cerrado  $\mathbb{R}^2 \setminus \Omega$ , resulta  $(t, \alpha u + (1 - \alpha)v) \in \mathbb{R}^2 \setminus \Omega$ , es decir  $(u, v, t) \notin \mathbb{R}^3 \setminus A$ .

Como  $A$  es abierto, dado  $(u_0, v_0, t_0) \in A$  existe  $r > 0$  tal que

$$W_r = [u_0 - r, u_0 + r] \times [v_0 - r, v_0 + r] \times [t_0 - r, t_0 + r] \subset A$$

Para cada  $(u, v, t) \in W_r$ , los tres puntos  $(u_0, v_0, t)$ ,  $(u_0, v, t)$ ,  $(u_0, v_0, t_0)$ , pertenecen a  $W_r \subset A$ , y podemos escribir

$$\varphi(u, v, t) - \varphi(u_0, v_0, t_0) =$$

$$(\varphi(u, v, t) - \varphi(u_0, v, t)) + (\varphi(u_0, v, t) - \varphi(u_0, v_0, t)) + (\varphi(u_0, v_0, t) - \varphi(u_0, v_0, t_0))$$

Los dos primeros sumandos se puede acotar en la forma

$$|\varphi(u, v, t) - \varphi(u_0, v, t)| \leq \int_{[u_0, u]} |f(t, x)|dx \leq C_1 |u - u_0|$$

$$|\varphi(u_0, v, t) - \varphi(u_0, v_0, t)| \leq \int_{[v_0, v]} |f(t, x)|dx \leq C_2 |v - v_0|$$

donde

$$C_1 = \max\{|f(t, x)| : |t - t_0| \leq r, |x - u_0| \leq r\}$$

$$C_2 = \max\{|f(t, x)| : |t - t_0| \leq r, |x - v_0| \leq r\}$$

y, según 12.9, el tercer sumando,  $|\varphi(u_0, v_0, t) - \varphi(u_0, v_0, t_0)|$ , tiende hacia, 0 cuando  $t \rightarrow t_0$ . Se sigue que para cada  $\epsilon$  existe  $\delta \in (0, r)$  tal que

$$\max\{|u - u_0|, |v - v_0|, |t - t_0|\} < \delta \Rightarrow |\varphi(u, v, t) - \varphi(u_0, v_0, t_0)| < \epsilon$$

Queda demostrado así que  $\varphi$  es continua en  $A$ , y con ello la continuidad de  $F$ . ■

**Ejercicio 12.18** En las condiciones de 12.17 si las funciones  $g, h : (a, b) \rightarrow \mathbb{R}$  son derivables y la derivada parcial  $\frac{\partial f}{\partial t} f(t, x)$  existe y es continua en todos los puntos de  $\Omega$ , demuestre que  $F$  es derivable en  $(a, b)$ , y vale la regla de derivación de Leibniz:

$$F'(t) = f(h(t), t)h'(t) - f(g(t), t)g'(t) + \int_{g(t)}^{h(t)} \frac{\partial f}{\partial t}(t, x)dx$$

SOLUCIÓN

Basta aplicar la regla de la cadena del cálculo diferencial a la función compuesta  $F(t) = \varphi(g(t), h(t), t)$ , donde  $\varphi(u, v, t) = \int_u^v f(t, x)dx$  está definida en el abierto  $A = \{(u, v, t) \in \mathbb{R}^3 : \{t\} \times [u, v] \subset \Omega\}$ , (véase 12.17) y tiene derivadas parciales continuas: Efectivamente, en virtud del teorema fundamental del cálculo y de 12.10

$$\frac{\partial \varphi}{\partial u}(u, v, t) = -f(u, t), \quad \frac{\partial \varphi}{\partial v}(u, v, t) = f(v, t), \quad \frac{\partial \varphi}{\partial t}(u, v, t) = \int_u^v \frac{\partial f}{\partial t}(t, x)dx$$

Se sigue que  $\varphi$  es diferenciable, y usando la regla de la cadena del cálculo diferencial

$$\begin{aligned} F'(t) &= D_1\varphi(g(t), h(t), t)g'(t) + D_2\varphi(g(t), h(t), t)h'(t) + D_3\varphi(g(t), h(t), t) = \\ &= -f(g(t), t)g'(t) + f(h(t), t)h'(t) + \int_{g(t)}^{h(t)} \frac{\partial f}{\partial t}(t, x)dx \end{aligned}$$

■

**Ejercicio 12.19** Sean  $K \subset \mathbb{R}^k$ ,  $M \subset \mathbb{R}^n$  compactos medibles Jordan. Dada una función continua  $f : K \times M \rightarrow \mathbb{R}$ , justifique que la integral sobre  $K$  de la función continua  $F(\mathbf{t}) = \int_M f(\mathbf{t}, \mathbf{x})d\mathbf{x}$ , viene dada por  $\int_K F(\mathbf{t})d\mathbf{t} = \int_M \left( \int_K f(\mathbf{t}, \mathbf{x})d\mathbf{t} \right) d\mathbf{x}$ .

SOLUCIÓN

$K \times M$  es compacto en  $\mathbb{R}^k \times \mathbb{R}^n$ , y medible Jordan en virtud del ejercicio 10.34. La función continua  $f$  es integrable Riemann sobre el conjunto medible Jordan  $K \times M$ , y aplicando el teorema de Fubini se obtiene que

$$\int_{K \times M} f(\mathbf{t}, \mathbf{x})d\mathbf{x}d\mathbf{t} = \int_K \left( \int_M f(\mathbf{t}, \mathbf{x})d\mathbf{x} \right) d\mathbf{t} = \int_M \left( \int_K f(\mathbf{t}, \mathbf{x})d\mathbf{t} \right) d\mathbf{x}$$

es decir

$$\int_K F(\mathbf{t})d\mathbf{t} = \int_M \left( \int_K f(\mathbf{t}, \mathbf{x})d\mathbf{t} \right) d\mathbf{x}$$

■

## 12.4. Ejercicios propuestos

◇ 12.4.1 Calcule la integral impropia

$$\int_E \frac{dx dy}{(x^2 + y^2)^2}$$

donde  $E = \{(x, y) : (x - 1)^2 + y^2 < 1\}$

◇ 12.4.2 Calcule la integral impropia  $\int_M \frac{dx dy dz}{z^2}$ , donde

$$M = \{(x, y, z) : \sqrt{x^2 + y^2} \leq z \leq \sqrt{1 - x^2 - y^2}\}$$

◇ 12.4.3 Considerando la integral de  $e^{-(x^2+y^2)}$  sobre los conjuntos

$$Q_R = \{(x, y) : |x| < R, |y| < R\} \quad y \quad D_R = \{(x, y) : x^2 + y^2 < R^2\}.$$

calcule el valor de la integral impropia  $\int_0^{+\infty} e^{-t^2} dt$ .

◇ 12.4.4 Criterio de Abel para la convergencia uniforme de integrales impropias. Sea  $T$  un conjunto arbitrario y  $f : [a, +\infty) \times T \rightarrow \mathbb{R}$  una función tal que todas las funciones parciales  $x \rightarrow f(x, t)$  son continuas.

Se dice que la integral  $\int_a^{+\infty} f(x, t) dx$  es uniformemente convergente sobre  $T$  cuando para cada sucesión  $u_n \geq a$  con  $\lim_n u_n = +\infty$  la sucesión de funciones  $F_n(t) = \int_a^{u_n} f(x, t) dx$  converge uniformemente sobre  $T$ . Demuestre que esto ocurre cuando  $f$  es de la forma  $f(x, t) = \alpha(x, t)\beta(x, t)$  donde todas las funciones parciales  $x \rightarrow \alpha(x, t)$ ,  $x \rightarrow \beta(x, t)$  son continuas y verifican:

a) Las funciones  $x \rightarrow \alpha(x, t)$  son decrecientes y tienden hacia 0, cuando  $x$  tiende hacia  $+\infty$ , uniformemente en  $T$ .

b) Las integrales  $S_u(t) = \int_a^u \beta(x, t) dx$  están uniformemente acotadas, e.d.

$$\sup\{|S_u(t)| : t \in T, u \geq a\} < +\infty$$

Aplicación: Justifique que la integral  $\int_0^{+\infty} te^{-(a^2+x^2)t^2} \cos x dx$  converge uniformemente, respecto al parámetro real  $t$ , en todo  $\mathbb{R}$ .

◇ 12.4.5 Si  $f : [a, +\infty) \times [c, d] \rightarrow \mathbb{R}$  es continua y la integral impropia  $F(t) = \int_a^{+\infty} f(x, t) dx$  converge uniformemente sobre  $[c, d]$  (véase la definición en 12.4.4) demuestre que

$$\int_c^d F(t) dt = \int_a^{+\infty} \left( \int_c^d f(x, t) dt \right) dx$$

◇ 12.4.6 Justifique las siguientes afirmaciones:

a)  $F(t, r) = \int_0^{+\infty} e^{-tx} \frac{\text{sen } rx}{x} dx$  converge para todo  $t \geq 0$  y todo  $r \in \mathbb{R}$ .

b) Para cada  $r \in \mathbb{R}$  la función  $t \rightarrow F(t, r)$  es continua en  $[0, +\infty)$ .

c) Para cada  $t > 0$  la función  $r \rightarrow F(t, r)$  es derivable y  $\frac{\partial F}{\partial r}(t, r) = \frac{t}{t^2 + r^2}$ .

Utilice a) y b) para calcular  $\int_0^{\infty} \frac{\operatorname{sen} x}{x} dx = \frac{\pi}{2}$ .

◇ **12.4.7** Para  $a > 0$  calcule la integral

$$\int_0^{+\infty} e^{-ax} \operatorname{sen} tx \, dx = \frac{t}{t^2 + a^2}$$

y utilice el resultado para obtener

$$a) \int_0^{+\infty} e^{-ax} \frac{\cos ux - \cos vx}{x} dx = \frac{1}{2} \log \frac{a^2 + v^2}{a^2 + u^2};$$

$$b) \int_0^{+\infty} \frac{\cos ux - \cos vx}{x} dx = \log \frac{v}{u}.$$

◇ **12.4.8** Establezca la igualdad  $\int_0^{+\infty} e^{-x^2} \frac{\operatorname{sen} ax}{x} dx = \frac{\sqrt{\pi}}{2} \int_0^a e^{-t^2/4} dt$ .

◇ **12.4.9** Justifique que la integral  $F(a) = \int_0^{+\infty} \frac{\log(1 + a^2 x^2)}{1 + x^2} dx$  define en  $\mathbb{R}$  una función continua que es derivable en cada  $a \neq 0$ . Calcule  $F'(a)$  y obtenga que el valor de la integral es  $F(a) = \pi \log(1 + |a|)$ .

◇ **12.4.10** Considerando la derivada de la función  $g(x) = \int_0^1 \frac{e^{-x^2(1+t^2)}}{1+t^2} dt$  obtenga el valor de la integral  $\int_0^{+\infty} e^{-t^2} dt = \frac{\sqrt{\pi}}{2}$ .

◇ **12.4.11** Derivando respecto al parámetro  $a \in \mathbb{R}$ , calcule el valor de las integrales:

$$\int_0^{+\infty} e^{-x^2} \cos ax \, dx; \quad \int_0^{+\infty} e^{-(x^2 + a^2/x^2)} dx.$$

◇ **12.4.12** Justifique la igualdad

$$\int_0^{+\infty} \left( \int_0^{+\infty} 2te^{-(a^2+x^2)t^2} \cos rx \, dx \right) dt = \int_0^{+\infty} \left( \int_0^{+\infty} 2te^{-(a^2+x^2)t^2} \cos rx \, dt \right) dx$$

y utilizando los resultados del problema 12.4.11 deduzca que

$$\int_0^{+\infty} \frac{\cos rx}{a^2 + x^2} dx = \frac{\pi}{2a} e^{-a|r|}$$

◇ **12.4.13** Utilice la igualdad obtenida en el problema 12.4.12 para calcular las integrales

$$\int_0^{+\infty} \frac{\operatorname{sen} rx}{x(a^2 + x^2)} dx; \quad \int_0^{+\infty} \frac{x \operatorname{sen} rx}{a^2 + x^2} dx.$$

# Capítulo 13

## Integral curvilínea

*Campos de vectores y formas diferenciales. Integración curvilínea: Independencia del camino y existencia de función potencial. Teorema de Green. Aplicaciones*

Para funciones reales de una variable real, toda función continua  $g : [a, b] \rightarrow \mathbb{R}$  es la derivada de su integral indefinida  $f(x) = \int_a^x g(t)dt$ . Además, si una derivada  $f' : [a, b] \rightarrow \mathbb{R}$  es integrable, la clásica fórmula de Barrow relaciona la integral de  $f'$  en el intervalo  $[a, b]$  con los valores de  $f$  en los extremos del mismo.

En el contexto de las funciones reales de varias variables si una función  $f$  es diferenciable en todos los puntos de su dominio, la alternativa a la función derivada es el campo de formas lineales  $\mathbf{x} \rightarrow df(\mathbf{x})$  (o el campo de vectores  $\mathbf{x} \rightarrow \nabla f(\mathbf{x})$ ). Ahora se plantean problemas análogos a los mencionados en el caso de las funciones de una sola variable: En primer lugar hay que averiguar cuando un campo de formas lineales (o un campo de vectores) es la diferencial (el gradiente) de alguna función real y en ese caso habrá que desarrollar mecanismos para calcularla. Los dos planteamientos, el de los campos de formas lineales y el de los campos de vectores conducen a dos lenguajes distintos para tratar el mismo problema. De momento usaremos el más familiar de los campos de vectores.

La integral curvilínea (o integral de línea) que se estudia con detalle en este capítulo, es la herramienta para calcular, en el caso de que exista, una primitiva de un campo de vectores, es decir una función real cuyo gradiente sea el campo dado. Con ella se obtienen versiones de los teoremas fundamentales del cálculo análogos a los mencionados al principio. La analogía consiste en que ahora la integral curvilínea también relaciona los valores de una función en los extremos de un camino con la integral curvilínea de su gradiente a lo largo del mismo. Como consecuencia de esto, cuando se sabe que un campo continuo de vectores es el gradiente de alguna función ésta se puede calcular mediante la integral curvilínea del campo de vectores a lo largo de un camino de origen fijo y extremo variable (una versión de la fórmula  $f(x) = \int_a^x g(t)dt$  para obtener una primitiva de la función continua  $g$ ).

Por otra parte, el problema de la existencia de primitiva de un campo de vectores no tiene una solución tan directa como en el caso de las funciones de una sola variable. Ahora no se puede asegurar que un campo continuo de vectores sea un gradiente:



Para que un campo de clase  $C^1$  sea un gradiente es necesario que las derivadas parciales de sus componentes estén relacionadas por las condiciones que se precisan en 13.13. Estas condiciones no son suficientes para dominios arbitrarios, pero la integral curvilínea sirve para demostrar que son suficientes para dominios especiales (los estrellados). Este resultado es muy útil en la práctica porque proporciona, para este tipo de dominios, una regla sencilla para saber cuando un campo de vectores de clase  $C^1$  es un gradiente.

La segunda parte del capítulo está dedicada a los aspectos especiales referentes a funciones de dos variables y a campos planos de vectores. En este contexto el resultado sobre los abiertos estrellados que se acaba de mencionar se extiende a la clase más amplia de los abiertos simplemente conexos del plano. En segundo lugar se demuestra una versión elemental del teorema de Green que tiene diversas aplicaciones. Este teorema puede considerarse como una generalización de la clásica regla de Barrow ya que relaciona una integral doble, en la que intervienen las derivadas parciales de las componentes del campo, con la integral del campo a lo largo del borde del dominio de integración.

### 13.1. Formas diferenciales e integral curvilínea

Una *forma diferencial* de grado 1 en un abierto  $\Omega \subset \mathbb{R}^n$  es un campo de formas lineales, es decir una aplicación  $\omega : \Omega \rightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R})$  donde  $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$  es el espacio vectorial de las aplicaciones lineales  $L : \mathbb{R}^n \rightarrow \mathbb{R}$ . Durante todo este capítulo sólo se van a considerar formas diferenciales de grado 1 y por ello cuando se hable de una forma diferencial deberá entenderse siempre que es de grado 1. Si en el espacio vectorial  $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$  se considera la base dual de la base canónica de  $\mathbb{R}^n$ , formada por las proyecciones  $dx_j : (x_1, x_2, \dots, x_n) \rightarrow x_j$ , entonces la forma diferencial  $\omega$  se escribe en la forma canónica

$$\omega(\mathbf{x}) = \sum_{j=1}^n F_j(\mathbf{x}) dx_j$$

donde  $F_j$  son las funciones, definidas en  $\Omega$ , que dan las coordenadas de  $\omega(\mathbf{x})$  respecto a esta base. Para cada  $\mathbf{x} \in \Omega$  y cada  $\mathbf{h} \in \mathbb{R}^n$ , la imagen del vector  $\mathbf{h}$  mediante la aplicación lineal  $\omega(\mathbf{x})$  viene dada por

$$\omega(\mathbf{x})\mathbf{h} = \left( \sum_{j=1}^n F_j(\mathbf{x}) dx_j \right) \mathbf{h} = \sum_{j=1}^n F_j(\mathbf{x}) h_j$$

Si las funciones coordenadas  $F_j$  son continuas (resp. de clase  $C^m$ ) en  $\Omega$  se dice que  $\omega$  es continua (resp. de clase  $C^m$ ). Esta definición es intrínseca, es decir, no depende de la base considerada en  $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$ .

Después de estas definiciones queda establecido el significado de una expresión como la siguiente:  $\sin(x+z)dx + zx^2y^3dy + \sin x dz$ .

Por otra parte, en virtud de la estructura euclídea de  $\mathbb{R}^n$ , a cada forma diferencial de grado 1,  $\omega : \Omega \rightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R})$  se le puede asociar un campo de vectores  $\mathbf{F} : \Omega \rightarrow \mathbb{R}^n$ , asignando a cada  $\mathbf{x} \in \Omega$ , el único vector  $\mathbf{F}(\mathbf{x}) \in \mathbb{R}^n$  que verifica

$$\omega(\mathbf{x})\mathbf{h} = \langle \mathbf{F}(\mathbf{x}) | \mathbf{h} \rangle \quad \text{para todo } \mathbf{h} \in \mathbb{R}^n.$$

Recíprocamente, a un campo de vectores  $\mathbf{F} : \Omega \rightarrow \mathbb{R}^n$  se le asocia la forma diferencial  $\omega$  cuyas funciones coordenadas respecto a la base canónica de  $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$  son las componentes del campo  $\mathbf{F}$ . Por las razones que se verán más adelante, a la forma diferencial  $\sum_{j=1}^n F_j(\mathbf{x})dx_j$  asociada al campo  $\mathbf{F}$  se le suele llamar *trabajo elemental* del campo de vectores  $\mathbf{F}$ .

**Ejemplo 13.1** Si  $f : \Omega \rightarrow \mathbb{R}$  es diferenciable en un abierto  $\Omega \subset \mathbb{R}^n$  entonces  $df$  es una forma diferencial cuya expresión canónica es  $df(\mathbf{x}) = \sum_{j=1}^n D_j f(\mathbf{x})dx_j$ . En particular, si  $f$  es la restricción a  $\Omega$  de una aplicación lineal  $L(\mathbf{x}) = \sum_j a_j x_j$ , entonces  $df$  es constante, y su expresión canónica es  $df(\mathbf{x}) = L = \sum_{j=1}^n a_j dx_j$ , para todo  $\mathbf{x} \in \Omega$ .

En el ejemplo anterior el campo vectorial asociado a la forma diferencial  $df$  es el gradiente,  $\nabla f : \Omega \rightarrow \mathbb{R}^n$ .

En lo que sigue  $\Lambda_1(\Omega)$  designará el conjunto de las formas diferenciales de grado 1 definidas en  $\Omega$ ,  $\Lambda_1^m(\Omega)$  el subconjunto de  $\Lambda_1(\Omega)$  formado por las formas diferenciables de clase  $C^m$  y  $\Lambda_1^0(\Omega)$  el conjunto de las formas diferenciales continuas. Obsérvese que  $\Lambda_1(\Omega)$  (resp.  $\Lambda_1^m(\Omega)$ ) es un espacio vectorial real con las operaciones naturales de suma y producto por un escalar

$$(\omega + \omega')(\mathbf{x}) = \omega(\mathbf{x}) + \omega'(\mathbf{x}); \quad (c\omega)(\mathbf{x}) = c\omega(\mathbf{x})$$

También se puede definir el producto de una función  $f : \Omega \rightarrow \mathbb{R}$  por una forma  $\omega \in \Lambda_1(\Omega)$  del modo natural:  $(f\omega)(\mathbf{x}) = f(\mathbf{x})\omega(\mathbf{x})$ . En particular, multiplicando las funciones  $F_j$  por las formas constantes  $dx_j$  resultan las formas diferenciales  $F_j dx_j$ , cuya suma es  $\omega$ .

Si  $\Lambda_0(\Omega)$  es el conjunto de las funciones diferenciables  $f : \Omega \rightarrow \mathbb{R}$  (se les llama también formas diferenciales de grado 0) entonces la diferencial  $d : f \rightarrow df$ , es una aplicación lineal  $d : \Lambda_0(\Omega) \rightarrow \Lambda_1(\Omega)$  que cumple  $d(fg) = fdg + gdf$ . Cuando  $\Omega$  es conexo su núcleo son las funciones constantes (véase 5.23).

**Definición 13.2** Si  $\omega$  es una forma diferencial en un abierto  $\Omega \subset \mathbb{R}^n$  y existe una función diferenciable  $f : \Omega \rightarrow \mathbb{R}$ , tal que  $\omega = df$  se dice que la forma diferencial  $\omega$  es exacta y que  $f$  es una primitiva de  $\omega$ . Si para cada  $\mathbf{a} \in \Omega$  existe una bola abierta  $B(\mathbf{a}, r) \subset \Omega$  tal que  $\omega|_{B(\mathbf{a}, r)}$  es exacta se dice que  $\omega$  es una forma diferencial cerrada.

Es decir, una forma diferencial  $\omega$  es exacta si está en la imagen de la aplicación lineal  $d : \Lambda_0(\Omega) \rightarrow \Lambda_1(\Omega)$ . Si  $\omega$  es exacta y  $\Omega$  es conexo, la primitiva de  $\omega$  queda unívocamente determinada salvo una constante aditiva.

Por las aplicaciones físicas conviene introducir también la terminología alternativa que corresponde al lenguaje de los campos de vectores: Si la forma diferencial  $\omega$  asociada a un campo de vectores  $\mathbf{F}$  es exacta y  $\omega = df$  entonces el campo es un gradiente,  $\mathbf{F} = \nabla f$ , y se dice que  $f$  es una función potencial del campo  $\mathbf{F}$ . Cuando  $\Omega$  es conexo, la función potencial de un campo de vectores, si existe, no es única, pero dos funciones potenciales del mismo campo difieren en una constante, de modo

que una función potencial concreta se determina especificando su valor en un punto.

La integral curvilínea que se estudia a continuación es la herramienta que permite obtener primitivas de formas diferenciales y caracterizar las formas diferenciales exactas. Con el fin de motivar la definición de la integral curvilínea comenzamos formulándola en términos de campos de vectores. Conceptos físicos importantes como el trabajo realizado por una fuerza al mover una partícula material a lo largo de una curva o la circulación de un campo de velocidades a lo largo de una trayectoria se definen mediante integrales curvilíneas.

Un campo de vectores en un abierto  $\Omega \subset \mathbb{R}^n$  es una aplicación continua  $\mathbf{F} : \Omega \rightarrow \mathbb{R}^n$  que se suele representar dibujando, para cada  $\mathbf{x} \in \Omega$ , un vector  $\mathbf{F}(\mathbf{x})$  aplicado en el punto  $\mathbf{x}$ . Para simplificar la cuestión de motivar la definición suponemos que  $\gamma$  es un camino en  $\Omega$  de clase  $C^1$  con derivada no nula en todo punto. Su abscisa curvilínea  $s = v(t)$ , es una función invertible de clase  $C^1$  y con la sustitución  $t = v^{-1}(s)$  se obtiene una representación paramétrica equivalente  $\tilde{\gamma}(s) = \gamma(v^{-1}(s))$  cuyo parámetro es el arco. Si  $L = \text{Long}(\gamma)$ , para cada  $s \in [0, L]$  la derivada  $\tilde{\gamma}'(s)$  es un vector tangente unitario y la componente del vector  $\mathbf{F}(\tilde{\gamma}(s))$  según este vector unitario viene dada por el producto escalar  $\langle \mathbf{F}(\tilde{\gamma}(s)) \mid \tilde{\gamma}'(s) \rangle$ . La integral de este producto escalar, sobre  $[0, L]$ , representa el trabajo realizado cuando la partícula material se mueve a lo largo de la trayectoria orientada  $\gamma$ , sometida al campo de fuerzas  $\mathbf{F}$ . Si se efectúa el cambio de variable  $s = v(t)$ , teniendo en cuenta que  $\gamma'(t) = \tilde{\gamma}'(s)v'(t)$ , resulta

$$\int_0^L \langle \mathbf{F}(\tilde{\gamma}(s)) \mid \tilde{\gamma}'(s) \rangle ds = \int_a^b \langle \mathbf{F}(\gamma(t)) \mid \gamma'(t) \rangle dt$$

Esta interpretación es la que motiva la siguiente definición

**Definición 13.3** *Dado un campo continuo  $\mathbf{F} : \Omega \rightarrow \mathbb{R}^n$ ,  $\mathbf{F} = (F_1, F_2, \dots, F_n)$  en un abierto  $\Omega \subset \mathbb{R}^n$ , y un camino regular a trozos  $\gamma : [a, b] \rightarrow \Omega$ ,  $\gamma = (\gamma_1, \gamma_2, \dots, \gamma_n)$ , la integral curvilínea de  $\mathbf{F}$  a lo largo de  $\gamma$  se define así:*

$$\int_{\gamma} \mathbf{F} = \int_a^b \langle \mathbf{F}(\gamma(t)) \mid \gamma'(t) \rangle dt = \sum_{j=1}^n \int_a^b F_j(\gamma(t)) \gamma'_j(t) dt$$

Obsérvese que  $\gamma$  es derivable en  $[a, b]$  salvo en un conjunto finito de puntos  $x_1 < x_2 < \dots < x_{n-1}$  del intervalo abierto  $(a, b)$ , por lo que la función

$$f(t) = \sum_{j=1}^n F_j(\gamma(t)) \gamma'_j(t)$$

está definida en  $[a, b]$  excepto en este conjunto finito. Sin embargo en todos los puntos  $x_1, x_2, \dots, x_{n-1}$  existen las derivadas laterales de  $\gamma$ , y  $f$  coincide en cada intervalo abierto  $(x_{i-1}, x_i)$  con la restricción de una función continua. Si se define  $f$  en los puntos  $x_1, x_2, \dots, x_{n-1}$ , asignándole valores arbitrarios, se obtiene una función integrable Riemann cuya integral no depende de los valores asignados a  $f$  en estos

puntos (recuérdese que si una función integrable se modifican en un conjunto finito de puntos se obtiene otra función integrable con la misma integral).

A la integral curvilínea también se le suele llamar *integral de línea* o *integral de contorno* y para ella también se suelen utilizar las notaciones

$$\int \langle \mathbf{F} \mid d\gamma \rangle = \int_{\gamma} \sum_{j=1}^n F_j dx_j = \sum_{j=1}^n \int F_j d\gamma_j$$

En el lenguaje de las formas diferenciales, la definición se formula así

**Definición 13.4** Sea  $\omega(\mathbf{x}) = \sum_{j=1}^n F_j(\mathbf{x})dx_j$  una forma diferencial de grado 1 definida y continua en un abierto  $\Omega \subset \mathbb{R}^n$ . Si  $\gamma : [a, b] \rightarrow \Omega$  es un camino regular a trozos, la integral curvilínea de  $\omega$  a lo largo de  $\gamma$  se define como la integral curvilínea del campo de vectores asociado:

$$\int_{\gamma} \omega = \int_a^b \omega(\gamma(t))\gamma'(t) dt = \int_a^b \sum_{j=1}^n F_j(\gamma(t))\gamma'_j(t) dt = \sum_{j=1}^n \int_a^b F_j(\gamma(t))\gamma'_j(t) dt$$

donde  $\gamma_j$ ,  $1 \leq j \leq n$ , son las componentes de  $\gamma$ .

Las consideraciones preliminares que han motivado la definición, la integral curvilínea de un campo de vectores  $\mathbf{F}$  a lo largo de un camino regular a trozos  $\gamma$  se puede interpretar como la integral respecto al arco de la componente de  $\mathbf{F}$  según la dirección de la tangente al camino. Por ello importantes conceptos físicos, como el trabajo realizado por una fuerza al mover una partícula material a lo largo de una curva o la circulación de un campo de velocidades a lo largo de una trayectoria se expresan mediante integrales curvilíneas de campos de vectores. Esta interpretación física es la que motiva el nombre de *trabajo elemental* del campo  $\mathbf{F}$  que se suele utilizar para designar la forma diferencial  $\sum_{j=1}^n F_j(\mathbf{x})dx_j$ .

Aunque para interpretaciones físicas conviene considerar las integrales curvilíneas en términos de campos de vectores, sin embargo, desde el punto de vista algorítmico del cálculo tienen ventaja las integrales curvilíneas expresadas en términos de formas diferenciales. Con ellas se pone de manifiesto la utilidad de la expresión canónica de una forma diferencial y la ventaja de la notación empleada para la base canónica de  $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$ : Para calcular la integral de una forma diferencial  $\omega$  sobre un camino  $\mathbf{x}(t) = (x_1(t), \dots, x_n(t))$  basta calcular la integral definida de la función que se obtiene sustituyendo formalmente,  $x_j = x_j(t)$ ,  $dx_j = x'_j(t) dt$  en la expresión canónica de la forma diferencial

$$\sum_{j=1}^n F_j(x_1, x_2, \dots, x_n)dx_j$$

**Orientación de un arco de curva regular a trozos.** Para caminos regulares a trozos, que son los que intervienen en la integral curvilínea, conviene considerar la siguiente relación de equivalencia: Dos caminos regulares a trozos  $\mathbf{f}$ ,  $\mathbf{g}$  son *equivalentes como caminos regulares a trozos* cuando se pueden expresar en la forma

$$\mathbf{f} = \mathbf{f}_1 \vee \mathbf{f}_2 \vee \dots \vee \mathbf{f}_m, \quad \mathbf{g} = \mathbf{g}_1 \vee \mathbf{g}_2 \vee \dots \vee \mathbf{g}_m$$

donde, para cada  $k \in \{1, \dots, m\}$  los caminos  $\mathbf{f}_k$  y  $\mathbf{g}_k$  son  $C^1$  equivalentes. Es fácil ver que, en este caso, un camino se obtiene efectuando en el otro un cambio de parámetro estrictamente monótono regular a trozos. Cuando sea creciente diremos que los dos caminos tienen la misma orientación y cuando sea decreciente que tienen orientaciones opuestas. En el primer caso los dos caminos tienen el mismo origen y el mismo extremo, pero en el segundo caso el origen de un camino coincide con el extremo del otro y los dos caminos recorren el mismo trayecto, pero en sentidos opuestos. La noción de caminos regulares a trozos equivalentes es una relación de equivalencia y cada clase de equivalencia se dice que es un *arco de curva regular a trozos*. Análogamente se define la noción de *arco de curva orientado regular a trozos* considerando como relación de equivalencia la de caminos regulares a trozos equivalentes con la misma orientación. Un arco de curva regular a trozos queda orientado cuando se elige una de sus representaciones paramétricas regulares a trozos.

Las siguientes propiedades de la integral curvilínea son consecuencia inmediata de la definición y de las propiedades básicas de la integral de Riemann:

**Proposición 13.5** Sean  $\mathbf{F}_1, \mathbf{F}_2$  campos de vectores definidos y continuos en un abierto  $\Omega \subset \mathbb{R}^n$  y  $\gamma, \gamma_1, \gamma_2$  caminos regulares a trozos en  $\Omega$ . Se verifica:

$$i) \int_{\gamma} (\mathbf{F}_1 + \mathbf{F}_2) = \int_{\gamma} \mathbf{F}_1 + \int_{\gamma} \mathbf{F}_2$$

$$ii) \int_{\gamma} \mathbf{F} = \int_{\gamma_1} \mathbf{F} + \int_{\gamma_2} \mathbf{F} \text{ si } \gamma = \gamma_1 \vee \gamma_2$$

$$iii) \int_{\sim\gamma} \mathbf{F} = - \int_{\gamma} \mathbf{F}$$

iv) Si  $\gamma_1, \gamma_2$  son caminos regulares a trozos equivalentes entonces  $\int_{\gamma_1} \mathbf{F} = \epsilon \int_{\gamma_2} \mathbf{F}$  donde  $\epsilon = 1$  (resp.  $\epsilon = -1$ ) si los caminos tienen la misma orientación (resp. orientaciones opuestas).

La propiedad iv) de la proposición 13.5 permite definir la integral curvilínea de un campo de vectores continuo sobre un arco de curva orientado regular a trozos, a través de una cualquiera de sus representaciones paramétricas admisibles. Es decir, la integral curvilínea es realmente una noción asociada al arco de curva orientado, que cambia de signo cuando se cambia su orientación.

Aunque no se acostumbra a hacer énfasis en este hecho, sin embargo se hace uso frecuente del mismo sin advertirlo explícitamente. Así por ejemplo, como cualquier camino regular a trozos es equivalente a otro, con la misma orientación, cuyo dominio es un intervalo prefijado, frecuentemente se asume que el camino que interviene en una integral curvilínea está definido en el intervalo que convenga en cada caso. Esto es lo que se hace cuando se considera la yuxtaposición de dos caminos, que a priori no están definidos en intervalos contiguos, siempre que el extremo del primero coincida con el origen del segundo (para definir explícitamente la yuxtaposición sería preciso reparametrizar los caminos en intervalos contiguos).

**Proposición 13.6** Sea  $\mathbf{F}(\mathbf{x}) = (F_1(\mathbf{x}), \dots, F_n(\mathbf{x}))$  un campo vectorial continuo en un abierto  $\Omega \subset \mathbb{R}^n$  y  $\gamma$  un camino regular a trozos en  $\Omega$ . Entonces

$$\left| \int_{\gamma} \mathbf{F} \right| \leq M \text{Long}(\gamma) \quad \text{con} \quad M = \sup\{\|\mathbf{F}(\mathbf{x})\|_2 : \mathbf{x} \in \gamma([a, b])\}$$

DEM: En virtud de la desigualdad de Cauchy, si  $\mathbf{x} \in \gamma([a, b])$

$$|\langle \mathbf{F}(\mathbf{x}) \mid \mathbf{h} \rangle| \leq \|\mathbf{F}(\mathbf{x})\|_2 \|\mathbf{h}\|_2 \leq M \|\mathbf{h}\|_2$$

luego

$$\left| \int_{\gamma} \mathbf{F} \right| \leq \int_a^b |\langle \mathbf{F}(\gamma(t)) \mid \gamma'(t) \rangle| dt \leq M \int_a^b \|\gamma'(t)\|_2 dt = M \text{Long}(\gamma)$$

■

Se deja al cuidado del lector el enunciado de los resultados anteriores en términos de formas diferenciales. En lo que sigue, por comodidad de notación, consideraremos preferentemente integrales curvilíneas de formas diferenciales.

### Independencia del camino.

**Definición 13.7** Si  $\omega$  es una forma diferencial de grado 1 definida y continua en un abierto  $\Omega \subset \mathbb{R}^n$ , se dice que la integral curvilínea  $\int_{\gamma} \omega$  no depende del camino en  $\Omega$  si para cada par de caminos regulares a trozos  $\gamma_1, \gamma_2$  en  $\Omega$ , con el mismo origen y el mismo extremo, se verifica  $\int_{\gamma_1} \omega = \int_{\gamma_2} \omega$ .

**Proposición 13.8** Si  $\omega$  es una forma diferencial de grado 1 continua en un abierto  $\Omega \subset \mathbb{R}^n$  son equivalentes

- La integral curvilínea  $\int_{\gamma} \omega$  no depende del camino en  $\Omega$ .
- $\int_{\gamma} \omega = 0$  para cada camino  $\gamma$  en  $\Omega$ , cerrado y regular a trozos.

DEM: a)  $\Rightarrow$  b) es inmediato pues todo camino cerrado tiene los mismos extremos que un camino constante.

b)  $\Rightarrow$  a) Si  $\gamma_1, \gamma_2$  son caminos regulares a trozos en  $\Omega$  con los mismos extremos entonces  $\gamma = \gamma_1 \vee (\sim \gamma_2)$  es un camino cerrado en  $\Omega$  y por hipótesis

$$0 = \int_{\gamma} \omega = \int_{\gamma_1} \omega - \int_{\gamma_2} \omega$$

■

**Ejemplo 13.9** La forma diferencial  $\omega(x, y) = ydx + 2xdy$ , está definida en todo el plano, y su integral curvilínea  $\int_{\gamma} \omega$  depende del camino.

DEM: Obsérvese  $\gamma_1(t) = (t, t)$  y  $\gamma_2(t) = (t, t^2)$ ,  $t \in [0, 1]$ , son caminos en  $\mathbb{R}^2$  con el mismo origen y el mismo extremo, que proporcionan distinta integral curvilínea. ■

La siguiente proposición que da una condición suficiente para que la integral de línea  $\int_{\gamma} \omega$  sea independiente del camino en  $\Omega$ , proporciona el procedimiento estándar para conseguir una primitiva de una forma diferencial exacta.

**Proposición 13.10** *Sea  $\omega$  una forma diferencial de grado 1 continua en un abierto  $\Omega \subset \mathbb{R}^n$ . Si  $\omega$  es exacta y  $f$  es una primitiva de  $\omega$  entonces para todo camino regular a trozos  $\gamma$  en  $\Omega$  de origen  $\mathbf{x}$  y extremo  $\mathbf{y}$ , se verifica*

$$\int_{\gamma} \omega = f(\mathbf{y}) - f(\mathbf{x})$$

DEM: Sea  $\omega = df$  donde  $f : \Omega \rightarrow \mathbb{R}$  es diferenciable. Consideremos una subdivisión  $a = x_0 < x_1 < x_2 < \dots < x_n = b$  de  $[a, b]$  tal que cada  $\gamma|_{[x_{j-1}, x_j]}$  es de clase  $C^1$ . Entonces

$$\int_{\gamma} \omega = \int_{\gamma} df = \sum_{j=1}^n \int_{x_{j-1}}^{x_j} df(\gamma(t))\gamma'(t) dt = \sum_{j=1}^n \int_{x_{j-1}}^{x_j} (f \circ \gamma)'(t) dt$$

Utilizando el teorema fundamental del cálculo y recordando que  $\gamma(a) = \mathbf{x}$ ,  $\gamma(b) = \mathbf{y}$  se obtiene

$$\int_{\gamma} \omega = \sum_{j=1}^n [f(\gamma(x_j)) - f(\gamma(x_{j-1}))] = f(\gamma(b)) - f(\gamma(a)) = f(\mathbf{y}) - f(\mathbf{x})$$

■

Esta última proposición pone de manifiesto que cuando se sabe que una forma diferencial  $\omega$  es exacta,  $\omega = df$ , la integral curvilínea es la herramienta adecuada para determinar (salvo una constante) la primitiva  $f$ : Si  $\Omega$  es conexo, se obtiene una primitiva  $f$  de  $\omega$  fijando un punto  $\mathbf{a} \in \Omega$  y definiendo  $f(\mathbf{x}) = \int_{\gamma_{\mathbf{x}}} \omega$  donde  $\gamma_{\mathbf{x}}$  es cualquier camino en  $\Omega$ , regular a trozos, con origen fijo en  $\mathbf{a}$  y extremo variable  $\mathbf{x} \in \Omega$ . (Si  $\Omega$  no es conexo se obtiene la primitiva procediendo como se acaba de indicar en cada una de sus componentes conexas).

En el lenguaje de los campos de vectores, e interpretando la integral curvilínea como trabajo, la proposición anterior se traduce en el principio físico que dice que si un campo de fuerzas  $\mathbf{F}$  admite función potencial  $f$ , entonces el trabajo realizado cuando una partícula recorre la trayectoria  $\gamma$  sometida al campo de fuerzas  $\mathbf{F}$  es igual a la diferencia del potencial del campo entre los extremos de la trayectoria. En este caso el trabajo realizado no depende de la trayectoria que ha seguido la partícula; sólo depende de la posición final y de la posición inicial de la misma. Por esta razón se llaman conservativos a los campos de fuerzas cuya integral curvilínea no depende del camino, es decir el trabajo que realizan a lo largo de un camino sólo depende de los extremos del camino. En particular, no se realiza trabajo al cuando la partícula recorre una trayectoria cerrada. Con el siguiente teorema quedan caracterizados los campos conservativos como aquellos que tienen función potencial.

**Teorema 13.11** Si  $\omega$  es una forma diferencial continua, de grado 1, definida en un abierto  $\Omega \subset \mathbb{R}^n$  son equivalentes

a)  $\omega$  es exacta.

b)  $\int_{\gamma} \omega = 0$  para cada camino cerrado y regular a trozos  $\gamma$  en  $\Omega$ .

DEM: a)  $\Rightarrow$  b) Es consecuencia inmediata de la proposición 13.10

b)  $\Rightarrow$  a) No es restrictivo suponer que  $\Omega$  es conexo y por lo tanto conexo por poligonales, (si no es así se aplica el siguiente razonamiento en cada componente conexa). Fijado  $\mathbf{a} \in \Omega$ , para cada  $\mathbf{x} \in \Omega$  existe un camino regular a trozos  $\gamma_{\mathbf{x}} : [0, 1] \rightarrow \Omega$  de origen  $\mathbf{a} = \gamma_{\mathbf{x}}(0)$  y extremo  $\mathbf{x} = \gamma_{\mathbf{x}}(1)$  y se define

$$f(\mathbf{x}) = \int_{\gamma_{\mathbf{x}}} \omega.$$

En virtud de la proposición 13.8 la definición de  $f(\mathbf{x})$  sólo depende del extremo  $\mathbf{x}$  del camino. El objetivo es demostrar que  $f$  es diferenciable en  $\Omega$  con  $df(\mathbf{x}) = \omega(\mathbf{x})$  para todo  $\mathbf{x} \in \Omega$ . Es decir, hay que demostrar que  $\epsilon(\mathbf{h}) = f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) - \omega(\mathbf{x})\mathbf{h}$  verifica

$$\lim_{\mathbf{h} \rightarrow 0} \epsilon(\mathbf{h}) / \|\mathbf{h}\|_2 = 0$$

Fijado un punto  $\mathbf{x} \in \Omega$  comenzamos eligiendo  $\rho > 0$  tal que  $B(\mathbf{x}, \rho) \subset \Omega$ . De esta forma, si  $\|\mathbf{h}\| < \rho$ , podemos asegurar que el segmento  $\sigma(t) = \mathbf{x} + t\mathbf{h}$ ,  $t \in [0, 1]$  está contenido en  $\Omega$  y con ello que el camino regular a trozos  $\gamma_{\mathbf{x}} \vee \sigma$ , de origen  $\mathbf{a}$  y extremo  $\mathbf{x} + \mathbf{h}$ , está contenido en  $\Omega$ , de modo que podemos utilizarlo para calcular  $f(\mathbf{x} + \mathbf{h})$ . Usando las propiedades de la integral curvilínea

$$f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) = \int_{\gamma_{\mathbf{x}} \vee \sigma} \omega - \int_{\gamma_{\mathbf{x}}} \omega = \int_{\sigma} \omega$$

Si  $\mathbf{F}$  es el campo vectorial asociado a  $\omega$ , como  $\sigma'(t) = \mathbf{h}$ , resulta

$$\epsilon(\mathbf{h}) = \int_{\sigma} \omega - \omega(\mathbf{x})\mathbf{h} = \int_0^1 \langle \mathbf{F}(\sigma(t)) - \mathbf{F}(\mathbf{x}) \mid \mathbf{h} \rangle dt$$

Como  $\mathbf{F}$  es continuo en  $\mathbf{x} \in \Omega$  existe  $B(\mathbf{x}, r) \subset B(\mathbf{x}, \rho)$  tal que  $\|\mathbf{F}(\mathbf{y}) - \mathbf{F}(\mathbf{x})\|_2 < \epsilon$  para todo  $\mathbf{y} \in B(\mathbf{x}, r)$ . Entonces, si  $\|\mathbf{h}\|_2 < r$ , en virtud de la desigualdad de Cauchy,

$$|\langle \mathbf{F}(\sigma(t)) - \mathbf{F}(\mathbf{x}) \mid \mathbf{h} \rangle| \leq \|\mathbf{F}(\sigma(t)) - \mathbf{F}(\mathbf{x})\|_2 \|\mathbf{h}\|_2 \leq \epsilon \|\mathbf{h}\|_2$$

y se obtiene

$$|\epsilon(\mathbf{h})| \leq \int_0^1 |\langle \mathbf{F}(\sigma(t)) - \mathbf{F}(\mathbf{x}) \mid \mathbf{h} \rangle| dt \leq \epsilon \|\mathbf{h}\|_2$$

es decir

$$\|\mathbf{h}\|_2 < r \Rightarrow |\epsilon(\mathbf{h})| < \epsilon \|\mathbf{h}\|_2$$

y esto termina la prueba. ■



Si  $\omega(\mathbf{x}) = df(\mathbf{x}) = \sum_{j=1}^n D_j f(\mathbf{x}) dx_j$  es una forma diferencial exacta de clase  $C^1$ , sus funciones coordenadas  $F_j(\mathbf{x}) = D_j f(\mathbf{x})$  son de clase  $C^1$ , lo que significa que  $f$  es de clase  $C^2$  y aplicando el teorema de Young 6.4 se obtiene que para todo  $i, j \in \{1, 2, \dots, n\}$  y todo  $\mathbf{x} \in \Omega$  se cumple  $D_i F_j(\mathbf{x}) = D_{ij} f(\mathbf{x}) = D_{ji} f(\mathbf{x}) = D_j F_i(\mathbf{x})$ . El recíproco se cumple cuando el abierto  $\Omega$  es estrellado:

**Definición 13.12** Un abierto  $\Omega$  de  $\mathbb{R}^n$  se dice que es estrellado si hay un punto  $\mathbf{a} \in \Omega$  tal que para cada  $\mathbf{x} \in \Omega$  el segmento  $[\mathbf{a}, \mathbf{x}]$  está contenido en  $\Omega$ .

**Teorema 13.13** Si  $\omega(\mathbf{x}) = \sum_{j=1}^n F_j(\mathbf{x}) dx_j$  es una forma diferencial de clase  $C^1$  definida en un abierto estrellado  $\Omega \subset \mathbb{R}^n$ , son equivalentes:

i)  $\omega$  es exacta.

ii)  $D_i F_j(\mathbf{x}) = D_j F_i(\mathbf{x})$  para cada  $i, j \in \{1, 2, \dots, n\}$  y cada  $\mathbf{x} \in \Omega$ .

DEM: Ya hemos visto que i)  $\Rightarrow$  ii) aunque  $\Omega$  no sea estrellado,

ii)  $\Rightarrow$  i) Supongamos, para simplificar la escritura, que  $\Omega$  es estrellado respecto al origen. Para cada  $\mathbf{x} \in \Omega$  sea  $\sigma_{\mathbf{x}}(t) = t\mathbf{x}$ ,  $t \in [0, 1]$ , el segmento de origen 0 y extremo  $\mathbf{x}$ , que por hipótesis está contenido en  $\Omega$ , lo que permite definir la función

$$f(\mathbf{x}) = \int_{\sigma_{\mathbf{x}}} \omega = \int_0^1 \langle \mathbf{F}(t\mathbf{x}) \mid \mathbf{x} \rangle dt = \int_0^1 h(\mathbf{x}, t) dt$$

donde  $\mathbf{F}$  es el campo vectorial asociado a  $\omega$ . La función

$$h(\mathbf{x}, t) = \sum_{j=1}^n x_j F_j(t\mathbf{x})$$

posee derivadas parciales continuas respecto a las variables  $x_1, x_2, \dots, x_n$

$$\frac{\partial h}{\partial x_k}(\mathbf{x}, t) = F_k(t\mathbf{x}) + \sum_{j=1}^n t x_j D_k F_j(t\mathbf{x})$$

En virtud de la hipótesis ii),  $D_k F_j = D_j F_k$  luego

$$\frac{\partial h}{\partial x_k}(\mathbf{x}, t) = F_k(t\mathbf{x}) + \sum_{j=1}^n t x_j D_j F_k(t\mathbf{x}) = \frac{d}{dt}(t F_k(t\mathbf{x}))$$

Utilizando 12.10 se concluye que  $f$  posee derivadas parciales continuas en  $\Omega$  que se obtienen derivando bajo la integral

$$D_k f(\mathbf{x}) = \int_0^1 \frac{\partial h}{\partial x_k}(\mathbf{x}, t) dt = \int_0^1 \frac{d}{dt}(t F_k(t\mathbf{x})) dt = F_k(\mathbf{x}), \quad 1 \leq k \leq n$$

luego  $f$  es diferenciable en  $\Omega$  y  $df = \omega$  ■

Como las bolas son conjuntos estrellados, aplicando el teorema anterior sobre cada bola  $B(\mathbf{a}, r) \subset \Omega$  se caracterizan las formas diferenciales cerradas de clase  $C^1$

**Corolario 13.14** Si  $\omega(\mathbf{x}) = \sum_{j=1}^n F_j(\mathbf{x})dx_j$  es una forma diferencial de clase  $C^1$  definida en un abierto  $\Omega \subset \mathbb{R}^n$ , son equivalentes:

- i)  $\omega$  es cerrada.
- ii)  $D_i F_j(\mathbf{x}) = D_j F_i(\mathbf{x})$  para cada  $i, j \in \{1, 2, \dots, n\}$  y cada  $\mathbf{x} \in \Omega$ .

El siguiente ejemplo pone de manifiesto que el resultado expuesto en el teorema 13.13 no se cumple cuando  $\Omega$  es un abierto arbitrario.

### Ejemplo 13.15

En virtud del corolario 13.14 la forma diferencial

$$\omega(x, y) = \frac{-y}{x^2 + y^2} dx + \frac{x}{x^2 + y^2} dy$$

es cerrada en el abierto  $\Omega = \mathbb{R}^2 \setminus \{0\}$ . Sin embargo no es exacta porque  $C(t) = (\cos t, \sin t)$ ,  $0 \leq t \leq 2\pi$  es un camino cerrado en  $\Omega$  y  $\int_C \omega = 2\pi \neq 0$ . ■

### Ejemplo 13.16

En virtud del teorema 13.13 la forma diferencial  $\omega(x, y) = 2xydx + (x^2 + 2y)dy$ , definida en  $\mathbb{R}^2$ , es exacta. Si  $f : \mathbb{R} \rightarrow \mathbb{R}$  es una primitiva de  $\omega$  debe cumplir  $D_1 f(x, y) = 2xy$ ,  $D_2 f(x, y) = x^2 + 2y$ . De la primera condición se sigue que  $f$  es de la forma  $f(x, y) = x^2 y + \varphi(y)$  donde  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$  es una función derivable, y utilizando la segunda condición se llega a que  $x^2 + 2y = x^2 + \varphi'(y)$  luego  $\varphi(y) = y^2 + c$ . Se obtiene así la primitiva  $f(x, y) = x^2 y + y^2 + c$ .

## 13.2. Formas diferenciales en el plano

En esta sección se consideran aspectos particulares de las formas diferenciales de dos variables que escribiremos en la forma  $\omega(x, y) = P(x, y)dx + Q(x, y)dy$ .

Mediante el corolario 13.14 han quedado caracterizadas las formas diferenciales cerrada de clase  $C^1$  como aquellas que verifican la condición  $D_2 P = D_1 Q$ . Cuando  $n = 2$  las formas diferenciales cerradas también se pueden caracterizar mediante una condición de distinta naturaleza, que tiene la ventaja de aplicarse a formas diferenciales que sólo se suponen continuas (véase 13.17). También se estudia en esta sección el problema general de la independencia del camino para el caso de las formas diferenciales de dos variables.

En lo que sigue cuando se hable de rectángulos en el plano se supondrá que son cerrados de lados paralelos a los ejes, es decir, de la forma

$$R = \{(x, y) : x \in [a, b], y \in [c, d]\}$$

En ese caso  $\partial R$  denota al camino poligonal cerrado que recorre la frontera en el sentido  $(a, c) \rightarrow (b, c) \rightarrow (b, d) \rightarrow (a, d) \rightarrow (a, c)$ .

Diremos que  $\Omega$  es un *abierto especial* si existe  $(a, b) \in \Omega$  tal que para cada  $(x, y) \in \Omega$  el rectángulo  $R$  de vértices opuestos  $(a, b)$ ,  $(x, y)$  está contenido en  $\Omega$ . (Nótese que  $R$  puede degenerar en un segmento si  $x = a$  o si  $y = b$ ). Son abiertos especiales los discos, los semiplanos abiertos determinados por rectas paralelas a uno de los ejes y los cuadrantes abiertos. Para los abiertos especiales se verifica

**Teorema 13.17** *Si  $\omega(x, y) = P(x, y)dx + Q(x, y)dy$  es una forma diferencial definida y continua en un abierto especial  $\Omega \subset \mathbb{R}^2$ , son equivalentes*

a)  $\omega$  es exacta.

b)  $\int_{\partial R} \omega = 0$  para cada rectángulo  $R \subset \Omega$ .

DEM: a)  $\Rightarrow$  b) está probado en 13.11.

b)  $\Rightarrow$  a) Sea  $(a, b) \in \Omega$  un punto tal que para todo  $(x, y) \in \Omega$  el rectángulo  $R$  determinado por  $(a, b)$ ,  $(x, y)$  está contenido en  $\Omega$ . Su borde orientado es la yuxtaposición de cuatro segmentos

$$\partial R = \sigma_1 \vee \sigma_2 \vee \sigma_3 \vee \sigma_4$$

donde  $\sigma_j$  son los segmentos que se indican a continuación

$\sigma_1$  es el segmento horizontal de origen  $(a, b)$  y extremo  $(x, b)$ .

$\sigma_2$  es el segmento vertical de origen  $(x, b)$  y extremo  $(x, y)$ .

$\sigma_3$  es el segmento horizontal de origen  $(x, y)$  y extremo  $(a, y)$ .

$\sigma_4$  es el segmento vertical de origen  $(a, y)$  y extremo  $(a, b)$ .

Consideremos los dos caminos  $\gamma_1 = \sigma_1 \vee \sigma_2$ ,  $\gamma_2 = (\sim \sigma_3) \vee (\sim \sigma_4)$  de origen  $(a, b)$  y extremo  $(x, y)$ . Como  $\partial R = \gamma_1 \vee (\sim \gamma_2)$ , en virtud de la hipótesis b) se cumple

$$\int_{\gamma_1} \omega - \int_{\gamma_2} \omega = \int_{\partial R} \omega = 0$$

Para cada  $(x, y) \in \Omega$  sea  $f(x, y) = \int_{\gamma_{(x,y)}} \omega$  donde  $\gamma_{(x,y)}$  es uno de los dos caminos  $\gamma_1$ ,  $\gamma_2$  que se acaban de considerar. Bastará demostrar que para todo  $(x, y) \in \Omega$  se cumple  $D_1 f(x, y) = P(x, y)$  y  $D_2 f(x, y) = Q(x, y)$  pues de aquí se sigue, usando la continuidad de  $P$  y  $Q$ , que  $f$  es diferenciable en  $\Omega$  con  $df = \omega$ .

Fijado  $(x, y) \in \Omega$ , sea  $r > 0$  tal que  $\overline{B((x, y), r)} \subset \Omega$ . Entonces si  $|h| \leq r$  podemos asegurar que el segmento  $\sigma_h(t) = (x + th, y)$ ,  $0 \leq t \leq 1$ , está contenido en  $\overline{B((x, y), r)}$ . Si usamos el camino  $\gamma_2$  para calcular  $f(x, y)$  y el camino  $\gamma_2 \vee \sigma_h$  para calcular  $f(x + h, y)$  obtenemos la siguiente expresión del cociente incremental

$$\Delta(h) = \frac{f(x + h, y) - f(x, y)}{h} = \frac{1}{h} \int_{\sigma_h} \omega = \frac{1}{h} \int_0^1 P(x + th, y) h dt$$

En virtud del teorema 12.9 la función  $\Delta(h) = \int_0^1 P(x + th, y) dt$  es continua en  $[-r, r]$  y se sigue que  $\lim_{h \rightarrow 0} \Delta(h) = \Delta(0) = P(x, y)$ .

Con un razonamiento análogo se demuestra que  $D_2f = Q$ . En este caso, hay que considerar cociente incremental

$$\Delta(h) = \frac{f(x, y+h) - f(x, y)}{h}$$

y para calcular  $f(x, y+h)$  (resp.  $f(x, y)$ ) debemos usar el camino  $\gamma_1 \vee \sigma_h$  (resp.  $\gamma_1$ ) con  $\sigma_h(t) = (x, y+th)$ ,  $0 \leq t \leq 1$ . ■

NOTA: En las condiciones del teorema anterior, haciendo explícitas las dos integrales de línea que se pueden usar para obtener la primitiva  $f$  se llega a las siguientes fórmulas para una primitiva de  $P(x, y)dx + Q(x, y)dy$  en  $\Omega$ .

$$f(x, y) = \int_{y_0}^y Q(x_0, t) dt + \int_{x_0}^x P(t, y) dt = \int_{x_0}^x P(t, y_0) dt + \int_{y_0}^y Q(x, t) dt$$

En la definición de forma diferencial cerrada sólo se exige que fijado un punto  $\mathbf{a} \in \Omega$  haya una bola suficientemente pequeña  $B(\mathbf{a}, r) \subset \Omega$  donde la forma diferencial tenga primitiva. Cuando  $n = 2$  ocurre lo mismo cuando la bola se toma todo lo grande que se pueda. Esto es consecuencia de la siguiente proposición, según la cual en los abiertos especiales toda forma diferencial continua y cerrada es exacta.

**Proposición 13.18** *Si  $\omega$  es una forma diferencial de grado 1 definida y continua en un abierto  $\Omega \subset \mathbb{R}^2$ , son equivalentes:*

- $\omega$  es cerrada.
- $\int_{\partial R} \omega = 0$  para cada rectángulo  $R \subset \Omega$ .
- $\omega$  posee primitiva en cada abierto especial  $V \subset \Omega$  (y en particular en cada bola  $B(\mathbf{a}, r) \subset \Omega$ )

DEM: b)  $\Rightarrow$  c) está probado en 13.17 y c)  $\Rightarrow$  a) es evidente.

a)  $\Rightarrow$  b) Se puede probar por reducción al absurdo, suponiendo que  $\int_{\partial R} \omega \neq 0$  para algún rectángulo cerrado  $R \subset \Omega$ . Sea  $\Delta = \text{diámetro}(R)$ . Trazando los segmentos que unen los puntos medios de los lados opuestos se descompone  $R$  en cuatro rectángulos congruentes  $R^1, R^2, R^3, R^4$ . Teniendo en cuenta las cancelaciones de la integral curvilínea sobre segmentos opuestos resulta:

$$0 \neq \int_{\partial R} \omega = \sum_{j=1}^4 \int_{\partial R_j} \omega$$

luego  $\int_{\partial R^i} \omega \neq 0$  para algún  $i \in \{1, 2, 3, 4\}$ . Si  $R_1 = R^i$  se tiene,  $\text{diámetro}(R_1) = \Delta/2$ . Repitiendo con  $R_1$  el razonamiento que se acaba de hacer con  $R$  se obtiene un rectángulo cerrado  $R_2 \subset R_1$  tal que  $\text{diámetro}(R_2) = \Delta/2$  y  $\int_{\partial R_2} \omega \neq 0$ . De modo recurrente se obtiene una sucesión decreciente de rectángulos cerrados  $R_n$  tal que para todo  $n \in \mathbb{N}$  se cumple

$$\text{diámetro}(R_n) = \Delta/2^n \quad \text{y} \quad \int_{\partial R_n} \omega \neq 0$$

La intersección de la sucesión decreciente de compactos  $R_n$  no es vacía (de hecho se reduce a un punto). Si  $(x_0, y_0) \in \bigcap_{n \in \mathbb{N}} R_n$ , por hipótesis  $\omega$  tiene primitiva en alguna bola  $B = B((x_0, y_0), r) \subset \Omega$ . Sin embargo para  $n$  suficientemente grande es  $R_n \subset B$  y  $\int_{\partial R_n} \omega \neq 0$ . Con esta contradicción concluye la demostración ■

**Homotopía e independencia del camino.** Diremos provisionalmente que un abierto  $\Omega \subset \mathbb{R}^2$  tiene la propiedad  $P$  si todas las formas diferenciales cerradas y continuas definidas en  $\Omega$  son exactas. Según el ejemplo 13.15 el abierto  $\Omega = \mathbb{R}^2 \setminus \{(0, 0)\}$  no tiene la propiedad  $P$  y la proposición 13.18 dice que todos los abiertos especiales la tienen. Los abiertos con la propiedad  $P$  se pueden caracterizar en términos topológicos mediante la noción de homotopía que estudiamos a continuación.

**Definición 13.19** *Dos caminos cerrados  $\gamma_0, \gamma_1 : [0, 1] \rightarrow \Omega$  en un abierto  $\Omega \subset \mathbb{R}^2$  se dice que son  $\Omega$ -homotópicos (como caminos cerrados) si existe una función continua  $H : [0, 1] \times [0, 1] \rightarrow \Omega$  que verifica:*

- i)  $H(0, t) = \gamma_0(t)$ ,  $H(1, t) = \gamma_1(t)$ , para todo  $t \in [0, 1]$ .
- ii)  $H(s, 0) = H(s, 1)$  para todo  $s \in [0, 1]$ .

*Dos caminos  $\gamma_0, \gamma_1 : [0, 1] \rightarrow \Omega$ , con los mismos extremos:  $\gamma_0(0) = \gamma_1(0) = \mathbf{x}_0$ ,  $\gamma_0(1) = \gamma_1(1) = \mathbf{x}_1$ , se dice que son  $\Omega$ -homotópicos (con los extremos fijos) si existe una función continua  $H : [0, 1] \times [0, 1] \rightarrow \Omega$  que cumple:*

- i)  $H(0, t) = \gamma_0(t)$ ,  $H(1, t) = \gamma_1(t)$ , para todo  $t \in [0, 1]$ .
- ii)  $H(s, 0) = \mathbf{x}_0$ ,  $H(s, 1) = \mathbf{x}_1$  para todo  $s \in [0, 1]$ .

Para interpretar el significado de la  $\Omega$ -homotopía de caminos cerrados consideremos el conjunto  $\Lambda(\Omega)$  formado por los caminos cerrados  $\gamma : [0, 1] \rightarrow \Omega$ , dotado de la distancia de la convergencia uniforme

$$d_\infty(\gamma, \eta) = \max\{\|\gamma(t) - \eta(t)\|_2 : 0 \leq t \leq 1\}$$

Si  $H : [0, 1] \times [0, 1] \rightarrow \Omega$  es una homotopía entre los caminos cerrados  $\gamma_0, \gamma_1$ , para cada  $s \in [0, 1]$  la función parcial  $H_s : [0, 1] \rightarrow \Omega$ ,  $H_s(t) = H(s, t)$  es un camino cerrado en  $\Omega$ , con  $H_0 = \gamma_0$  y  $H_1 = \gamma_1$ .

Como  $H$  es uniformemente continua en el compacto  $[0, 1] \times [0, 1]$ , dado  $\epsilon > 0$  existe  $\delta > 0$  tal que si  $s, s' \in [0, 1]$  y  $t, t' \in [0, 1]$  verifican  $|s - s'| < \delta$ ,  $|t - t'| < \delta$  entonces  $\|H(s, t) - H(s', t')\|_2 < \epsilon$ . Entonces, si  $|s - s'| < \delta$ , para cada  $t \in [0, 1]$  se verifica  $d_\infty(H_s, H_{s'}) \leq \epsilon$ , lo que significa que la aplicación  $s \rightarrow H_s$  de  $[0, 1]$  en el espacio métrico  $(\Lambda(\Omega), d_\infty)$  es continua. Vemos así que el hecho de que dos caminos cerrados  $\gamma_0$  y  $\gamma_1$  sean  $\Omega$ -homotópicos (como caminos cerrados) significa que existe una familia uniparamétrica  $H_s$  de caminos cerrados en  $\Omega$ , que depende continuamente de  $s \in [0, 1]$ , mediante la cual el camino  $\gamma_0 = H_0$  se va deformando continuamente, dentro de  $\Omega$ , hasta transformarse en  $\gamma_1 = H_1$ .

La interpretación de la  $\Omega$ -homotopía de caminos con extremos fijos es similar considerando el conjunto  $\Lambda_{\mathbf{x}_0 \mathbf{x}_1}(\Omega)$  formado por los caminos  $\gamma : [0, 1] \rightarrow \Omega$ , con

$\gamma(0) = \mathbf{x}_0$ ,  $\gamma(1) = \mathbf{x}_1$ , dotado de la métrica de la convergencia uniforme. Ahora todos los caminos intermedios  $H_s$  tienen los mismos extremos que  $\gamma_0$  y  $\gamma_1$ .

Comenzamos con algunas observaciones preliminares que ayudarán a redactar la prueba de teorema 13.22. Si  $\omega$  es una forma diferencial cerrada definida en un abierto  $\Omega \subset \mathbb{R}^2$ , dados dos caminos regulares a trozos con los mismos extremos  $\gamma, \tilde{\gamma} : [0, 1] \rightarrow \Omega$ , diremos que  $\tilde{\gamma}$  es una  $\omega$ -modificación elemental de  $\gamma$  si existe un intervalo  $[t_0, t_1] \subset [0, 1]$  tal que  $\gamma(t) = \tilde{\gamma}(t)$  para todo  $t \in [0, 1] \setminus [t_0, t_1]$  y además

$$\gamma([t_0, t_1]) \subset D, \quad \tilde{\gamma}([t_0, t_1]) \subset D$$

donde  $D \subset \Omega$  es un disco abierto tal que  $\omega|_D$  es exacta. Si  $\tilde{\gamma}$  se obtiene a partir de  $\gamma$  mediante una cadena finita de modificaciones elementales sucesivas diremos que  $\tilde{\gamma}$  es una  $\omega$ -modificación de  $\gamma$ .

**Lema 13.20** *Sea  $\omega$  una forma diferencial cerrada y continua, definida en un abierto  $\Omega \subset \mathbb{R}^2$  y  $\gamma, \tilde{\gamma} : [0, 1] \rightarrow \Omega$  dos caminos regulares a trozos en  $\Omega$ , con los mismos extremos. Si  $\tilde{\gamma}$  es una  $\omega$ -modificación de  $\gamma$  se cumple*

$$\int_{\gamma} \omega = \int_{\tilde{\gamma}} \omega$$

DEM: Basta demostrarlo cuando  $\tilde{\gamma}$  es una  $\omega$ -modificación elemental de  $\gamma$ . En este caso basta observar que las integrales de  $\omega$  a lo largo de  $\gamma|_{[t_0, t_1]}$  y  $\tilde{\gamma}|_{[t_0, t_1]}$  coinciden en virtud del teorema 13.11. ■

**Lema 13.21** *Sea  $\omega$  una forma diferencial cerrada, definida en un abierto  $\Omega \subset \mathbb{R}^2$  y  $H : Q \rightarrow \Omega$  una función continua definida en  $Q = [0, 1] \times [0, 1]$ . Entonces existe una subdivisión  $p \in \mathcal{P}(Q)$  tal que para cada rectángulo  $S \in \Delta(p)$  hay un disco abierto  $D_S \subset \Omega$  tal que  $H(S) \subset D_S$  y  $\omega|_{D_S}$  es exacta.*

DEM: Consideremos una sucesión  $p_n \in \mathcal{P}(Q)$  tal que cada  $p_{n+1}$  es más fina que  $p_n$  y  $\|p_n\| \rightarrow 0$ . Para cada  $n$ , diremos que  $S \in \Delta(p_n)$  es aceptable si cumple la condición requerida en el enunciado (e.d. existe un disco abierto  $D_S \subset \Omega$  tal que  $H(S) \subset D_S$  y  $\omega|_{D_S}$  es exacta). Sea  $K_n$  la unión de los rectángulos no aceptables  $S \in \Delta(p_n)$ . Al refinar una subdivisión, los rectángulos aceptables se descomponen en rectángulos aceptables luego  $K_n$  es una sucesión decreciente de compactos. La prueba habrá terminado cuando probemos que algún  $K_n$  es vacío (ya que, en ese caso, todos los rectángulos de  $\Delta(p_n)$  serán aceptables). Esto lo haremos por reducción al absurdo. Si suponemos lo contrario la intersección de la sucesión decreciente de compactos  $K_n$  será no vacía y si  $\mathbf{a} = (s_0, t_0)$  es un punto de esta intersección, para cada  $n$  existirá un rectángulo no aceptable  $S_n \in \Delta(p_n)$  tal que  $\mathbf{a} \in S_n$ . Por otra parte, como  $\omega$  es cerrada habrá un disco  $D = B(H(\mathbf{a}), r) \subset \Omega$  tal que  $\omega|_D$  es exacta. Entonces, en virtud de la continuidad de  $H$ , existirá  $\delta > 0$  tal que  $H(Q \cap B(\mathbf{a}, \delta)) \subset D$ . Como  $\mathbf{a} \in S_n$  y  $\text{diam}(S_n) \leq \|p_n\| \rightarrow 0$ , para algún  $n$  se cumplirá  $S_n \subset Q \cap B(\mathbf{a}, \delta)$  luego  $H(S_n) \subset D$  y por lo tanto  $S_n$  será aceptable. Con esta contradicción queda demostrado que algún  $K_n$  es vacío ■

**Teorema 13.22** Sea  $\omega : \Omega \rightarrow \mathcal{L}(\mathbb{R}^2, \mathbb{R})$  una forma diferencial cerrada, definida y continua en un abierto  $\Omega \subset \mathbb{R}^2$ . Si  $\gamma_0, \gamma_1 : [0, 1] \rightarrow \Omega$  son caminos regulares a trozos  $\Omega$ -homotópicos con los extremos fijos se cumple

$$\int_{\gamma_0} \omega = \int_{\gamma_1} \omega$$

DEM: En virtud del lema 13.20 basta demostrar que  $\gamma_1$  es una  $\omega$ -modificación de  $\gamma_0$ . Sea  $Q = [0, 1] \times [0, 1]$  y  $H : Q \rightarrow \Omega$  una homotopía de caminos con extremos fijos entre  $\gamma_0$  y  $\gamma_1$ . Según el lema 13.21 existe una subdivisión de  $Q$ ,  $p = ((s_0, s_1, \dots, s_n), (t_0, t_1, t_2, \dots, t_m))$  tal que para cada rectángulo  $S_{ij} = [s_{i-1}, s_i] \times [t_{j-1}, t_j]$ , existe un disco abierto  $D_{ij} \subset \Omega$  tal que  $H(S_{ij}) \subset D_{ij}$  y  $\omega|_{D_{ij}}$  es exacta. .

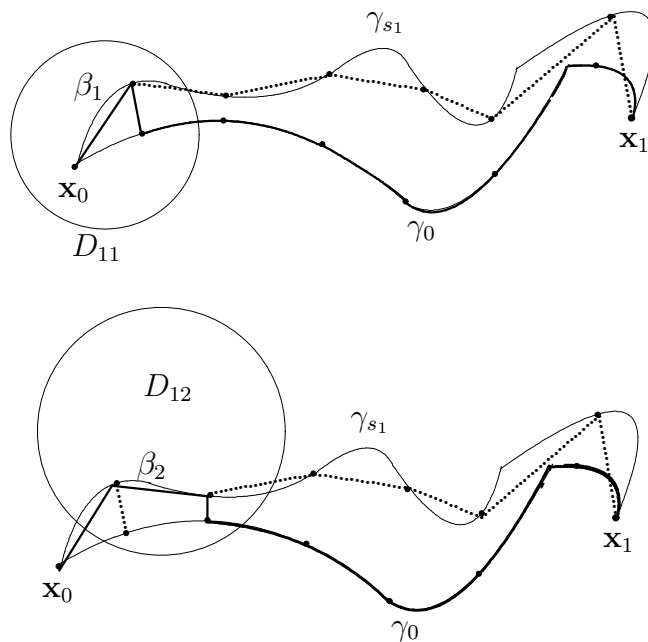
Consideremos los caminos continuos  $\gamma_{s_k}(t) = H(s_k, t)$ ,  $t \in [0, 1]$ . Para  $1 \leq k < n$  sea  $\chi_{s_k}$  el camino poligonal de origen  $\mathbf{x}_0 = \gamma_0(0)$  y extremo  $\mathbf{x}_1 = \gamma_0(1)$ , inscrito en  $\gamma_{s_k}$ , con vértices en los puntos  $\gamma_{s_k}(t_i)$ ,  $1 \leq i < m$ ; es decir,  $\chi_{s_k}$  se obtiene mediante yuxtaposición sucesiva de los segmentos  $[\gamma_{s_k}(t_{i-1}), \gamma_{s_k}(t_i)]$   $i = 1, \dots, m$ .

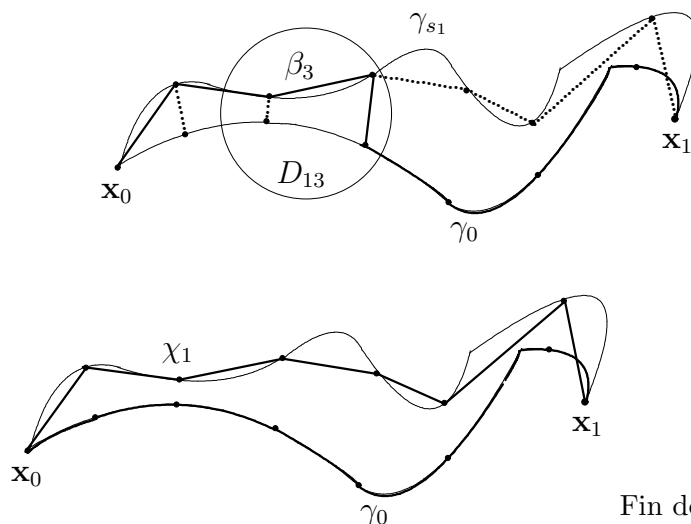
En una primera etapa el camino  $\gamma_0$  se transforma en la poligonal  $\chi_{s_1}$  mediante una cadena de  $m$  modificaciones elementales, sucesivas  $\beta_1, \beta_2, \dots, \beta_m$ , realizadas en la forma indicada en la figura.

La primera modificación  $\beta_1$  de  $\gamma_0$  se realiza dentro del disco  $D_{11}$ , sustituyendo el trozo del camino  $\gamma_0|_{[t_0, t_1]}$  por la yuxtaposición de dos segmentos contenidos en este disco. Análogamente la modificación  $\beta_{j+1}$  de  $\beta_j$  se realiza dentro del disco  $D_{1j}$ , donde  $\omega$  tiene primitiva.

En una segunda etapa, mediante otra cadena de  $m$  modificaciones elementales sucesivas se transforma la poligonal  $\chi_{s_1}$  en la poligonal  $\chi_{s_2}$ .

Finalmente, en la última etapa se obtiene  $\gamma_1$  como una  $\omega$ -modificación de la poligonal  $\chi_{s_{n-1}}$ . Queda demostrado así que  $\gamma_1$  es una  $\omega$ -modificación de  $\gamma_0$ .





■ .

Con una demostración similar se obtiene

**Teorema 13.23** Sea  $\omega : \Omega \rightarrow \mathcal{L}(\mathbb{R}^2, \mathbb{R})$  una forma diferencial cerrada, definida y continua en un abierto  $\Omega \subset \mathbb{R}^2$ . Si  $\gamma_0, \gamma_1 : [0, 1] \rightarrow \Omega$  son caminos cerrados regulares a trozos  $\Omega$ -homotópicos (como caminos cerrados) se cumple

$$\int_{\gamma_0} \omega = \int_{\gamma_1} \omega$$

**Definición 13.24** Un subconjunto abierto y conexo  $\Omega \subset \mathbb{R}^2$  se dice que es simplemente conexo si cada camino cerrado  $\gamma$  en  $\Omega$  es  $\Omega$ -homotópico a un camino constante.

Los abiertos estrellados son simplemente conexos: Todo abierto estrellado  $\Omega \subset \mathbb{R}^2$  es conexo porque es conexo por poligonales y si  $\gamma : [0, 1] \rightarrow \Omega$  es un camino cerrado en  $\Omega$ , que se supone estrellado respecto al punto  $a \in \Omega$ , entonces la función

$$H(s, t) = sa + (1 - s)\gamma(t) \in \Omega, \quad (s, t) \in [0, 1] \times [0, 1]$$

establece una homotopía en  $\Omega$  mediante la cual  $\gamma = H_0$  se transforma en el camino constante  $H_1 = a$ . También es inmediato que todo abierto homeomorfo a un abierto simplemente conexo es simplemente conexo. Se sigue de esto que son simplemente conexos todos los abiertos  $\Omega \subset \mathbb{R}^2$  que son homeomorfos al disco  $D(0, 1)$ .

El siguiente resultado topológico, que no será demostrado, proporciona una caracterización útil de los abiertos simplemente conexos como los abiertos conexos que no tienen orificios. Esta noción se formula de modo preciso utilizando la compactificación por un punto del plano euclídeo  $\mathbb{R}^2$ , denotada  $\mathbb{R}^2_\infty$ .



**Proposición 13.25** *Las siguientes propiedades de un abierto conexo  $\Omega \subset \mathbb{R}^2$  son equivalentes*

- a)  $\Omega$  es homeomorfo al disco  $D(0, 1)$ .
- b)  $\Omega$  es simplemente conexo.
- c) Toda pareja de caminos en  $\Omega$  con los mismos extremos, son  $\Omega$ -homotópicos como caminos con extremos fijos.
- d) Para toda curva cerrada simple (curva de Jordan)  $C$  en  $\Omega$  la región interior a  $C$  está contenida en  $\Omega$ .
- e)  $\mathbb{R}_\infty^2 \setminus \Omega$  es conexo.

**Teorema 13.26** *Si  $\omega$  es una forma diferencial cerrada definida y continua en un abierto simplemente conexo  $\Omega \subset \mathbb{R}^2$  se verifica:*

- a)  $\int_\gamma \omega = 0$  para cada camino cerrado y regular a trozos  $\gamma$  en  $\Omega$
- b)  $\int_{\gamma_0} \omega = \int_{\gamma_1} \omega$  para cada par de caminos  $\gamma_0, \gamma_1$  en  $\Omega$  regulares a trozos y con los mismos extremos.

*Es decir toda forma diferencial cerrada  $\omega$  definida y continua en un abierto simplemente conexo es exacta.*

DEM: a) Como  $\Omega$  es simplemente conexo  $\gamma$  es  $\Omega$ -homotópico a un camino constante  $\gamma_1$ , para el que es obvio que  $\int_{\gamma_1} \omega = 0$ , luego, en virtud del teorema 13.22,  $\int_\gamma \omega = 0$ .  
 b) Si se aplica a) al camino cerrado  $\gamma = \gamma_0 \vee (\sim \gamma_1)$  resulta  $0 = \int_\gamma \omega = \int_{\gamma_0} \omega - \int_{\gamma_1} \omega$ . (también se puede obtener como consecuencia de 13.25 y 13.22) y con el teorema 13.11 se concluye que  $\omega$  es exacta. ■

**Corolario 13.27** *Sea  $\omega = P(x, y)dx + Q(x, y)dy$  una forma diferencial continua en un abierto simplemente conexo  $\Omega \subset \mathbb{R}^2$ . Son equivalentes*

- a)  $\omega$  es exacta.
- b)  $\int_{\partial R} \omega = 0$  para todo rectángulo cerrado  $R \subset \Omega$ .

*Cuando  $P, Q$  son de clase  $C^1(\Omega)$  también es equivalente*

- c)  $D_2P(x, y) = D_1Q(x, y)$  para todo  $(x, y) \in \Omega$ .

DEM: Es consecuencia inmediata de 13.26, 13.18 y 13.14. ■

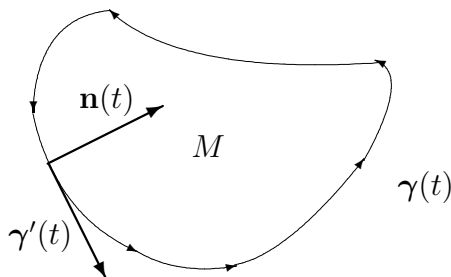
### 13.3. El teorema de Green

La fórmula de Green relaciona una integral doble sobre un recinto plano  $M$  con una integral de línea a lo largo de su frontera  $\partial M$ :

$$\int_M (D_1Q - D_2P) dx dy = \int_{\partial M} P dx + Q dy$$

Las hipótesis para la validez de esta fórmula son las naturales para que tengan sentido las integrales que figuran en ella: Por una parte  $M \subset \mathbb{R}^2$  es un compacto medible Jordan cuya frontera  $\partial M$  es una curva cerrada simple, regular a trozos (brevemente, región de Green). En la integral curvilínea de la derecha se supone que la frontera  $\partial M$  está orientada positivamente (es decir, en el sentido opuesto al de las manecillas del reloj). Por otra parte, para asegurar la existencia de las integrales involucradas, se supone que  $P$  y  $Q$  son funciones continuas en un abierto  $\Omega \supset M$  donde existen y son continuas las derivadas parciales  $D_1Q$ ,  $D_2P$ .

La condición de que  $\partial M$  sea una curva cerrada simple regular a trozos significa que existe un camino cerrado simple y regular a trozos  $\gamma : [0, 1] \rightarrow \mathbb{R}^2$ , tal que  $\partial M = \gamma([0, 1])$ . Las curvas cerradas simples se suelen llamar curvas de Jordan, debido al famoso teorema de Camille Jordan (1838-1922) que asegura que toda curva plana cerrada simple descompone al plano en dos abiertos conexos que tienen a la curva como frontera común. Uno de ellos es acotado y se llama región interior a la curva y el otro, que no es acotado, se llama región exterior. La orientación positiva de una curva cerrada simple es la que se obtiene al recorrerla en sentido opuesto al de las manecillas del reloj, de modo que la región interior quede siempre a la izquierda (se supone que usan los criterios habituales para representar los ejes de coordenadas en el plano). Esta definición, que no es rigurosa pero tiene la virtud de ser muy clara a nivel intuitivo, se puede formular de modo más formal pero más oscuro (que a lo mejor tranquiliza a algún lector muy escrupuloso con el rigor): Una parametrización regular a trozos  $\gamma(t) = (x(t), y(t))$  de la curva cerrada simple  $C$  tiene orientación positiva cuando para cada  $t \in [0, 1]$  donde existe y no es nulo el vector tangente  $\gamma'(t)$  se cumple que el vector normal  $\mathbf{n}(t) = (-y'(t), x'(t))$  (obtenido girando  $\pi/2$  el vector tangente) entra en  $M$ , región interior a  $C$ , es decir, existe  $\delta > 0$  tal que  $0 < s < \delta \Rightarrow \gamma(t) + s\mathbf{n}(t) \in M$ .



No demostraremos la versión general de la fórmula de Green. Solo veremos la demostración para regiones de Green que son de uno de los dos tipos siguientes

$$\text{I) } M = \{(x, y) : a \leq x \leq b, f_1(x) \leq y \leq f_2(x)\}$$

$$\text{II) } M = \{(x, y) : c \leq y \leq d, g_1(y) \leq x \leq g_2(y)\}$$

donde las funciones que las determinan  $f_1, f_2 : [a, b] \rightarrow \mathbb{R}$ ,  $g_1, g_2 : [c, d] \rightarrow \mathbb{R}$  son de clase  $C^1$  a trozos. Con esta definición conviene advertir que una región tan sencilla como el disco  $M := \{(x, y) : x^2 + y^2 \leq 1\}$  no es ni de tipo I) ni de tipo II) porque al describirlo en la forma  $\{(x, y) : -1 \leq x \leq 1, -\sqrt{1-x^2} \leq y \leq \sqrt{1-x^2}\}$  las funciones involucradas en la descripción no son derivables en los extremos del intervalo  $[-1, 1]$ . La mayor parte de los textos que demuestran la fórmula de Green sólo lo hacen para regiones de Green que son simultáneamente de los tipos I) y II) pero no advierten que con esta hipótesis una región tan simple como un disco compacto queda excluida de la clase de regiones para las que demuestran la fórmula.

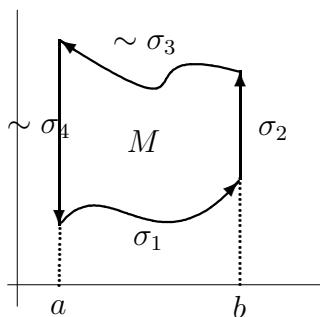
Obsérvese que para una región de tipo I) la frontera se recorre en sentido positivo mediante la curva cerrada simple regular a trozos  $\gamma = \sigma_1 \vee \sigma_2 \vee (\sim \sigma_3) \vee (\sim \sigma_4)$ .

$$\text{i) } \sigma_1(t) = (t, f_1(t)), \quad a \leq t \leq b.$$

$$\text{ii) } \sigma_2(t) = (b, t), \quad f_1(b) \leq t \leq f_2(b).$$

$$\text{iii) } \sigma_3(t) = (t, f_2(t)), \quad a \leq t \leq b.$$

$$\text{iv) } \sigma_4(t) = (a, t), \quad f_1(a) \leq t \leq f_2(a).$$



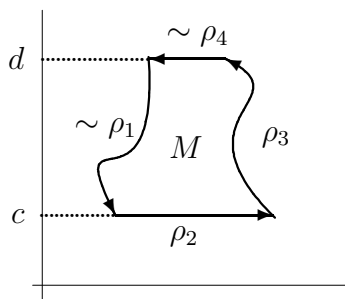
Análogamente, para una región de tipo II) la frontera se recorre en sentido positivo mediante la curva cerrada simple, regular a trozos  $\gamma = (\sim \rho_1) \vee \rho_2 \vee (\rho_3) \vee (\sim \rho_4)$ .

$$\text{i) } \rho_1(t) = (g_1(t), t), \quad c \leq t \leq d.$$

$$\text{ii) } \rho_2(t) = (t, c), \quad g_1(c) \leq t \leq g_2(c).$$

$$\text{iii) } \rho_3(t) = (t, g_2(t)), \quad c \leq t \leq d.$$

$$\text{iv) } \rho_4(t) = (t, d), \quad g_1(d) \leq t \leq g_2(d).$$



**Teorema 13.28** (Versión elemental del teorema de Green) Sean  $P, Q : \Omega \rightarrow \mathbb{R}$  funciones continuas en un abierto  $\Omega \subset \mathbb{R}^2$ , tales que las derivadas parciales  $D_1Q, D_2P$  existen y son continuas en todo  $\Omega$ . Si  $M \subset \Omega$  es una región de Green que simultáneamente es de tipo I) y de tipo II) se verifica

$$\int_M (D_1Q - D_2P) dx dy = \int_{\partial M} P dx + Q dy$$

donde la frontera  $\partial M$  se supone orientada positivamente.

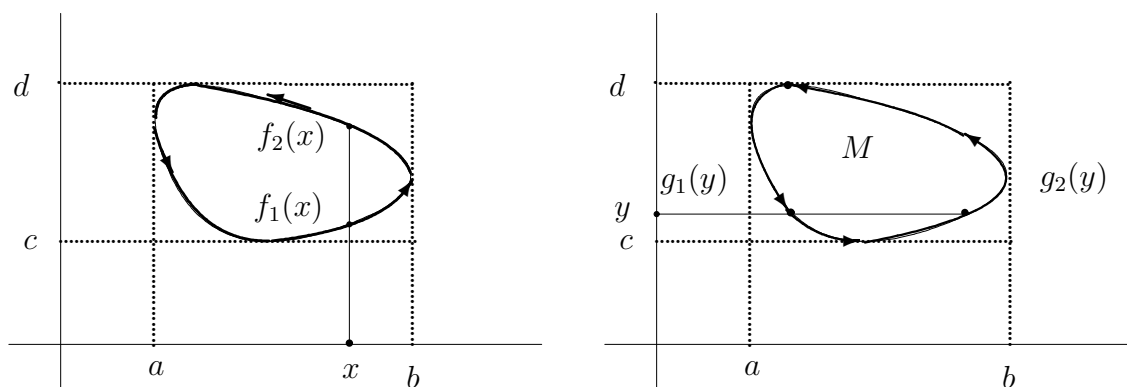
DEM: Utilizando la descripción de  $M$  como región de tipo I) se probará la igualdad

$$-\int_M D_2 P dx dy = \int_{\partial M} P dx \quad (\text{I})$$

Análogamente, usando la descripción de  $M$  como región de tipo II) resultará

$$\int_M D_1 Q dx dy = \int_{\partial M} Q dy, \quad (\text{II})$$

y sumando miembro a miembro las dos igualdades se obtendrá el resultado. Bastará hacer con detalle la prueba de (I) pues la prueba de (II) es análoga.



Si  $M$  es de tipo I), utilizando la parametrización de  $\partial M$  descrita anteriormente para las regiones de este tipo resulta

$$\begin{aligned} \int_{\partial M} P dx &= \int_{\sigma_1} P dx + \int_{\sigma_2} P dx - \int_{\sigma_3} P dx - \int_{\sigma_4} P dx = \\ &= \int_{\sigma_1} P dx - \int_{\sigma_3} P dx = \int_a^b P(t, f_1(t)) dt - \int_a^b P(t, f_2(t)) dt \end{aligned}$$

Por otra parte, utilizando el teorema de Fubini y el teorema fundamental del cálculo

$$\begin{aligned} \iint_M D_2 P dx dy &= \int_a^b dx \int_{f_1(x)}^{f_2(x)} D_2 P dy = \\ &= \int_a^b [P(x, f_2(x)) - P(x, f_1(x))] dx = - \int_{\partial M} P dx \end{aligned}$$

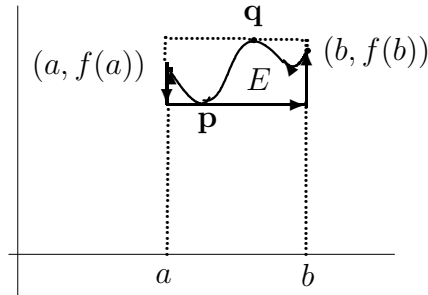
■

La versión elemental del teorema de Green que acabamos de demostrar se aplica en particular a los rectángulos  $M = [a, b] \times [c, d]$  y esto será la clave para la demostración cuando  $M$  es un recinto que sólo se supone de tipo I (o de tipo II).

Antes de emprender la demostración de este resultado conviene hacer algunas observaciones preliminares que recogemos en forma de lemas

**Lema 13.29** Sea  $E = \{(x, y) : x \in [a, b], m \leq y \leq f(x)\}$ , donde  $f : [a, b] \rightarrow \mathbb{R}$  es regular a trozos y  $m = \inf\{f(t) : t \in [a, b]\}$ . Entonces  $\text{Long}(\partial E) \leq 4\text{Long}(\gamma)$  donde  $\gamma(t) = (t, f(t))$ ,  $t \in [a, b]$ .

DEM:  $\partial E$  está formado por cuatro trozos: Dos segmentos verticales, un segmento horizontal y el camino  $\gamma$ . Basta ver que los tres segmentos tienen longitudes menores o iguales que  $\text{Long}(\gamma)$ .



La longitud del segmento horizontal es  $b - a$  y teniendo en cuenta que  $\gamma$  pasa por  $(a, f(a))$  y  $(b, f(b))$  resulta  $b - a \leq \text{Long}(\gamma)$ .

Por otra parte, si  $M = \max\{f(t) : t \in [a, b]\}$ , la longitud de cada segmento vertical es menor o igual que  $\leq M - m$ . Como existen  $\alpha, \beta \in [a, b]$  con  $m = f(\alpha)$  y  $M = f(\beta)$ , y el camino  $\gamma$  pasa por  $\mathbf{p} = (\alpha, m)$  y  $\mathbf{q} = (\beta, M)$  resulta

$$M - m \leq \|\mathbf{p} - \mathbf{q}\|_2 \leq \text{Long}(\gamma)$$

■ .

**Lema 13.30** En las condiciones del lema 13.29 si la forma diferencial

$$\omega = P dx + Q dy$$

está definida y es continua en un abierto  $\Omega \supset \partial E$ , se verifica

$$\left| \int_{\partial E} \omega \right| \leq M \text{Long}(\partial E)$$

donde

$$M = \sup\{\|\mathbf{F}(x, y) - \mathbf{F}(s, t)\|_2 : (x, y), (s, t) \in \partial E\}$$

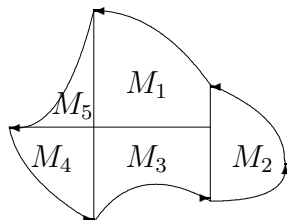
es la oscilación sobre  $\partial E$  de  $\mathbf{F} = (P, Q)$ .

DEM: Fijado un punto  $(x_0, y_0) \in E$ , como la forma diferencial constante  $\omega_0 = P(x_0, y_0)dx + Q(x_0, y_0)dy$  es exacta, se cumple  $\int_{\partial E} \omega_0 = 0$ , y utilizando la desigualdad 13.6 se obtiene

$$\left| \int_{\partial E} \omega \right| = \left| \int_{\partial E} (\omega - \omega_0) \right| \leq M \text{Long}(\partial E)$$

■

La siguiente observación, que se aplicará varias veces durante la prueba del teorema 13.31, también es útil en la práctica para justificar que la fórmula de Green es válida para una región concreta. Sea  $M$  una región que se puede descomponer, sin solapamiento, en un número finito de regiones  $M_j$ ,  $1 \leq j \leq m$ , para cada una de las cuales vale la fórmula de Green.



Se supone que la descomposición tiene la propiedad de que la curva orientada  $\partial M$  se deduce de las curvas orientadas  $\partial M_j$  efectuando las cancelaciones de los trozos de estas curvas que intervienen dos veces, pero con orientaciones opuestas. Entonces es inmediato que la fórmula de Green también se verifica para la región  $M$ .

**Teorema 13.31** (Teorema de Green para regiones de tipo I) Sean  $P, Q : \Omega \rightarrow \mathbb{R}$  funciones continuas en un abierto  $\Omega \subset \mathbb{R}^2$ , en el que existen y son continuas las derivadas parciales  $D_1Q, D_2P$ . Si  $M \subset \Omega$  es una región de tipo I o de tipo II se verifica

$$\int_M (D_1Q - D_2P) dx dy = \int_{\partial M} P dx + Q dy$$

donde  $\partial M$  se supone con la orientación positiva.

DEM: Bastará hacer la prueba para regiones de tipo I

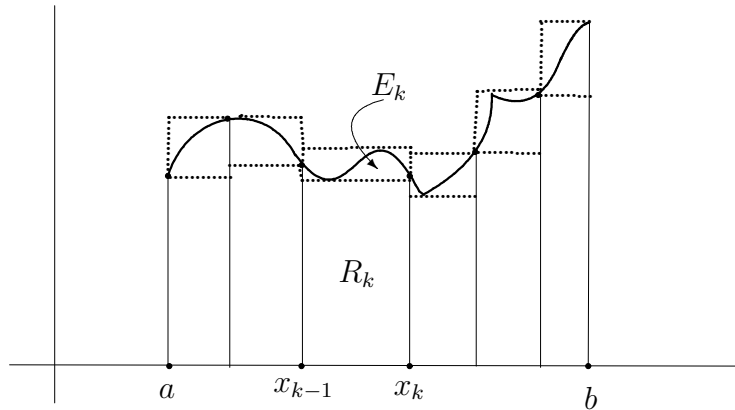
$$M = \{(x, y) : x \in [a, b], g(x) \leq y \leq f(x)\}$$

donde  $f$  y  $g$  son regulares a trozos (la prueba para regiones de tipo II es similar.) Consideraremos primero el caso en que una de las funciones que intervienen en la definición de  $M$  es constante; Si suponemos que  $g$  es la función constante 0, se tendrá

$$M = \{(x, y) : x \in [a, b], 0 \leq y \leq f(x)\}$$

Sea  $L = \text{Long}(\gamma)$  donde  $\gamma(t) = (t, f(t))$ ,  $t \in [a, b]$ . Utilizando la sobreyectividad de la función abscisa curvilínea  $v(t) = V(\gamma, [a, t])$  podemos obtener una subdivisión  $p_n = (t_0 < t_1 < \dots < t_n)$  de  $[a, b]$  tal que todos los trozos  $\gamma_k = \gamma|_{[x_{k-1}, x_k]}$  tienen la misma longitud  $\text{Long}(\gamma_k) = L/n$ . Obsérvese que  $x_k - x_{k-1} \leq L/n$  luego  $\|p_n\| \leq L/n$ . Entonces, dado  $\epsilon > 0$  existe un  $n \in \mathbb{N}$  que cumple

$$\int_a^b f(x) dx - s(f, p_n) < \epsilon$$



Si  $R_k = [t_{k-1}, t_k] \times [m_k, M_k]$ , con  $m_k = \inf f[t_{k-1}, t_k]$  y  $M_k = \sup f[t_{k-1}, t_k]$ , la última desigualdad se traduce en los siguientes términos

$$\text{Área}(M) - \sum_{k=1}^n \text{Área}(R_k) < \epsilon$$

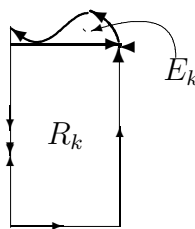
La fórmula de Green es cierta para rectángulos, y en virtud de la observación previa al enunciado del teorema, también lo es para la región  $M_\epsilon = \bigcup_{k=1}^n R_k$ .

$$\int_{M_\epsilon} (D_1Q - D_2P) dx dy = \int_{\partial M_\epsilon} P dx + Q dy$$

Como  $\text{Área}(M \setminus M_\epsilon) = \text{Área}(M) - \sum_{k=1}^n \text{Área}(R_k) < \epsilon$  resulta

$$\left| \int_M (D_1Q - D_2P) dx dy - \int_{M_\epsilon} (D_1Q - D_2P) dx dy \right| \leq K \text{Área}(M \setminus M_\epsilon) \leq K\epsilon$$

donde  $K > 0$  es el máximo de la función continua  $|D_1Q - D_2P|$  sobre el compacto  $M$ .



En virtud del lema 13.29, cada  $E_k = \{(x, y) : x \in [t_{k-1}, t_k], m_k \leq y \leq f(x)\}$  cumple  $\text{Long}(\partial E_k) < 4L/n$  luego  $\text{diam}(E_k) \leq 4\sqrt{2}L/n$ . En virtud de la continuidad uniforme de  $\mathbf{F} = (P, Q)$  sobre el compacto  $M$ , tomando  $n$  suficientemente grande podemos garantizar que la oscilación de  $\mathbf{F}$  sobre cada  $E_k$  es menor que  $\epsilon$  y aplicando el lema 13.30 se obtiene

$$\left| \int_{\partial E_k} \omega \right| \leq \epsilon \text{Long}(\partial E_k) \leq \epsilon 4L/n$$

Sea  $G_k = \{(x, y) : x \in [x_{k-1}, x_k], 0 \leq y \leq f(x)\} = R_k \cup E_k$  Teniendo en cuenta las cancelaciones de la integral sobre segmentos opuestos podemos escribir

$$\int_{\partial M} \omega = \sum_{k=1}^n \int_{\partial G_k} \omega = \sum_{k=1}^n \left( \int_{\partial R_k} \omega + \int_{\partial E_k} \omega \right) = \int_{\partial M_\epsilon} \omega + \sum_{k=1}^n \int_{\partial E_k} \omega$$

y se obtiene

$$\left| \int_{\partial M} \omega - \int_{\partial M_\epsilon} \omega \right| \leq \sum_{k=1}^n \left| \int_{\partial E_k} \omega \right| \leq n\epsilon 4L/n = \epsilon 4L$$

Combinando la igualdad  $\int_{M_\epsilon} (D_1Q - D_2P)dxdy = \int_{\partial M_\epsilon} Pdx + Qdy$  con las dos desigualdades que hemos obtenido resulta

$$\left| \int_M (D_1Q - D_2P)dxdy - \int_{\partial M} \omega \right| \leq (K + 4L)\epsilon$$

y como  $\epsilon > 0$  es arbitrario, la demostración ha terminado para el caso particular que hemos considerado.

Supongamos ahora que  $M = \{(x, y) : a \leq x \leq b, g(x) \leq y \leq f(x)\}$  es de tipo I, pero con la condición adicional:  $f(x) < g(x)$  para todo  $x \in [a, b]$ . En este caso  $\mu = \min\{f(x) - g(x) : x \in [a, b]\}$  se alcanza en algún punto, luego  $\mu > 0$ , y en virtud de la continuidad uniforme de  $g$  sobre  $[a, b]$  existe  $\delta > 0$  tal que todo par  $s, t \in [a, b]$  con  $|s - t| < \delta$  cumple  $-\mu < g(s) - g(t) < \mu$ .

Entonces, para una subdivisión ( $a = x_0 < x_1 < \dots < x_m = b$ ) que cumpla la condición  $\max\{x_k - x_{k-1} : 1 \leq k \leq m\} < \delta$  se verifica

$$c_k = \max\{g(t) : t \in [x_{k-1}, x_k]\} \leq \min\{f(t) : t \in [x_{k-1}, x_k]\}$$

(Efectivamente, si  $s_k \in [x_{k-1}, x_k]$  es tal que  $f(s_k) = \min\{f(t) : t \in [x_{k-1}, x_k]\}$ , entonces para todo  $t \in [x_k, x_{k-1}]$  se cumple

$$f(s_k) \geq \mu + g(s_k) = \mu + (g(s_k) - g(t)) + g(t) \geq \mu - \mu + g(t) = g(t)$$

. Ahora, si descomponemos  $M$  en las regiones

$$A_k = \{(x, y) : x \in [x_{k-1}, x_k], g(x) \leq y \leq c_k\}$$

$$B_k = \{(x, y) : x \in [x_{k-1}, x_k], c_k \leq y \leq f(x)\}$$

para las que ya hemos demostrado que se cumple la fórmula de Green, obtenemos que dicha fórmula se sigue cumpliendo para  $M$ .

Finalmente, sea  $M$  una región de tipo I sin condición adicional. Si  $r > 0$  es suficientemente pequeño la región  $M_r = \{(x, y) : a \leq x \leq b, g(x) \leq y \leq f(x) + r\}$  está contenido en  $\Omega$  y cumple la condición adicional bajo la que tenemos demostrada la igualdad

$$\int_{M_r} (D_1Q - D_2P)dxdy = \int_{\partial M_r} Pdx + Qdy$$



Es fácil ver que

$$\lim_{r \rightarrow 1} \int_{M_r} (D_1Q - D_2P) dx dy = \int_M (D_1Q - D_2P) dx dy$$

$$\lim_{r \rightarrow 1} \int_{\partial M_r} P dx + Q dy = \int_{\partial M} P dx + Q dy$$

(esto se deja como ejercicio) y así queda establecida la fórmula de Green para las regiones de tipo I. ■

La validez de la fórmula de Green para un recinto circular  $\{(x, y) : x^2 + y^2 \leq R\}$  no se obtiene con una aplicación directa del teorema 13.31. Se puede justificar a posteriori descomponiendo el disco en tres regiones, trazando dos cuerdas paralelas al eje de abscisas, una por encima y otra por debajo del centro. Para las tres regiones se tiene demostrada la validez de la fórmula: Obsérvese que la que contiene al centro es de tipo II mientras que las otras dos (segmentos circulares) son de tipo I.

**Aplicaciones del teorema de Green.** La caracterización de las formas diferenciales cerradas obtenida en el teorema 13.13, bastante útil en la práctica, tiene el inconveniente de que sólo se aplica a formas diferenciales de clase  $C^1$ . Por otra parte, en la proposición 13.18 se obtuvo otra caracterización, usando una condición de distinta naturaleza, que tiene la ventaja de aplicarse a todas las formas diferenciales continuas. El teorema de Green, en su versión elemental para rectángulos, permite aclarar la relación que hay entre las condiciones que intervienen las dos caracterizaciones mencionadas. Al mismo tiempo proporciona otra demostración de una versión algo más general de la proposición 13.14, que no utiliza el teorema de derivación de integrales dependientes de un parámetro. Para demostrar el teorema 13.33 se necesita el siguiente lema que se deja como ejercicio

**Lema 13.32** *Si  $f : \Omega \rightarrow \mathbb{R}$  es continua en un abierto  $\Omega \subset \mathbb{R}^2$  y  $\int_R f(x, y) dx dy = 0$  para todo rectángulo cerrado  $R \subset \Omega$  entonces  $f$  es idénticamente nula.*

**Teorema 13.33** *Sea  $\omega(x, y) = P(x, y)dx + Q(x, y)dy$  una forma diferencial definida y continua en un abierto  $\Omega \subset \mathbb{R}^2$  tal que en todo punto  $(x, y) \in \Omega$  las derivadas parciales  $D_2P(x, y)$ ,  $D_1Q(x, y)$  existen y son continuas. Entonces son equivalentes*

- a)  $D_2P(x, y) = D_1Q(x, y)$  para todo  $(x, y) \in \Omega$ .
- b)  $\int_{\partial R} \omega = 0$  para cada rectángulo  $R \subset \Omega$ .
- c)  $\omega$  es cerrada.

*Si  $\Omega$  es simplemente conexo, también es equivalente*

- d)  $\omega$  es exacta.

DEM: a)  $\Rightarrow$  b) es consecuencia inmediata del teorema de Green en su versión elemental para rectángulos.

b)  $\Rightarrow$  a) se obtiene combinando el teorema de Green con el lema 13.32 aplicado a la función  $f = D_2P - D_1Q$ . El resto de las afirmaciones del enunciado han sido probadas anteriormente. ■

El siguiente corolario proporciona una nueva demostración, basada en el teorema de Green, de un resultado obtenido en el capítulo 6.

**Corolario 13.34** *Sea  $f : \Omega \rightarrow \mathbb{R}$  una función de clase  $C^1(\Omega)$  tal que en todo punto  $(x, y) \in \Omega$  existen y son continuas las derivadas parciales  $D_{21}f(x, y)$ ,  $D_{12}f(x, y)$ . Entonces  $D_{21}f(x, y) = D_{12}f(x, y)$  para todo  $(x, y) \in \Omega$ .*

DEM: Basta aplicar el teorema anterior a la forma diferencial  $df(x, y) = D_1f(x, y)dx + D_2f(x, y)dy$ . ■

El teorema de Green se aplica tanto para el cálculo de integrales curvilíneas, (transformándolas en integrales dobles) como para el cálculo de integrales dobles (transformándolas en integrales curvilíneas). En particular se puede aplicar para calcular áreas:

**Proposición 13.35** *Sea  $M \subset \mathbb{R}^2$  una región de Green y  $\gamma(t) = (x(t), y(t))$ ,  $t \in [a, b]$  un camino regular a trozos que recorre la frontera  $\partial M$ , con la orientación positiva. Entonces*

$$\text{Área}(M) = \frac{1}{2} \int_{\gamma} xdy - ydx = \int_a^b (y'(t)x(t) - x'(t)y(t))dt$$

DEM: Basta aplicar el teorema de Green con  $P(x, y) = -y/2$ ,  $Q(x, y) = x/2$ . ■

## 13.4. Ejercicios resueltos

**Ejercicio 13.36** *Estudie si la forma diferencial  $\omega(x, y) = \frac{-2xy}{x^4 + y^2}dx + \frac{x^2}{x^4 + y^2}dy$  es cerrada o exacta en el abierto  $\Omega = \mathbb{R}^2 \setminus \{(0, 0)\}$ . Calcule  $\int_{\gamma} \omega$ , donde  $\gamma$  es un camino regular a trozos en  $\Omega$ , de origen  $(-a, 0)$  y extremo  $(a, 0)$ .*

SOLUCIÓN

Como las funciones

$$P(x, y) = \frac{-2xy}{x^4 + y^2}, \quad Q(x, y) = \frac{x^2}{x^4 + y^2}$$

son de clase  $C^1(\Omega)$  y  $D_2P(x, y) = D_1Q(x, y)$  para todo  $(x, y) \in \Omega$  podemos afirmar que la forma diferencial es cerrada en  $\Omega$ . Obsérvese que  $\Omega$  no es estrellado, de modo que sólo podemos asegurar que  $\omega|_G$  es exacta sobre cada abierto estrellado  $G \subset \Omega$ .

En particular, sobre el abierto  $A = \{(x, y) : x > 0\}$ , la restricción  $\omega|_A$  es exacta y es fácil encontrar una primitiva  $f$  de  $\omega|_A$ : Buscamos una función diferenciable  $f : A \rightarrow \mathbb{R}$ , que verifique  $D_1f(x, y) = P(x, y)$ ,  $D_2f(x, y) = Q(x, y)$ . Para cada  $x > 0$  la función parcial

$$y \rightarrow Q(x, y) = \frac{x^2}{x^4 + y^2}$$

tiene primitivas inmediatas. Son las funciones de la forma  $\text{Arctg}(y/x^2) + \varphi(x)$  donde  $\text{Arctg} : \mathbb{R} \rightarrow (-\pi/2, \pi/2)$  es la rama principal de la función multivaluada  $\text{arc tg}$ , y  $\varphi$  es una función derivable que hay que se determinar imponiendo la condición

$$P(x, y) = \frac{\partial}{\partial x} (\text{Arctg}(y/x^2) + \varphi(x))$$

es decir

$$\frac{-2xy}{x^4 + y^2} = \frac{-2xy}{x^4 + y^2} + \varphi'(x)$$

Se concluye que  $\varphi$  es constante y con ello que  $f(x, y) = \text{Arctg}(y/x^2)$  es una primitiva de  $\omega|_A$ . Se aprecia que la primitiva  $f$ , definida inicialmente en  $A$ , se puede extender una primitiva  $F$  definida en  $\Omega$ :

$$F(x, y) = \text{Arctg}(y/x^2) \text{ si } x \neq 0;$$

$$F(0, y) = -\pi/2 \text{ si } y < 0; \quad F(0, y) = \pi/2 \text{ si } y > 0.$$

Es fácil comprobar que en cada punto de la forma  $(0, b)$ , con  $b \neq 0$ , la función  $F$  tiene derivadas parciales,  $D_1F(0, b) = 0 = P(0, b)$ ,  $D_2F(0, b) = 0 = Q(0, b)$ , y que las funciones  $D_1F = P$ ,  $D_2F = Q$  son continuas en dicho punto. Por lo tanto  $F$  es de clase  $C^1(\Omega)$  y  $dF = \omega$ .

Como la forma diferencial  $\omega$  es exacta, la integral  $\int_\gamma \omega$  no depende del camino regular a trozos  $\gamma$  en  $\Omega$  de origen  $(-a, 0)$  y extremo  $(a, 0)$ . Para calcular su valor podemos elegir un camino particular con el que los cálculos sean sencillos. Por ejemplo un camino poligonal formado por tres segmentos de lados paralelos a los ejes que pase por los puntos  $(-a, 0)$ ,  $(-a, 1)$ ,  $(a, 1)$ ,  $(a, 0)$ , en este orden. Con este camino se obtiene fácilmente que  $\int_\gamma \omega = 0$ .

NOTA: También se puede razonar de modo alternativo comenzando con el cálculo de las primitivas de la función de una variable  $x \rightarrow P(x, y)$  que para  $y > 0$  son de la forma

$$-\text{Arctg}(x^2/y) + \psi(y)$$

donde  $\psi$  es una función derivable que queda determinada por la condición

$$Q(x, y) = \frac{\partial}{\partial y} (-\text{Arctg}(x^2/y) + \psi(y))$$

Ahora también se obtiene que la función  $\psi$  es constante y se llega a que  $g(x, y) = -\text{Arc tg}(x^2/y)$  es una primitiva de  $\omega|_B$  en el abierto  $B = \{(x, y) : y > 0\}$ . Esta primitiva  $g$ , definida inicialmente en  $B$  se puede extender a una primitiva  $G$  definida en todo  $\Omega$  mediante la fórmula

$$G(x, y) = -\operatorname{Arctg}(x^2/y) \text{ si } y > 0; \quad G(x, y) = -\operatorname{Arctg}(x^2/y) - \pi \text{ si } y < 0;$$

$$G(x, 0) = -\pi/2$$

Se deja al cuidado del lector la comprobación de que  $G$  también es una primitiva de  $\omega$  clase  $C^1(\Omega)$ . Como  $\Omega$  es conexo y  $F(1, 0) - G(1, 0) = \pi/2$  podemos afirmar que  $F(x, y) - G(x, y) = \pi/2$  para todo  $(x, y) \in \Omega$ . ■

**Ejercicio 13.37** Sea  $f : \Omega \rightarrow \mathbb{R}$  de clase  $C^1$  en  $\Omega = \mathbb{R}^2 \setminus \{(0, 0)\}$ . Demuestre que la forma diferencial  $\omega(x, y) = yf(x, y)dx - xf(x, y)dy$  es cerrada si y sólo si la función  $r^2 f(r \cos \theta, r \sin \theta)$  no depende de  $r$ .

En el caso particular  $f(x, y) = x^2/(x^2 + y^2)^2$  justifique que  $\omega$  no tiene primitiva en  $\Omega$  pero tiene primitiva en  $A = \mathbb{R}^2 \setminus \{(x, 0) : x \leq 0\}$ .

Sea  $g$  la primitiva de  $\omega$  en  $A$  determinada por  $g(1, 0) = 0$ . Calcule  $g(r \cos \theta, r \sin \theta)$  donde  $r > 0$  y  $-\pi < \theta < \pi$ .

SOLUCIÓN

Las funciones  $P(x, y) = yf(x, y)$ ,  $Q(x, y) = -xf(x, y)$  son de clase  $C^1$  luego la forma diferencial  $\omega(x, y) = P(x, y)dx + Q(x, y)dy$  es cerrada si y sólo si  $D_2P = D_1Q$  es decir  $2f(x, y) + xf_x(x, y) + yf_y(x, y) = 0$ . Con el cambio de variable a coordenadas polares la condición anterior se expresa en la forma  $2F(r, \theta) + rF_r(r, \theta) = 0$  donde  $F(r, \theta) = f(r \cos \theta, r \sin \theta)$ . Multiplicando por  $r > 0$  resulta la condición

$$0 = 2rF(r, \theta) + r^2F_r(r, \theta) = \frac{d}{dr}[r^2F(r, \theta)]$$

que equivale a que  $r^2F(r, \theta)$  no depende de  $r$ .

En el caso particular  $f(x, y) = x^2/(x^2 + y^2)^2$  se comprueba fácilmente que la circunferencia  $C(t) = (\cos t, \sin t)$ ,  $t \in [0, 2\pi]$  proporciona una integral no nula  $\int_C \omega \neq 0$ , luego  $\omega$  no es exacta en  $\Omega$ . Sin embargo, como  $A$  es estrellado, se puede asegurar que  $\omega|_A$  es exacta. Si  $g$  es la primitiva de  $\omega$  en  $A$  que se anula en  $(0, 1)$  para calcular  $g(r \cos \theta, r \sin \theta)$  basta calcular la integral curvilínea de  $\omega$  a lo largo de cualquier camino regular a trozos de origen  $(1, 0)$  y extremo  $(r \cos \theta, r \sin \theta)$ . Tomamos el camino compuesto del segmento  $\sigma$  de origen  $(1, 0)$  y extremo  $(r, 0)$ , seguido de un arco del circunferencia  $\gamma$  de origen  $(r, 0)$  y extremo  $(r \cos \theta, r \sin \theta)$ , contenido en  $A$ :  $\gamma(t) = (r \cos t\theta, r \sin t\theta)$ ,  $t \in [0, 1]$ . Este camino regular a trozos conduce al valor  $g(r \cos \theta, r \sin \theta) = \int_\sigma \omega + \int_\gamma \omega$ . Es inmediato que  $\int_\sigma \omega = 0$ , luego  $g(r \cos \theta, r \sin \theta) = \int_\gamma \omega = -\int_0^\theta \cos^2 t dt = \dots$ . ■

**Ejercicio 13.38** Calcule la integral curvilínea  $\int_C \omega$  donde

$$\omega(x, y) = \sqrt{x^2 + y^2} dx + y[xy + \log(x + \sqrt{x^2 + y^2})] dy$$

y  $C$  es el borde de  $M = \{(x, y) : 0 \leq x \leq 1, 0 \leq y \leq x, x^2 + y^2 \geq 1\}$ , orientado en sentido positivo.

## SOLUCIÓN

Las funciones  $P(x, y) = \sqrt{x^2 + y^2}$ ,  $Q(x, y) = y(xy + \log(x + \sqrt{x^2 + y^2}))$  son de clase  $C^1$  en  $\Omega = \{(x, y) : x > 0\}$  y  $M \subset \Omega$  es una región simple

$$M = \{(x, y) : \alpha \leq x \leq 1, \sqrt{1 - x^2}\} \text{ con } \alpha = \sqrt{2}/2$$

Según la versión elemental del teorema de Green,

$$\begin{aligned} \int_C \omega &= \int_M [Q_x(x, y) - P_y(x, y)] dx dy = \int_M y^2 dx dy = \int_\alpha^1 dx \int_{\sqrt{1-x^2}}^x y^2 dy = \\ &= \frac{1}{3} \int_\alpha^1 [x^3 - (1 - x^2)^{3/2}] dx = \frac{1}{12}(1 - \alpha^4) - \frac{1}{3}\beta \end{aligned}$$

donde

$$\alpha = \sqrt{2}/2; \quad \beta = \int_\alpha^1 (1 - x^2)\sqrt{1 - x^2} dx = \int_0^{\pi/4} (\sin s)^4 ds = \frac{3\pi - 4}{16}$$

luego  $I = (5 - 3\pi)/48$ . ■

## 13.5. Ejercicios propuestos

◇ **13.5.1** Obtenga la forma canónica de las siguientes formas diferenciales:

a)  $\omega = d(f^2) - 3dg$ , donde  $f(x, y) = \sqrt{xy}$ ,  $g(x, y) = (x + 2y) \log(x^2 + y^2)$  se suponen definidas en  $\Omega = \{(x, y) : x > 0, y > 0\}$ .

b)  $x\omega(x, y) + \cos y df(x, y)$ , donde  $\omega(x, y) = e^y dx + dy/x$ ,  $f(x, y) = \operatorname{sen} \sqrt{x + y^2}$ , se suponen definidas en  $\Omega = \{(x, y) : x > 0\}$ .

c)  $d\left(\sum_{i \neq j} x_i x_j\right)$ .

◇ **13.5.2** Sea  $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$  una base de  $\mathbb{R}^n$  y  $\{d\alpha_1, d\alpha_2, \dots, d\alpha_n\}$  su base dual en  $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$ . Obtenga el campo vectorial asociado a la forma diferencial  $\omega = \sum_{j=1}^n A_j d\alpha_j$

◇ **13.5.3** Calcule las siguientes integrales curvilíneas

a)  $\int_{\gamma} \operatorname{sen} z dx + \cos z dy - \sqrt[3]{xy} dz$  donde  $\gamma : [0, 1] \rightarrow \mathbb{R}^3$  es el camino definido por  $\gamma(t) = (\cos^3 t, \operatorname{sen}^3 t, t)$ ,  $t \in [0, 7\pi/2]$ .

b)  $\int_{\gamma} y^2 \cos(xy^2) dx + 2xy \cos(xy^2) dy$  donde  $\gamma : [0, 1] \rightarrow \mathbb{R}^2$  es el camino definido por  $x(t) = t^4$ ,  $y(t) = \operatorname{sen}^3(\pi t/2)$ .

◇ **13.5.4** Compruebe que el campo de vectores  $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,

$$F(x, y) = (3x^2 + 2y \operatorname{sen} 2x, 2(\sin x)^2 + 6y^2)$$

es un gradiente y obtenga un potencial del campo.

◇ **13.5.5** Demuestre que existe una función diferenciable  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  que verifica

$$D_1 F(x, y) = e^{x^2 - y^2} \operatorname{sen}(2xy), \quad D_2 F(x, y) = e^{x^2 - y^2} \cos(2xy)$$

Obtenga una fórmula integral para la función  $F$ .

◇ **13.5.6** Sea  $\omega(x, y) = P(x, y)dx + Q(x, y)dy$ , donde

$$P(x, y) = Q(y, x) = \frac{(x^2 + y^2)(3x^2 - y^2)}{x^2 y}$$

Compruebe que la forma diferencial es exacta en  $\Omega = \{(x, y) : x > 0, y > 0\}$  y obtenga una primitiva. Calcule  $\int_{\gamma} \omega$  donde  $\gamma(t) = (t + \cos 3t, 1 + \operatorname{sen}^2 t)$ ,  $0 \leq t \leq \pi/2$ .

◇ **13.5.7** Compruebe que la forma diferencial

$$\omega(x, y, z) = 2xy dx + (x^2 + \log z) dy + (y/z) dz$$

es exacta en  $\Omega = \{(x, y, z) : z > 0\}$  y obtenga una primitiva.

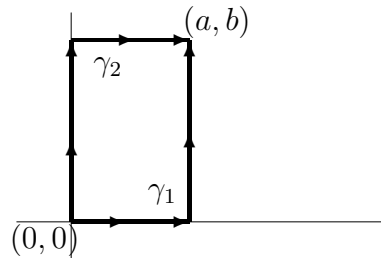
◇ **13.5.8** Considerando la integral curvilínea de

$$\omega(x,y) = e^{x^2-y^2} \operatorname{sen}(2xy) dx + e^{x^2-y^2} \operatorname{cos}(2xy) dy$$

a lo largo de los caminos poligonales  $\gamma_1, \gamma_2$ , que se indican en la figura, cacule la integral impropia

$$J = \int_0^{+\infty} e^{-t^2} \operatorname{cos}(2at) dt$$

en función de  $I = \int_0^{+\infty} e^{-t^2} dt$ . ( $a > 0, b > 0$ ).



◇ **13.5.9** Sea  $\omega(x,y) = P(x,y) dx + Q(x,y) dy$ , donde

$$P(x,y) = \frac{e^{-y}}{x^2+y^2}(x \operatorname{sen} x - y \operatorname{cos} x); \quad Q(x,y) = \frac{e^{-y}}{x^2+y^2}(x \operatorname{cos} x + y \operatorname{sen} x)$$

- Compruebe que  $\omega$  es una forma diferencial cerrada en  $\mathbb{R}^2 \setminus \{(0,0)\}$ .
- Calcule el valor de la integral  $\int_{\gamma} \omega$  donde  $\gamma$  es un camino cerrado que recorre la frontera de  $K(\epsilon, R) = \{(x,y) : \epsilon^2 \leq x^2 + y^2 \leq R^2, y \geq 0\}$ .
- Deduzca de b), tomando límites cuando  $\epsilon \rightarrow 0, R \rightarrow +\infty$ , el valor de la integral impropia

$$\int_0^{+\infty} \frac{\operatorname{sen} x}{x} dx = \frac{\pi}{2}$$

◇ **13.5.10** En cada uno de los siguientes casos, compruebe que el campo de vectores:  $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  es un gradiente y obtenga una función potencial del campo:

- $F(x,y,z) = (2xyz + z^2 - 2y^2 + 1, x^2z - 4xy, x^2y + 2xz - 2)$ .
- $F(x,y,z) = (2xyze^{x^2}, ze^{x^2} + 2ye^z, y^2e^z + ye^{x^2})$ .
- $F(x,y,z) = (e^{-xy}(y - xy^2 + yz), e^{-xy}(x - x^2y + xz), e^{-xy})$ .

◇ **13.5.11** Si  $g : U \rightarrow \mathbb{R}$  es continua en un abierto  $U \subset \mathbb{R}$ , demuestre que la forma diferencial

$$\omega(\mathbf{x}) = \sum_{j=1}^n g(\|\mathbf{x}\|_2) x_j dx_j$$

es exacta en el abierto  $\Omega = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_2 \in U\}$ , y obtenga una primitiva.

Sea  $r = \sqrt{x^2 + y^2 + z^2}$ . Justifique que el campo de vectores

$$F(x, y, z) = \frac{1}{r^3}(x, y, z)$$

admite función potencial en  $\mathbb{R}^3 \setminus \{(0, 0, 0)\}$ .

◇ **13.5.12** Si  $f : \mathbb{R}^3 \rightarrow \mathbb{R}$  es diferenciable, demuestre que la forma diferencial

$$\omega(x, y, z) = xf(x, y, z) dx + yf(x, y, z) dy + zf(x, y, z) dz$$

es exacta si y sólo si  $f$  es constante sobre cada esfera centrada en  $(0, 0, 0)$ .

◇ **13.5.13** El campo de fuerzas ejercido por una masa  $M$  colocada en  $\mathbf{0} = (0, 0, 0)$  sobre una masa  $m$  colocada en  $\mathbf{p} = (x, y, z) \neq (0, 0, 0)$ , viene dado por

$$\mathbf{F}(x, y, z) = -\frac{GMm}{r^3}(x, y, z)$$

donde  $G$  es una constante y  $r = \|\mathbf{p}\|_2 = \sqrt{x^2 + y^2 + z^2}$ .

Demuestre que el trabajo realizado por este campo de fuerzas cuando la partícula de masa  $m$  se mueve desde el punto  $\mathbf{p}_1 = (x_1, y_1, z_1)$  hasta el punto  $\mathbf{p}_2 = (x_2, y_2, z_2)$ , (a lo largo un camino regular a trozos que no pasa por  $\mathbf{0}$ ) sólo depende de  $r_1 = \|\mathbf{p}_1\|_2$  y  $r_2 = \|\mathbf{p}_2\|_2$ .

◇ **13.5.14** Una partícula de masa  $m$  se mueve en el espacio  $\mathbb{R}^3$  a lo largo de una curva  $\gamma$  bajo la acción de un campo de fuerzas  $\mathbf{F}$ .

Su energía cinética en el instante  $t$  es  $e(t) = \frac{1}{2}m\mathbf{v}(t)^2$  donde  $\mathbf{v}(t)$  es la velocidad de la partícula en el instante  $t$ . Demuestre que la variación de la energía cinética en un intervalo de tiempo  $[t_0, t_1]$  es igual al trabajo realizado durante dicho intervalo.

Si se supone que el campo de fuerzas  $\mathbf{F}$  tiene una función potencial  $f$ , se dice que  $-f(\mathbf{x})$  es la energía potencial en el punto  $\mathbf{x}$ . La energía mecánica de la partícula en el instante  $t$  es la suma de la energía cinética y de la energía potencial, es decir  $E(t) = e(t) - f(\gamma(t))$ . Demuestre la ley de conservación de la energía mecánica: Si una partícula se mueve sometida a un campo de fuerzas conservativo entonces la energía mecánica  $E(t)$  permanece constante.

◇ **13.5.15** Utilice el teorema de Green para hallar el área de los recintos planos que se indican

i) Recinto encerrado por la elipse  $(x/a)^2 + (y/2)^2 = 1$ .

ii) Recinto encerrado por la hipocicloide  $x^{2/3} + y^{2/3} = a^{2/3}$ .

iii) Un lazo de la rosa de cuatro hojas de ecuación (en polares)  $r = 3 \operatorname{sen} 2t$ .

iv) Recinto limitado por el eje de abscisas y un arco de cicloide

$$x(t) = a(t - \cos t); \quad y(t) = a(1 - \operatorname{sen} t); \quad 0 \leq t \leq 2\pi; \quad (a > 0)$$



◇ **13.5.16** Utilice el teorema de Green para deducir la fórmula que da el área de una región plana descrita en coordenadas polares:

$$M = \{(r \cos t, r \sin t) : \alpha < t < \beta, 0 < r < f(t)\}$$

donde  $f$  se supone de clase  $C^1$ .

◇ **13.5.17** Enuncie y demuestre la versión del teorema de Green para una región de la forma  $M = \overline{B(0, R)} \setminus (B(a, r) \cup B(-a, r))$  donde  $a + r < R$ .

◇ **13.5.18** Sea  $B = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1, x^2 + y^2 - 2x \geq 0, x \geq 0, y \geq 0\}$  y  $\gamma$  el borde de  $B$  orientado positivamente. Calcule:

$$\int_{\gamma} y^2 dx - x^2 dy$$

(Nota: La solución se puede dar en términos de  $S = \text{área}(B)$ ).

# Capítulo 14

## Integrales de superficie

*Área de una superficie. Integral respecto al elemento de área Flujo de un campo de vectores.*

El principal objetivo de este capítulo es formular, con una motivación razonable, la definición de área para una parametrización  $\varphi : U \rightarrow \mathbb{R}^3$ , de clase  $C^1$ , definida en un abierto  $U \subset \mathbb{R}^2$ . Conviene advertir que la definición no proporciona un número ligado a la superficie paramétrica  $S = \varphi(U) \subset \mathbb{R}^3$ , sino un número asociado a la aplicación  $\varphi$  que mide el área recorrida o 'barrida' por el punto  $\varphi(u_1, u_2)$  cuando  $(u_1, u_2)$  recorre una vez el dominio  $U$ . Se consigue una definición intrínseca del área de  $S = \varphi(U)$  cuando  $\varphi : U \rightarrow \mathbb{R}^n$  es una parametrización regular (definición 9.2).

En la sección 14.5 se formula la definición general de área  $k$ -dimensional para una parametrización  $\varphi : U \rightarrow \mathbb{R}^n$  de clase  $C^1$  en un abierto  $U \subset \mathbb{R}^k$ , ( $1 \leq k \leq n$ ), que también proporciona una definición intrínseca de área  $k$ -dimensional de la imagen  $S = \varphi(U) \subset \mathbb{R}^n$  cuando  $\varphi : U \rightarrow \mathbb{R}^n$  es regular.

Sin embargo, hemos preferido comenzar considerando el caso estandar de una superficie paramétrica, que corresponde al caso  $n = 3$  y  $k = 2$ , asumiendo algunos resultados establecidos en las secciones K.1 y 9.1. En el caso de las superficies, para motivar la definición de su área sólo se requiere aceptar que el área del paralelogramo determinado por dos vectores  $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^3$  viene dada por la norma euclídea de su producto vectorial. Los ejemplos que se consideran en la sección 14.2 ponen de manifiesto que la definición 14.16 asigna a superficies sencillas el área que prescribe la geometría elemental. En ellos  $U$  suele ser un recinto plano bastante simple, como un rectángulo, un disco, o un sector circular, o más generalmente un abierto simple medible Jordan, y a veces ocurre que la integral doble que proporciona el área es una genuina integral impropia para cuyo cálculo suele ser útil el teorema del cambio de variable J.15.

Para una función escalar definida sobre la imagen de un camino rectificable, en el capítulo 4 se formuló la definición de integral respecto al arco. Análogamente se define en este capítulo la integral, respecto al elemento de área, de una función escalar definida sobre una superficie que se supone dada en forma paramétrica. Diversas nociones físicas, como la masa y el centro de masa de una lámina delgada

o el flujo de un campo de vectores a través de una superficie, se formulan mediante integrales de superficie apropiadas.

## 14.1. Preliminares geométricos

Para el lector que sólo esté interesado en la definición de área de una superficie resumimos a continuación los resultados elementales de geometría euclídea tridimensional que intervienen en la definición. Estos resultados fueron establecidos con detalle, en una situación general, en la sección K.1.

Tres vectores  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in \mathbb{R}^3$  generan un paralelepípedo  $P(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = L([0, 1]^3)$ , donde  $L : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  es la aplicación lineal  $L(t_1, t_2, t_3) = \sum_{i=1}^3 t_i \mathbf{v}_i$ . Según el teorema J.8 el paralelepípedo es medible Jordan y  $c_3(P(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)) = \det L$ . Si  $\mathbf{v}_i = (v_{i1}, v_{i2}, v_{i3})$ , la matriz de la aplicación lineal  $L$  es la traspuesta de la matriz  $(v_{ij})$  y se obtiene que

$$c_3(P(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)) = \det [(v_{ij})_{1 \leq i, j \leq 3}]$$

Dos vectores  $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^3$  generan un paralelogramo  $P(\mathbf{v}_1, \mathbf{v}_2) = R([0, 1]^2)$ , donde  $R : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  es la aplicación lineal  $R(t_1, t_2) = t_1 \mathbf{v}_1 + t_2 \mathbf{v}_2$ . Este paralelogramo lo podemos considerar sumergido en un subespacio  $E \subset \mathbb{R}^3$  de dimensión 2, (que habitualmente será un espacio tangente a una superficie) y nuestro primer objetivo es mostrar  $P(\mathbf{v}_1, \mathbf{v}_2)$  tiene un área dentro de  $E$  y obtener una fórmula para ella. Debemos comenzar definiendo la clase  $\mathcal{M}_E$  formada por los subconjuntos de  $E$  que tienen contenido, y la medida de sus áreas  $c_E : \mathcal{M}_E \rightarrow [0, +\infty)$ . Para ello utilizamos que  $E$  es un espacio euclídeo con el producto escalar inducido por el producto escalar de  $\mathbb{R}^3$ . Mediante una base ortonormal  $\beta = \{\mathbf{u}_1, \mathbf{u}_2\}$ , este espacio euclídeo queda identificado con  $\mathbb{R}^2$ , a través de la aplicación lineal  $T_\beta((x_1, x_2)) = x_1 \mathbf{u}_1 + x_2 \mathbf{u}_2$ . Utilizando esta identificación se define la familia  $\mathcal{M}_E$  formada por los conjuntos  $M \subset E$  tales que  $M_\beta = T_\beta^{-1}(M)$  es medible Jordan en  $\mathbb{R}^2$ , y para ellos se define el contenido  $c_E(M) = c_2(M_\beta)$ . El hecho de que estas definiciones no dependen de la base ortonormal elegida se puede ver con detalle en la sección K.1.

Dados dos vectores  $\mathbf{v}_1, \mathbf{v}_2 \in E$ , es fácil ver que el paralelogramo  $P(\mathbf{v}_1, \mathbf{v}_2)$  es un subconjunto medible Jordan de  $E$  cuya área viene dada por la fórmula

$$c_E(P(\mathbf{v}_1, \mathbf{v}_2)) = |\det_\beta(\mathbf{v}_1, \mathbf{v}_2)|$$

donde  $\beta = \{\mathbf{u}_1, \mathbf{u}_2\}$  es una base ortonormal de  $E$ , y  $\det_\beta(\mathbf{v}_1, \mathbf{v}_2)$  es el determinante de la matriz formada con las coordenadas de  $\mathbf{v}_1, \mathbf{v}_2 \in E$  respecto a esta base.

El inconveniente de esta fórmula reside en que hay que comenzar eligiendo una base ortonormal y calcular luego las coordenadas de  $\mathbf{v}_1$  y  $\mathbf{v}_2$  respecto a esta base. Por ello es interesante disponer de otra fórmula alternativa como la siguiente

$$c_E(P(\mathbf{v}_1, \mathbf{v}_2)) = \sqrt{|\det(\langle \mathbf{v}_i | \mathbf{v}_j \rangle)_{1 \leq i, j \leq 2}|}$$

en la que sólo intervienen las coordenadas de los vectores  $\mathbf{v}_1, \mathbf{v}_2$  respecto a la base canónica de  $\mathbb{R}^3$  (véase la proposición K.2).

*Producto mixto y producto vectorial.* Dada una terna de vectores  $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ , donde  $\mathbf{v}_i = (v_{i1}, v_{i2}, v_{i3})$ ,  $(i = 1, 2, 3)$  su *producto mixto*, denotado  $[\mathbf{v}_1 \cdot \mathbf{v}_2 \cdot \mathbf{v}_3]$ , se define como el valor del determinante  $\det(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = \det(v_{ij})_{1 \leq i, j \leq 3}$ , cuyo valor absoluto proporciona el volumen del paralelepípedo  $P(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ .

Si a un par ordenado de vectores  $(\mathbf{v}_1, \mathbf{v}_2) \in \mathbb{R}^3 \times \mathbb{R}^3$  le asociamos la aplicación lineal  $L : \mathbb{R}^3 \rightarrow \mathbb{R}$ , definida por  $L(\mathbf{x}) = \det(\mathbf{x}, \mathbf{v}_1, \mathbf{v}_2)$ , en virtud de la proposición B.8 existe un único  $\mathbf{z} \in \mathbb{R}^3$  tal que  $L(\mathbf{x}) = \langle \mathbf{x} | \mathbf{z} \rangle$  para todo  $\mathbf{x} \in \mathbb{R}^3$ , es decir

$$\langle \mathbf{x} | \mathbf{z} \rangle = \det(\mathbf{x}, \mathbf{v}_1, \mathbf{v}_2) = [\mathbf{x} \cdot \mathbf{v}_1 \cdot \mathbf{v}_2]$$

Este vector, denotado  $\mathbf{z} = \mathbf{v}_1 \times \mathbf{v}_2$ , recibe el nombre de *producto vectorial* del par ordenado de vectores  $(\mathbf{v}_1, \mathbf{v}_2)$ . De la definición se deduce que  $\mathbf{z}$  es un vector ortogonal a los vectores  $\mathbf{v}_1, \mathbf{v}_2$ , no nulo si y sólo si estos vectores son linealmente independientes. En este caso, sea  $E \subset \mathbb{R}^3$  el subespacio generado por  $\{\mathbf{v}_1, \mathbf{v}_2\}$  y  $\mathbf{n}$  el vector unitario ortogonal a  $E$  determinado  $\det(\mathbf{n}, \mathbf{v}_1, \mathbf{v}_2) > 0$ . Como  $\langle \mathbf{n} | \mathbf{z} \rangle = \det(\mathbf{n}, \mathbf{v}_1, \mathbf{v}_2) > 0$ , se sigue que  $\mathbf{z}$  tiene la dirección y el sentido de  $\mathbf{n}$ , y su norma euclídea vale

$$\|\mathbf{z}\|_2 = \langle \mathbf{n} | \mathbf{z} \rangle = \det(\mathbf{n}, \mathbf{v}_1, \mathbf{v}_2) = c_3(P(\mathbf{n}, \mathbf{v}_1, \mathbf{v}_2))$$

Como  $\mathbf{n}$  es un vector unitario ortogonal a los vectores  $\mathbf{v}_1, \mathbf{v}_2$ , es geoméricamente evidente que el volumen del paralelepípedo  $P(\mathbf{n}, \mathbf{v}_1, \mathbf{v}_2)$  coincide con el área de su base  $P(\mathbf{v}_1, \mathbf{v}_2)$  (véase el ejercicio K.3) luego la norma euclídea del producto vectorial  $\mathbf{z} = \mathbf{v}_1 \times \mathbf{v}_2$  proporciona el área del paralelogramo generado por los vectores  $(\mathbf{v}_1, \mathbf{v}_2)$

$$\|\mathbf{z}\|_2 = c_E(P(\mathbf{v}_1, \mathbf{v}_2)) = \sqrt{|\det(\langle \mathbf{v}_i | \mathbf{v}_j \rangle)_{1 \leq i, j \leq 2}|}$$

Dados los vectores  $\mathbf{v}_1 = (v_{11}, v_{12}, v_{13})$ ,  $\mathbf{v}_2 = (v_{21}, v_{22}, v_{23})$ , para obtener las coordenadas de  $\mathbf{z} = \mathbf{v}_1 \times \mathbf{v}_2$  respecto a la base canónica de  $\mathbb{R}^3$  basta calcular los productos escalares  $z_j = \langle \mathbf{e}_j | \mathbf{z} \rangle = \det(\mathbf{e}_j, \mathbf{v}_1, \mathbf{v}_2)$ , luego  $\mathbf{z} = \mathbf{v}_1 \times \mathbf{v}_2$  es el vector que se obtiene desarrollando formalmente el determinante

$$\mathbf{z} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \end{vmatrix}$$

$$z_1 = v_{12}v_{23} - v_{13}v_{22}, \quad z_2 = v_{13}v_{21} - v_{11}v_{23}, \quad z_3 = v_{11}v_{22} - v_{12}v_{21}$$

y usando esta fórmula se obtiene que  $\mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_3$ ;  $\mathbf{e}_2 \times \mathbf{e}_3 = \mathbf{e}_1$ ;  $\mathbf{e}_3 \times \mathbf{e}_1 = \mathbf{e}_2$ .

Para ver el significado geométrico de las coordenadas del producto vectorial consideramos la primera coordenada  $z_1 = \det(\mathbf{e}_1, \mathbf{v}_1, \mathbf{v}_2)$ . Desarrollando este determinante por la primera fila se obtiene que  $|z_1| = |\det(\mathbf{v}'_1, \mathbf{v}'_2)|$  donde  $\mathbf{v}'_i = (v_{i2}, v_{i3})$ .

Identificando  $\mathbf{v}'_i$  con la proyección ortogonal de  $\mathbf{v}_i$  sobre  $E_1 = \{\mathbf{x} \in \mathbb{R}^3 : x_1 = 0\}$  podemos interpretar la fórmula  $|z_1| = |\det(\mathbf{v}'_1, \mathbf{v}'_2)|$  diciendo que  $|z_1|$  es el área de la proyección ortogonal del paralelogramo  $P(\mathbf{v}_1, \mathbf{v}_2)$  sobre el plano  $x_1 = 0$ . Esto se recuerda fácilmente teniendo en cuenta que  $z_1 = \langle \mathbf{e}_1 | \mathbf{z} \rangle$ , y por ello  $|z_1| = \|\mathbf{z}\|_2 \cos \theta$ , donde  $\theta$  es el ángulo agudo formado por los vectores  $\mathbf{e}_1$  y  $\mathbf{z}$ .

Análogamente,  $|z_j|$  es el área, medida en el plano  $E_j = \{\mathbf{x} \in \mathbb{R}^3 : x_j = 0\}$ , de la proyección ortogonal de  $P(\mathbf{v}_1, \mathbf{v}_2)$  sobre este plano.

## 14.2. Área de una superficie

Comenzamos recordando las nociones que intervienen en lo que sigue. Una *parametrización* de clase  $C^m$  ( $m \geq 1$ ) y dimensión 2 es una aplicación  $\varphi : U \rightarrow \mathbb{R}^3$  de clase  $C^m$  definida en un abierto  $U \subset \mathbb{R}^2$ . A su imagen  $S = \varphi(U)$  se le suele llamar superficie paramétrica o superficie parametrizada. Todas las parametrizaciones consideradas en este capítulo, aunque no hagamos referencia explícita a su clase (para agilizar la redacción) siempre supondremos que son por lo menos de clase  $C^1$ .

Una parametrización  $\varphi$  de clase  $C^m$  ( $m \geq 1$ ) y dimensión 2 se dice que es *regular* cuando establece un homeomorfismo entre su dominio  $U$  y la imagen  $S = \varphi(U)$  y además, para cada  $\mathbf{u} \in U$ , los vectores  $D_1\varphi(\mathbf{u})$ ,  $D_2\varphi(\mathbf{u})$  son linealmente independientes. En este caso, según el ejemplo 9.6 a)  $S = \varphi(U)$  es una subvariedad diferenciable de  $\mathbb{R}^3$ , de clase  $C^m$  y dimensión 2. Este tipo de subvariedades diferenciables  $S \subset \mathbb{R}^3$ , que son imagen de una parametrización regular de clase  $C^1$  y dimensión 2 las llamaremos más brevemente *superficies paramétricas regulares*.

NOTA: Conviene advertir que algunos textos llaman regulares a las parametrizaciones  $\varphi : U \rightarrow \mathbb{R}^3$  que son de clase  $C^1$  en un abierto  $U \subset \mathbb{R}^2$ , y cumplen que los vectores  $D_1\varphi(\mathbf{u})$ ,  $D_2\varphi(\mathbf{u})$ , son linealmente independientes para cada  $\mathbf{u} \in U$ .

Dos parametrizaciones,  $\varphi : U \rightarrow \mathbb{R}^3$ ,  $\Psi : V \rightarrow \mathbb{R}^3$  de clase  $C^m$  y dimensión 2, se dice que son  $C^m$ -equivalentes cuando existe un  $C^m$ -difeomorfismo  $\mathbf{g} : V \rightarrow U$ , tal que  $\Psi = \varphi \circ \mathbf{g}$ . En este caso es claro que si una de las dos parametrizaciones es regular la otra también lo es. El interés de las parametrizaciones regulares se debe, entre otras cosas, al siguiente resultado.

**Proposición 14.1** *Dos parametrizaciones regulares  $\varphi : U \rightarrow \mathbb{R}^3$ ,  $\Psi : V \rightarrow \mathbb{R}^3$  de clase  $C^m$  y dimensión 2, con la misma imagen son  $C^m$ -equivalentes.*

DEM: Es un caso particular de la proposición H.9. ■

Con el fin de motivar la definición de área de una superficie consideramos una parametrización regular  $\varphi : U \rightarrow \mathbb{R}^3$  definida en un abierto  $U \subset \mathbb{R}^2$ . Si  $\mathbf{u} \in U$  es un punto genérico del dominio, la derivada parcial  $D_1\varphi(\mathbf{u})$  se puede interpretar como el vector velocidad de la curva  $t \rightarrow \varphi(t, u_2)$  en el instante  $t = u_1$ , de modo que un pequeño incremento  $h > 0$  en la variable  $u_1$  hace que el punto  $\varphi(u_1, u_2)$  se desplace a la nueva posición  $\varphi(u_1 + h, u_2)$ , cercana al punto  $\varphi(u_1, u_2) + hD_1\varphi(u_1, u_2)$ . Análogamente, un pequeño incremento  $k > 0$  en la segunda variable  $u_2$  conduce a un punto  $\varphi(u_1, u_2 + k)$  cercano al punto  $\varphi(u_1, u_2) + kD_2\varphi(u_1, u_2)$ . Vemos así que la imagen del rectángulo  $R = [u_1, u_1 + h] \times [u_2, u_2 + k]$  es un trozo de superficie  $\varphi(R)$  que tendrá un área (en sentido intuitivo) próxima a la del paralelogramo determinado por los vectores  $hD_1\varphi(\mathbf{u})$ ,  $kD_2\varphi(\mathbf{u})$ . Según la notación introducida anteriormente, se trata del paralelogramo  $P(hD_1\varphi(\mathbf{u}), kD_2\varphi(\mathbf{u})) = d\varphi(\mathbf{u})(R)$ , que está situado en el espacio tangente  $E = E(\varphi, \mathbf{u})$ . Sabemos que su área, dentro de este plano, viene dada por la norma euclídea del producto vectorial de los vectores que lo determinan

$$c_E[P(hD_1\varphi(\mathbf{u}), kD_2\varphi(\mathbf{u}))] = \|hD_1\varphi(\mathbf{u}) \times kD_2\varphi(\mathbf{u})\|_2 = hk \|\mathbf{N}(\mathbf{u})\|_2$$

donde  $\mathbf{N}(\mathbf{u}) = D_1\boldsymbol{\varphi}(\mathbf{u}) \times D_2\boldsymbol{\varphi}(\mathbf{u})$ , es el llamado *producto vectorial fundamental* de la parametrización  $\boldsymbol{\varphi}$ . Se trata de un vector normal al plano tangente  $E(\boldsymbol{\varphi}, \mathbf{u})$ , cuya norma, según lo que acabamos de ver, es el factor de proporcionalidad entre el área  $hk$  de un rectángulo  $R = [u_1, u_1 + h] \times [u_2, u_2 + k] \subset U$  y el área de su imagen mediante la diferencial,  $d\boldsymbol{\varphi}(\mathbf{u})(R)$ . Esto motiva la consideración de la función

$$P_{\boldsymbol{\varphi}}(\mathbf{u}) := \|\mathbf{N}(\mathbf{u})\|_2 = \sqrt{|\det(\langle D_i\boldsymbol{\varphi}(\mathbf{u}) \mid D_j\boldsymbol{\varphi}(\mathbf{u}) \rangle)_{1 \leq i, j \leq 2}|}$$

que proporciona el área  $c_E[P(D_1\boldsymbol{\varphi}(\mathbf{u}), D_2\boldsymbol{\varphi}(\mathbf{u}))]$ , dentro del plano  $E = E(\boldsymbol{\varphi}, \mathbf{u})$ , del paralelogramo generado por los vectores  $D_1\boldsymbol{\varphi}(\mathbf{u})$ ,  $D_2\boldsymbol{\varphi}(\mathbf{u})$ . Obsérvese que, al ser  $\boldsymbol{\varphi}$  de clase  $C^1(U)$ , la función  $P_{\boldsymbol{\varphi}}$  es continua en  $U$ .

NOTA: Cuando la parametrización  $\boldsymbol{\varphi}$  no es regular los vectores  $D_1\boldsymbol{\varphi}(\mathbf{u})$ ,  $D_2\boldsymbol{\varphi}(\mathbf{u})$  pueden ser linealmente dependientes en algún punto  $\mathbf{u} \in U$ , y en ese caso el valor  $P_{\boldsymbol{\varphi}}(\mathbf{u}) = 0$  lo podemos seguir interpretando como el área del paralelogramo degenerado engendrado por estos vectores (medida en un plano  $E$  que los contenga).

Sea  $p \in \mathcal{P}(A)$  una partición de un rectángulo cerrado  $A \supset U$ . Según hemos visto antes, para cada rectángulo  $R = [u_1, u_1 + h] \times [u_2, u_2 + k] \in \Delta(p)$ , contenido en  $U$ , el número

$$\|\mathbf{N}(\mathbf{u})\|_2 hk = P_{\boldsymbol{\varphi}}(\mathbf{u})hk = P_{\boldsymbol{\varphi}}(\mathbf{u})v_2(R)$$

proporciona una aproximación razonable del área (considerada en sentido intuitivo) del trozo de superficie  $\boldsymbol{\varphi}(R)$ , y podemos tomar las sumas  $\sum_{U \supset R \in \Delta(p)} P_{\boldsymbol{\varphi}}(\mathbf{u})v_2(R)$  como valores que, al refinar  $p$ , aproximan cada vez más el área que deseamos definir.

Si  $U$  es medible Jordan y la función continua  $P_{\boldsymbol{\varphi}}$  está acotada sobre  $U$  las sumas anteriores son sumas de Riemann que al refinar la partición  $p$  aproximan a la integral  $\int_U P_{\boldsymbol{\varphi}}$ , por lo que es razonable formular la siguiente definición

**Definición 14.2** Si  $\boldsymbol{\varphi} : U \rightarrow \mathbb{R}^3$  es una parametrización de clase  $C^1$ , definida en un abierto  $U \subset \mathbb{R}^2$ , se define su área como la integral doble, en sentido impropio

$$\text{Área}(\boldsymbol{\varphi}) = \int_U P_{\boldsymbol{\varphi}}(u_1, u_2) du_1 du_2 \leq +\infty \quad (14.1)$$

de la función continua  $P_{\boldsymbol{\varphi}} = \|D_1\boldsymbol{\varphi} \times D_2\boldsymbol{\varphi}\|_2 = \sqrt{|\det(\langle D_i\boldsymbol{\varphi} \mid D_j\boldsymbol{\varphi} \rangle)_{1 \leq i, j \leq 2}|}$ .

Cuando el abierto  $U \subset \mathbb{R}^2$  es medible Jordan y la función continua  $P_{\boldsymbol{\varphi}}$  está acotada sobre  $U$  la integral que interviene en la definición anterior es una genuina integral de Riemann. En otro caso, según la proposición 12.3, el valor de la integral impropia (posiblemente infinito) viene dado por el límite

$$\text{Área}(\boldsymbol{\varphi}) = \lim_j \int_{K_j} P_{\boldsymbol{\varphi}}(\mathbf{u}) d\mathbf{u}$$

donde  $K_j \subset U$  es cualquier sucesión expansiva en  $U$  formada por compactos medibles Jordan (la definición de sucesión expansiva en  $U$  se ha dado en el capítulo 12).

Obsérvese que en la definición 14.2 no se ha supuesto que  $\varphi$  sea regular, aunque al principio de esta sección habíamos considerado esta condición para motivar la interpretación geométrica de la integral 14.1 como una medida del área de la superficie  $S = \varphi(U)$ , lo que quedará confirmado con el corolario 14.4. Por otra parte, cuando la parametrización  $\varphi$  no es inyectiva, el significado geométrico de  $\text{Área}(\varphi)$  es el de área 'barrida' por  $\varphi(\mathbf{u})$  cuando  $\mathbf{u}$  recorre  $U$  (véase el ejemplo 14.6).

**Proposición 14.3** Si  $\varphi : U \rightarrow \mathbb{R}^3$ ,  $\Psi : V \rightarrow \mathbb{R}^3$  son parametrizaciones  $C^1$ -equivalentes definidas en abiertos  $U, V \subset \mathbb{R}^2$ , entonces  $\text{Área}(\varphi) = \text{Área}(\Psi)$ , es decir

$$\int_U P_\varphi(u_1, u_2) du_1 du_2 = \int_V P_\Psi(v_1, v_2) dv_1 dv_2$$

DEM: La hipótesis significa que existe un  $C^1$ -difeomorfismo  $\mathbf{g} : V \rightarrow U$  tal que para cada  $(v_1, v_2) \in V$  se cumple  $\Psi(v_1, v_2) = \varphi(u_1, u_2)$  donde  $(u_1, u_2) = \mathbf{g}(v_1, v_2) \in U$ . Según la regla de la cadena para las derivadas parciales de una función compuesta

$$\frac{\partial \Psi}{\partial v_1} = \frac{\partial \varphi}{\partial u_1} \frac{\partial g_1}{\partial v_1} + \frac{\partial \varphi}{\partial u_2} \frac{\partial g_2}{\partial v_1}$$

$$\frac{\partial \Psi}{\partial v_2} = \frac{\partial \varphi}{\partial u_1} \frac{\partial g_1}{\partial v_2} + \frac{\partial \varphi}{\partial u_2} \frac{\partial g_2}{\partial v_2}$$

El producto vectorial de estos vectores viene dado por

$$\frac{\partial \Psi}{\partial v_1} \times \frac{\partial \Psi}{\partial v_2} = \frac{\partial \varphi}{\partial u_1} \times \frac{\partial \varphi}{\partial u_2} \left( \frac{\partial g_1}{\partial v_1} \frac{\partial g_2}{\partial v_2} - \frac{\partial g_2}{\partial v_1} \frac{\partial g_1}{\partial v_2} \right)$$

y calculando su norma euclídea se llega a la igualdad

$$P_\Psi(v_1, v_2) = P_\varphi(u_1, u_2) |\det \mathbf{g}'(v_1, v_2)|$$

Entonces, con el cambio de variable  $(u_1, u_2) = \mathbf{g}(v_1, v_2)$ , se obtiene

$$\int_U P_\varphi(\mathbf{u}) d\mathbf{u} = \int_{\mathbf{g}(V)} P_\varphi(\mathbf{u}) d\mathbf{u} = \int_V P_\varphi(\mathbf{g}(\mathbf{v})) |\det \mathbf{g}'(\mathbf{v})| d\mathbf{v} = \int_V P_\Psi(\mathbf{v}) d\mathbf{v}$$

■

**Corolario 14.4** Si  $\varphi : U \rightarrow \mathbb{R}^3$ ,  $\Psi : V \rightarrow \mathbb{R}^3$  son parametrizaciones regulares con la misma imagen, definidas en abiertos  $U, V \subset \mathbb{R}^2$ , entonces  $\text{Área}(\varphi) = \text{Área}(\Psi)$ .

DEM: Es consecuencia directa de las proposiciones H.9 y 14.3. ■

Este corolario permite formular la siguiente definición

**Definición 14.5** El área de una superficie paramétrica regular  $S \subset \mathbb{R}^3$  se define como el área de cualquier parametrización regular  $\varphi : U \rightarrow \mathbb{R}^3$  con  $\varphi(U) = S$ :

$$\text{Area}(S) = \text{Area}(\varphi) = \int_U P_\varphi(u_1, u_2) du_1 du_2$$

Dada una superficie paramétrica regular  $S \subset \mathbb{R}^3$ , para cada abierto  $W \subset \mathbb{R}^n$  con  $W \cap S \neq \emptyset$ , es claro que el abierto relativo  $B = W \cap S$  sigue siendo una superficie paramétrica regular (si  $\varphi : U \rightarrow \mathbb{R}^3$  es una parametrización regular de  $S$ , entonces  $U_0 = \varphi^{-1}(W) \subset \mathbb{R}^2$  es abierto y  $\varphi|_{U_0}$  es una parametrización regular de  $B$ ) luego está definida el área

$$\text{Área}(B) = \int_{\varphi^{-1}(B)} P_{\varphi}(u_1, u_2) du_1 du_2$$

Por razones de simplicidad nos hemos limitado a dar la definición de área para los subconjuntos  $B \subset S$  que son abiertos la topología relativa. Sin embargo, el lector que tenga nociones básicas sobre de la integral de Lebesgue puede adoptar la misma fórmula para definir el área de cualquier subconjunto de Borel  $B \subset S$ .

Por otra parte, si  $S$  es una superficie paramétrica regular y  $\varphi : U \rightarrow \mathbb{R}^3$  y  $\Psi : V \rightarrow \mathbb{R}^3$ , son parametrizaciones regulares de  $S$ , en virtud de las proposiciones 14.1 y 14.3 dado  $\mathbf{p} = \varphi(\mathbf{u}) = \Psi(\mathbf{v}) \in S$  se verifica  $\int_U P_{\varphi}(\mathbf{u}) d\mathbf{u} = \int_V P_{\Psi}(\mathbf{v}) d\mathbf{v}$ . Esto justifica que podemos hablar del *elemento de área* de la superficie regular, denotaremos brevemente  $d\sigma$ , sin especificar la parametrización, interpretando que  $d\sigma$  como un símbolo que actúa sobre cada parametrización concreta  $\varphi$  de  $S$  en el punto genérico  $\mathbf{p} = \varphi(\mathbf{u})$ , dando lugar a la expresión  $d\sigma(\mathbf{p}) = P_{\varphi}(\mathbf{u}) d\mathbf{u}$ , que se suele llamar *elemento de área* de la parametrización  $\varphi$  en el punto  $\mathbf{p} = \varphi(\mathbf{u})$ .

**Ejemplos** En los ejemplos concretos, para el cálculo efectivo de la función  $P_{\varphi}$  se puede elegir entre las dos fórmulas

$$P_{\varphi} = \|D_1\varphi \times D_2\varphi\|_2 = \sqrt{A^2 + B^2 + C^2}$$

$$P_{\varphi} = \sqrt{\det (\langle D_i\varphi | D_j\varphi \rangle)_{1 \leq i, j \leq 2}} = \sqrt{EG - F^2}$$

donde  $A(\mathbf{u}), B(\mathbf{u}), C(\mathbf{u})$  son las componentes del producto vectorial  $\mathbf{N}(\mathbf{u})$ :

$$A = \frac{D(\varphi_2, \varphi_3)}{D(u_1, u_2)} = D_1\varphi_2 D_2\varphi_3 - D_1\varphi_3 D_2\varphi_2$$

$$B = \frac{D(\varphi_3, \varphi_1)}{D(u_1, u_2)} = D_1\varphi_3 D_2\varphi_1 - D_1\varphi_1 D_2\varphi_3$$

$$C = \frac{D(\varphi_1, \varphi_2)}{D(u_1, u_2)} = D_1\varphi_1 D_2\varphi_2 - D_1\varphi_2 D_2\varphi_1$$

y  $E(\mathbf{u}), F(\mathbf{u}), G(\mathbf{u})$  son las funciones definidas por los productos escalares

$$E = \langle D_1\varphi | D_1\varphi \rangle = (D_1\varphi_1)^2 + (D_1\varphi_2)^2 + (D_1\varphi_3)^2$$

$$F = \langle D_1\varphi | D_2\varphi \rangle = D_1\varphi_1 D_2\varphi_1 + D_1\varphi_2 D_2\varphi_2 + D_1\varphi_3 D_2\varphi_3$$

$$G = \langle D_2\varphi | D_2\varphi \rangle = (D_2\varphi_1)^2 + (D_2\varphi_2)^2 + (D_2\varphi_3)^2$$



Con el siguiente ejemplo se pone de manifiesto que, en general, el área de una parametrización no inyectiva  $\varphi : U \rightarrow \mathbb{R}^3$  no proporciona el área de la imagen  $S = \varphi(U)$ , y que su significado geométrico es el de área 'barrida' por el punto  $\varphi(u_1, u_2)$  cuando  $\mathbf{u} = (u_1, u_2)$  recorre el dominio  $U$ .

**Ejemplo 14.6** *Área de un trozo de esfera*

Consideremos el trozo de esfera

$$S_\alpha \subset \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = R^2\}$$

obtenido como imagen del abierto

$$U_\alpha = \{(s, t) \in \mathbb{R}^2 : -\pi/2 < s < \pi/2, 0 < t < \alpha\}$$

mediante la parametrización  $\varphi : U_\alpha \rightarrow \mathbb{R}^3$  definida por

$$\varphi(s, t) = (R \cos s \cos t, R \cos s \sin t, R \sin s)$$

donde  $s$  (resp.  $t$ ) representa la latitud (resp. longitud) de un punto de la esfera.

Para esta parametrización, de clase  $C^\infty$  y dimensión 2, se tiene

$$D_1\varphi(s, t) = (-R \sin s \cos t, -R \sin s \sin t, R \cos s)$$

$$D_2\varphi(s, t) = (-R \cos s \sin t, R \cos s \cos t, 0)$$

$$E(s, t) = \|D_1\varphi(s, t)\|_2^2 = R^2$$

$$G(s, t) = \|D_2\varphi(s, t)\|_2^2 = R^2 \cos^2 s$$

$$F(s, t) = \langle D_1\varphi(s, t) | D_2\varphi(s, t) \rangle = 0$$

$$P_\varphi(s, t) = \sqrt{E(s, t)G(s, t) - F(s, t)^2} = R^2 |\cos s|,$$

y teniendo en cuenta que  $\cos s > 0$  cuando  $s \in (-\pi/2, \pi/2)$  se obtiene

$$\text{Área}(\varphi) = \int_U R^2 \cos^2 s \, ds \, dt = R^2 \int_0^\alpha \left( \int_{-\pi/2}^{+\pi/2} \cos s \, ds \right) dt = 2\alpha R^2$$

Para  $\alpha > 2\pi$ , todas las parametrizaciones  $\varphi : U_\alpha \rightarrow \mathbb{R}^3$  tienen la misma imagen  $S_\alpha = S \setminus \{(0, 0, R), (0, 0, -R)\}$  y sin embargo las áreas son distintas. Cuando  $\alpha \leq \pi$  la parametrización  $\varphi : U_\alpha \rightarrow \mathbb{R}^3$  es regular (véase el ejemplo H.7) y el número  $\text{Área}(\varphi)$ , que sólo depende de la imagen  $S_\alpha = \varphi(U_\alpha)$ , coincide con el área que asigna la geometría elemental. Así, el área de  $S_{2\pi}$  (esfera completa, salvo el meridiano  $\{(x, y, z) \in S : x \geq 0, y = 0\}$ ) es  $4\pi R^2$ , y el área de la semiesfera  $S_\pi$  es  $2\pi R^2$ . ■

Con los ejemplos que siguen se pone de manifiesto que la definición 14.16 asigna a superficies sencillas el área que prescribe la geometría elemental.

**Ejemplo 14.7** *Área de una superficie dada en forma explícita*

Sea  $f : U \rightarrow \mathbb{R}^3$  una función de clase  $C^1(U)$  en un abierto  $U \subset \mathbb{R}^2$ . La superficie en forma explícita  $S = \{(x, y, f(x, y)) : (x, y) \in U\}$  es la imagen de  $U$  mediante la parametrización  $\varphi : U \rightarrow \mathbb{R}^3$  definida por  $\varphi(x, y) = (x, y, f(x, y))$ . Es fácil ver que esta parametrización es regular y así podemos considerar el área de su imagen  $S$ , que vendrá dada por la fórmula 14.1. Ahora  $D_1\varphi = (1, 0, D_1f)$ ,  $D_2\varphi = (0, 1, D_2f)$ , y el producto vectorial fundamental es  $\mathbf{N} = (-D_1\varphi, -D_2\varphi, 1)$ , luego  $P_\varphi = \|\mathbf{N}\|_2 = \sqrt{1 + (D_1f)^2 + (D_2f)^2}$ , y se llega a la fórmula

$$\text{Área}(S) = \text{Área}(\varphi) = \int_U \sqrt{1 + (D_1f(x, y))^2 + (D_2f(x, y))^2} dx dy \quad (14.2)$$

*Área de un trozo de plano:*

Aplicando la fórmula 14.2 podemos ver que el área de un trozo de plano no paralelo al eje  $Oz$ ,  $E = \{(x, y, z) : ax + by = z\} \subset \mathbb{R}^3$ , es igual al área de su proyección en el plano  $(x, y)$  multiplicada por  $1/\cos \alpha$  donde  $\alpha$  es el ángulo agudo que determina el eje  $Oz$  con la normal al plano (regla del coseno). Obsérvese que al ser  $\mathbf{N} = (-a, -b, 1)$  normal al plano, el ángulo agudo  $\alpha$  es el que cumple  $1 = \langle \mathbf{e}_3 | \mathbf{N} \rangle = \|\mathbf{N}\|_2 \cos \alpha$ .

Si consideramos un abierto  $U \subset \mathbb{R}^2$  de área finita, según la fórmula anterior el área del trozo de plano  $S = \{(x, y, z) \in E : (x, y) \in U\}$  viene dada por la integral

$$\text{Área}(S) = \int_U \sqrt{1 + a^2 + b^2} dx dy = \|\mathbf{N}\|_2 \text{Área}(U) = (1/\cos \alpha) \text{Área}(U)$$

El lector interesado puede comprobar que si el abierto  $U \subset \mathbb{R}^2$  es medible Jordan en entonces  $S$  es un subconjunto medible Jordan de  $E$  y  $c_E(S) = (1/\cos \alpha) \text{Área}(U)$ .

*Área de la semiesfera:* Otra aplicación de la fórmula 14.2 permite obtener fácilmente el área de la semiesfera  $S = \{(x, y, z) : x^2 + y^2 + z^2 = R^2, z > 0\}$ , que es la gráfica de la función  $f : U \rightarrow \mathbb{R}$  definida en el abierto  $U = \{(x, y) : x^2 + y^2 < R^2\}$ . Con un cálculo sencillo se obtiene

$$1 + (D_1f(x, y))^2 + (D_2f(x, y))^2 = \frac{R^2}{R^2 - x^2 - y^2}$$

luego

$$\text{Área}(S) = R \int_U \frac{dx dy}{\sqrt{R^2 - x^2 - y^2}}$$

Obsérvese que en este caso la fórmula 14.2 (que es un caso especial de 14.1) conduce a una genuina integral impropia pues la función que aparece bajo la integral no está acotada en  $U$ . Su valor se puede calcular con un cambio de variable a coordenadas polares (véase el teorema J.15) con el que se obtiene

$$\text{Área}(S) = \int_0^{2\pi} d\theta \int_0^R \frac{r dr}{\sqrt{R^2 - r^2}} = 2\pi R^2$$

■

**Ejemplo 14.8** *Área de un trozo de superficie cónica*

Sea  $\varphi : U \rightarrow \mathbb{R}^3$  definida en  $U = \{(r, \theta) : 0 < r < R, 0 < \theta < \alpha\}$ , por

$$\varphi(r, \theta) = (r \cos \theta, r \sin \theta, ar)$$

donde  $0 < \alpha \leq 2\pi$ , y  $a > 0$ . Es fácil comprobar que  $\varphi$  es una parametrización regular de un trozo  $S = \varphi(U)$  del cono  $\{(x, y, z) : x^2 + y^2 = (z/a)^2\}$  que se desarrolla según un sector circular determinado por un arco de circunferencia de radio  $\rho = \sqrt{1 + a^2}R$  y longitud  $2\pi R$ . Con razonamientos de geometría elemental se obtiene que el área del sector es  $\pi\rho = \pi R^2\sqrt{1 + a^2}$ , luego esta debe ser el área del trozo de cono que proporciona la fórmula 14.1. Efectivamente, con un cálculo elemental se obtiene el producto vectorial fundamental

$$\mathbf{N}(r, \theta) = D_1\varphi(r, \theta) \times D_2\varphi(r, \theta) = (ar \cos \theta, -ar \sin \theta, r)$$

luego

$$\text{Área}(\varphi) = \int_U \|\mathbf{N}(r, \theta)\|_2 \, dr \, d\theta = \int_U r\sqrt{1 + a^2} \, dr \, d\theta = \pi R^2\sqrt{1 + a^2}$$

**Ejemplo 14.9** *Área de un trozo de superficie cilíndrica*

En el plano  $(x, y)$  se considera una curva  $C$  dada en forma paramétrica como imagen de un camino  $\gamma(t) = (x(t), y(t))$ , de clase  $C^1$  en un intervalo abierto  $U = (a, b) \subset \mathbb{R}^2$ . A lo largo de la curva se levanta una valla cuya altura en el punto  $(x, y) \in C$  viene dada por una función de dos variables  $h(x, y)$  que se supone definida y de clase de clase  $C^1$  en un abierto  $\Omega \supset C$ . La intuición nos dice que el área de la valla debe ser igual a la integral, respecto al arco, de la función  $h$  sobre la curva  $C$ . Esta conjetura queda avalada con los cálculos que siguen: La valla  $S$  es un trozo de superficie cilíndrica que se parametriza en

$$U = \{(s, t) : a < t < b, 0 < s < h(x(t), y(t))\}$$

mediante la función de clase  $C^1$ ,  $\varphi(s, t) = (x(t), y(t), s)$ . (Obsérvese que la continuidad de la función compuesta  $t \rightarrow g(t) = h(x(t), y(t))$  garantiza que  $U$  es abierto). Es inmediato que  $D_1\varphi(s, t) = (0, 0, 1)$ ,  $D_2\varphi(s, t) = (x'(t), y'(t), 0)$ , luego el producto vectorial fundamental vale  $\mathbf{N}(s, t) = (-y'(t), x'(t), 0)$  y así se obtiene que

$$\begin{aligned} \text{Área}(\varphi) &= \int_U \sqrt{x'(t)^2 + y'(t)^2} \, ds \, dt = \\ &= \int_a^b dt \int_0^{g(t)} \|\gamma'(t)\|_2 \, ds = \int_a^b h(x(t), y(t)) \|\gamma'(t)\|_2 \, dt \end{aligned}$$

y la última integral no es otra cosa que la integral respecto al arco de la función  $h$ .

Si  $\gamma$  es regular se comprueba fácilmente que  $\varphi$  es regular, y por lo tanto tenemos derecho a decir que  $\text{Área}(\varphi)$  es el área de la valla  $S = \varphi(U)$ . En cualquier caso es razonable admitir que el número  $\text{Área}(\varphi)$  mide el área de la valla. ■

### 14.3. Integral respecto al elemento de área

Para motivar la definición consideremos una lámina delgada de un material no homogéneo cuya forma queda descrita mediante una superficie  $S = \varphi(U)$  que viene dada como imagen de una parametrización  $\varphi : U \rightarrow \mathbb{R}^3$  definida en un abierto  $U \subset \mathbb{R}^2$ . Para cada  $\mathbf{p} \in S$  sea  $f(\mathbf{p}) \geq 0$  la densidad del material con el que se ha construido la lámina. Supongamos, para simplificar el asunto, que la función  $f$  es continua sobre  $S$  y que  $\varphi$  es una parametrización regular de clase  $C^1$  cuyo dominio  $U$  es un rectángulo. Mediante una partición  $p \in \mathcal{P}(U)$  descomponemos  $U$  en un número finito de pequeños rectángulos elementales  $\{U_j : 1 \leq j \leq m\}$ , cuyas imágenes  $S_j = \varphi(U_j)$  son trozos de superficie de área

$$\text{Área}(S_j) = \int_{U_j} P_\varphi(\mathbf{u}) d\mathbf{u}$$

Como  $P_\varphi$  es continua existe  $\mathbf{u}_j \in U_j$  tal que  $\text{Área}(S_j) = P_\varphi(\mathbf{u}_j) \text{Área}(U_j)$  luego un valor aproximado de la masa del trozo  $S_j$  será  $\text{Área}(S_j)f(\mathbf{p}_j)$ , donde  $\mathbf{p}_j = \varphi(\mathbf{u}_j) \in S_j$ . Así podemos asumir que una aproximación razonable de la masa total de la lámina viene dada por las sumas de Riemann

$$\sum_{j=1}^m f(\mathbf{p}_j) \text{Área}(S_j) = \sum_{j=1}^m f(\varphi(\mathbf{u}_j)) P_\varphi(\mathbf{u}_j) \text{Área}(U_j)$$

que aproximan el valor de la integral  $\int_U f(\varphi(\mathbf{u})) P_\varphi(\mathbf{u}) d\mathbf{u}$ .

Refinando la partición  $p \in \mathcal{P}(U)$  cabe esperar que se obtengan aproximaciones cada vez mejores de la masa de la lámina, por lo que es razonable definir la masa total de la lámina mediante esta integral. Esta integral también se puede interpretar como límite de sumas de tipo de Riemann  $\sum_{j=1}^m f(\mathbf{p}_j) \text{Área}(S_j)$ , donde  $\mathbf{p}_j \in S_j$  y  $S_j$  recorre los elementos de una 'partición' finita de  $S$ , en trozos de superficie, generados por una partición adecuada del dominio  $U$ .

**Definición 14.10** Sea  $\varphi : U \rightarrow \mathbb{R}^3$  de clase  $C^1$  definida en un abierto  $U \subset \mathbb{R}^2$  y  $f : S \rightarrow \mathbb{R}$  una función definida sobre  $S = \varphi(U)$ . Si  $(f \circ \varphi)P_\varphi$  es absolutamente integrable sobre  $U$  se dice que  $f$  es integrable respecto a  $\varphi$  y se define

$$\int_\varphi f = \int_U f(\varphi(\mathbf{u})) P_\varphi(\mathbf{u}) d\mathbf{u} = \int_U f(\varphi(\mathbf{u})) \|\mathbf{N}(\mathbf{u})\|_2 d\mathbf{u}$$

donde  $\mathbf{N}(\mathbf{u}) = D_1\varphi(\mathbf{u}) \times D_2\varphi(\mathbf{u})$  es el producto vectorial fundamental asociado a la parametrización  $\varphi$ .

Para la integral de una función  $f$  respecto a una parametrización  $\varphi$  también se suelen utilizar las notaciones

$$\int_\varphi f d\sigma; \int_\varphi f dS; \int_\varphi dA$$

En lo que sigue adoptamos la primera de ellas que sugiere la siguiente regla para obtener la definición 14.10: Se sustituye  $\mathbf{p} = \boldsymbol{\varphi}(\mathbf{u})$  en la expresión  $f(\mathbf{p})d\sigma(\mathbf{p})$  recordando que el símbolo  $d\sigma$  al actuar en un punto genérico  $\mathbf{p} = \boldsymbol{\varphi}(\mathbf{u})$  produce la expresión  $d\sigma(\mathbf{p}) = P_{\boldsymbol{\varphi}}(\mathbf{u})d\mathbf{u}$ .

Según lo que se ha visto en la sección 14.2 al motivar la definición de área de una superficie,  $P_{\boldsymbol{\varphi}}(\mathbf{u}) = P_{\boldsymbol{\varphi}}(u_1, u_2)$  es el coeficiente por el que hay que multiplicar el área  $d\mathbf{u} = du_1 du_2$  del rectángulo elemental  $[u_1, u_1 + du_1] \times [u_2, u_2 + du_2]$ , situado en el plano  $(u_1, u_2)$  para obtener el área  $d\sigma(\mathbf{p})$  de su imagen sobre la superficie (un cuadrilátero curvilíneo que para valores pequeños de los incrementos  $du_1, du_2$  se confunde con un paralelogramo de área  $P_{\boldsymbol{\varphi}}(u_1, u_2)du_1du_2 = P_{\boldsymbol{\varphi}}(\mathbf{u})d\mathbf{u}$  situado en el plano tangente a  $S$  en  $\mathbf{p}$ ).

**Proposición 14.11** Sean  $\boldsymbol{\varphi}_j : U_j \rightarrow \mathbb{R}^3$  parametrizaciones  $C^1$ -equivalentes de clase  $C^1$  definidas en los abiertos  $U_j \subset \mathbb{R}^2$ ,  $j = 1, 2$ . Una función  $f : S \rightarrow \mathbb{R}$  definida sobre  $S = \boldsymbol{\varphi}_1(U_1) = \boldsymbol{\varphi}_2(U_2)$ , es integrable respecto a  $\boldsymbol{\varphi}_1$  si y sólo si es integrable respecto a  $\boldsymbol{\varphi}_2$ , y en ese caso

$$\int_{\boldsymbol{\varphi}_1} f d\sigma = \int_{\boldsymbol{\varphi}_2} f d\sigma$$

DEM: La hipótesis significa que existe un  $C^1$ -difeomorfismo  $\mathbf{g} : U_1 \rightarrow U_2$  tal que  $\boldsymbol{\varphi}_1 = \boldsymbol{\varphi}_2 \circ \mathbf{g}$ . Según la demostración de la proposición 14.16 se verifica

$$P_{\boldsymbol{\varphi}_1} = (P_{\boldsymbol{\varphi}_2} \circ \mathbf{g})|\det \mathbf{g}'|$$

Según el teorema del cambio de variable J.15 podemos afirmar que  $(f \circ \boldsymbol{\varphi}_2)P_{\boldsymbol{\varphi}_2}$  es integrable sobre  $U_2 = \mathbf{g}(U_1)$  si y sólo si

$$(f \circ \boldsymbol{\varphi}_2 \circ \mathbf{g})(P_{\boldsymbol{\varphi}_2} \circ \mathbf{g})|\det \mathbf{g}'| = (f \circ \boldsymbol{\varphi}_1)P_{\boldsymbol{\varphi}_1}$$

es integrable sobre  $U_1$  y en ese caso las integrales coinciden

$$\int_{U_2} (f \circ \boldsymbol{\varphi}_2)P_{\boldsymbol{\varphi}_2} = \int_{U_1} (f \circ \boldsymbol{\varphi}_1)P_{\boldsymbol{\varphi}_1}$$

■

Puesto que dos parametrizaciones regulares con la misma imagen  $S$  son  $C^1$  equivalentes (véase la proposición 14.1) tomando como base la proposición 14.11 se puede definir la integral de una función  $f : S \rightarrow \mathbb{R}$ , sobre una superficie paramétrica regular  $S \subset \mathbb{R}^3$

**Definición 14.12** Una función  $f : S \rightarrow \mathbb{R}$  definida sobre una superficie paramétrica regular  $S \subset \mathbb{R}^3$  se dice que es integrable sobre  $S$  cuando  $f$  es integrable respecto a una (o cualquier) parametrización regular  $\boldsymbol{\varphi} : U \rightarrow \mathbb{R}^3$  tal que  $\boldsymbol{\varphi}(U) = S$  (lo que significa que  $(f \circ \boldsymbol{\varphi})P_{\boldsymbol{\varphi}}$  es absolutamente integrable sobre  $U$ ). En ese caso la integral de  $f$  sobre  $S$ , denotada  $\int_S f d\sigma = \int_S f(\mathbf{p})d\sigma(\mathbf{p})$ , se define en términos de alguna parametrización regular  $\boldsymbol{\varphi}$  de  $S$ :

$$\int_S f(\mathbf{p})d\sigma(\mathbf{p}) := \int_{\boldsymbol{\varphi}} f = \int_U (f \circ \boldsymbol{\varphi})(\mathbf{u})P_{\boldsymbol{\varphi}}(\mathbf{u})d\mathbf{u}$$

Cuando  $f$  es la función constante 1, según la notación introducida en la definición 14.12, podemos escribir la sugestiva fórmula  $\text{Área}(S) = \int_S d\sigma$  que se suele leer diciendo que el área total de  $S$  es la suma de sus elementos de área.

NOTA: Si  $S = \varphi(U) \subset \mathbb{R}^3$  donde  $\varphi$  es una parametrización inyectiva de clase  $C^1$  no regular, también se suele utilizar la notación  $\int_S f d\sigma = \int_S f(\mathbf{p}) d\sigma(\mathbf{p})$ , para designar el valor de la integral  $\int_{\varphi} f$ , cuando por el contexto se sobreentiende cual es la parametrización  $\varphi$  (en el supuesto de que esta integral exista).

Las consideraciones previas a la definición 14.10 motivan la siguiente:

**Definición 14.13** *La masa de una lámina que tiene la forma de una superficie paramétrica regular  $S \subset \mathbb{R}^3$ , se dice que está, distribuida según la función de densidad  $\rho : S \rightarrow [0, +\infty)$ , cuando  $\rho$  es integrable sobre  $S$  y para porción abierta  $B \subset S$  (en la topología relativa de  $S$ ) la masa de  $B$  viene dada por la integral  $\mu(B) = \int_B \rho(\mathbf{x}) d\sigma(\mathbf{x})$ .*

*En estas condiciones, el centro de masa, o centro de gravedad de la lámina, es el punto  $\mathbf{b} = (b_1, b_2, b_3) \in \mathbb{R}^3$  de coordenadas*

$$b_j = \frac{1}{\mu(S)} \int_{\varphi} x_j \rho(\mathbf{x}) d\sigma(\mathbf{x}) = \frac{1}{\mu(S)} \int_U \varphi_j(s, t) \rho(\varphi(s, t)) P_{\varphi}(s, t) ds dt$$

*y el momento de inercia de la lámina respecto a un eje  $e$  viene dado por la integral*

$$I_e = \int_S \delta^2(\mathbf{x}) \rho(\mathbf{x}) d\sigma(\mathbf{x})$$

*donde  $\delta(\mathbf{p})$  es la distancia del punto  $\mathbf{p}$  a la recta  $e$ .*

En el caso de una lámina homogénea de densidad constante  $\rho(\mathbf{x}) \equiv \rho$ , al centro de masa se le suele llamar baricentro. En este caso  $\mu(S) = \rho \text{Área}(S)$ , y las coordenadas del baricentro  $\mathbf{b} = (b_1, b_2, b_3)$  vienen dadas por las integrales

$$b_j = \frac{1}{\text{Área}(S)} \int_{\varphi} x_j d\sigma(\mathbf{x}) = \frac{1}{\text{Área}(S)} \int_U \varphi_j(s, t) P_{\varphi}(s, t) ds dt$$

Con el ejercicio resuelto 14.24 se pone de manifiesto que, en las condiciones de la definición 14.13, si la función de densidad  $\rho$  es continua en  $\mathbf{p} \in S$ , entonces  $\rho(\mathbf{p})$  es el límite, cuando  $r \rightarrow 0$ , de los cocientes entre la masa y el área de las porciones de superficie  $S_r(\mathbf{p}) = S \cap B(\mathbf{p}, r)$ .

## 14.4. Flujo de un campo de vectores

En las aplicaciones a la Física las funciones de tres variables con valores vectoriales se suelen llamar campos de vectores, (y a las funciones con valores reales, campos escalares). Esta sección está dedicada a la noción de flujo de un campo de vectores a través de una superficie. Esta noción, que desempeña un papel importante en el

estudio del movimiento de los fluidos, sirve para medir la cantidad neta de fluido que pasa a través de una superficie en un sentido determinado. Para explicar el significado preciso de la última frase debemos comenzar con la noción de orientación de una superficie. De momento sólo consideraremos superficies paramétricas regulares para las que la noción de orientación se formula en una forma bastante simple.

**Orientación de una superficie paramétrica regular.** Si  $S \subset \mathbb{R}^3$  es una superficie paramétrica regular y  $\varphi_j : U_j \rightarrow S$ ,  $j = 1, 2$ , son parametrizaciones regulares de  $S$ , según la proposición H.9,  $\mathbf{g} = \varphi_2^{-1} \circ \varphi_1 : U_1 \rightarrow U_2$  es un difeomorfismo. Se dice que  $\varphi_1$  y  $\varphi_2$  tienen la misma orientación (resp. orientaciones opuestas) cuando  $\det \mathbf{g}'(\mathbf{u}) > 0$  (resp.  $\det \mathbf{g}'(\mathbf{u}) < 0$ ) para todo  $\mathbf{u} \in U_1$ . Es fácil ver que así queda definida una relación de equivalencia en la familia de todas parametrizaciones regulares de  $S$ . Se dice que la superficie paramétrica regular  $S$  está orientada cuando ha sido elegida una clase de equivalencia, declarando como positivas a todas las parametrizaciones de la clase. Habitualmente esta clase positiva se suele determinar eligiendo uno de sus representantes  $\varphi$ , y en ese caso se dice que  $S$  está orientada mediante la parametrización  $\varphi$ .

En general, si  $\varphi_j : U_j \rightarrow S$ ,  $j = 1, 2$ , son dos parametrizaciones regulares de  $S$ , el difeomorfismo  $\mathbf{g} = \varphi_2^{-1} \circ \varphi_1 : U_1 \rightarrow U_2$  cumple que  $\det \mathbf{g}'(\mathbf{u}) \neq 0$  para cada  $\mathbf{u} \in U_1$ . Cuando la superficie  $S$  es conexa,  $U_1$  también lo es (porque  $\varphi_1 : U_1 \rightarrow S$  es un homeomorfismo) y así la función continua  $\det \mathbf{g}'(\mathbf{u}) \neq 0$  no puede cambiar de signo en  $U_1$ , luego o bien  $\det \mathbf{g}'(\mathbf{u}) > 0$  para todo  $\mathbf{u} \in U_1$ , o bien  $\det \mathbf{g}'(\mathbf{u}) < 0$  para todo  $\mathbf{u} \in U_1$ . Es decir, cuando la superficie  $S$  es conexa, dos parametrizaciones regulares de  $S$  o tienen la misma orientación o tienen orientaciones opuestas, luego  $S$  sólo admite dos orientaciones, una opuesta de la otra.

Un método alternativo para orientar una superficie paramétrica regular  $S$  se basa en la consideración de vectores unitarios normales a la superficie: Si  $\varphi : U \rightarrow S$  es una parametrización regular de  $S$ , en cada punto  $\mathbf{p} = \varphi(\mathbf{u}) \in S$  se puede definir un vector normal unitario  $\mathbf{n}(\mathbf{p}) = \mathbf{N}(\mathbf{u}) / \|\mathbf{N}(\mathbf{u})\|_2$  donde  $\mathbf{N}(\mathbf{u}) = D_1\varphi(\mathbf{u}) \times D_2\varphi(\mathbf{u})$  es el producto vectorial fundamental de la parametrización  $\varphi$ . Obsérvese que  $\mathbf{n}(\mathbf{p})$  depende continuamente de  $\mathbf{p}$  porque al ser  $\varphi$  una parametrización regular la transformación inversa  $\mathbf{u} = \varphi^{-1}(\mathbf{p})$  es continua. Además, este campo continuo de vectores normales unitarios  $\mathbf{n} : S \rightarrow \mathbb{R}^3$  sólo depende de la orientación porque si  $\varphi_j : U_j \rightarrow S$ ,  $1 \leq j \leq 2$ , son parametrizaciones regulares de  $S$  con la misma orientación dado un punto  $\mathbf{p} = \varphi_1(\mathbf{u}) = \varphi_2(\mathbf{v}) \in S$ , según la demostración de la proposición 14.3, los vectores normales  $\mathbf{N}_1(\mathbf{u}) = D_1\varphi_1(\mathbf{u}) \times D_2\varphi_1(\mathbf{u})$ ,  $\mathbf{N}_2(\mathbf{v}) = D_1\varphi_2(\mathbf{v}) \times D_2\varphi_2(\mathbf{v})$  cumplen la condición  $\mathbf{N}_2(\mathbf{v}) = \det \mathbf{g}'(\mathbf{v}) \mathbf{N}_1(\mathbf{u})$ , con  $\det \mathbf{g}'(\mathbf{v}) > 0$ , y por lo tanto

$$\mathbf{n}_1(\mathbf{p}) = \mathbf{N}_1(\mathbf{u}) / \|\mathbf{N}_1(\mathbf{u})\|_2 = \mathbf{N}_2(\mathbf{v}) / \|\mathbf{N}_2(\mathbf{v})\|_2 = \mathbf{n}_2(\mathbf{p})$$

En definitiva, cuando se orienta una superficie paramétrica regular  $S$  queda determinado un campo continuo de vectores normales unitarios  $\mathbf{n} : S \rightarrow \mathbb{R}^3$ , que calificaremos como positivos para la orientación de  $S$ . (la orientación opuesta en  $S$  produce el campo continuo de vectores normales unitarios  $-\mathbf{n}(\mathbf{p})$ ). Recíprocamente, si  $S \subset \mathbb{R}^3$  es una superficie paramétrica regular, sobre la que se ha definido un campo continuo de vectores normales unitarios  $\mathbf{n} : S \rightarrow \mathbb{R}^3$ , es fácil ver que  $S$  queda orientada

declarando como positivas a las parametrizaciones regulares cuyo campo de vectores normales unitarios coincide con  $\mathbf{n}$ . ■

Una superficie paramétrica regular orientada  $S$  conviene imaginarla como una lámina muy delgada que tiene dos caras, la positiva y la negativa, determinadas por la condición de que el campo continuo de vectores normales unitarios positivos  $\mathbf{n}(\mathbf{p})$  apunta siempre desde la cara negativa hacia la cara positiva.

NOTA: La noción de orientación para una superficie paramétrica regular es un caso particular de la que se formula en la sección H.3 para las subvariedades diferenciables de  $\mathbb{R}^n$ . Los dos métodos que hemos mencionado aquí para orientar una superficie paramétrica regular son versiones particulares de los considerados en la sección H.3.

**Flujo a través de una superficie paramétrica regular orientada.** La noción de flujo de un campo de vectores a través de una superficie paramétrica orientada se motiva considerando el movimiento de un fluido que ocupa un cierto recinto  $\Omega \subset \mathbb{R}^3$ . El movimiento durante un intervalo de tiempo  $[0, T]$  queda descrito mediante un campo de vectores  $\mathbf{F} : \Omega \times [0, T] \rightarrow \mathbb{R}^3$  que proporciona, en el instante  $t \in [0, T]$ , la velocidad de la corriente  $\mathbf{F}(t, \mathbf{p})$  en el punto  $\mathbf{p} \in \Omega$  (la velocidad de la partícula que en ese instante pasa por este punto). En lo que sigue supondremos, para simplificar el asunto, que el movimiento del fluido es estacionario, lo que significa que su campo de velocidades no depende del tiempo, de modo que todas las partículas que pasan por un punto  $(x, y, z)$  lo hacen siempre a la misma velocidad  $\mathbf{F}(x, y, z)$ . En este caso el movimiento del fluido se describe con un campo de vectores  $\mathbf{F} : \Omega \rightarrow \mathbb{R}^3$ .

Supongamos que  $S = \varphi(U) \subset \Omega$  es una superficie paramétrica regular orientada mediante una parametrización regular  $\varphi : U \rightarrow \mathbb{R}^3$  y que  $\mathbf{n} : S \rightarrow \mathbb{R}^3$  es el campo de vectores normales unitarios positivos para la orientación. Deseamos medir el volumen de fluido que atraviesa la superficie  $S$  en una unidad de tiempo, desde la cara negativa, hacia la cara positiva.

Al finalizar ese intervalo de tiempo la partícula que había pasado por el punto  $\mathbf{p} = \varphi(\mathbf{u})$  en el instante  $t$ , con velocidad  $\mathbf{F}(\mathbf{p})$ , se encuentra en el punto  $\mathbf{p} + \mathbf{F}(\mathbf{p})$ , y las partículas que, durante este intervalo de tiempo, han pasado a través del elemento de superficie  $d\sigma(\mathbf{p})$ , ocupan un paralelepípedo elemental, trasladado del generado por los vectores  $\mathbf{F}(\mathbf{p})$ ,  $D_1\varphi(\mathbf{u})du_1$ ,  $D_2\varphi(\mathbf{u})du_2$ . El volumen de este paralelepípedo elemental viene dado por el valor absoluto de  $\langle \mathbf{F}(\mathbf{p}) \mid \mathbf{N}(\mathbf{u}) \rangle du_1 du_2$ , donde  $\mathbf{N}(\mathbf{u}) = D_1\varphi(\mathbf{u}) \times D_2\varphi(\mathbf{u})$ . Como  $\mathbf{N}(\mathbf{u})$  es un vector normal a  $S$  en  $\mathbf{p} = \varphi(\mathbf{u})$  que tiene la dirección del vector unitario  $\mathbf{n}(\mathbf{p})$ , se cumple

$$\langle \mathbf{F}(\mathbf{p}) \mid \mathbf{N}(\mathbf{u}) \rangle du_1 du_2 = \langle \mathbf{F}(\mathbf{p}) \mid \mathbf{n}(\mathbf{p}) \rangle \|\mathbf{N}(\mathbf{u})\|_2 du_1 du_2$$

Obsérvese que en los puntos donde el producto escalar  $\langle \mathbf{F}(\mathbf{p}) \mid \mathbf{n}(\mathbf{p}) \rangle$  es positivo (resp. negativo) el fluido atraviesa la superficie desde la cara negativa (resp. positiva) hacia la cara positiva (resp. negativa). Si prescindimos del valor absoluto en el producto escalar entonces  $\langle \mathbf{F}(\mathbf{p}) \mid \mathbf{N}(\mathbf{u}) \rangle du_1 du_2$  proporciona el valor signado del elemento de volumen, con signo  $+$  en los puntos donde el fluido pasa por la superficie en el sentido determinado por su orientación, y con signo  $-$  en los puntos donde lo



hace en el sentido opuesto. Según esto, el volumen neto de fluido que pasa a través de la superficie  $S$  desde su cara negativa hacia su cara positiva, en la unidad de tiempo, será la suma de los volúmenes elementales (con signo)  $\langle \mathbf{F}(\mathbf{p}) | \mathbf{N}(\mathbf{u}) \rangle du_1 du_2$ , y vendrá dado por la integral de superficie

$$\int_U \langle \mathbf{F}(\boldsymbol{\varphi}(\mathbf{u})) | \mathbf{n}(\boldsymbol{\varphi}(\mathbf{u})) \rangle \|\mathbf{N}(\mathbf{u})\|_2 du_1 du_2 = \int_S \langle \mathbf{F}(\mathbf{p}) | \mathbf{n}(\mathbf{p}) \rangle d\sigma(\mathbf{p})$$

Es decir, el volumen de fluido que pasa a través de  $S$  en la unidad de tiempo, desde el lado al que apunta  $-\mathbf{n}$ , al lado contrario, es la integral de superficie, respecto al elemento de área, de la componente del vector velocidad en dirección de  $\mathbf{n}$ .

**Definición 14.14** Sea  $\mathbf{F} : S \rightarrow \mathbb{R}^3$  un campo de vectores definido sobre una superficie paramétrica regular orientada  $S$ , y  $\mathbf{n} : S \rightarrow \mathbb{R}^3$  el campo de vectores unitarios normales positivos para la orientación de  $S$ . El flujo de  $\mathbf{F}$  a través de  $S$  se define mediante la integral de superficie

$$\Phi = \int_S \langle \mathbf{F}(\mathbf{p}) | \mathbf{n}(\mathbf{p}) \rangle d\sigma(\mathbf{p})$$

en el supuesto de que la función escalar  $f(\mathbf{p}) = \langle \mathbf{F}(\mathbf{p}) | \mathbf{n}(\mathbf{p}) \rangle$  sea integrable sobre  $S$  respecto al elemento de área.

En las condiciones de la definición 14.14 si  $\boldsymbol{\varphi} : U \rightarrow S$  es una parametrización regular y positiva para la orientación de  $S$ , en cada punto  $\mathbf{p} = \boldsymbol{\varphi}(\mathbf{u}) \in S$ , el producto vectorial fundamental  $\mathbf{N}(\mathbf{u}) = D_1\boldsymbol{\varphi}(\mathbf{u}) \times D_2\boldsymbol{\varphi}(\mathbf{u})$  proporciona un vector normal a  $S$  en  $\mathbf{p}$  según la dirección del vector unitario  $\mathbf{n}(\mathbf{p})$ , luego

$$\mathbf{N}(\mathbf{u}) = \|\mathbf{N}(\mathbf{u})\|_2 \mathbf{n}(\boldsymbol{\varphi}(\mathbf{u})) = P_{\boldsymbol{\varphi}}(\mathbf{u}) \mathbf{n}(\boldsymbol{\varphi}(\mathbf{u}))$$

y se tiene

$$\Phi = \int_U \langle \mathbf{F}(\boldsymbol{\varphi}(\mathbf{u})) | \mathbf{n}(\boldsymbol{\varphi}(\mathbf{u})) \rangle P_{\boldsymbol{\varphi}}(\mathbf{u}) d\mathbf{u} = \int_U \langle \mathbf{F}(\boldsymbol{\varphi}(\mathbf{u})) | \mathbf{N}(\mathbf{u}) \rangle d\mathbf{u}$$

Si las ecuaciones de  $\boldsymbol{\varphi}$  las escribimos en forma explícita usando la notación

$$(x_1, x_2, x_3) = (\varphi_1(u_1, u_2), \varphi_2(u_1, u_2), \varphi_3(u_1, u_2))$$

las componentes del producto vectorial  $\mathbf{N}(\mathbf{u}) = D_1\boldsymbol{\varphi}(\mathbf{u}) \times D_2\boldsymbol{\varphi}(\mathbf{u})$  adoptan la forma

$$N_1 = \frac{D(\varphi_2, \varphi_3)}{D(u_1, u_2)}; \quad N_2 = \frac{D(\varphi_3, \varphi_1)}{D(u_1, u_2)}; \quad N_3 = \frac{D(\varphi_1, \varphi_2)}{D(u_1, u_2)}$$

luego el flujo  $\Phi$  viene dado por la integral doble:

$$\Phi = \int_U \left[ (F_1 \circ \boldsymbol{\varphi}) \frac{D(\varphi_2, \varphi_3)}{D(u_1, u_2)} + (F_2 \circ \boldsymbol{\varphi}) \frac{D(\varphi_3, \varphi_1)}{D(u_1, u_2)} + (F_3 \circ \boldsymbol{\varphi}) \frac{D(\varphi_1, \varphi_2)}{D(u_1, u_2)} \right] du_1 du_2$$

El lenguaje de las formas diferenciales de grado dos que introducimos a continuación proporciona un mecanismo formal para escribir la última integral en una forma fácil de recordar. Si  $1 \leq p < q \leq 3$ , sea  $dx_p \wedge dx_q = -dx_q \wedge dx_p : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$  la aplicación bilineal alternada que asigna a cada par de vectores

$$(\mathbf{u}, \mathbf{v}) = ((u_{11}, u_{12}, u_{13}), (u_{21}, u_{22}, u_{23})) \in \mathbb{R}^3 \times \mathbb{R}^3$$

el determinante de la matriz formada con las columnas que ocupan los lugares  $p, q$  (en este orden) en la matriz  $(u_{ij}) : 1 \leq i \leq 2, 1 \leq j \leq 3$ , de modo que

$$\frac{D(\varphi_2, \varphi_3)}{D(u_1, u_2)} = dx_2 \wedge dx_3 (D_1 \varphi, D_2 \varphi)$$

$$\frac{D(\varphi_3, \varphi_1)}{D(u_1, u_2)} = dx_3 \wedge dx_1 (D_1 \varphi, D_2 \varphi)$$

$$\frac{D(\varphi_1, \varphi_2)}{D(u_1, u_2)} = dx_1 \wedge dx_2 (D_1 \varphi, D_2 \varphi)$$

y así la expresión que figura bajo la integral última integral se escribe en la forma

$$[(F_1 \circ \varphi) dx_2 \wedge dx_3 + (F_2 \circ \varphi) dx_3 \wedge dx_1 + (F_3 \circ \varphi) dx_1 \wedge dx_2] (D_1 \varphi, D_2 \varphi)$$

Esto motiva la consideración, para cada  $\mathbf{x} \in S$  fijo, de la aplicación bilineal alternada  $\omega(\mathbf{x}) : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$  definida por

$$\omega(\mathbf{x}) = F_1(x_1, x_2, x_3) dx_2 \wedge dx_3 + F_2(x_1, x_2, x_3) dx_3 \wedge dx_1 + F_3(x_1, x_2, x_3) dx_1 \wedge dx_2$$

(En K.2 se puede ver que  $dx_2 \wedge dx_3, dx_3 \wedge dx_1, dx_1 \wedge dx_2$  forman una base del espacio vectorial de las aplicaciones bilineales alternadas  $\Gamma_2(\mathbb{R}^3)$ ).

Si en la expresión

$$F_1(x_1, x_2, x_3) dx_2 \wedge dx_3 + F_2(x_1, x_2, x_3) dx_3 \wedge dx_2 + F_3(x_1, x_2, x_3) dx_1 \wedge dx_2$$

y realizamos las sustituciones formales,

$$x_j = \varphi_j(u_1, u_2), \quad dx_j = d\varphi_j = D_1 \varphi_j(u_1, u_2) du_1 + D_2 \varphi_j(u_1, u_2) du_2$$

obtenemos

$$\begin{aligned} & (F_1 \circ \varphi) [D_1 \varphi_2 du_1 + D_2 \varphi_2 du_2] \wedge [D_1 \varphi_3 du_1 + D_2 \varphi_3 du_2] + \\ & (F_2 \circ \varphi) [D_1 \varphi_3 du_1 + D_2 \varphi_3 du_2] \wedge [D_1 \varphi_1 du_1 + D_2 \varphi_1 du_2] + \\ & (F_3 \circ \varphi) [D_1 \varphi_1 du_1 + D_2 \varphi_1 du_2] \wedge [D_1 \varphi_2 du_1 + D_2 \varphi_2 du_2] \end{aligned}$$

donde todas las funciones que intervienen se suponen evaluadas en  $\mathbf{u}$ . Utilizando las reglas formales del cálculo exterior

$$du_2 \wedge du_1 = -du_2 \wedge du_1, \quad du_1 \wedge du_1 = du_2 \wedge du_2 = 0$$

de un modo mecánico se llega a una expresión de la forma  $g(\mathbf{u})du_1 \wedge du_2$ , donde

$$\begin{aligned} g &= (F_1 \circ \varphi)(D_1\varphi_2 D_2\varphi_3 - D_2\varphi_2 D_1\varphi_3) + \\ &\quad (F_2 \circ \varphi)(D_1\varphi_3 D_2\varphi_1 - D_2\varphi_3 D_1\varphi_1) + \\ &\quad (F_3 \circ \varphi)(D_1\varphi_1 D_2\varphi_2 - D_2\varphi_1 D_1\varphi_2) = \end{aligned}$$

es la función que figura bajo la integral  $[*]$ . Con estos convenios de notación la integral doble que proporciona el flujo la escribiremos en la forma

$$\Phi = \int_U (F_1 \circ \varphi) d\varphi_2 \wedge d\varphi_3 + (F_2 \circ \varphi) d\varphi_3 \wedge d\varphi_1 + (F_3 \circ \varphi) d\varphi_1 \wedge d\varphi_2$$

que habitualmente se escribe así

$$\Phi = \int_{\varphi} F_1 dx_2 \wedge dx_3 + F_2 dx_3 \wedge dx_1 + F_3 dx_1 \wedge dx_2$$

## 14.5. Integración sobre variedades paramétricas $k$ -dimensionales

**Área de variedad paramétrica  $k$ -dimensional.** Para dar una interpretación geométrica de la fórmula 14.3 con la que se define el área  $k$ -dimensional de una parametrización  $\varphi : U \rightarrow \mathbb{R}^n$  de clase  $C^1$  definida en un abierto  $U \subset \mathbb{R}^k$ , ( $1 \leq k \leq n$ ), conviene tener presentes los resultados de la sección K.1 referentes a la definición del contenido de Jordan  $c_E$  en un subespacio  $E \subset \mathbb{R}^n$  de dimensión  $k$ , y las fórmulas obtenidas allí para calcular el contenido  $c_E(P)$  de un paralelepípedo  $P = P(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) \subset E$  generado por los vectores  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k \in E$ .

Comenzamos recordando las definiciones y los resultados que intervienen en lo que sigue. Una *parametrización* de clase  $C^m$  ( $m \geq 1$ ) y dimensión  $k$  ( $1 \leq k \leq n$ ) es una aplicación  $\varphi : U \rightarrow \mathbb{R}^n$  de clase  $C^m$  definida en un abierto  $U \subset \mathbb{R}^k$ . Si además  $\varphi$  es un homeomorfismo entre  $U$  y su imagen  $S = \varphi(U)$  y para cada  $\mathbf{u} \in U$ , los vectores  $D_j\varphi(\mathbf{u})$ ,  $1 \leq j \leq k$ , son linealmente independientes se dice que  $\varphi$  es una parametrización regular (de  $S = \varphi(U)$ ). En este caso, según el ejemplo 9.6 la imagen  $S = \varphi(U)$  es una subvariedad diferenciable de  $\mathbb{R}^n$ , de clase  $C^m$  y dimensión  $k$ . Este tipo de subvariedades diferenciables de  $\mathbb{R}^n$ , dadas como imagen de una parametrización regular, las llamamos  *$k$ -superficies paramétricas regulares*

Dos parametrizaciones,  $\varphi_j : U_j \rightarrow \mathbb{R}^n$ ,  $j = 1, 2$ , de clase  $C^m$  y dimensión  $k$ , se dice que son  $C^m$ -equivalentes cuando existe un  $C^m$ -difeomorfismo  $\mathbf{g} : U_1 \rightarrow U_2$ , tal que  $\varphi_1 = \varphi_2 \circ \mathbf{g}$ . Según la proposición H.9 dos parametrizaciones regulares de clase  $C^m$  y dimensión  $k$  con la misma imagen son  $C^m$ -equivalentes.

Si  $\varphi : U \rightarrow \mathbb{R}^n$  es una parametrización de clase  $C^1$  definida en un abierto  $U \subset \mathbb{R}^k$ , con  $1 \leq k \leq n$ , sea  $P_\varphi : U \rightarrow \mathbb{R}$  la función continua definida por

$$P_\varphi(\mathbf{u}) = \sqrt{|\det(\langle D_i\varphi(\mathbf{u}) \mid D_j\varphi(\mathbf{u}) \rangle)_{1 \leq i, j \leq k}|}$$

Según se ha visto en la sección K.1  $P_\varphi(\mathbf{u})$  proporciona el área  $k$ -dimensional del paralelepípedo  $P(D_1\varphi(\mathbf{u}), D_2\varphi(\mathbf{u}), \dots, D_k\varphi(\mathbf{u}))$ . En el caso particular  $k = n - 1$  esta función también viene dada por la fórmula  $P_\varphi(\mathbf{u}) = \|\mathbf{N}(\mathbf{u})\|_2$ , donde

$$\mathbf{N}(\mathbf{u}) = D_1\varphi(\mathbf{u}) \times D_2\varphi(\mathbf{u}) \times \dots \times D_k\varphi(\mathbf{u})$$

recibe el nombre de *producto vectorial fundamental* de la parametrización  $\varphi$ .

**Definición 14.15** Si  $\varphi : U \rightarrow \mathbb{R}^n$  es una parametrización de clase  $C^1$ , definida en un abierto  $U \subset \mathbb{R}^k$ , con  $1 \leq k \leq n$  su área  $k$ -dimensional se define como el valor de la integral impropia

$$k\text{-Área}(\varphi) = \int_U P_\varphi(\mathbf{u}) \, d\mathbf{u} \leq \infty \quad (14.3)$$

donde  $P_\varphi = \sqrt{|\det(\langle D_i\varphi \mid D_j\varphi \rangle)_{1 \leq i, j \leq k}|}$ .

La motivación de la definición es análoga a la del caso  $k = 2$  y  $n = 3$ .

Para simplificar la escritura el cubo  $[u_1, u_1 + h] \times \dots \times [u_n, u_n + h]$  lo denotamos  $Q[\mathbf{u}, h]$ . El abierto  $U$  lo aproximamos por dentro con una figura elemental formada por la unión de una familia finita de cubos de lado  $h$ , que no se solapan  $Q_j = Q[\mathbf{u}_j, h]$ ,  $1 \leq j \leq m$ . Una medida aproximada del área  $k$ -dimensional (en sentido intuitivo) del trozo  $\varphi(Q_j)$  la proporciona el área  $k$ -dimensional de  $d\varphi(Q_j)$  (dentro del espacio vectorial tangente  $E_j = E(\varphi, \mathbf{u}_j)$ , que suponemos de dimensión  $k$ ). Obsérvese que  $d\varphi(Q_j)$  es el paralelepípedo generado por los vectores  $hD_1\varphi(\mathbf{u}_j), hD_2\varphi(\mathbf{u}_j), \dots, hD_k\varphi(\mathbf{u}_j)$ , cuya área  $k$ -dimensional viene dada por  $h^k P_\varphi(\mathbf{u}_j) = P_\varphi(\mathbf{u}_j)v(Q_j)$ . La suma de estas áreas  $k$ -dimensionales

$$\sum_{j=1}^n P_\varphi(\mathbf{u}_j)v(Q_j)$$

es una suma de Riemann que aproxima a la integral  $\int_U P_\varphi(\mathbf{u}) \, d\mathbf{u}$  cuyo valor es razonable adoptar como medida del área  $k$ -dimensional (recorrida o barrida) por  $\varphi$ .

Es interesante observar que en los casos extremos  $k = 1$ , y  $k = n$ , la definición 14.15 está de acuerdo con los resultados previos referentes a esta situación: Cuando  $k = 1$ , y  $U = (a, b)$ , la definición 14.15 da lugar a la clásica fórmula para la longitud de un arco de curva de clase  $C^1$ ,

$$\text{Long}(\varphi) = \int_a^b \sqrt{\langle \varphi'(t) \mid \varphi'(t) \rangle} \, dt = \int_a^b \|\varphi'(t)\|_2 \, dt$$

En el caso  $k = n$ , si  $\varphi : U \rightarrow V$  es un difeomorfismo de clase  $C^1$  entre los abiertos  $U, V \subset \mathbb{R}^n$ , es fácil ver que  $P_\varphi(\mathbf{u}) = |\det \varphi'(\mathbf{u})|$ , de modo que, en este caso, en virtud de la fórmula del cambio de variable J.13, la fórmula de la definición 14.15 proporciona la medida usual en  $\mathbb{R}^n$  del volumen de la imagen  $V = \varphi(U)$ , es decir

$$\int_U P_\varphi(\mathbf{u}) \, d\mathbf{u} = \int_U |\det \varphi'(\mathbf{u})| \, d\mathbf{u} = \lambda(\varphi(U)) = \lambda(V)$$

donde  $\lambda(V)$  es la medida de Lebesgue del abierto  $V$ , definida en el capítulo ??.

Recordemos que dada una parametrización  $\varphi : U \rightarrow \mathbb{R}^n$  de clase  $C^m$  y dimensión  $k$ , para cada  $\mathbf{u} \in U$ , el subespacio vectorial  $E(\varphi, \mathbf{u}) := d\varphi(\mathbf{u})(\mathbb{R}^k) \subset \mathbb{R}^n$  está formado por vectores tangentes a  $S = \varphi(U)$  en  $\mathbf{p} = \varphi(\mathbf{u})$ , y se llama espacio tangente a la parametrización  $\varphi$ , en el punto  $\mathbf{p}$ , para el valor del parámetro  $\mathbf{u}$ .

El subespacio  $E(\varphi, \mathbf{u})$ , generado por los vectores  $D_1\varphi(\mathbf{u}), \dots, D_k\varphi(\mathbf{u})$ , tiene dimensión  $k$  cuando son linealmente independientes, lo que ocurre cuando  $\varphi$  es una parametrización regular. En la proposición H.10 se demostró que si las parametrizaciones  $\varphi : U \rightarrow \mathbb{R}^n$ ,  $\Psi : V \rightarrow \mathbb{R}^n$ , son  $C^m$ -equivalentes y  $\mathbf{g} : V \rightarrow U$ , es un  $C^m$  difeomorfismo tal que  $\Psi = \varphi \circ \mathbf{g}$ , entonces para cada  $\mathbf{v} = \mathbf{g}(\mathbf{u}) \in V$  se cumple  $E(\Psi, \mathbf{v}) = E(\varphi, \mathbf{u})$ .

**Proposición 14.16** *Si  $\varphi_1 : U_1 \rightarrow \mathbb{R}^n$ ,  $\varphi_2 : U_2 \rightarrow \mathbb{R}^n$  son parametrizaciones  $C^1$ -equivalentes (de clase  $C^1$  y dimensión  $k$ ), se verifica*

$$\int_{U_1} P_{\varphi_1}(\mathbf{u})d\mathbf{u} = \int_{U_2} P_{\varphi_2}(\mathbf{v})d\mathbf{v}$$

es decir,  $k$ -Área( $\varphi_1$ ) =  $k$ -Área( $\varphi_2$ ).

DEM: La hipótesis significa que existe un  $C^1$ -difeomorfismo  $\mathbf{g} : U_1 \rightarrow U_2$  tal que para cada  $\mathbf{u} \in U_1$  se cumple  $\varphi_1(\mathbf{u}) = \varphi_2(\mathbf{v})$  donde  $\mathbf{v} = \mathbf{g}(\mathbf{u}) \in U_2$ .

Sabemos que los subespacios  $d\varphi_1(\mathbf{u})(\mathbb{R}^k) = E(\varphi_1, \mathbf{u})$  y  $d\varphi_2(\mathbf{v})(\mathbb{R}^k) = E(\varphi_2, \mathbf{v})$ , son iguales y de dimensión  $\leq k$ . Para lo que sigue conviene fijar un subespacio  $k$ -dimensional  $E$  tal que  $E(\varphi_1, \mathbf{u}) = E(\varphi_2, \mathbf{v}) \subset E \subset \mathbb{R}^n$ .

Como las aplicaciones lineales  $d\varphi_1(\mathbf{u}), d\varphi_2(\mathbf{v}) : \mathbb{R}^k \rightarrow E \subset \mathbb{R}^n$  toman valores en el espacio euclídeo  $k$ -dimensional  $E$ , sus respectivas matrices, respecto a la base canónica de  $\mathbb{R}^k$  y a una base ortonormal  $\beta$  del espacio euclídeo  $E$ , son cuadradas y así podemos considerar sus determinantes,  $\det_{\beta} \varphi'_1(\mathbf{u}), \det_{\beta} \varphi'_2(\mathbf{v})$ . Obsérvese que en virtud de la regla de la cadena  $d\varphi_1(\mathbf{u}) = d\varphi_2(\mathbf{v}) \circ d\mathbf{g}(\mathbf{u})$ , se verifica

$$\det_{\beta} \varphi'_1(\mathbf{u}) = \det_{\beta} \varphi'_2(\mathbf{v}) \det \mathbf{g}'(\mathbf{u})$$

Por otra parte, según la definición,  $P_{\varphi_1}(\mathbf{u})$  es el contenido en  $E$  del paralelepípedo

$$P[D_1\varphi_1(\mathbf{u}), D_2\varphi_1(\mathbf{u}), \dots, D_k\varphi_1(\mathbf{u})]$$

cuyo valor, según K.2, viene dado por

$$P_{\varphi_1}(\mathbf{u}) = |\det_{\beta} \varphi'_1(\mathbf{u})| = |\det_{\beta} \varphi'_2(\mathbf{v})| \cdot |\det \mathbf{g}'(\mathbf{u})| = P_{\varphi_2}(\mathbf{v}) |\det \mathbf{g}'(\mathbf{u})|$$

donde  $\varphi'_1, \varphi'_2$ , son las matrices jacobianas de las correspondientes aplicaciones. Por lo tanto la igualdad del enunciado es una consecuencia directa de la fórmula del cambio de variable (que sigue valiendo para integrales en sentido impropio). ■

**Corolario 14.17** *Si  $\varphi_1 : U_1 \rightarrow \mathbb{R}^n$ ,  $\varphi_2 : U_2 \rightarrow \mathbb{R}^n$  son parametrizaciones regulares de clase  $C^1$  y dimensión  $k$  con la misma imagen entonces  $k$ -Área( $\varphi_1$ ) =  $k$ -Área( $\varphi_2$ ).*

DEM: Es consecuencia directa de las proposiciones H.9 y 14.16. ■

En virtud del corolario 14.17, si  $\varphi : U \rightarrow \mathbb{R}^n$  es una parametrización regular definida en un abierto  $U \subset \mathbb{R}^k$  y  $S = \varphi(U)$ , el número  $k$ -Área( $\varphi$ ) sólo depende de la imagen  $S = \varphi(U)$ , y no hay inconveniente en adoptarlo como medida de su área  $k$ -dimensional, definiendo  $k$ -Área( $S$ ) =  $k$ -Área( $\varphi$ ).

La noción de integral de una función escalar sobre una superficie paramétrica considerada en la sección 14.10 se extiende de modo natural al caso de funciones definidas sobre variedades paramétricas regulares en  $\mathbb{R}^n$ , clase  $C^1$  y dimensión  $k$  ( $1 \leq k \leq n$ ). La analogía es completa cuando  $k = n - 1$ , porque en este caso también se puede hacer que intervenga el producto vectorial fundamental de los  $n - 1$  vectores  $D_1\varphi(\mathbf{u}) \times \cdots \times D_{n-1}\varphi(\mathbf{u})$ , donde  $\varphi : U \rightarrow S$  es una parametrización regular de  $S$  definida en un abierto  $U \subset \mathbb{R}^{n-1}$ .

**Definición 14.18** *Sea  $\varphi : U \rightarrow \mathbb{R}^n$  una parametrización de clase  $C^1$  y dimensión  $k$  ( $1 \leq k \leq n$ ) definida en un abierto  $U \subset \mathbb{R}^k$  y  $P_\varphi(\mathbf{u}) = \sqrt{\det(\langle D_i\varphi(\mathbf{u}) | D_j\varphi(\mathbf{u}) \rangle)_{1 \leq i, j \leq k}}$ . Dada una función  $f : S \rightarrow \mathbb{R}$ , definida sobre  $S = \varphi(U)$ , si  $(f \circ \varphi)P_\varphi$  es absolutamente integrable sobre el abierto  $U$  se dice que  $f$  es integrable respecto a  $\varphi$  y se define*

$$\int_\varphi f = \int_U f(\varphi(\mathbf{u}))P_\varphi(\mathbf{u})d\mathbf{u}$$

donde

Según se ha visto en la sección 14.5, cuando  $k = n - 1$  la función  $P_\varphi$  que interviene en la definición anterior también viene dada por  $P_\varphi(\mathbf{u}) = \|\mathbf{N}(\mathbf{u})\|_2$ , donde

$$\mathbf{N}(\mathbf{u}) = D_1\varphi(\mathbf{u}) \times D_2\varphi(\mathbf{u}) \times \cdots \times D_{n-1}\varphi(\mathbf{u})$$

Por otra parte, en el caso particular  $k = 1$ , la definición 14.18, proporciona la fórmula para la integral de una función respecto al arco considerada en el capítulo 4.

**Proposición 14.19** *Sean  $\varphi_j : U_j \rightarrow \mathbb{R}^n$ ,  $j = 1, 2$ , parametrizaciones  $C^1$ -equivalentes de clase  $C^1$  y dimensión  $k$ , ( $1 \leq k \leq n$ ). Una función  $f : S \rightarrow \mathbb{R}$  definida sobre la  $k$ -superficie  $S = \varphi_1(U_1) = \varphi_2(U_2)$ , es integrable respecto a  $\varphi_1$  si y sólo si es integrable respecto a  $\varphi_2$ , y en ese caso*

$$\int_{\varphi_1} f = \int_{\varphi_2} f$$

DEM: La hipótesis significa que existe un  $C^1$ -difeomorfismo  $\mathbf{g} : U_1 \rightarrow U_2$  tal que  $\varphi_1 = \varphi_2 \circ \mathbf{g}$ . Según la demostración de la proposición 14.16 se verifica

$$P_{\varphi_1} = (P_{\varphi_2} \circ \mathbf{g})|\det \mathbf{g}'|$$

Según el teorema del cambio de variable J.15 podemos afirmar que  $(f \circ \varphi_2)P_{\varphi_2}$  es integrable sobre  $U_2 = \mathbf{g}(U_1)$  si y sólo si

$$(f \circ \varphi_2 \circ \mathbf{g})(P_{\varphi_2} \circ \mathbf{g})|\det \mathbf{g}'| = (f \circ \varphi_1)P_{\varphi_1}$$

es integrable sobre  $U_1$  y en ese caso las integrales coinciden

$$\int_{U_2} (f \circ \varphi_2)(\mathbf{v}) P_{\varphi_2}(\mathbf{v}) d\mathbf{v} = \int_{U_1} (f \circ \varphi_1)(\mathbf{u}) P_{\varphi_1}(\mathbf{u}) d\mathbf{u}$$

■

La proposición 14.19 se puede tomar como base para definir la integral  $\int_S f(\mathbf{p}) d\sigma(\mathbf{p})$  de una función  $f$  definida sobre una variedad paramétrica regular clase  $C^1$  y dimensión  $k$ ,  $S \subset \mathbb{R}^n$ , en términos de una de sus parametrizaciones regulares  $\varphi : U \rightarrow S$ ,

$$\int_S f(\mathbf{p}) d\sigma(\mathbf{p}) = \int_U f(\varphi(\mathbf{u})) P_{\varphi}(\mathbf{u}) d\mathbf{u}$$

ya que, según la proposición H.9, todas las parametrizaciones regulares de  $S$  son  $C^1$ -equivalentes y por lo tanto proporcionan el mismo valor de la integral.

La noción de flujo de un campo de vectores  $\mathbf{F}(x_1, x_2, \dots, x_n)$  a través de una variedad paramétrica regular orientada de dimensión  $k$  sólo tiene sentido cuando  $k = n - 1$ , pues sólo en este caso se puede formar el producto vectorial fundamental  $\mathbf{N}(\mathbf{u}) = D_1\varphi(\mathbf{u}) \times \dots \times D_{n-1}\varphi(\mathbf{u})$  asociado a una parametrización  $\varphi$  regular y positiva para la orientación de  $S$ . Igual que se hizo en el caso  $k = 2$ ,  $n = 3$ , normalizando los vectores  $\mathbf{N}(\varphi^{-1}(\mathbf{p}))$  se consigue un campo continuo de vectores normales unitarios  $\mathbf{n} : S \rightarrow \mathbb{R}^n$  que son positivos para la orientación de  $S$ , y se define el flujo

$$\Phi = \int_S \langle \mathbf{F}(\mathbf{p}) \mid \mathbf{n}(\mathbf{p}) \rangle d\sigma(\mathbf{p}) = \int_U \langle \mathbf{F}(\varphi(\mathbf{u})) \mid \mathbf{N}(\mathbf{u}) \rangle d\mathbf{u}$$

en el supuesto de que la función  $f(\mathbf{p}) = \langle \mathbf{F}(\mathbf{p}) \mid \mathbf{n}(\mathbf{p}) \rangle$  sea integrable sobre  $S$ .

## 14.6. Ejercicios resueltos

**Ejercicio 14.20** Calcule el área del trozo de cilindro  $x^2 + (y - a)^2 = a^2$  que queda por encima del plano  $z = 0$  y dentro de la esfera  $x^2 + y^2 + z^2 = 4a^2$ , ( $a > 0$ ).

SOLUCIÓN

Se trata de calcular el área de  $S = \{(x, y, z) : (x, y) \in D, 0 < z < \sqrt{4a^2 - x^2 - y^2}\}$  donde  $D = \{(x, y) : x^2 + (y - a)^2 \leq a^2\}$ . Prescindiendo del segmento  $L = \{(0, 0, z) : 0 \leq z \leq 2a\} \subset S$  es fácil obtener una parametrización regular  $\varphi : U_1 \rightarrow \mathbb{R}^3$  de trozo  $S_0 = S \setminus L$ : La intersección del cilindro con el plano horizontal  $z = 0$  es la circunferencia  $C = \{(x, y) : x^2 + (y - a)^2 = a^2\}$ , cuya ecuación en coordenadas polares  $r = 2a \sin \theta$ , conduce a la parametrización

$$\gamma(\theta) = (x(\theta), y(\theta)), \quad \text{donde } x(\theta) = 2a \sin \theta \cos \theta, \quad y(\theta) = 2a \sin \theta \sin \theta$$

Como  $\gamma(0, \pi) = C \setminus \{(0, 0)\}$ , se obtiene que  $S_0 = \varphi(U_1)$  donde

$$U = \{(\theta, z) : 0 < \theta < \pi, 0 < z < \sqrt{4a^2 - x(\theta)^2 - y(\theta)^2}\}$$

y  $\varphi : U \rightarrow \mathbb{R}^3$  es la parametrización regular definida por  $\varphi(\theta, z) = (x(\theta), y(\theta), z)$ .

Es razonable asumir que  $S$  tiene la misma área que  $S_0$  la cual, según el ejemplo 14.9, viene dada por la integral

$$\text{Área}(\varphi) = \int_0^\pi \sqrt{4a^2 - x(\theta)^2 - y(\theta)^2} \|\gamma'(\theta)\|_2 d\theta$$

Como  $\|\gamma'(\theta)\|_2 = 2a$ , resulta

$$\text{Área}(\varphi) = 2a \int_0^\pi \sqrt{4a^2 - 4a^2 \sin^2 \theta} d\theta = 4a^2 \int_0^\pi |\cos \theta| d\theta = 8a^2 \int_0^{\pi/2} \cos \theta = 8a^2$$

■

**Ejercicio 14.21** Calcule el área del trozo de esfera  $x^2 + y^2 + z^2 = 4a^2$  que queda en el semiespacio  $z > 0$ , dentro del cilindro  $x^2 + (y - a)^2 = a^2$ , ( $a > 0$ ).

SOLUCIÓN

Se trata de calcular el área de la gráfica de la función  $f(x, y) = \sqrt{4a^2 - x^2 - y^2}$  sobre el abierto  $U = \{(x, y) : x^2 + (y - a)^2 < a^2\}$  Según la fórmula 14.2 el área de  $S = \{(x, y, f(x, y)) : (x, y) \in U\}$  viene dada por la integral

$$I = \int_U \sqrt{1 + (D_1 f(x, y))^2 + (D_2 f(x, y))^2} dx dy = \int_U \frac{2a}{\sqrt{4a^2 - x^2 - y^2}} dx dy$$

Obsérvese que se trata de una integral impropia, ya que el integrando tiende hacia  $+\infty$  cuando  $U \ni (0, y) \rightarrow (0, 2a)$ . Podemos calcularla con un cambio de variable a coordenadas polares (véase el teorema J.15) con el que se obtiene

$$\begin{aligned} I &= 2a \int_0^\pi \int_0^{2a \sin \theta} \frac{r}{\sqrt{4a^2 - r^2}} dr = 2a \int_0^\pi (2a - \sqrt{4a^2 - 4a^2 \sin^2 \theta}) d\theta = \\ &= 4a^2 \int_0^\pi (1 - |\cos \theta|) d\theta = 8a^2 \int_0^{\pi/2} (1 - \cos \theta) d\theta = 4a^2(\pi - 2) \end{aligned}$$

**Ejercicio 14.22** Sea  $L : \mathbb{R}^k \rightarrow \mathbb{R}^n$  una aplicación lineal inyectiva y  $U \subset \mathbb{R}^k$  un conjunto abierto medible Jordan. Demuestre que  $S = L(U)$  es un subconjunto medible Jordan de  $E = L(\mathbb{R}^k) \subset \mathbb{R}^n$  cuyo contenido de Jordan en  $E$  viene dado por la fórmula 14.1, es decir

$$c_E(S) = \text{Área}(L|_U) = \int_U P_L(\mathbf{u}) d\mathbf{u}$$

SOLUCIÓN

Como la aplicación lineal  $L$  es inyectiva su imagen  $E = L(\mathbb{R}^k)$  es un subespacio vectorial  $k$ -dimensional en el que elegimos una base ortonormal  $\beta = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ .



Con el isomorfismo  $T_\beta : \mathbb{R}^k \rightarrow E$ ,  $T_\beta(\mathbf{x}) = \sum_{j=1}^k x_j \mathbf{u}_j$ , el subespacio  $E$  queda identificado con  $\mathbb{R}^k$ , y según la definición  $S = L(U)$  es medible Jordan en  $E$  porque  $T_\beta^{-1}(S) = (T_\beta^{-1} \circ L)(U)$  es la imagen, mediante la aplicación lineal  $T_\beta^{-1} \circ L$ , del conjunto medible Jordan  $U$  (véase J.8). Además, en virtud de la definición de  $c_E$  y de la proposición J.8 se verifica

$$\begin{aligned} c_E(S) &= c_k(T_\beta^{-1}(S)) = c_k((T_\beta^{-1} \circ L)(U)) = \\ &= |\det(T_\beta^{-1} \circ L)| c_k(U) = |\det T_\beta^{-1}| |\det L| c_k(U) \end{aligned}$$

donde los determinantes se refieren a las matrices de las aplicaciones lineales respecto a la base  $\beta$  de  $E$  y a la base canónica de  $\mathbb{R}^k$ . Es claro que  $|\det T_\beta^{-1}| = 1$ , luego  $c_E(S) = |\det L| c_k(U)$ . Por otra parte, si  $\mathbf{v}_j = L(\mathbf{e}_j)$ ,  $1 \leq j \leq k$ , para cada  $\mathbf{u} \in U$  se cumple  $D_j L(\mathbf{u}) = L(\mathbf{e}_j) = \mathbf{v}_j$  luego la función

$$P_L(\mathbf{u}) = c_E(P(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)) = |\det_\beta(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)| = |\det L|$$

es constante, y se sigue que  $\text{Área}(L|_U) = \int_U |\det L| = |\det L| c_k(U) = c_E(S)$ . ■

**Ejercicio 14.23** Sea  $\gamma : (a, b) \rightarrow \mathbb{R}^2$  una parametrización regular de clase  $C^1$  y longitud finita  $L = \int_a^b \|\gamma'(t)\|_2 dt$ , cuya imagen es una curva plana situada en el semiplano  $\{(x, y) : y > 0\}$ . Sea  $S$  la superficie de revolución que engendra la curva al girar alrededor del eje  $Ox$ . Demuestre que su área vale  $2\pi y_0 L$ , donde  $y_0$  es el radio de la circunferencia que describe el centro de masa de la curva (teorema de Pappus).

SOLUCIÓN

Consideramos la curva sumergida en  $\mathbb{R}^3$ , dentro del plano  $z = 0$ , mediante la parametrización  $\mathbf{p}(t) = (x(t), y(t), 0)$ , donde  $(x(t), y(t)) = \gamma(t)$ . Cuando el punto  $\mathbf{p}(t)$  gira un ángulo  $\theta \in (0, 2\pi)$  alrededor del eje  $Ox$ , pasa a ocupar la posición

$$\varphi(t, \theta) = (x(t), y(t) \cos \theta, y(t) \sin \theta)$$

Usando que  $\gamma(t) = (x(t), y(t))$  es una parametrización regular de clase  $C^1$  se puede comprobar que  $\varphi : U \rightarrow \mathbb{R}^3$  también es regular y de clase  $C^1$  (los detalles se dejan al cuidado del lector), luego podemos considerar el área de la imagen  $S_0 = \varphi(U)$ ,

$$\text{Área}(S_0) = \text{Área}(\varphi) = \int_U \|\mathbf{N}(t, \theta)\|_2 dt d\theta$$

donde  $\mathbf{N}(t, \theta) = D_1 \varphi(t, \theta) \times D_2 \varphi(t, \theta)$  es el producto vectorial fundamental. Con un cálculo rutinario se obtiene que  $\mathbf{N}(t, \theta) = (y(t)y'(t), -x'(t)y(t) \cos \theta, -x'(t)y(t) \sin \theta)$ . Como  $y(t) > 0$  para todo  $t \in (a, b)$ , resulta  $\|\mathbf{N}(t, \theta)\|_2 = y(t) \sqrt{x'(t)^2 + y'(t)^2}$ , luego

$$\text{Área}(S_0) = \int_0^{2\pi} \left( \int_a^b y(t) \|\gamma'(t)\|_2 dt \right) d\theta = 2\pi \int_a^b y(t) \|\gamma'(t)\|_2 dt$$

Según el ejercicio 4.7.10 la coordenada  $y_0$  del centro de masa de la curva plana  $\gamma(a, b)$  viene dado por  $y_0 = \frac{1}{L} \int_a^b y(t) \|\gamma'(t)\|_2 dt$ , luego  $\text{Área}(S_0) = 2\pi y_0 L$ , donde  $y_0$  es

la distancia del centro de masa de la curva  $C$  al eje de giro. Obsérvese que  $S_0 = S \setminus C$  donde  $C = \mathbf{p}(a, b)$  es la curva plana que genera la superficie al girar. Admitiendo que  $S$  y  $S_0$  tienen la misma área se obtiene el resultado. ■

**Ejercicio 14.24** Sea  $S \subset \mathbb{R}^3$  una superficie paramétrica regular y  $f : S \rightarrow \mathbb{R}$  una función integrable sobre  $S$ . Se considera la función de conjunto  $\mu(B) = \int_B f(\mathbf{x}) d\sigma(\mathbf{x})$ , definida sobre las partes abiertas  $B \subset S$  (en la topología relativa). Si  $f$  es continua en  $\mathbf{p} \in S$ , y  $S_r(\mathbf{p}) = \{\mathbf{x} \in S : \|\mathbf{x} - \mathbf{p}\| < r\}$ , demuestre que

$$\lim_{r \rightarrow 0} \frac{\mu(S_r(\mathbf{p}))}{\text{Área}(S_r(\mathbf{p}))} = f(\mathbf{p})$$

SOLUCIÓN

Por hipótesis  $S = \varphi(U)$  donde  $\varphi : U \rightarrow \mathbb{R}^3$  es una parametrización regular de clase  $C^1$ , definida en un abierto  $U \subset \mathbb{R}^2$ . Es claro que  $S_r(\mathbf{p})$  es una superficie paramétrica regular, parametrizada mediante la restricción de  $\varphi$  al abierto  $U_r = \varphi^{-1}(S_r(\mathbf{p}))$ .

Dado  $\epsilon > 0$  en virtud de la continuidad de  $f \circ \varphi$  en  $\mathbf{u}_0 = \varphi^{-1}(\mathbf{p}) \in U$ , existe una bola  $B(\mathbf{u}_0, \delta) \subset U$  tal que  $\mathbf{u} \in B(\mathbf{u}_0, \delta) \Rightarrow |f(\varphi(\mathbf{u})) - f(\varphi(\mathbf{u}_0))| < \epsilon$ .

Por otra parte, como  $\varphi : U \rightarrow S$  es un homeomorfismo (porque  $\varphi$  es una parametrización regular), usando la continuidad de  $\varphi^{-1} : S \rightarrow U$  en el punto  $\mathbf{p} \in S$  podemos encontrar  $\eta > 0$  tal que  $\mathbf{x} = \varphi(\mathbf{u}) \in S_\eta(\mathbf{p}) \Rightarrow \|\mathbf{u} - \mathbf{u}_0\| < \delta$ , lo que significa que el abierto  $U_\eta = \varphi^{-1}(S_\eta(\mathbf{p}))$  está contenido en la bola  $B(\mathbf{u}_0, \delta)$ .

Por consiguiente, cuando  $0 < r < \eta$ , podemos afirmar que para todo  $\mathbf{u} \in U_r \subset U_\eta$  se cumple  $|f(\varphi(\mathbf{u})) - f(\varphi(\mathbf{u}_0))| < \epsilon$ , y con ello se obtiene que

$$\begin{aligned} |\mu(S_r) - f(\mathbf{p})\text{Área}(S_r(\mathbf{p}))| &= \left| \int_{U_r} [f(\varphi(\mathbf{u})) - f(\varphi(\mathbf{u}_0))] P_\varphi(\mathbf{u}) d\mathbf{u} \right| \leq \\ &\leq \int_{U_r} |f(\varphi(\mathbf{u})) - f(\varphi(\mathbf{u}_0))| P_\varphi(\mathbf{u}) d\mathbf{u} \leq \int_{U_r} \epsilon P_\varphi(\mathbf{u}) d\mathbf{u} = \epsilon \text{Área}(S_r(\mathbf{p})) \end{aligned}$$

Así queda demostrado que  $0 < r < \eta \Rightarrow \left| \frac{\mu(S_r(\mathbf{p}))}{\text{Área}(S_r(\mathbf{p}))} - f(\mathbf{p}) \right| < \epsilon$  ■

**Ejercicio 14.25** Área de un trozo de superficie en forma implícita: Sea  $F : \Omega \rightarrow \mathbb{R}$  de clase  $C^1$  en un abierto  $\Omega \subset \mathbb{R}^3$  y  $S \subset \{(x, y, z) \in \Omega : F(x, y, z) = 0\}$  un trozo de superficie que se proyecta de modo biyectivo sobre un abierto  $U$  del plano  $(x, y)$ , donde queda determinada una función implícita  $z = z(x, y)$  de clase  $C^1(U)$ .

Obtenga la fórmula

$$\text{Área}(S) = \int_U \frac{\|\nabla F(x, y, z)\|_2}{|D_3 F(x, y, z)|} dx dy$$

donde en el integrando se supone realizada la sustitución  $z = z(x, y)$ . Utilícela para volver a calcular el área de la semiesfera  $S = \{(x, y, z) : x^2 + y^2 + z^2 = R^2, z > 0\}$ .

## 14.7. Ejercicios propuestos

◇ 14.7.1 Obtenga el área de los siguientes trozos de superficie:

- i) Hemisferio esférico  $S = \{(x, y, z) : x^2 + y^2 + z^2 = a^2, z \geq 0\}$ .
- ii) Trozo del plano  $x + y + z = a$  determinado por el cilindro  $x^2 + y^2 = a^2$ .
- iii) Trozo de plano  $x + y + z = 1$  determinado por el cilindro  $x^2 + 2y^2 = 1$ .
- iv) Trozo de la esfera  $x^2 + y^2 + z^2 = a^2$  dentro del cilindro  $x^2 + y^2 = ay$ , ( $a > 0$ ).
- v) Trozo de la superficie cónica  $x^2 + y^2 = z^2$ , situada sobre el plano  $xy$  y limitada por la esfera  $x^2 + y^2 + z^2 = 2ax$

◇ 14.7.2 Halle el área de los siguientes trozos de superficie dados en forma paramétrica

- i) Cono  $S = \{(r \cos t, r \sin t, r) : 0 \leq t \leq 2\pi, 0 \leq r \leq 1\}$ ;
- ii) Helicoide  $S = \{(r \cos t, r \sin t, t) : 0 \leq t \leq 2\pi, 0 \leq r \leq 1\}$ ;

◇ 14.7.3 Halle el área del toro

$$T = \{(a + b \cos u) \sin v, (a + b \cos u) \cos v, b \sin u) : 0 \leq u \leq 2\pi, 0 \leq v \leq 2\pi\} \quad 0 < b < a$$

Utilice el teorema de Pappus para comprobar el resultado.

◇ 14.7.4 El cilindro  $x^2 + y^2 = x$  divide a la superficie esférica  $x^2 + y^2 + z^2 = 1$  en dos trozos  $S_1$  y  $S_2$ , donde  $S_1$  está dentro del cilindro y  $S_2$  afuera. Halle la razón de las áreas  $\text{área}(S_2)/\text{área}(S_1)$ .

◇ 14.7.5 Una esfera está inscrita en un cilindro circular recto y es cortada por dos planos paralelos perpendiculares al eje del cilindro. Demuestre que las porciones de esfera y de cilindro comprendidas entre estos planos tienen la misma área.

◇ 14.7.6 Exprese, mediante integrales, el área de las siguientes superficies:

- i)  $x^2 - y^2 = 1, x > 0, -1 \leq y \leq 1$ ;
- ii)  $(x/a)^2 + (y/b)^2 + (z/c)^2 = 1$

◇ 14.7.7 Muestre que la superficie  $x = 1/\sqrt{y^2 + z^2}, 1 \leq x < +\infty$  se puede llenar pero no se puede pintar.

◇ 14.7.8 Obtenga una fórmula para el área de la superficie generada al girar la gráfica de una función  $y = f(x), a \leq x \leq b$  alrededor del eje  $OX$  y alrededor del eje  $OY$ . (Se supone que  $f$  es de clase  $C^1$ ).

# A

## Sucesiones y series de funciones

*Convergencia puntual y convergencia uniforme. Condición de Cauchy y criterio de Weierstrass. Teoremas sobre continuidad, derivabilidad e integrabilidad del límite de una sucesión de funciones. Versiones para series*

En este capítulo, que se desarrolla en el ámbito de las funciones reales de una variable real, se estudia cuando el límite de una sucesión de funciones continuas, integrables o derivables hereda la correspondiente propiedad. Ejemplos sencillos muestran que la convergencia puntual es insuficiente para este propósito, y este inconveniente motiva la introducción de la convergencia uniforme, con la que se consigue la conservación de la continuidad, de la integrabilidad, así como el paso al límite bajo la integral (teoremas A.6 y A.7). El tercer resultado central de este capítulo (teorema A.11) se refiere a la derivabilidad del límite de una sucesión de funciones, y a la validez de la derivabilidad término a término (la derivada del límite es el límite de las derivadas). Para este resultado la hipótesis adecuada es la convergencia uniforme de la sucesión de derivadas junto con la convergencia de la sucesión en algún punto.

Estos resultados tienen sus correspondientes versiones para series de funciones y para establecer la convergencia uniforme de estas series son muy útiles el criterio de Weierstrass, y los criterios de Abel y Dirichlet. La trascendencia del criterio de Weierstrass se pone de manifiesto al utilizarlo para definir funciones patológicas, como el célebre ejemplo de Weierstrass de una función continua que no es derivable en ningún punto.

En relación con el problema del paso al límite bajo la integral se mencionan en este capítulo, sin demostración, otros resultados más generales que garantizan el paso al límite bajo una integral impropia en términos de la existencia de una función dominadora (un anticipo modesto de los potentes resultados que proporciona la integral de Lebesgue). Aunque no se demuestren estos resultados se ven algunos ejemplos de aplicación y se proponen algunos ejercicios sobre este asunto.

## A.1. Convergencia puntual y uniforme

Una sucesión de funciones  $f_n : T \rightarrow \mathbb{R}$  definidas en un conjunto  $T \subset \mathbb{R}$  se dice que converge puntualmente cuando para cada  $t \in T$  la sucesión de números reales  $f_n(t)$  es convergente. En este caso el límite puntual de la sucesión  $f_n$  es la función  $f : T \rightarrow \mathbb{R}$  definida por  $f(t) = \lim_n f_n(t)$ .

**Ejemplo A.1** La sucesión  $f_n : [0, 1] \rightarrow \mathbb{R}$ ,  $f_n(t) = t^n$ , converge puntualmente hacia la función discontinua  $f : [0, 1] \rightarrow \mathbb{R}$ , que vale 0 si  $0 \leq t < 1$ , y  $f(1) = 1$ . (Véase *Figura 1*).

Si  $f$  es el límite puntual de  $f_n$ , dados  $t \in T$ , y  $\epsilon > 0$  existe  $n(\epsilon, t) \in \mathbb{N}$  tal que  $n \geq n(\epsilon, t) \Rightarrow |f_n(t) - f(t)| \leq \epsilon$ . Es decir, la  $\epsilon$ -aproximación al límite se consigue a partir de un valor de  $n$  que depende de  $t$ . Al considerar otro punto  $t' \in T$ , puede ocurrir que con este valor de  $n$  no se logre la aproximación  $|f_n(t') - f(t')| \leq \epsilon$ , y sea necesario avanzar más en la sucesión hasta conseguirla. En el ejemplo A.1 se aprecia gráficamente que al tomar puntos  $t$  cada vez más próximos a 1 la sucesión  $f_n(t)$  va tardando más tiempo en entrar en el entorno  $(-\epsilon, \epsilon)$  de su límite  $f(t) = 0$ . Con este ejemplo se pone de manifiesto que la convergencia puntual no garantiza la continuidad del límite de una sucesión de funciones continuas.

El ejemplo que sigue muestra que la convergencia puntual tampoco garantiza la integrabilidad del límite de una sucesión de funciones integrables.

**Ejemplo A.2** Sea  $\{r_n : n \in \mathbb{N}\}$  una enumeración de  $\mathbb{Q} \cap [0, 1]$ , y  $f_n : \mathbb{R} \rightarrow \mathbb{R}$  definida por  $f_n(x) = 1$  si  $x \in \{r_k : 1 \leq k \leq n\}$ ,  $f_n(x) = 0$  si  $x \notin \{r_k : 1 \leq k \leq n\}$ . Cada  $f_n$  es integrable Riemann en  $[0, 1]$  con  $\int_0^1 f_n(t) dt = 0$ , pero la sucesión  $f_n$  converge puntualmente hacia la función no integrable

$$f(x) = 1 \text{ si } x \in \mathbb{Q}, \quad f(x) = 0 \text{ si } x \notin \mathbb{Q}$$

Con el siguiente ejemplo (véase ([5] prob.12, pág. 222) queda patente que el paso al límite bajo la integral tampoco es lícito cuando la función límite es integrable y la convergencia es puntual.

**Ejemplo A.3** Si  $p \geq 1$ , en el intervalo  $[0, 1]$  la sucesión  $f_n(x) = n^p x(1 - x^2)^n$  converge puntualmente hacia la función idénticamente nula  $f \equiv 0$ . Sin embargo no converge hacia  $0 = \int_0^1 f(x) dx$  la sucesión de las integrales, ya que

$$\int_0^1 n^p x(1 - x^2)^n dx = \left[ -\frac{n^p}{2} \frac{(1 - x^2)^{n+1}}{n + 1} \right]_0^1 = \frac{n^p}{2(n + 1)}$$

(Véase *Figura 2* .)

En los teoremas A.6 y A.7 veremos que con la noción de convergencia uniforme, formulada en la siguiente definición, se evitan las patologías de los ejemplos anteriores

**Definición A.4** Se dice que la sucesión  $f_n : T \rightarrow \mathbb{R}$  converge uniformemente hacia  $f : T \rightarrow \mathbb{R}$  si para cada  $\epsilon > 0$  existe  $n(\epsilon) \in \mathbb{N}$  (que depende sólo de  $\epsilon$ ) tal que para todo  $n \geq n(\epsilon)$  y todo  $t \in T$  se cumple  $|f_n(t) - f(t)| \leq \epsilon$ .

Es inmediato que la convergencia uniforme implica la convergencia puntual y el ejemplo A.1 pone de manifiesto que el recíproco es falso. La convergencia uniforme es más fuerte que la convergencia puntual porque en ella el valor de  $n$  a partir del cual se consigue la aproximación prefijada  $|f_n(t) - f(t)| \leq \epsilon$ , es independiente del punto  $t \in T$ , es decir, se exige aproximación uniforme al límite en todos los puntos.

Si  $K \subset T$  y la sucesión  $f_n|_K$  converge puntualmente (resp. uniformemente) se dice, más brevemente, que la sucesión  $f_n$  converge puntualmente (resp. uniformemente) sobre  $K$ . Con el fin de formular la condición de convergencia uniforme de modo más conciso conviene introducir la siguiente notación: Si  $K \subset T$ , dadas  $f, g : T \rightarrow \mathbb{R}$ , definimos  $\rho_K(f, g) = \sup\{|f(t) - g(t)| : t \in K\} \leq +\infty$ . Ahora, el hecho de que la sucesión  $f_n : T \rightarrow \mathbb{R}$  sea uniformemente convergente hacia  $f : T \rightarrow \mathbb{R}$  se escribe en la forma  $\lim_n \rho_T(f_n, f) = 0$ . Análogamente, la convergencia uniforme sobre  $K \subset T$  se expresa mediante la condición  $\lim_n \rho_K(f_n, f) = 0$ .

A veces ocurre que una sucesión de funciones  $f_n : T \rightarrow \mathbb{R}$ , no converge uniformemente sobre todo  $T$ , pero la convergencia es uniforme sobre cada conjunto  $A$  de cierta familia  $\mathcal{A}$  de subconjuntos de  $T$ . En ese caso se dice que la sucesión converge uniformemente sobre los conjuntos de  $\mathcal{A}$ . Como caso particular, cuando  $\mathcal{A}$  es la familia de los subconjuntos compactos de  $T$ , se habla de convergencia uniforme sobre compactos.

**Proposición A.5** [Condición de Cauchy] Una sucesión de funciones  $f_n : T \rightarrow \mathbb{R}$  converge uniformemente sobre  $K \subset T$  si y sólo si cumple:

Para cada  $\epsilon > 0$  existe  $n(\epsilon) \in \mathbb{N}$  tal que  $[k > n \geq n(\epsilon), t \in K] \Rightarrow |f_n(t) - f_k(t)| \leq \epsilon$ .

DEM: La demostración de que la condición es necesaria es inmediata y se deja al cuidado del lector. La condición es suficiente: La sucesión es puntualmente convergente porque, para cada  $t \in K$ , la sucesión  $f_n(t)$  cumple la condición de Cauchy. Sea  $f : K \rightarrow \mathbb{R}$  el límite puntual de la sucesión. Veamos que la convergencia es uniforme. Dado  $\epsilon > 0$ , si  $k > n \geq n(\epsilon)$ , para todo  $t \in K$  se cumple  $|f_n(t) - f_k(t)| \leq \epsilon$ . Fijando  $t \in K$  y pasando al límite cuando  $k \rightarrow +\infty$  la desigualdad se convierte en  $|f_n(t) - f(t)| \leq \epsilon$ , que resulta válida para todo  $t \in K$  y todo  $n \geq n(\epsilon)$ . ■

OBSERVACIÓN: La condición de Cauchy para la convergencia uniforme sobre  $K \subset T$  se puede expresar de modo conciso asociando a la sucesión de funciones  $f_n : T \rightarrow \mathbb{R}$  la sucesión numérica  $\alpha_n = \sup_{k \geq n} \sup_{t \in K} |f_n(t) - f_k(t)| \leq +\infty$ . Así la condición de Cauchy para la convergencia uniforme sobre  $K$  equivale a que  $\lim_n \alpha_n = 0$ . Basta observar que la implicación  $[k > n \geq n(\epsilon), t \in K] \Rightarrow |f_n(t) - f_k(t)| \leq \epsilon$  se traduce en la forma siguiente:  $n \geq n(\epsilon) \Rightarrow 0 \leq \alpha_n \leq \epsilon$ .

## A.2. Continuidad, derivabilidad e integrabilidad del límite

**Teorema A.6** *Si la sucesión  $f_n : T \rightarrow \mathbb{R}$  converge uniformemente hacia  $f : T \rightarrow \mathbb{R}$  y cada  $f_n$  es continua en  $a \in T$  entonces el límite  $f$  también lo es. En particular, si las funciones  $f_n$  son continuas en todo punto, el límite uniforme  $f$  también lo es.*

DEM: Dado  $\epsilon > 0$ , en virtud de la convergencia uniforme, existe  $m \in \mathbb{N}$  tal que para todo  $t \in T$  se cumple  $|f_m(t) - f(t)| \leq \epsilon/3$ . Por la continuidad de  $f_m$  en  $a$ , existe  $r > 0$  tal que si  $|t - a| < r$  y  $t \in T$  se cumple  $|f_m(t) - f_m(a)| \leq \epsilon/3$ , luego  $|f(t) - f(a)| \leq |f(t) - f_m(t)| + |f_m(t) - f_m(a)| + |f_m(a) - f(a)| \leq \epsilon$ . ■

Obsérvese que, en las condiciones del teorema anterior, para conseguir la continuidad del límite  $f$  en un punto concreto  $a \in T$  basta suponer que las funciones de la sucesión son continuas en  $a$  y que la convergencia de la sucesión es uniforme en  $V_a \cap T$  donde  $V_a$  es un entorno de  $a$ . Por lo tanto la continuidad global del límite  $f$  se conseguirá cuando las funciones de la sucesión sean continuas en todo punto y cada  $a \in T$  tenga un entorno abierto  $V_a$  tal que la sucesión sea uniformemente convergente sobre  $T \cap V_a$ . Cuando ocurra esto diremos que hay *convergencia uniforme local*. Es claro que la convergencia uniforme sobre todo  $T$  implica la convergencia uniforme local pero la afirmación recíproca es falsa: La sucesión considerada en el ejemplo A.1 no converge uniformemente sobre  $T = (0, 1)$ , pero para cada  $a \in (0, 1)$ , la sucesión converge uniformemente en  $(a - r, a + r) \subset (0, 1)$ , donde  $0 < a - r < a + r < 1$ .

Es fácil ver que la convergencia uniforme local implica la convergencia uniforme sobre compactos y que el recíproco es cierto cuando  $T \subset \mathbb{R}$  es un intervalo.

**Teorema A.7** *Sea  $f_n : [a, b] \rightarrow \mathbb{R}$  una sucesión de funciones integrables Riemann que converge uniformemente hacia  $f : [a, b] \rightarrow \mathbb{R}$ . Entonces  $f$  es integrable Riemann en  $[a, b]$  y  $\int_a^b f(t)dt = \lim_n \int_a^b f_n(t)dt$ .*

DEM: Sabemos que la sucesión  $\rho_n = \sup\{|f_n(t) - f(t)| : t \in [a, b]\} \leq +\infty$ , converge hacia 0, luego existe  $n_0$  tal que  $\rho_n < +\infty$ , para todo  $n \geq n_0$ . Para  $n > n_0$  y todo  $t \in [a, b]$  se cumple  $f_n(t) - \rho_n \leq f(t) \leq f_n(t) + \rho_n$ , luego  $f$  es acotada en  $[a, b]$ . Además, para todo  $n \in \mathbb{N}$ , en virtud de la monotonía de la integral inferior y de la integral superior se cumple

$$\int_a^b (f_n(t) - \rho_n)dt \leq \underline{\int_a^b} f \leq \overline{\int_a^b} f \leq \int_a^b (f_n(t) + \rho_n)dt$$

luego

$$0 \leq \overline{\int_a^b} f - \underline{\int_a^b} f \leq 2\rho_n(b - a)$$

y pasando al límite se obtiene  $\overline{\int_a^b} f = \underline{\int_a^b} f$ , es decir,  $f$  es integrable sobre  $[a, b]$ .

Por otra parte, usando la desigualdad  $|f(t) - f_n(t)| \leq \rho_n$ , válida para todo  $t \in [a, b]$ , y todo  $n \in \mathbb{N}$ , resulta

$$\left| \int_a^b f_n(t) dt - \int_a^b f(t) dt \right| \leq \int_a^b |f(t) - f_n(t)| dt \leq \rho_n(b-a)$$

luego,  $\lim_n \int_a^b f_n(t) dt = \int_a^b f(t) dt$ . ■

La sucesión del ejemplo A.2 pone de manifiesto que en el teorema anterior la hipótesis de convergencia uniforme es esencial para conseguir la integrabilidad de la función límite. Por otra parte, el ejemplo A.3 muestra que el paso al límite bajo la integral tampoco es lícito cuando el límite es integrable y sólo se supone convergencia puntual. Cuando la función límite es integrable Riemann los siguientes resultados (teoremas A.8 y A.9) garantizan el paso al límite bajo la integral con hipótesis más débiles que la convergencia uniforme.

**Teorema A.8** *Sea  $f_n : [a, b] \rightarrow \mathbb{R}$  una sucesión de funciones integrables Riemann que converge puntualmente hacia una función integrable Riemann  $f : [a, b] \rightarrow \mathbb{R}$ . Si la sucesión  $f_n$  es uniformemente acotada, (existe  $C > 0$  tal que  $|f_n(t)| \leq C$  para todo  $t \in [a, b]$ , y todo  $n \in \mathbb{N}$ ) entonces,  $\int_a^b f = \lim_n \int_a^b f_n$ .*

Recordemos que  $f : (\alpha, \beta) \rightarrow \mathbb{R}$  se dice que es localmente integrable (Riemann) cuando es integrable Riemann sobre cada  $[a, b] \subset (\alpha, \beta)$ . En lo que sigue diremos que la sucesión  $f_n : (\alpha, \beta) \rightarrow \mathbb{R}$  está dominada por la función  $g : (\alpha, \beta) \rightarrow [0, +\infty)$  cuando para todo  $t \in (\alpha, \beta)$  y todo  $n \in \mathbb{N}$  se cumple  $|f_n(t)| \leq g(t)$ .

**Teorema A.9** *Sea  $f_n : (\alpha, \beta) \rightarrow \mathbb{R}$  una sucesión de funciones localmente integrables que converge puntualmente hacia una función  $f : (\alpha, \beta) \rightarrow \mathbb{R}$  localmente integrable. Se supone que*

*i) Las integrales impropias  $\int_\alpha^\beta f_n(t) dt$  son absolutamente convergentes.*

*ii) La sucesión  $f_n$  está dominada por una función localmente integrable Riemann  $g : (\alpha, \beta) \rightarrow [0, +\infty)$  con  $\int_\alpha^\beta g(t) dt < +\infty$ .*

*Entonces la integral impropia  $\int_\alpha^\beta f(t) dt$  es absolutamente convergente y se verifica  $\int_\alpha^\beta f(t) dt = \lim_n \int_\alpha^\beta f_n(t) dt$ .*

La demostración directa de los teoremas A.8, A.9, con los recursos propios de la integral de Riemann es técnicamente complicada y no la expondremos aquí. Estos dos teoremas son versiones particulares de resultados generales sobre la integral de Lebesgue que el lector interesado puede consultar en el capítulo 10 de [2]. Esperamos que estos resultados sirvan de motivación para que el lector se interese por la integral de Lebesgue, más potente y flexible que la de Riemann.

Con los teoremas A.6 y A.7 ha quedado establecido que la continuidad y la integrabilidad Riemann se conservan por convergencia uniforme. No ocurre lo mismo con la derivabilidad, como se verá más adelante en el ejemplo A.17. Incluso cuando el límite es derivable, no se puede garantizar que la derivada del límite de una sucesión uniformemente convergente sea el límite de las derivadas:



**Ejemplo A.10** La sucesión  $f_n(x) = x/(1 + n^2x^2)$  converge uniformemente hacia la función idénticamente nula,  $f(x) \equiv 0$ , pero en el punto  $x = 0$ , las derivadas  $f'_n(0) = 1$  no convergen hacia  $f'(0) = 0$ .

En efecto, es claro que la sucesión de este ejemplo converge puntualmente hacia la función nula  $f(x) \equiv 0$ , y es fácil ver que la función  $|f_n(x) - f(x)| = |f_n(x)|$  alcanza un máximo absoluto en  $x = 1/n$ , luego  $\sup\{|f_n(x) - f(x)| : x \in \mathbb{R}\} = |f_n(1/n)| = \frac{1}{2n}$ , de donde se sigue que la sucesión  $f_n$  es uniformemente convergente. Sin embargo la sucesión  $f'_n(0) = 1$  no converge hacia  $f'(0) = 0$ . ■

Según los ejemplos A.17 y A.10 para conseguir un resultado sobre derivación término a término de una sucesión de funciones, la convergencia uniforme de la sucesión no es la hipótesis adecuada. Según el siguiente teorema las hipótesis adecuadas son la convergencia de la sucesión en algún punto y la convergencia uniforme de la sucesión de derivadas

**Teorema A.11** Sea  $f_n : (a, b) \rightarrow \mathbb{R}$  una sucesión de funciones derivables en un intervalo acotado  $(a, b) \subset \mathbb{R}$ , que converge en algún  $x_0 \in (a, b)$ . Si la sucesión de derivadas  $f'_n$  converge uniformemente en  $(a, b)$  entonces la sucesión  $f_n$  converge uniformemente en  $(a, b)$  hacia una función derivable  $f : (a, b) \rightarrow \mathbb{R}$ , y para todo  $x \in (a, b)$  se cumple  $\lim_n f'_n(x) = f'(x)$ .

DEM: Consideremos la sucesión de funciones continuas  $g_n : (a, b) \rightarrow \mathbb{R}$ ,

$$g_n(x) = \frac{f_n(x) - f_n(x_0)}{x - x_0} \text{ si } x \neq x_0, g_n(x_0) = f'_n(x_0)$$

(la continuidad de  $g_n$  en  $x_0$  es consecuencia de la definición de derivada, y la continuidad en los restantes puntos es inmediata).

a) La sucesión  $g_n$  converge uniformemente en  $(a, b)$ , pues cumple la condición de Cauchy para la convergencia uniforme A.5:

Si  $p > q$ , y  $x_0 \neq x \in (a, b)$ , aplicando el teorema del valor medio a la función derivable  $f_p - f_q$  en el intervalo de extremos  $x, x_0$  podemos escribir

$$g_p(x) - g_q(x) = \frac{(f_p(x) - f_q(x)) - (f_p(x_0) - f_q(x_0))}{x - x_0} = f'_p(\xi) - f'_q(\xi)$$

donde  $\xi$  es un punto del intervalo de extremos  $x, x_0$ . Por otra parte, cuando  $x = x_0$ , se tiene  $g_p(x_0) - g_q(x_0) = f'_p(x_0) - f'_q(x_0)$ , luego, para todo  $x \in (a, b)$  se cumple

$$|g_p(x) - g_q(x)| \leq \sup\{|f'_p(t) - f'_q(t)| : t \in (a, b)\}$$

Como la sucesión  $f'_n$  verifica la condición de Cauchy para la convergencia uniforme en  $(a, b)$ , esta desigualdad implica que la sucesión  $g_n$  también la cumple.

b) Sea  $y_0 = \lim_n f_n(x_0)$  y  $g : (a, b) \rightarrow \mathbb{R}$  la función continua que se obtiene como límite uniforme de la sucesión de funciones continuas  $g_n$ . Utilizando que la función  $(x - x_0)$  es acotada en el intervalo acotado  $(a, b)$  se obtiene fácilmente que la sucesión  $f_n(x) = f_n(x_0) + (x - x_0)g_n(x)$  converge uniformemente en  $(a, b)$  hacia la función

$f(x) = y_0 + (x - x_0)g(x)$ . Como  $f(x_0) = y_0$ , y  $g$  es continua en  $x_0$  se sigue que existe el límite

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \rightarrow x_0} g(x) = g(x_0)$$

luego  $f$  es derivable en  $x_0$  y  $f'(x_0) = g(x_0) = \lim_n g_n(x_0) = \lim_n f'_n(x_0)$ .

Queda demostrado que  $f$  es derivable en  $x_0$ , y que  $f'(x_0) = \lim_n f'_n(x_0)$ . Como ya hemos visto que  $f_n(t)$  converge en cada  $t \in (a, b)$ , reemplazando  $x_0$  por  $t$  en toda la demostración anterior, se obtiene que también existe la derivada  $f'(t)$ , y que  $f'(t) = \lim_n f'_n(t)$ .

NOTA: Añadiendo la hipótesis de que las derivadas  $f'_n$  son continuas, siguiendo el siguiente esquema se puede dar una demostración más breve: Sea  $y_0 = \lim_n f_n(x_0)$ . Según el teorema A.6, la función  $\varphi(t) = \lim_n f'_n(t)$  es continua en  $(a, b)$ , luego  $f(x) = y_0 + \int_{x_0}^x \varphi(t)dt$  es una función derivable en  $(a, b)$ , con derivada  $f'(x) = \varphi(x)$ . Por otra parte, usando la representación integral  $f_n(x) = f_n(x_0) + \int_{x_0}^x f'_n(t)dt$  se demuestra fácilmente que la sucesión  $f_n$  converge uniformemente hacia la función  $f$ . Entonces, en virtud del teorema fundamental del cálculo, se concluye que en cada  $x \in (a, b)$ ,  $f$  es derivable y  $f'(x) = \varphi(x) = \lim_n f'_n(x)$ . ■

La demostración del último teorema muestra que para una sucesión  $f_n$  de funciones derivables en un intervalo  $(a, b)$ , la convergencia uniforme de la sucesión de derivadas  $f'_n$  se transmite a la sucesión  $f_n$ , bajo la hipótesis de que esta sucesión sea convergente en algún punto. Por otra parte, ejemplos sencillos muestran que la convergencia uniforme de una sucesión de funciones derivables no garantiza la convergencia uniforme de la sucesión de derivadas (véase el ejercicio resuelto A.21). El siguiente ejemplo es más sorprendente: Una sucesión de funciones derivables uniformemente convergente tal que la sucesión de derivadas no converge en ningún punto.

**Ejemplo A.12** La sucesión  $f_n(x) = [\text{sen}(2\pi nx)]/\sqrt{n}$  converge uniformemente en  $\mathbb{R}$  hacia la función nula, pero la sucesión de las derivadas  $f'_n(x) = 2\pi\sqrt{n} \cos(2\pi nx)$  no converge en ningún punto.

DEM: La sucesión de las derivadas  $f'_n(x) = 2\pi\sqrt{n} \cos(2\pi nx)$  no es convergente cuando  $x$  es racional, pues si  $x = p/q$ , donde  $p, q \in \mathbb{Z}$ ,  $q > 0$ , con  $n_k = kq$  se obtiene la subsucesión  $f'_{n_k}(x) = 2\pi\sqrt{n_k} \cos(2\pi kp) = 2\pi\sqrt{n_k}$  que no es convergente. Consideremos ahora el caso  $x \notin \mathbb{Q}$ . Dado  $\epsilon \in (0, 1)$ , usando la continuidad uniforme de la función  $\cos t$  podemos encontrar  $\delta > 0$  que cumple

$$|s - t| < \delta \Rightarrow |\cos s - \cos t| < \epsilon$$

Utilizamos ahora la siguiente propiedad de los números irracionales cuya demostración se verá después: Si  $x \notin \mathbb{Q}$  el conjunto  $A_\beta(x) = \{n \in \mathbb{N} : \exists m \in \mathbb{Z} |nx - m| < \beta\}$  es infinito para cada  $\beta > 0$ . Usando esta propiedad con  $\beta = \delta/2\pi$ , obtenemos la subsucesión  $f_{n_k}(x)$ , donde  $\{n_1 < n_2 < n_3 < \dots\} = A_\beta(x)$ . Según la definición de  $A_\beta(x)$  para cada  $k \in \mathbb{N}$  existe  $m_k \in \mathbb{Z}$  verificando  $|n_k x - m_k| < \beta$ , es decir,  $|2\pi n_k x - 2\pi m_k| < 2\pi\beta = \delta$ , luego  $|\cos(2\pi n_k x) - 1| < \epsilon$ , de donde se sigue que  $\cos(2\pi n_k x) > 1 - \epsilon > 0$ . Por lo tanto la sucesión  $f'_n(x)$  no es convergente porque tiene una subsucesión  $f'_{n_k}(x) = 2\pi\sqrt{n_k} \cos(2\pi n_k x) > 2\pi(1 - \epsilon)\sqrt{n_k}$  que no converge.

Para terminar demostramos la propiedad de los irracionales que hemos usado. Sea  $x \notin \mathbb{Q}$  y  $\beta > 0$ . Es claro que para cada  $k \in \mathbb{N}$  existe  $q_k \in \mathbb{Z}$  tal que  $\alpha_k = q_k + kx \in [0, 1)$ . Si  $m \in \mathbb{N}$  y  $1/m < \beta$ , descomponiendo el intervalo  $[0, 1]$  en  $m$  subintervalos contiguos de longitud  $1/m$ , es claro que alguno de los subintervalos contiene dos puntos distintos  $\alpha_i, \alpha_j$  con  $1 \leq i < j \leq m+1$ , luego  $|\alpha_i - \alpha_j| \leq 1/m < \beta$ , es decir  $|q_i - q_j + (i - j)x| < \beta$ , y esto demuestra que  $(i - j) \in A_\beta(x)$ . Así queda justificado que  $A_\beta(x) \neq \emptyset$  para cada  $\beta > 0$ . Para ver que  $A_\beta(x)$  es infinito no es restrictivo suponer la condición  $0 < \beta < 1/2$  y así tenemos garantizado que para cada  $n_k \in A_\beta(x)$  existe un único  $m_k \in \mathbb{Z}$  verificando  $|n_k x - m_k| < \beta$ . Razonamos por reducción al absurdo suponiendo que el conjunto  $A_\beta(x) = \{n_1 < n_2 < \dots < n_p\}$  es finito. Como  $x \notin \mathbb{Q}$  podemos elegir un número  $0 < \eta < \min\{|n_k x - m_k| : 1 \leq k \leq p\}$  para el que se cumple que  $A_\eta(x) \neq \emptyset$ . Obsérvese que, en virtud de la unicidad de los  $m_k$  antes mencionada, la elección de  $\eta$  garantiza que  $A_\beta(x)$  y  $A_\eta(x)$  son disjuntos. Por otra parte, al ser  $\eta < \beta$  se debe cumplir que  $\emptyset \neq A_\eta(x) \subset A_\beta(x)$  y con esta contradicción termina la demostración. ■

### A.3. Series de funciones

Hasta ahora sólo hemos considerado sucesiones de funciones reales definidas en un subconjunto  $T$  de la recta real. Es claro que las nociones de convergencia puntual y uniforme se extienden de forma natural al caso de funciones con valores complejos  $f_n : T \rightarrow \mathbb{C}$  definidas en un conjunto arbitrario  $T$ . En esta situación más general es obvio que sigue valiendo la condición de Cauchy para la convergencia uniforme. También sigue valiendo el teorema de conservación de la continuidad A.6, siempre que tenga sentido hablar de continuidad, como ocurre cuando  $T$  es un subconjunto de  $\mathbb{C}$  (o más generalmente, un espacio métrico). En lo que sigue, con el fin de poder considerar más adelante las series de potencias de variable compleja, consideraremos siempre series  $\sum_{n=1}^{\infty} f_n(t)$  de funciones  $f_n : T \rightarrow \mathbb{C}$ , definidas en un conjunto  $T$ , que habitualmente será un subconjunto de  $\mathbb{R}$  ó  $\mathbb{C}$ .

En esta situación la definición de convergencia uniforme tiene su correspondiente versión para series  $\sum_{n=1}^{\infty} f_n(t)$ , formulada en términos de la sucesión de sumas parciales  $S_n = \sum_{j=1}^n f_j$ . Se dice que una serie converge uniformemente sobre  $K \subset T$  cuando la sucesión de sus sumas parciales converge uniformemente sobre  $K$ . El siguiente resultado es muy útil a la hora de establecer la convergencia uniforme de una serie:

**Teorema A.13** [Criterio de Weierstrass] *Una condición suficiente para que la serie  $\sum_{n=1}^{\infty} f_n(t)$  de funciones  $f_n : T \rightarrow \mathbb{C}$  sea uniformemente convergente sobre  $K \subset T$  es que exista una serie numérica convergente  $\sum_{n=1}^{\infty} \rho_n$  verificando:  $|f_n(t)| \leq \rho_n$  para todo  $t \in K$  y todo  $n \in \mathbb{N}$ .*

DEM: Basta demostrar que la sucesión de sumas parciales  $S_n(t) = \sum_{j=1}^n f_j(t)$  cumple la condición de Cauchy para la convergencia uniforme sobre  $K$ , es decir, que la sucesión numérica  $\alpha_n := \sup_{k > n} \sup_{t \in K} |S_n(t) - S_k(t)|$  converge hacia 0. Obsérvese

que para todo  $k > n$  y todo  $t \in K$  se cumple

$$|S_n(t) - S_k(t)| = \left| \sum_{j=n+1}^k f_j(t) \right| \leq \sum_{j=n+1}^k |f_j(t)| \leq \sum_{j=n+1}^k \rho_j$$

luego  $0 \leq \alpha_n \leq \sum_{j=n+1}^{\infty} \rho_j$ , de donde se sigue que  $\lim_n \alpha_n = 0$ . ■

Cuando se aplica el criterio de Weierstras, además de la convergencia uniforme se obtiene la convergencia absoluta de la serie, de modo que este criterio no sirve para obtener convergencia uniforme de series que no son absolutamente convergentes. Para establecer la convergencia uniforme de series de funciones que no son absolutamente convergentes son muy útiles los criterios de Dirichlet y Abel, recogidos en los siguientes teoremas cuya demostración se basa en la siguiente fórmula de sumación parcial, cuya comprobación se deja al cuidado del lector:

*Dadas dos sucesiones finitas de números reales (o complejos)  $\{a_j : 1 \leq j \leq n\}$ ,  $\{b_j : 1 \leq j \leq n\}$ , para  $n \geq 2$  se verifica*

$$S_n = a_n B_n + \sum_{j=1}^{n-1} B_j(a_j - a_{j+1}), \quad \text{donde } S_n = \sum_{k=1}^n a_k b_k, \quad B_j = \sum_{k=1}^j b_k$$

**Teorema A.14** [Dirichlet] *Una serie de la forma  $\sum_{n=1}^{+\infty} a_n(t)b_n(t)$ , con  $a_n : T \rightarrow \mathbb{R}$ ,  $b_n : T \rightarrow \mathbb{C}$ , converge uniformemente sobre  $K \subset T$  cuando se cumple a) y b):*

- a) *La sucesión  $B_n(t) = \sum_{j=1}^n b_j(t)$  está uniformemente acotada sobre  $K \subset T$ .*  
 b) *La sucesión  $a_n(t)$  es monótona decreciente para cada  $t \in K$  y converge uniformemente hacia 0 sobre  $K$ .*

DEM: Por hipótesis existe  $M > 0$  tal que  $|B_n(t)| \leq M$  para todo  $n \in \mathbb{N}$  y todo  $t \in K$  y la sucesión  $\rho_n = \sup_{t \in K} |a_n(t)|$  converge hacia 0. Según la fórmula de sumación parcial las sumas  $S_n(t) = \sum_{k=1}^n a_k(t)b_k(t)$  se pueden escribir en la forma

$$S_n(t) = a_n(t)B_n(t) + \sum_{j=1}^{n-1} B_j(t)(a_j(t) - a_{j+1}(t))$$

Para cada  $t \in K$  la sucesión  $a_n(t)B_n(t)$  converge hacia 0 (porque es el producto de una sucesión acotada por una sucesión que converge hacia 0) y la serie  $\sum_{j=1}^{\infty} B_j(t)(a_j(t) - a_{j+1}(t))$  es absolutamente convergente porque

$$\sum_{j=1}^{\infty} |B_j(t)(a_j(t) - a_{j+1}(t))| \leq M \sum_{j=1}^{\infty} (a_j(t) - a_{j+1}(t)) = Ma_1(t)$$

Se sigue que la sucesión de sumas parciales  $S_n(t)$  converge puntualmente en  $K$  hacia la función  $S(t) = \sum_{j=1}^{\infty} B_j(t)(a_j(t) - a_{j+1}(t))$  que verifica  $|S(t)| \leq Ma_1(t)$ .

Para terminar debemos demostrar que la sucesión  $S_m(t)$  converge hacia  $S(t)$  uniformemente sobre  $K$ . La serie  $\sum_{j=m+1}^{\infty} a_j(t)b_j(t)$  cumple las mismas hipótesis que

la serie original, la única diferencia es que ahora las sumas  $B_n^*(t) = \sum_{j=m+1}^n b_j(t)$  están uniformemente acotadas sobre  $K$  por la constante  $2M$ . Según el razonamiento anterior esta serie converge puntualmente sobre  $K$  y para todo  $t \in K$  se verifica

$$\left| \sum_{j=m+1}^{\infty} a_j(t)b_j(t) \right| \leq 2Ma_{m+1}(t) \leq 2M\rho_{m+1}$$

luego

$$|S(t) - S_m(t)| = \left| \sum_{j=m+1}^{\infty} a_j(t)b_j(t) \right| \leq 2M\rho_{m+1}$$

y así se obtiene que la sucesión  $S_m$  converge uniformemente sobre  $K$ . ■

**Teorema A.15** [Abel] *Una serie de la forma  $\sum_{n=1}^{+\infty} a_n(t)b_n(t)$ , con  $a_n : T \rightarrow \mathbb{R}$ ,  $b_n : T \rightarrow \mathbb{C}$ , converge uniformemente sobre  $K \subset T$  cuando se cumple a) y b):*

a) *La serie  $\sum_{n=1}^m b_n(t)$  converge uniformemente sobre  $K \subset T$ .*

b) *La sucesión  $a_n(t)$  es monótona decreciente para cada  $t \in K$  y está uniformemente acotada sobre  $K$ .*

DEM: La idea de la demostración consiste en utilizar la fórmula de sumación parcial para ver que la sucesión de sumas parciales  $S_n(t) = \sum_{j=1}^n a_j(t)b_j(t)$  cumple la condición de Cauchy para la convergencia uniforme sobre  $K$ . Así, para  $m > n$  la suma  $S_m(t) - S_n(t) = \sum_{j=n+1}^m a_j(t)b_j(t)$  la podemos escribir en la forma

$$S_m(t) - S_n(t) = a_m(t)B_{nm}(t) + \sum_{j=n+1}^{m-1} B_{nj}(t)(a_j(t) - a_{j+1}(t))$$

donde  $B_{nj}(t) = \sum_{k=n+1}^j b_k(t)$ . Según las hipótesis existe  $C > 0$  tal que  $|a_j(t)| \leq C$  para todo  $t \in K$  y todo  $j \in \mathbb{N}$  y además la serie  $\sum_{j=1}^{\infty} b_j(t)$  converge uniformemente sobre  $K$ , lo que significa (según la condición de Cauchy) que para cada  $\epsilon > 0$  existe  $n(\epsilon) \in \mathbb{N}$  tal que  $[j > n \geq n(\epsilon), t \in K] \Rightarrow |B_{nj}(t)| \leq \epsilon$ . Entonces, usando la desigualdad triangular, se obtiene

$$|S_m(t) - S_n(t)| \leq \epsilon|a_m(t)| + \sum_{j=n+1}^{m-1} \epsilon(a_j(t) - a_{j+1}(t))$$

Teniendo en cuenta  $\sum_{j=n+1}^{m-1} (a_j(t) - a_{j+1}(t)) = a_{n+1}(t) - a_m(t)$  se obtiene que

$$|S_m(t) - S_n(t)| \leq 3C\epsilon$$

Como esta desigualdad es válida para  $m > n \geq n(\epsilon)$  y todo  $t \in K$  queda establecido que la sucesión de sumas parciales  $S_n(t)$  cumple la condición de Cauchy para la convergencia uniforme sobre  $K$ . ■

Los resultados sobre continuidad, integrabilidad y derivabilidad del límite de una sucesión de funciones A.6, A.7, A.11 tienen su correspondiente versión para series

$\sum_{n=1}^{\infty} f_n(t)$  de funciones reales  $f_n : [a, b] \rightarrow \mathbb{R}$ . Las versiones para series se obtienen de modo inmediato considerando la sucesión de las sumas parciales. A título de ejemplo estableceremos el resultado referente a la integral de la suma de una serie dejando al cuidado del lector los referentes a continuidad y derivabilidad de la suma.

**Proposición A.16** *Sea  $\sum_{n=1}^{\infty} f_n(t)$  una serie de funciones  $f_n : [a, b] \rightarrow \mathbb{R}$  integrables Riemann. Si la serie converge uniformemente entonces la suma  $f(t) = \sum_{n=1}^{\infty} f_n(t)$  es integrable en  $[a, b]$  y se cumple  $\int_a^b f = \sum_{n=1}^{\infty} \int_a^b f_n$ .*

DEM: La sucesión  $S_n = \sum_{j=1}^n f_j$ , converge uniformemente sobre  $[a, b]$  hacia  $f$ , y en virtud de A.7 la sucesión  $\int_a^b S_n = \sum_{j=1}^n (\int_a^b f_j)$  converge hacia  $\int_a^b f$  luego  $\sum_{n=1}^{\infty} \int_a^b f_n = \int_a^b f$ . ■

**Funciones patológicas definidas por series.** En 1875 Weierstrass descubrió el siguiente ejemplo una serie uniformemente convergente de funciones indefinidamente derivables cuya suma es continua pero no es derivable en ningún punto.

**Ejemplo A.17** [Weierstrass] *Si  $m \in \mathbb{N}$  es impar y  $2m > 2mb > 2 + 3\pi$ , entonces la serie  $f(x) = \sum_{k=0}^{\infty} b^k \cos(m^k \pi x)$  converge uniformemente y define una función continua acotada  $f : \mathbb{R} \rightarrow \mathbb{R}$  que no es derivable en ningún punto. (Véase **Figura 3**)*

La convergencia uniforme de la serie que interviene en el ejemplo anterior es consecuencia directa del criterio de Weierstrass A.13 ya que, al ser  $0 < b < 1$ , la serie geométrica  $\sum_{k=0}^{\infty} b^k$  es convergente, y es claro que para todo  $n \in \mathbb{N}$ , y todo  $x \in \mathbb{R}$  se cumple  $|b^k \cos(m^k \pi x)| \leq b^k$ . Como la serie está formada por funciones continuas, aplicando el teorema A.6 a la sucesión de sumas parciales  $S_n(x) = \sum_{k=1}^n f_k(x)$  se obtiene la continuidad de  $f$ . El hecho sorprendente de que esta función no sea derivable en ningún punto es más difícil de establecer, y remitimos a la página 258 del libro [14], donde el lector interesado puede encontrar una demostración.

En 1916 Hardy logró demostrar que lo que ocurre en el ejemplo A.17 se sigue cumpliendo cuando sólo se supone que  $m > mb > 1$ . Hardy también proporcionó otro ejemplo, similar al de Weierstrass, que resolvía una conjetura de Riemann: La suma de la serie uniformemente convergente  $\sum_{k=1}^{\infty} n^{-2} \text{sen}(\pi n^2 x)$  define una función continua que no es derivable en ningún punto. Las sucesivas sumas parciales de esta serie se pueden visualizar en H 1 )H 2) H 3 )H 4 ) H 5 )H 6 )

El siguiente es el clásico ejemplo de Peano de una trayectoria continua y plana cuya imagen llena un cuadrado. Los detalles se pueden ver en [2] pág 225.

**Ejemplo A.18** *Sea  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$  la función continua periódica de periodo 2, cuya restricción al intervalo  $[0, 2]$  viene dada por*

$$\begin{array}{ll} \varphi(t) = 0 & \text{si } t \in [0, 1/3] \cup [5/3, 2] \\ \varphi(t) = 1 & \text{si } t \in [2/3, 4/3] \end{array} \quad \begin{array}{ll} \varphi(t) = 3t - 1 & \text{si } t \in [1/3, 2/3] \\ \varphi(t) = 5 - 3t & \text{si } t \in [4/3, 5/3] \end{array}$$

Para cada  $t \in [0, 1]$  sea  $x(t) = \sum_{n=1}^{\infty} 2^{-n} \varphi(3^{2n-2}t)$ ;  $y(t) = \sum_{n=1}^{\infty} 2^{-n} \varphi(3^{2n-1}t)$ . Entonces  $\mathbf{f}(t) = (x(t), y(t))$  define una función continua  $\mathbf{f} : [0, 1] \rightarrow \mathbb{R}^2$  cuya imagen es el cuadrado  $[0, 1] \times [0, 1]$ .

En [5] págs. 238 y 240 se pueden ver los siguientes ejemplos:

**Ejemplo A.19** Sea  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$  periódica de periodo 4 determinada por los valores  $\varphi(x) = |x|$  para  $|x| \leq 2$ . La serie  $f(x) = \sum_{n \geq 0} 4^{-n} \varphi(4^n x)$  define una función continua  $f : \mathbb{R} \rightarrow \mathbb{R}$  que no es derivable en ningún punto

**Ejemplo A.20** Para cada  $x \in \mathbb{R}$  sea  $f_n(x) = nx - [nx]$  (donde  $[nx]$  es la parte entera de  $nx$ ). La serie  $f(x) = \sum_{n \geq 0} f_n(x) n^{-2}$  define una función  $f : \mathbb{R} \rightarrow \mathbb{R}$  continua en cada  $x$  irracional y discontinua en cada  $x$  racional.

## A.4. Ejercicios resueltos

**Ejercicio A.21** Compruebe que la sucesión de funciones  $f_n(x) = e^{-n^2 x^2} / n$  converge uniformemente hacia 0, en  $\mathbb{R}$ , pero la sucesión de sus derivadas  $f'_n(x) = -2nx e^{-n^2 x^2}$  no converge uniformemente en ningún entorno de 0.

SOLUCIÓN

([5] pág. 222) La primera afirmación es obvia, pues  $\max\{f_n(x) : x \in \mathbb{R}\} = f_n(0) = 1/n$ . Por otra parte, es fácil ver que la sucesión de derivadas  $f'_n(x)$  converge hacia 0 en todo  $x \in \mathbb{R}$ . Con un esquema de la gráfica de  $f'_n$  se observa que  $|f'_n|$  alcanza un máximo absoluto en el punto  $x_n = 1/(n\sqrt{2})$ , cuyo valor es  $|f'_n(x_n)| = \sqrt{2}/e$ . Si  $V \subset \mathbb{R}$  es un entorno de 0, sea  $m \in \mathbb{N}$  tal que  $n \geq m \Rightarrow x_n \in V$ . Entonces, para todo  $n \geq m$  se cumple  $\sup\{|f'_n(x)| : x \in V\} = |f'_n(x_n)| = \sqrt{2}/e$ , luego la sucesión  $f'_n$  no converge uniformemente sobre  $V$ . ■

**Ejercicio A.22** Estudie la convergencia uniforme, en  $[0, +\infty)$  de la sucesión

$$f_n(x) = \frac{\log(x+n)}{ne^x}$$

SOLUCIÓN

Si  $x \geq 0$  la sucesión  $\log(x+n)/n$  converge hacia 0, pues según la regla de l'Hôpital,

$$\lim_{t \rightarrow +\infty} \frac{\log(x+t)}{t} = \lim_{t \rightarrow +\infty} \frac{1}{x+t} = 0$$

Se sigue que para cada  $x \geq 0$  existe  $\lim_n f_n(x) = 0$ , luego la sucesión  $f_n$  converge puntualmente, en  $[0, +\infty)$ , hacia la función idénticamente nula  $f \equiv 0$ .

Para estudiar la convergencia uniforme sobre  $[0, +\infty)$  consideramos la sucesión numérica  $\rho_n = \sup\{f_n(x) : x \geq 0\}$  y para calcularla comenzamos estudiando el signo de la derivada

$$f'_n(x) = \frac{1 - (n+x) \log(x+n)}{(n+x)ne^x}$$

Con este fin consideramos la función auxiliar  $\varphi(t) = 1 - t \log t$ , que crece en  $(0, 1/e)$ , tiene un máximo absoluto para  $t = 1/e$  y decrece en  $(1/e, +\infty)$ . Como  $\varphi(1) = 1 > 0$  y  $\varphi(e) = 1 - e < 0$  con el teorema de Bolzano se obtiene que  $\varphi$  se anula en un punto  $\alpha \in (1, e)$  y se sigue que  $\varphi(t) < 0$  para todo  $t > \alpha$ .

Cuando  $n \geq 3$ , para todo  $x \geq 0$  se cumple  $n + x \geq 3 > \alpha$ , luego  $\varphi(n + x) < 0$  y por lo tanto  $f'_n(x) < 0$ . Es decir, para  $n \geq 3$ , la función  $f_n$  es decreciente en  $[0, +\infty)$  y por lo tanto  $\rho_n = f_n(0) = (\log n)/n$ . Como  $\lim_n \rho_n = 0$ , se concluye que  $f_n$  converge hacia 0 uniformemente sobre  $[0, +\infty)$ . ■

**Ejercicio A.23** Demuestre que la sucesión  $f_n(x) = (n/x) \log(1 + x/n)$  converge uniformemente sobre  $(0, b]$  para todo  $b > 0$  pero no converge uniformemente sobre  $(0, +\infty)$ .

SOLUCIÓN

Como  $\lim_{t \rightarrow 0} \frac{\log(1+t)}{t} = 1$  es claro que para cada  $x > 0$  existe el límite

$$\lim_n f_n(x) = \lim_n \frac{\log(1 + x/n)}{x/n} = 1$$

luego la sucesión converge puntualmente en  $(0, +\infty)$  hacia la función constante 1. Para estudiar la convergencia uniforme en un intervalo  $I \subset (0, +\infty)$  hemos de considerar la sucesión numérica  $\rho_n(I) = \sup_{x \in I} |f_n(x) - 1|$ . Para calcular este supremo conviene estudiar el comportamiento (crecimiento, decrecimiento) de  $f_n$  en el intervalo  $I$ . Este comportamiento lo proporciona el signo de la derivada  $f'_n(x)$  que coincide con el de la expresión

$$\frac{x/n}{1 + x/n} - \log(1 + x/n)$$

Para estudiarlo consideramos la función auxiliar  $\varphi(t) = t/(1+t) - \log(1+t)$ . Como  $\varphi$  es decreciente en  $[0, +\infty)$  (porque  $\varphi'(t) \leq 0$ ) y  $\varphi(0) = 0$ , se cumple que  $\varphi(t) \leq 0$  para todo  $t \geq 0$ . Se sigue de esto que para todo  $n \in \mathbb{N}$  y todo  $x \geq 0$  es  $f'_n(x) \leq 0$ , luego todas las funciones  $f_n$  son decrecientes en  $(0, +\infty)$ . Como  $\lim_{x \rightarrow 0} f_n(x) = 1$ , se sigue que  $|f_n(x) - 1| = 1 - f_n(x)$ . Como  $1 - f_n(x)$  es creciente en  $(0, +\infty)$  se sigue que para  $I = (0, b]$  se cumple

$$\rho_n(I) = \sup_{x \in I} |f_n(x) - 1| = \sup_{x \in I} (1 - f_n(x)) = 1 - f_n(b)$$

luego  $\lim_n \rho_n(I) = 0$ , y la sucesión  $(f_n)$  converge uniformemente sobre  $I = (0, b]$ . Por otra parte, para  $J = (0, +\infty)$  se cumple

$$\rho_n(J) = \sup_{x > 0} |f_n(x) - 1| = \sup_{x > 0} (1 - f_n(x)) = \lim_{x \rightarrow +\infty} (1 - f_n(x)) = 1$$

y por ello la sucesión  $(f_n)$  no converge uniformemente sobre  $J = (0, +\infty)$ . ■



**Ejercicio A.24** Se considera la sucesión  $f_n : [0, 1] \rightarrow \mathbb{R}$ , definida por  $f_n(x) = n^p x(1 - x^2)^n$ , Estudie los valores de  $p > 0$  para los que la sucesión es uniformemente convergente y los valores de  $p > 0$  para los que se cumplen las hipótesis del teorema A.9.

SOLUCIÓN

([5] prob.12, pág. 222) Si  $0 < x < 1$  y  $r = 1 - x^2$  entonces  $0 < r < 1$ , luego  $\log r < 0$  y por lo tanto la sucesión  $n^p r^n = n^p e^{n \log r}$  tiene límite 0 para todo  $p \in \mathbb{R}$ . Por lo tanto  $f_n(x)$  converge hacia 0 para todo  $x \in (0, 1)$ . Como las sucesiones  $f_n(0)$  y  $f_n(1)$  también convergen hacia 0, queda establecido que la sucesión  $f_n$  converge puntualmente, en  $[0, 1]$ , hacia la función constante 0.

En virtud del teorema A.7, y teniendo en cuenta el ejemplo A.3, podemos asegurar que para  $p \geq 1$  la sucesión  $f_n$  no converge uniformemente sobre  $[0, 1]$ . Veamos directamente que la sucesión converge uniformemente si y sólo si  $p < 1/2$ . Con un cálculo rutinario que se deja al cuidado del lector se obtiene que el máximo de  $f_n(x) \geq 0$  en  $[0, 1]$  se alcanza en  $x_n = 1/\sqrt{2n+1}$ , y vale

$$f_n(x_n) = \frac{n^p}{\sqrt{2n+1}} \left(1 - \frac{1}{2n+1}\right)^n$$

y es claro que esta sucesión converge hacia 0 si y sólo si  $p < 1/2$ .

Obsérvese que, para  $p \in [1/2, 1)$ , la sucesión  $f_n$  no es uniformemente convergente, y sin embargo, según los cálculos del ejemplo A.3 se cumple que  $\lim_n \int_0^1 f_n = \int_0^1 f$ . Veamos si en este caso existe una función dominadora de la sucesión  $f_n$  que justifique, de acuerdo con el teorema A.9, el paso al límite bajo la integral. Buscamos una función localmente integrable  $g : (0, 1] \rightarrow [0, +\infty)$ , con integral finita  $\int_0^1 g(x) dx < +\infty$ , que verifique

$$f_n(x) = n^p x(1 - x^2)^n \leq g(x), \text{ para todo } x \in [0, 1] \text{ y todo } n \in \mathbb{N}.$$

Si  $p > 1/2$  el máximo de  $f_n$  en  $[0, 1]$  tiende hacia infinito y se alcanza en un punto  $x_n$ , cada vez más próximo 0. Por lo tanto, la función dominadora, si la hay, no está acotada en los entornos de 0 por lo que es natural buscarla de la forma  $g(t) = C/t^\alpha$ , con  $\alpha < 1$ , ya que así se cumplirá la condición  $\int_0^1 g(x) dx < +\infty$ . En definitiva, basta encontrar  $\alpha < 1$ , de modo que la sucesión

$$\varphi_n(x) = n^p x^{\alpha+1} (1 - x^2)^n$$

esté uniformemente acotada por alguna constante  $C > 0$ . Calculando el máximo de  $\varphi_n$  en  $[0, 1]$  se observa que con  $\alpha \in [2p - 1, 1)$  y  $C = 1$  se consigue una función dominadora (recuérdese que en el caso que estamos considerando es  $2p - 1 < 1$ ). ■

**Ejercicio A.25** Sea  $f_n : [a, b] \rightarrow \mathbb{R}$  una sucesión de funciones continuas que converge uniformemente hacia una función  $f$  tal que  $0 \notin f([a, b])$ . Demuestre que para  $n$  suficientemente grande  $0 \notin f_n([a, b])$  y la sucesión  $1/f_n$  converge uniformemente sobre  $[a, b]$ .

## SOLUCIÓN

Las funciones  $f_n$  son continuas y la convergencia es uniforme, luego la función límite  $f$  también es continua. Como  $[a, b]$  es cerrado y acotado existe  $z \in [a, b]$  donde la función continua  $|f|$  alcanza el mínimo absoluto  $\min\{|f(x)| : x \in [a, b]\} = |f(z)|$ . Por la hipótesis  $\mu = |f(z)| > 0$  y en virtud de la convergencia uniforme, existe  $n_0$  tal que para  $n \geq n_0$  y todo  $x \in [a, b]$  se cumple  $|f_n(x) - f(x)| \leq \mu/2$ . Esto implica que  $0 \notin f_n([a, b])$  cuando  $n \geq n_0$  (si fuese  $f_n(x) = 0$  para algún  $y \in [a, b]$ , sería  $|f(y)| < \mu/2$ , ¡absurdo!).

Si  $n \geq n_0$  y  $x \in [a, b]$  se cumple  $|f_n(x)| \geq |f(x)| - |f(x) - f_n(x)| \geq \mu - \mu/2 = \mu/2$ , luego

$$\left| \frac{1}{f_n(x)} - \frac{1}{f(x)} \right| \leq \frac{|f(x) - f_n(x)|}{|f(x)||f_n(x)|} \leq \frac{2}{\mu^2} |f(x) - f_n(x)| \leq \frac{2}{\mu^2} \rho_n$$

donde  $\rho_n = \sup_{x \in [a, b]} |f_n(x) - f(x)|$  converge hacia 0. Se sigue que

$$r_n = \sup_{x \in [a, b]} \left| \frac{1}{f_n(x)} - \frac{1}{f(x)} \right| \leq 2\mu^{-2} \rho_n$$

converge hacia 0, lo que significa que  $1/f_n$  converge hacia  $1/f$  uniformemente sobre  $[a, b]$ . ■

**Ejercicio A.26** Sea  $g : \mathbb{R} \rightarrow \mathbb{R}$  una función continua tal que  $g(x) > 0$  para todo  $x \in \mathbb{R}$ . Demuestre que la sucesión de funciones  $f_n(x) = ng(x)/(1 + ng(x))$  converge uniformemente sobre cada intervalo acotado  $[a, b] \subset \mathbb{R}$ . Estudie la convergencia uniforme sobre intervalos no acotados cuando  $g(x) = e^x$ .

## SOLUCIÓN

Para todo  $x \in \mathbb{R}$  existe  $\lim_n f_n(x) = 1$ , es decir, la sucesión  $f_n$  converge puntualmente hacia la función constante 1. Para estudiar la convergencia uniforme en un intervalo  $I \subset \mathbb{R}$ , se considera la sucesión numérica

$$\rho_n(I) = \sup\{|f_n(x) - 1| : x \in I\} = \sup\{1/(1 + ng(x)) : x \in I\}$$

Cuando  $I = [a, b] \subset \mathbb{R}$ , es claro que  $\rho_n([a, b]) = 1/(1 + n\alpha)$  donde  $\alpha > 0$  es el mínimo absoluto de la función continua  $g$  sobre el intervalo compacto  $[a, b]$  (obsérvese que  $\alpha = g(x_0)$  para algún  $x_0 \in [a, b]$ , luego  $\alpha > 0$ ). Como  $\lim_n \rho_n([a, b]) = 0$ , podemos afirmar que la sucesión  $f_n$  converge uniformemente sobre  $[a, b]$ .

Cuando  $g(x) = e^x$ , se verifica

$$\rho_n([a, +\infty)) = 1/(1 + ne^a), \quad \rho_n((-\infty, b]) = 1$$

luego la sucesión  $f_n$  converge uniformemente sobre los intervalos  $[a, +\infty)$ , pero no converge uniformemente sobre los intervalos  $(-\infty, b]$ . ■

Para los ejercicios que siguen se suponen conocidas las funciones elementales de variable compleja: La función exponencial  $e^z$  y la validez de la ecuación funcional  $e^{z+w} = e^z e^w$ , así como la definición habitual de las funciones de variable compleja

$$\operatorname{sen} z = \frac{e^{iz} - e^{-iz}}{2i}, \quad \cos z = \frac{e^{iz} + e^{-iz}}{2}, \quad \operatorname{tg} z = \frac{\operatorname{sen} z}{\cos z}, \quad \operatorname{cot} z = \frac{\cos z}{\operatorname{sen} z}$$

**Ejercicio A.27** *Se considera la función exponencial de variable compleja*

$$e^z = \sum_{n=0}^{+\infty} \frac{z^n}{n!}$$

Si  $|z| \leq m \in \mathbb{N}$ , establezca las desigualdades

$$\left| e^z - \left(1 + \frac{z}{m}\right)^m \right| \leq e^{|z|} - \left(1 + \frac{|z|}{m}\right)^m \leq \frac{|z|^2 e^{|z|}}{m}$$

Deduzca de ellas que, para cada  $R > 0$ , la sucesión  $(1 + z/n)^n$  converge hacia  $e^z$  uniformemente sobre  $\{z : |z| \leq R\}$ .

SOLUCIÓN

$e^z - (1 + z/m)^m = D_m + R_m$  donde

$$D_m(z) = \sum_{n=0}^m \frac{z^n}{n!} - \left(1 + \frac{z}{m}\right)^m, \quad R_m(z) = \sum_{n=m+1}^{+\infty} \frac{z^n}{n!}.$$

Usando la fórmula del binomio de Newton

$$D_m(z) = \frac{z^2}{2!} \left(1 - \frac{m-1}{m}\right) + \frac{z^3}{3!} \left(1 - \frac{(m-1)(m-2)}{m^2}\right) + \cdots + \frac{z^m}{m!} \left(1 - \frac{m!}{m^m}\right)$$

Aplicando la desigualdad triangular y teniendo en cuenta que en la expresión anterior los paréntesis son positivos se obtiene que  $|D_m(z)| \leq D_m(|z|)$ .

Por otra parte, es inmediato que  $|R_m(z)| \leq R_m(|z|)$ , luego

$$\left| e^z - \left(1 + \frac{z}{m}\right)^m \right| \leq D_m(|z|) + R_m(|z|) = e^{|z|} - \left(1 + \frac{|z|}{m}\right)^m$$

En virtud de la desigualdad  $1 + x \leq e^x$ , válida para todo  $x \in \mathbb{R}$ , se cumple

$$\left(1 + \frac{x}{m}\right)^m \leq e^x, \quad \left(1 - \frac{x}{m}\right)^m \leq e^{-x},$$

y cuando  $0 \leq x \leq m$  se obtienen las desigualdades

$$\begin{aligned} 0 &\leq e^x - \left(1 + \frac{x}{m}\right)^m \leq e^x \left[1 - e^{-x} \left(1 + \frac{x}{m}\right)^m\right] \leq \\ &\leq e^x \left[1 - \left(1 - \frac{x}{m}\right)^m \left(1 + \frac{x}{m}\right)^m\right] = e^x \left[1 - \left(1 - \frac{x^2}{m^2}\right)^m\right] = \\ &= e^x \frac{x^2}{m^2} \left[1 + \left(1 - \frac{x^2}{m^2}\right) + \left(1 - \frac{x^2}{m^2}\right)^2 + \cdots + \left(1 - \frac{x^2}{m^2}\right)^{m-1}\right] \leq \\ &\leq e^x \frac{x^2}{m^2} m = \frac{x^2 e^x}{m} \end{aligned}$$

Con  $x = |z|$  se obtiene la segunda desigualdad del enunciado. En virtud de las desigualdades establecidas, si  $|z| \leq R$ , se verifica

$$\left| e^z - \left(1 + \frac{z}{m}\right)^m \right| \leq \frac{R^2 e^R}{m}$$

luego

$$\lim_m \left(1 + \frac{z}{m}\right)^m = e^z \quad \text{uniformemente en } \{z : |z| \leq R\}.$$

■

**Ejercicio A.28** Se supone que la sucesión  $f_n : K \rightarrow \mathbb{C}$  converge uniformemente sobre  $K$  hacia una función  $f = u + iv$  cuya parte real  $u$  está acotada superiormente sobre  $K$ . Demuestre que la sucesión  $e^{f_n(z)}$  converge uniformemente sobre  $K$ .

SOLUCIÓN

Se supone que  $u(z) \leq M$  para todo  $z \in K$ . Entonces cuando  $z \in K$  se cumple

$$\begin{aligned} |e^{f_n(z)} - e^{f(z)}| &= |e^{f(z)}| |e^{f_n(z)-f(z)} - 1| \leq \\ &\leq e^{u(z)} |e^{f_n(z)-f(z)} - 1| \leq e^M |e^{f_n(z)-f(z)} - 1| \end{aligned}$$

Como  $e^z$  es continua en  $z = 0$ , dado  $\epsilon > 0$  existe  $\delta > 0$  tal que

$$|w| < \delta \Rightarrow |e^w - 1| < \epsilon e^{-M}.$$

Por la convergencia uniforme de  $f_n$  existe  $n(\delta) \in \mathbb{N}$  tal que si  $n \geq n(\delta)$  entonces para todo  $z \in K$  se cumple  $|f_n(z) - f(z)| < \delta$ . Combinando las dos afirmaciones anteriores se concluye que para todo  $n \geq n(\delta)$  y todo  $z \in K$  se verifica

$$|e^{f_n(z)} - e^{f(z)}| \leq e^M |e^{f_n(z)-f(z)} - 1| \leq e^M \epsilon e^{-M} = \epsilon$$

■

**Ejercicio A.29** Demuestre que  $\lim_{n \rightarrow \infty} \operatorname{tg} nz = -i$ , y que para cada  $\epsilon > 0$  el límite es uniforme sobre el semiplano  $H_\epsilon := \{z : \operatorname{Im} z < -\epsilon\}$ .

SOLUCIÓN

$$\operatorname{tg} nz = \frac{\operatorname{sen} nz}{\operatorname{cos} nz} = \frac{1}{i} \frac{e^{inz} - e^{-inz}}{e^{inz} + e^{-inz}} = \frac{1}{i} \frac{e^{i2nz} - 1}{e^{i2nz} + 1}$$

luego

$$|\operatorname{tg} nz + i| = \left| \operatorname{tg} nz - \frac{1}{i} \right| = \left| \frac{e^{i2nz} - 1}{e^{i2nz} + 1} - 1 \right| = \left| \frac{2}{e^{i2nz} + 1} \right|$$

de donde se sigue que para todo  $z \in H_\epsilon$  se verifica

$$|\operatorname{tg} nz + i| \leq \frac{2}{|e^{i2nz}| - 1} = \frac{2}{e^{-2ny} - 1} \leq \frac{2}{e^{2n\epsilon} - 1}$$

Como la sucesión  $2/(e^{2n\epsilon} - 1)$  converge hacia 0, la última desigualdad nos asegura que  $\lim_n \operatorname{tg} nz = -i$ , uniformemente sobre  $H_\epsilon$ . ■

**Ejercicio A.30** Demuestre que  $\lim_n \cotg(x+in) = -i$ , y que el límite es uniforme respecto de  $x \in \mathbb{R}$ .

SOLUCIÓN

Para todo  $z = x + iy$  se cumple

$$|\cotg z + i| = \left| i \frac{e^{iz} + e^{-iz}}{e^{iz} - e^{-iz}} + i \right| = \left| \frac{2e^{i2z}}{e^{i2z} - 1} \right| \leq \frac{2e^{-2y}}{1 - e^{-2y}}$$

donde la función  $h(y) = 2e^{-2y}/(1 - e^{-2y})$  converge hacia 0 cuando  $y \rightarrow +\infty$ . Como para todo  $x \in \mathbb{R}$  se cumple la desigualdad  $|\cot(x + in) + i| \leq h(n)$  se concluye que la sucesión  $f_n(x) = \cot(x + in)$  converge hacia  $-i$  uniformemente respecto de  $x \in \mathbb{R}$ .

■

## A.5. Ejercicios propuestos

◇ **A.5.1** Muestre que la sucesión  $f_n(x) = 1/(1+n^2x^2)$  no converge uniformemente sobre  $[0, 1]$  pero su límite puntual  $f$  verifica  $\int_0^1 f(x)dx = \lim_n \int_0^1 f_n(x)dx$ . (Obsérvese que esta sucesión cumple las hipótesis del teorema A.8)

◇ **A.5.2** Si la sucesión  $f_n : T \rightarrow \mathbb{R}$  converge uniformemente sobre  $T$  demuestre que la sucesión  $\text{sen } f_n(t)$  también converge uniformemente sobre  $T$ .

◇ **A.5.3** Se considera la sucesión de funciones  $f_n : [0, 1] \rightarrow \mathbb{R}$  definida por:

$$f_n(x) = n^2x(1-nx) \text{ si } x \in [0, 1/n]; \quad f_n(x) = 0 \text{ si } x \in (1/n, 1]$$

Demuestre que la sucesión converge puntualmente hacia 0, pero no converge uniformemente sobre  $[0, 1]$ . ¿Sobre qué intervalos  $I \subset [0, 1]$  la convergencia es uniforme?

◇ **A.5.4** Dada una sucesión estrictamente creciente  $a_n \in [0, 1]$  estudie la convergencia puntual y uniforme de la sucesión de funciones  $f_n : [0, 1] \rightarrow \mathbb{R}$ , definida así:

$$f_n(x) = \frac{(x-a_n)(x-a_{n+1})}{(a_{n+1}-a_n)^2} \quad \text{si } x \in [a_n, a_{n+1}]$$

$$f_n(x) = 0 \quad \text{si } x \notin [a_n, a_{n+1}].$$

◇ **A.5.5** Estudie la convergencia puntual y uniforme de la sucesión

$$g_n(x) = x^{2n}/(1+x^{2n})$$

sobre  $\mathbb{R}$  y sobre  $\{x \in \mathbb{R} : |x| \geq a\}$ , con  $a > 0$ .

◇ **A.5.6** En cada uno de los siguientes casos estudie los intervalos  $I \subset \mathbb{R}$  sobre los que la sucesión de funciones  $f_n : \mathbb{R} \rightarrow \mathbb{R}$  es uniformemente convergente.

$$a) \quad f_n(x) = \frac{1}{1+x^{2n}}; \quad b) \quad f_n(x) = \frac{x}{1+x^{2n}};$$

$$c) \quad f_n(x) = \frac{n^2x}{1+n^3x^2}; \quad d) \quad f_n(x) = \frac{x^2}{x^2+(x-n)^2};$$

$$e) \quad f_n(x) = \frac{x^2}{1+n|x|}; \quad f) \quad f_n(x) = \frac{x}{1+nx^2};$$

$$g) \quad f_n(x) = \frac{1}{1+(x-n)^2}; \quad h) \quad f_n(x) = \frac{|x-n|+|x|}{n};$$

Para las sucesiones de los apartados f) y g) estudie la validez de la derivación término a término.

◇ **A.5.7** Se consideran las sucesiones  $s_n(t) = \operatorname{sen}(\lambda nt)e^{-nt}$ ,  $c_n(t) = \operatorname{cos}(\lambda nt)e^{-nt}$  definidas en  $[0, +\infty)$ , donde  $\lambda \neq 0$  es un parámetro real.

a) Obtenga los límites puntuales de ambas sucesiones, y justifique que, para cada  $a > 0$  las dos sucesiones convergen uniformemente sobre  $[a, +\infty)$ .

b) Estimando la sucesión  $d_n := \sup\{|s_n(t)| : t > 0\}$ , deduzca que la sucesión  $s_n$  no converge uniformemente sobre  $[0, +\infty)$ . Justifique sin cálculos que la sucesión  $c_n$  tampoco converge uniformemente sobre  $[0, +\infty)$ .

◇ **A.5.8** Estudie, según los valores del parámetro real  $a > 0$ , los intervalos  $I \subset \mathbb{R}$  sobre los que la sucesión  $f_n(x) = \frac{nx}{1+n^ax^2}$  es uniformemente convergente.

◇ **A.5.9** Para  $p = 1, 2$ , estudie los intervalos  $I \subset \mathbb{R}$  sobre los que es uniformemente convergente la sucesión  $f_n(x) = \frac{nx^p}{1+n^2x^2}$ .

◇ **A.5.10** Si  $g : [0, 1] \rightarrow \mathbb{R}$  es continua, demuestre que la sucesión  $x^n g(x)$  converge uniformemente en  $[0, 1]$  si y sólo si  $g(1) = 0$ .

◇ **A.5.11** Se supone que  $f_n : [0, 1] \rightarrow \mathbb{R}$  es una sucesión de funciones continuas que converge uniformemente hacia  $f$ . Demuestre que

$$\int_0^1 f(x)dx = \lim_n \int_0^{1-1/n} f_n(x)dx$$

◇ **A.5.12** Si una sucesión de funciones continuas  $f_n : \mathbb{R} \rightarrow \mathbb{R}$  converge uniformemente sobre  $(a, b)$  demuestre que también converge uniformemente sobre  $[a, b]$ . Demuestre que la sucesión  $f_n(x) = x^2/(1+x^{2n})$  converge uniformemente sobre cada intervalo  $[-r, r] \subset (-1, 1)$  pero no converge uniformemente sobre  $(-1, 1)$ .

◇ **A.5.13** Demuestre que la sucesión

$$f_n(x) = \frac{x^2}{x^2 + (1-nx)^2}$$

converge puntualmente hacia 0 pero no posee subsucesiones uniformemente convergentes.

◇ **A.5.14** Sean  $f_n, g_n : T \rightarrow \mathbb{R}$  sucesiones uniformemente convergentes hacia  $f, g : T \rightarrow \mathbb{R}$ , respectivamente. Si  $f$  y  $g$  son acotadas, demuestre que la sucesión producto  $f_n g_n$  converge uniformemente hacia  $fg$ .

◇ **A.5.15** Se considera la sucesión de funciones  $f_n : \mathbb{R} \rightarrow \mathbb{R}$  definida por

$$f_n(x) = 1/n \text{ si } x = 0 \text{ o si } x \text{ es irracional}$$

$$f_n(x) = 1/n + q \text{ si } x = p/q, \text{ fracción irreducible, con } p, q \in \mathbb{Z}, q > 0.$$

Sea  $g_n(x) = (1 + 1/n)x$ . Compruebe que las sucesiones  $f_n$ ,  $g_n$  convergen uniformemente sobre  $[-R, R]$ , pero el producto  $f_n g_n$  no converge uniformemente sobre  $[-R, R]$ .

(Ejercicio 9.2) de [2])

◇ **A.5.16** Sea  $f_n : T \rightarrow \mathbb{R}$  una sucesión de funciones continuas, definidas en un intervalo  $T \subset \mathbb{R}$  que converge puntualmente hacia la función  $f : T \rightarrow \mathbb{R}$ . Demuestre que son equivalentes:

- $f_n$  converge uniformemente sobre cada intervalo cerrado y acotado  $[a, b] \subset T$ .
- $f$  es continua y para cada sucesión  $x_n \in T$  convergente hacia un punto  $x \in T$ , existe el límite  $\lim_n f_n(x_n)$

◇ **A.5.17** Compruebe que para cada  $m \in \mathbb{N}$  y cada  $x \in \mathbb{R}$  existe el límite puntual  $f_m(x) = \lim_n (\cos m! \pi x)^{2^n}$ . Demuestre que cada  $f_m$  es integrable Riemann sobre  $[0, 1]$ , pero su límite puntual  $f(x) = \lim_m f_m(x)$  no lo es.

◇ **A.5.18** Demuestre que la serie  $\sum_{n=1}^{+\infty} a_n(t)b_n(t)$  converge uniformemente sobre  $K \subset T$  cuando las sucesiones de funciones  $a_n, b_n : T \rightarrow \mathbb{R}$  verifican:

- La serie  $\sum_{n=1}^{\infty} b_n(t)$  converge uniformemente sobre  $K \subset T$ .
- Existe  $C > 0$  tal que  $|a_1(t)| + \sum_{n=1}^{\infty} |a_n(t) - a_{n+1}(t)| \leq C$  para todo  $t \in K$ .

Obtenga como corolario el criterio de Abel A.15.

◇ **A.5.19** Demuestre que la serie  $\sum_{n=1}^{+\infty} a_n(t)b_n(t)$  converge puntualmente sobre  $K \subset T$  cuando las sucesiones de funciones  $a_n, b_n : T \rightarrow \mathbb{R}$  verifican:

- La sucesión  $B_n(t) = \sum_{j=1}^n b_j(t)$  está uniformemente acotada sobre  $K \subset T$ .
- La sucesión de funciones  $a_n(t)$  converge puntualmente hacia 0 sobre  $K$ , y la serie  $\sum_{n=1}^{\infty} |a_n(t) - a_{n+1}(t)|$  converge uniformemente sobre  $K$ .

Obtenga como corolario el criterio de Dirichlet A.14.

◇ **A.5.20** Demuestre que la serie  $\sum_{n=1}^{\infty} x^n(1-x)$  no converge uniformemente sobre  $[0, 1]$ , pero la serie  $\sum_{n=1}^{\infty} (-x)^n(1-x)$  sí converge uniformemente sobre  $[0, 1]$ ,

◇ **A.5.21** Demuestre que la serie  $\sum_{n=1}^{\infty} (-1)^n \frac{1+x^n}{n}$  converge uniformemente sobre cada intervalo  $[a, b] \subset (-1, 1)$ , pero no converge absolutamente en ningún punto del intervalo  $(-1, 1)$ .

◇ **A.5.22** Sea  $f_n(x) = 0$  si  $x < 1/(n+1)$ ;  $f_n(x) = \sin^2(\pi/x)$  si  $x \in [1/(n+1), 1/n]$ ,  $f_n(x) = 0$  si  $x > 1/n$ . Demuestre que la serie  $\sum_{n=1}^{\infty} f_n(x)$  es absolutamente convergente pero la convergencia no es uniformemente en ningún entorno de 0.

◇ **A.5.23** Compruebe que la serie  $\sum_{n=1}^{+\infty} ne^{-nx}$  converge uniformemente sobre  $[a, +\infty)$ , para cada  $a > 0$ . Utilice el teorema de integración término a término de series funcionales para obtener su suma.



◇ **A.5.24** Compruebe que la serie  $\sum_{n=1}^{+\infty} e^{-nx}/(1+n^2)$  converge para  $x \geq 0$  y define en  $[0, +\infty)$  una función continua que es derivable en cada  $x > 0$ .

◇ **A.5.25** Demuestre que la serie  $\sum_{n=1}^{+\infty} \left(\frac{1}{n} - \frac{1}{x+n}\right)$  converge para todo  $x \geq 0$  y que su suma  $S(x)$  es una función continua estrictamente creciente en  $[0, +\infty)$ .

◇ **A.5.26** Se considera la serie de funciones  $\sum_{n=1}^{+\infty} \frac{1}{1+n^2|x|}$ .

Determine los valores de  $x$  para los que la serie converge. ¿En qué intervalos la convergencia de la serie no es uniforme? ¿En qué puntos es continua la función  $f$  definida por la suma de la serie? ¿Es  $f$  acotada?

◇ **A.5.27** Estudie la convergencia puntual y la convergencia uniforme sobre intervalos de las series

$$\sum_{n=1}^{+\infty} \frac{1}{1+x^n}, \quad \sum_{n=1}^{+\infty} \frac{x}{1+x^n}$$

En cada caso estudie la derivabilidad de la suma de la serie en el interior de su dominio de convergencia.

◇ **A.5.28** Justifique que la serie de funciones

$$\sum_{n=1}^{+\infty} \frac{x^2}{(1+x^2)^n}$$

converge puntualmente en todo  $\mathbb{R}$ , y que para cada  $\epsilon > 0$  hay convergencia uniforme en  $\{x : |x| > \epsilon\}$  y no hay convergencia uniforme en  $\{x : |x| < \epsilon\}$ .

◇ **A.5.29** Se  $(a_n)$  una sucesión decreciente de números reales con  $\lim_n a_n = 0$ . Justifique que, para cada  $\delta \in (0, 1)$  la serie  $\sum_{n=1}^{+\infty} a_n x^n$  converge uniformemente en  $A_\delta = [-1, 1 - \delta]$ . Muestre que la serie converge uniformemente sobre  $[-1, 1)$  si  $\sum_{n=1}^{+\infty} a_n < +\infty$ ,

◇ **A.5.30** Estudie la convergencia uniforme de las series

$$\sum_{n=1}^{+\infty} \frac{1}{n^x}; \quad \sum_{n=1}^{+\infty} \frac{(-1)^n}{n^x}$$

y demuestre que la suma de la primera define en  $(1, +\infty)$  una función derivable

$$S(x) = \sum_{n=1}^{+\infty} \frac{1}{n^x} \text{ con derivada } S'(x) = - \sum_{n=1}^{+\infty} \frac{\log n}{n^x}.$$

◇ **A.5.31** Sea  $x_n \in (a, b)$  una sucesión de puntos distintos y  $f_n : (a, b) \rightarrow \mathbb{R}$  la función definida por  $f_n(x) = 0$  si  $x \leq x_n$ ,  $f_n(x) = 1$  si  $x > x_n$ . Demuestre que la suma serie  $f(x) = \sum_n 2^{-n} f_n(x)$ , define en  $(a, b)$  una función, que es continua en  $x \in (a, b)$  si y sólo si  $x \notin \{x_n : n \in \mathbb{N}\}$ . Deduzca de ello que existe una función estrictamente creciente  $f : \mathbb{R} \rightarrow \mathbb{R}$ , que es continua en cada  $x \notin \mathbb{Q}$  y discontinua en cada  $x \in \mathbb{Q}$ .

◇ **A.5.32** Demuestre que la serie  $\sum_{n=0}^{+\infty} (-x^2)^n (\log x)^2$  converge uniformemente sobre cada  $[a, b] \subset (0, 1)$  y que su suma  $f(x)$  posee una integral impropia convergente, cuyo valor es

$$\int_0^1 f(x) dx = 2 \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n+1)^3}$$

# B

## Complementos al capítulo 2

### B.1. La recta real

Realizamos aquí un breve repaso de las propiedades topológicas de la recta real, haciendo énfasis en aquellos aspectos donde interviene el orden.

Las propiedades que caracterizan al cuerpo  $\mathbb{R}$  de los números reales se resumen diciendo que  $\mathbb{R}$ , con la relación de desigualdad usual  $\leq$ , es un cuerpo ordenado completo respecto al orden. Esto significa que la relación de orden es compatible con las operaciones del cuerpo:

$$x, y \in \mathbb{R}, \quad 0 \leq x, \quad 0 \leq y \quad \Rightarrow \quad 0 \leq x + y, \quad 0 \leq xy$$

y que todo conjunto no vacío acotado superiormente  $M \subset \mathbb{R}$  tiene extremo superior, es decir, existe una cota superior mínima de  $M$ , denotada  $\sup M$ .

Con la función valor absoluto  $|x|$  se define la distancia entre dos números  $d(x, y) = |x - y|$ , y con ella la noción de sucesión convergente y de sucesión de Cauchy: Una sucesión de números reales  $x_n \in \mathbb{R}$  es de Cauchy si para cada  $\epsilon > 0$  existe  $n(\epsilon) \in \mathbb{N}$  tal que

$$p, q \in \mathbb{N}, \quad p \geq q \geq n(\epsilon) \Rightarrow |x_p - x_q| < \epsilon$$

Usando que  $\mathbb{R}$  es completo respecto al orden se demuestra fácilmente:

- i) Toda sucesión de Cauchy de números reales es convergente
- ii)  $\mathbb{R}$  es un cuerpo ordenado arquimediano: Para cada par de números reales  $x > 0, y > 0$  existe  $n \in \mathbb{N}$  tal que  $nx > y$ .

Estas dos propiedades caracterizan al cuerpo de los números reales: Si  $(\mathbb{K}, \leq)$  es un cuerpo ordenado que contiene a  $\mathbb{Q}$  como subcuerpo ordenado, entonces las tres propiedades que siguen son equivalentes:

- a)  $(\mathbb{K}, \leq)$  es un cuerpo ordenado completo respecto al orden.
- b)  $(\mathbb{K}, \leq)$  es un cuerpo ordenado arquimediano en el que toda sucesión de Cauchy es convergente.

c)  $(\mathbb{K}, \leq)$  es isomorfo, como cuerpo ordenado, al cuerpo de los números reales  $\mathbb{R}$ .

Una consecuencia notable de estas propiedades es el hecho de que  $\mathbb{R}$  no es numerable. Usando la distancia  $d(x, y) = |x - y|$  se definen las nociones topológicas usuales de la recta real que se consideran más adelante en el contexto general de los espacios métricos. De momento nos limitamos a recordar brevemente aquellas propiedades topológicas que son características de  $\mathbb{R}$ .

La primera de ellas se refiere a la estructura de los conjuntos abiertos: Todo abierto  $\Omega \subset \mathbb{R}$  se puede descomponer, de modo único, como unión de una familia numerable de intervalos abiertos (en sentido amplio) disjuntos, es decir,  $\Omega = \bigcup_{n \in M} I_n$  donde  $M \subset \mathbb{N}$  y los  $I_n$  son intervalos disjuntos de la forma  $I_n = (a_n, b_n)$  con  $-\infty \leq a_n < b_n \leq +\infty$ . Si  $\Omega$  no está acotado superiormente (resp. inferiormente) habrá un intervalo de la forma  $I_p = (a_p, +\infty)$ , (resp.  $I_q = (-\infty, b_q)$ ). Esta propiedad, en el lenguaje de la topología general, se expresa diciendo que en  $\mathbb{R}$  las componentes conexas de los abiertos son intervalos abiertos (sólo puede haber una cantidad numerable de estos intervalos porque cada intervalo abierto no vacío contiene un número racional).

Nociones específicas para sucesiones de números reales son las de límite superior y límite inferior: El límite superior (resp. inferior) de una sucesión acotada de números reales  $x_n$  es el límite de la sucesión decreciente acotada  $b_k = \sup\{x_n : n \geq k\}$  (resp. creciente acotada  $a_k = \inf\{x_n : n \geq k\}$ ), es decir

$$\overline{\lim}_n x_n = \lim_{k \rightarrow \infty} \sup_{n \geq k} x_n, \quad \underline{\lim}_n x_n = \lim_{k \rightarrow \infty} \inf_{n \geq k} x_n$$

Si  $x_n$  es una sucesión acotada de números reales,  $a = \underline{\lim}_n x_n$  y  $b = \overline{\lim}_n x_n$  son puntos de aglomeración de la sucesión  $x_n$ , es decir, son límites de subsucesiones convergentes extraídas de la sucesión. Además  $a$  y  $b$  son, respectivamente, el menor y el mayor punto de aglomeración de la sucesión y la sucesión  $x_n$  es convergente si y sólo si  $\underline{\lim}_n x_n = \overline{\lim}_n x_n$ , lo que ocurre si y sólo si la sucesión tiene un único punto de aglomeración.

Por último, recordemos que para un conjunto  $K \subset \mathbb{R}$  son equivalentes

- a)  $K$  es cerrado y acotado.
- b) De cada sucesión  $x_n \in K$  se puede extraer una subsucesión convergente hacia un punto de  $K$
- c) De todo cubrimiento abierto de  $K$  se puede extraer un subcubrimiento finito.

Los conjuntos  $K \subset \mathbb{R}$  que cumplen estas propiedades equivalentes se llaman compactos. La propiedad c) es la que se adopta para dar la definición general de conjunto compacto en un espacio topológico general.

Como  $\mathbb{R}$  no es compacto, para diversas cuestiones conviene ampliar la recta real y considerar la recta real ampliada  $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\}$  que se obtiene añadiendo un primer elemento,  $-\infty$ , y un último elemento,  $+\infty$ , es decir, se adopta la validez, para todo  $x \in \mathbb{R}$ , de las desigualdades  $-\infty \leq x \leq +\infty$  y se escribe  $\overline{\mathbb{R}} = [-\infty, +\infty]$ .

En la recta real ampliada  $\overline{\mathbb{R}}$  todo subconjunto tiene extremo superior y extremo inferior. Si  $M \subset \mathbb{R}$  no está acotado superiormente (resp. inferiormente) se verifica  $\sup M = +\infty$  (resp.  $\inf M = -\infty$ ).

Los entornos de  $+\infty$  (resp.  $-\infty$ ) en  $\overline{\mathbb{R}}$  son los conjuntos que contienen a un intervalo de la forma  $(a, +\infty]$  (resp.  $[-\infty, b)$ ) y los subconjuntos de  $\overline{\mathbb{R}}$  que son entornos de todos sus puntos se llaman abiertos. De esta forma  $\overline{\mathbb{R}}$  resulta compacto, es decir, de todo cubrimiento abierto de  $\overline{\mathbb{R}}$  se puede extraer un subcubrimiento finito. Ahora toda sucesión en  $\overline{\mathbb{R}}$  tiene una subsucesión convergente. Si la sucesión está contenida en  $\mathbb{R}$  y no está acotada superiormente (resp. inferiormente) en  $\mathbb{R}$  entonces hay una subsucesión que converge hacia  $+\infty$  (resp.  $-\infty$ ).

## B.2. Completitud y compacidad

En un espacio métrico  $(E, d)$  todo conjunto compacto  $M \subset E$  es completo ya que, según el teorema 2.7, toda sucesión de Cauchy en el espacio métrico  $(M, d_M)$ , posee una subsucesión convergente hacia un punto de  $M$ , y por lo tanto converge en este espacio. El siguiente objetivo es demostrar la validez del recíproco cuando se cumple una propiedad que se define a continuación:

**Definición B.1** *Un subconjunto  $M$  del espacio métrico  $(E, d)$  se dice que es totalmente acotado cuando para cada  $\epsilon > 0$  existe un conjunto finito  $H \subset E$  tal que  $M \subset \bigcup_{\mathbf{a} \in H} B(\mathbf{a}, \epsilon)$ .*

La definición de conjunto totalmente acotado es equivalente a la que resulta considerando bolas cerradas en lugar de bolas abiertas, y se sigue de esto que si  $M$  es totalmente acotado, su clausura  $\overline{M}$  también lo es. Es inmediato que todo conjunto compacto es totalmente acotado y que todo conjunto totalmente acotado es acotado. El problema 2.6.28 sirve para ver que el recíproco es falso: En el espacio  $(C([0, 1], \|\cdot\|_\infty)$  la bola cerrada  $\{f \in C([0, 1]) : \|f\|_\infty \leq 1\}$  es un conjunto acotado y completo que no compacto luego, en virtud del siguiente teorema, no es totalmente acotado.

**Teorema B.2** *Un subconjunto  $K$  de un espacio métrico  $(E, d)$  es compacto si y sólo si es completo y totalmente acotado.*

DEM: Basta demostrar que todo conjunto completo y totalmente acotado  $K \subset E$  es compacto y esto lo obtendremos viendo que todo conjunto infinito  $M \subset K$  tiene un punto de acumulación en  $K$  (teorema 2.7).

Como  $K$  es totalmente acotado el subconjunto  $M \subset K$  se puede recubrir con un número finito de bolas de radio 1, y alguna de ellas, que denotamos  $B(\mathbf{a}_1, 1)$ , cumple que  $M_1 = M \cap B(\mathbf{a}_1, 1)$  es infinito. Razonando en forma similar con el conjunto infinito  $M_1 \subset K$  y con bolas de radio  $1/2$  se obtiene una bola  $B(\mathbf{a}_2, 1/2)$  con la propiedad de que  $M_2 = M \cap B(\mathbf{a}_2, 1/2)$  es infinito. De modo recurrente obtenemos una sucesión de bolas  $B(\mathbf{a}_n, 1/n)$  y una sucesión decreciente de conjuntos infinitos  $M_n \subset B(\mathbf{a}_n, 1/n)$ . Consideramos la sucesión decreciente de conjuntos cerrados no vacíos  $C_n = \overline{M_n} \subset K$  que, en virtud del lema 2.13, verifica

$\text{diam}(C_n) = \text{diam}(M_n) \leq 2/n$ . Como estamos suponiendo que el espacio métrico  $(K, d_K)$  es completo, según el teorema 2.14, existe un punto  $\mathbf{x} \in \bigcap_n C_n \subset K$ . Es claro que  $\mathbf{x}$  es punto de acumulación de  $M$  (cada bola  $B(\mathbf{x}, \epsilon)$  contiene infinitos puntos de  $M$  pues si  $2/k < \epsilon$  el conjunto infinito  $M_k$  está contenido en  $C_k \subset \overline{B(\mathbf{x}, 2/k)} \subset B(\mathbf{x}, \epsilon)$ ). ■

La noción de conjunto totalmente acotado no es topológica:  $\mathbb{R}$  con la distancia usual no es totalmente acotado, pero lo es con la distancia equivalente  $d'$  considerada en el problema 2.6.5. Es fácil comprobar que toda aplicación uniformemente continua  $\mathbf{f} : (E, d) \rightarrow (F, \rho)$  transforma conjuntos totalmente acotados en conjuntos totalmente acotados. En particular, en el contexto de los espacios normados se verifica:

**Proposición B.3** *Dos normas equivalentes en un espacio vectorial  $E$  (real o complejo) producen los mismos conjuntos totalmente acotados.*

DEM: Es una consecuencia sencilla de la proposición 2.4 y se deja como ejercicio. ■

Las tres normas que hemos considerado en  $\mathbb{R}^n$  producen los mismos conjuntos totalmente acotados porque son equivalentes. Con la norma  $\|\cdot\|_\infty$  es fácil ver que las bolas cerradas  $\overline{B_\infty(\mathbf{0}, r)}$  son totalmente acotadas: Todo cubo  $n$ -dimensional  $[-r, r] \times [-r, r] \times \cdots \times [-r, r]$  se puede descomponer en un número finito de cubos de lados iguales, tan pequeños como se quiera. Se sigue de esto que un subconjunto de  $\mathbb{R}^n$  es totalmente acotado si y sólo si es acotado. El teorema B.2, que se puede contemplar como una generalización del teorema 2.9, pone de manifiesto que los conjuntos totalmente acotados desempeñan en los espacios métricos generales un papel similar al que desempeñan los conjuntos acotados en  $\mathbb{R}^n$ .

**Corolario B.4** *Un subconjunto  $M$  de un espacio métrico completo es totalmente acotado si y sólo si  $\overline{M}$  es compacto. Un subconjunto  $K$  de un espacio métrico completo es compacto si y sólo si es cerrado y totalmente acotado.*

DEM: Si  $M$  es totalmente acotado  $\overline{M}$  también lo es y basta aplicar el teorema B.2 para obtener que  $\overline{M}$  es compacto. El recíproco es inmediato. La segunda afirmación es consecuencia directa de la primera. ■

### B.3. Espacios de sucesiones

*El espacio  $l^2$ .* Es el formado por las sucesiones de números reales  $\mathbf{x} = (x(k))$  que cumplen  $\sum_{k=1}^{\infty} x(k)^2 < +\infty$ . En  $l^2$  están definidas las operaciones naturales

$$\mathbf{x} + \mathbf{y} = (x(1) + y(1), x(2) + y(2), \dots); \quad \mu\mathbf{x} = (\mu x(1), \mu x(2), \dots).$$

y con ellas  $l^2$  es un espacio vectorial real infinito dimensional.

Efectivamente, si  $\mu \in \mathbb{R}$  y  $\mathbf{x} \in l^2$  es inmediato que  $\mu\mathbf{x} \in l^2$ . Para ver que la suma de dos sucesiones  $\mathbf{x}, \mathbf{y} \in l^2$  es una sucesión de  $l^2$  observemos en primer lugar que para cada  $m \in \mathbb{N}$ , según la desigualdad de Cauchy-Schwarz en  $\mathbb{R}^m$ , se cumple

$$\sum_{k=1}^m |x(k)y(k)| \leq \sqrt{\sum_{k=1}^m x(k)^2} \sqrt{\sum_{k=1}^m y(k)^2} \leq \|\mathbf{x}\|_2 \|\mathbf{y}\|_2.$$

Como esta desigualdad es cierta para todo  $m \in \mathbb{N}$  se obtiene la desigualdad:

$$[C] \quad \sum_{k=1}^{\infty} |x(k)y(k)| \leq \|\mathbf{x}\|_2 \|\mathbf{y}\|_2$$

luego

$$\sum_{k=1}^{\infty} (x(k) + y(k))^2 \leq \sum_{k=1}^{\infty} (x(k)^2 + y(k)^2 + 2|x(k)y(k)|) < +\infty$$

y queda demostrado que  $\mathbf{x} + \mathbf{y} \in l^2$ .

Por otra parte, si  $\mathbf{x}, \mathbf{y} \in l^2$ , en virtud de la desigualdad [C] es convergente la serie que interviene en la fórmula

$$\langle \mathbf{x} | \mathbf{y} \rangle = \sum_{k=1}^{\infty} x(k)y(k)$$

y es fácil comprobar que así queda definido un producto escalar en  $l^2$  cuya norma asociada es  $\|\mathbf{x}\|_2 = \sqrt{\sum_{k=1}^{\infty} x(k)^2}$ .

Seguidamente demostramos que el espacio normado  $(l^2, \|\cdot\|_2)$  es completo:

Sea  $\mathbf{x}_n = (x_n(1), x_n(2), \dots, x_n(k), \dots)$  una sucesión de Cauchy en  $l^2$ . Para cada  $k \in \mathbb{N}$  se cumple  $|x_p(k) - x_q(k)| \leq \|\mathbf{x}_p - \mathbf{x}_q\|_2$  luego todas las sucesiones de números reales  $(x_n(k))_{n=1}^{\infty}$  son de Cauchy y podemos asegurar que existen los límites  $\lim_n x_n(k) = x(k)$ . Entonces  $\mathbf{x} = (x(1), x(2), \dots, x(k), \dots)$  pertenece a  $l^2$  y  $\lim_n \|\mathbf{x}_n - \mathbf{x}\|_2 = 0$ . Efectivamente, dado  $\epsilon > 0$  existe  $n(\epsilon) \in \mathbb{N}$  tal que

$$p > q \geq n(\epsilon) \Rightarrow \|\mathbf{x}_p - \mathbf{x}_q\|_2^2 < \epsilon^2$$

Si  $p > q \geq n(\epsilon)$ , para cada  $m \in \mathbb{N}$  se cumple  $\sum_{k=1}^m (x_p(k) - x_q(k))^2 < \epsilon^2$ . En esta suma finita podemos pasar al límite, cuando  $p \rightarrow +\infty$ , y obtenemos

$$q \geq n(\epsilon) \Rightarrow \sum_{k=1}^m (x(k) - x_q(k))^2 < \epsilon^2$$

Si  $q \geq n(\epsilon)$ , la desigualdad de la derecha se cumple para todo  $m \in \mathbb{N}$ , luego  $\|\mathbf{x} - \mathbf{x}_q\|_2^2 \leq \epsilon^2$ . Se sigue que  $\mathbf{x} = \mathbf{x}_q + (\mathbf{x} - \mathbf{x}_q) \in l^2$  y  $\lim_q \|\mathbf{x} - \mathbf{x}_q\|_2 = 0$ . ■

*El espacio  $l^1$ .* Es el espacio vectorial real formado por las sucesiones de números reales  $\mathbf{x} = (x(n))$  que cumplen  $\sum_{k=1}^{\infty} |x(k)| < +\infty$ .

Es inmediato que  $\|\mathbf{x}\|_1 = \sum_{k=1}^{\infty} |x(k)|$  es una norma sobre  $l^1$ , y con un razonamiento similar al efectuado con  $l^2$  se demuestra que el espacio normado  $(l^1, \|\cdot\|_1)$  es completo.

*El espacio  $l^\infty$ .* Es el espacio vectorial real formado por las sucesiones acotadas de números reales. Es inmediato que  $\|\mathbf{x}\|_\infty = \sup\{|x(k)| : k \in \mathbb{N}\}$  es una norma sobre  $l^\infty$  y se demuestra fácilmente que el espacio normado  $(l^\infty, \|\cdot\|_\infty)$  es completo.

**Ejercicio B.5** Sean  $p > 1$ , números reales que verifican  $1/p + 1/q = 1$ . Establezca las siguientes desigualdades:

- a) (Young) Si  $a \geq 0, b \geq 0$  entonces  $ab \leq a^p/p + b^q/q$ .
- b) (Hölder) Si  $(a_n), (b_n)$  son sucesiones de números reales no negativos, tales que  $\sum_n a_n^p < +\infty, \sum_n b_n^q < +\infty$  entonces

$$\sum_n a_n b_n \leq \left( \sum_n a_n^p \right)^{1/p} \left( \sum_n b_n^q \right)^{1/q}$$

Estudie las condiciones para que la desigualdad anterior sea una igualdad.

- c) (Minkowski) Si  $(a_n), (b_n)$  son sucesiones en  $\mathbb{R}$  tales que  $\sum_n |a_n|^p < +\infty, \sum_n |b_n|^p < +\infty$  entonces

$$\left( \sum_n |a_n + b_n|^p \right)^{1/p} \leq \left( \sum_n |a_n|^p \right)^{1/p} + \left( \sum_n |b_n|^p \right)^{1/p}$$

Estudie las condiciones para que la desigualdad anterior sea una igualdad.

Para  $1 \leq p < +\infty$  sea  $l^p$  el conjunto de las sucesiones de números reales  $\mathbf{x} = (x_n)$  tales que  $\sum_n |x_n|^p < +\infty$ . Demuestre que, con las operaciones naturales,  $l^p$  es un espacio vectorial y que

$$\|\mathbf{x}\|_p = \left( \sum_n |x_n|^p \right)^{1/p}$$

es una norma  $l^p$ . Demuestre que el espacio normado  $(l^p, \|\cdot\|_p)$  es completo

SOLUCIÓN

Véase [6], ejercicio resuelto 2.14, en pág. 87. ■

**Ejercicio B.6** Dada una sucesión de números positivos  $r_n > 0$ , en el espacio  $(l^2, \|\cdot\|_2)$  se considera el conjunto  $Q_r$  formado por las sucesiones  $\mathbf{x} = (x_n)_{n \in \mathbb{N}}$  tales que  $\{n : x_n \neq 0\}$  es finito y  $|x_n| \leq r_n$  para todo  $n \in \mathbb{N}$ . Demuestre que  $Q_r$  es relativamente compacto si y sólo si  $\sum_n r_n^2 < +\infty$

SOLUCIÓN

Véase [6], ejercicio 2.31 propuesto en pág. 118 y resuelto en pág. 134. ■

## B.4. Formas lineales y producto escalar

El resultado que se expone en la siguiente proposición B.8 será la clave para entender la noción de gradiente al caso de funciones diferenciables  $f : \Omega \rightarrow \mathbb{R}$  definidas en un abierto  $\Omega$  de un espacio normado completo  $(E, \|\cdot\|)$  cuya norma procede de un producto escalar, aunque el espacio no sea de dimensión finita.

Observemos en primer lugar que para cada  $\mathbf{z} \in E$  la aplicación lineal  $L_{\mathbf{z}} : E \rightarrow \mathbb{R}$  definida por  $L_{\mathbf{z}}(\mathbf{h}) = \langle \mathbf{z} | \mathbf{h} \rangle$  es continua ya que, en virtud de la desigualdad de Cauchy-Schwarz, para todo  $\mathbf{h} \in E$  se cumple  $|L_{\mathbf{z}}(\mathbf{h})| \leq \|\mathbf{z}\| \|\mathbf{h}\|$ .



**Proposición B.7** Sea  $(E, \| \cdot \|)$  un espacio normado completo cuya norma procede de un producto escalar y  $A \subset E$  un subconjunto es cerrado convexo. Entonces existe un único  $\mathbf{a} \in A$  que verifica  $\|\mathbf{a}\| = \min\{\|\mathbf{x}\| : \mathbf{x} \in A\}$ .

DEM: La demostración, que se basa en la identidad del paralelogramo (problema 2.6.1):

$$\|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2 = 2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2 \quad \text{para todo } \mathbf{x}, \mathbf{y} \in E$$

se puede ver en el Capítulo 4 del texto *Análisis real y complejo* de W. Rudin ■

**Teorema B.8** Sea  $(E, \| \cdot \|)$  un espacio normado completo, cuya norma  $\| \cdot \|$  es la asociada a un producto escalar  $\langle \cdot | \cdot \rangle$ . Entonces, para cada aplicación lineal continua  $L : E \rightarrow \mathbb{R}$  existe un (único) vector  $\mathbf{z} \in E$  que verifica  $L(\mathbf{h}) = \langle \mathbf{z} | \mathbf{h} \rangle$  para todo  $\mathbf{h} \in E$ .

DEM: Como  $L$  es lineal y continua  $M = \{\mathbf{x} \in E : L(\mathbf{x}) = 0\}$  es un subespacio cerrado de  $E$ . Si  $M = E$ , el vector  $\mathbf{z} = 0$  cumple la condición deseada. En otro caso, si  $M \neq E$ , existe  $\mathbf{y} \in E$  tal que  $L(\mathbf{y}) \neq 0$ . El subespacio afín  $\mathbf{y} + M$  es un subconjunto cerrado convexo de  $E$  y según existe un elemento  $\mathbf{w} \in \mathbf{y} + M$  de norma mínima (obsérvese que la condición  $\mathbf{y} \notin M$  implica que  $\mathbf{0} \notin \mathbf{y} + M$ , luego  $\mathbf{w} \neq 0$ ). Entonces para todo  $t \in \mathbb{R}$  y todo  $\mathbf{u} \in M$  se cumple  $\mathbf{w} + t\mathbf{u} \in \mathbf{y} + M$  luego  $0 < \|\mathbf{w}\|^2 \leq \|\mathbf{w} + t\mathbf{u}\|^2$  es decir

$$\langle \mathbf{w} | \mathbf{w} \rangle \leq \langle \mathbf{w} + t\mathbf{u} | \mathbf{w} + t\mathbf{u} \rangle \quad \text{para todo } t \in \mathbb{R}$$

Usando la bilinealidad del producto escalar resulta

$$0 \leq 2t\langle \mathbf{w} | \mathbf{u} \rangle + t^2\langle \mathbf{u} | \mathbf{u} \rangle \quad \text{para todo } t \in \mathbb{R}$$

y esta última condición implica que  $\langle \mathbf{w} | \mathbf{u} \rangle = 0$ .

Hemos probado así que existe un vector  $\mathbf{0} \neq \mathbf{w} \in E$  ortogonal a  $M$ , que podemos suponer normalizado con  $\|\mathbf{w}\| = 1$ . Para cada  $\mathbf{x} \in E$ , el vector  $\mathbf{v} = L(\mathbf{x})\mathbf{w} - L(\mathbf{w})\mathbf{x}$  pertenece a  $M$ , luego

$$0 = \langle \mathbf{w} | \mathbf{v} \rangle = L(\mathbf{x})\|\mathbf{w}\|^2 - L(\mathbf{w})\langle \mathbf{w} | \mathbf{x} \rangle$$

es decir  $L(\mathbf{x}) = L(\mathbf{w})\langle \mathbf{w} | \mathbf{x} \rangle$  luego  $\mathbf{z} = L(\mathbf{w})\mathbf{w}$  cumple la condición requerida. ■

## B.5. Espacios complejos con producto interior

Un producto interior sobre un espacio vectorial complejo  $E$  es una aplicación

$$\langle \cdot | \cdot \rangle : E \times E \rightarrow \mathbb{C}, \quad (\mathbf{x}, \mathbf{y}) \rightarrow \langle \mathbf{x} | \mathbf{y} \rangle$$

que verifica

$$\text{i) } \langle \mathbf{x} | \mathbf{y} \rangle = \overline{\langle \mathbf{y} | \mathbf{x} \rangle} \quad \text{para cada } (\mathbf{x}, \mathbf{y}) \in E \times E.$$

ii) Para cada  $\mathbf{y} \in E$  la aplicación  $\mathbf{x} \rightarrow \langle \mathbf{x} | \mathbf{y} \rangle$  es lineal.

iii)  $\langle \mathbf{x} | \mathbf{x} \rangle \geq 0$  para cada  $\mathbf{x} \in E$  y  $\langle \mathbf{x} | \mathbf{x} \rangle = 0$  si y sólo si  $\mathbf{x} = 0$ .

Obsérvese que, en virtud de i) y ii), fijado  $\mathbf{x} \in E$ , la aplicación  $\mathbf{y} \rightarrow \langle \mathbf{x} | \mathbf{y} \rangle$  es antilineal, es decir para cada  $\mathbf{x}, \mathbf{y} \in E$  y cada  $\mu \in \mathbb{C}$  se verifica

$$\langle \mathbf{x} | \mathbf{y} + \mathbf{z} \rangle = \langle \mathbf{x} | \mathbf{y} \rangle + \langle \mathbf{x} | \mathbf{z} \rangle, \quad \langle \mathbf{x} | \mu \mathbf{y} \rangle = \bar{\mu} \langle \mathbf{x} | \mathbf{y} \rangle$$

De manera similar a como se ha visto en 2.2 al producto interior se le puede asociar una norma de espacio vectorial complejo  $\|\mathbf{x}\|_2 = \sqrt{\langle \mathbf{x} | \mathbf{x} \rangle}$ , que cumple la desigualdad de Cauchy-Schwarz:

$$|\langle \mathbf{x} | \mathbf{y} \rangle| \leq \|\mathbf{x}\| \|\mathbf{y}\| \quad \text{para cada } \mathbf{x}, \mathbf{y} \in E$$

Efectivamente, dados  $\mathbf{x}, \mathbf{y} \in E$ , y  $\mu \in \mathbb{C}$ , para todo  $t \in \mathbb{R}$  se cumple

$$0 \leq h(t) = \langle \mathbf{x} + t\mu\mathbf{y} | \mathbf{x} + t\mu\mathbf{y} \rangle$$

y usando las propiedades del producto interior se obtiene:

$$\begin{aligned} 0 \leq h(t) &= \langle \mathbf{x} | \mathbf{x} \rangle + t\mu \langle \mathbf{y} | \mathbf{x} \rangle + t\bar{\mu} \langle \mathbf{x} | \mathbf{y} \rangle + t^2 \mu \bar{\mu} \langle \mathbf{y} | \mathbf{y} \rangle = \\ &= \|\mathbf{x}\|^2 + 2t \operatorname{Real}(\bar{\mu} \langle \mathbf{x} | \mathbf{y} \rangle) + t^2 |\mu|^2 \|\mathbf{y}\|^2 \end{aligned}$$

Eligiendo  $\mu \in \mathbb{C}$  de modo que  $\mu \langle \mathbf{x} | \mathbf{y} \rangle = |\langle \mathbf{x} | \mathbf{y} \rangle|$  se consigue que para todo  $t \in \mathbb{R}$  se cumpla la desigualdad

$$0 \leq h(t) = \|\mathbf{x}\|^2 + 2|\langle \mathbf{x} | \mathbf{y} \rangle| t + t^2 \|\mathbf{y}\|^2$$

y razonando como en la proposición 2.2 se termina la demostración. ■

En el espacio vectorial complejo  $\mathbb{C}^n$  se define el producto interior de los vectores  $\mathbf{z} = (z_1, z_2, \dots, z_n)$ ,  $\mathbf{w} = (w_1, w_2, \dots, w_n)$  mediante la fórmula

$$\langle \mathbf{z} | \mathbf{z} \rangle = \sum_{i=1}^n z_i \bar{y}_i$$

que lleva asociada la norma  $\|\mathbf{z}\|_2 = \sqrt{\langle \mathbf{z} | \mathbf{z} \rangle} = \sqrt{\sum_{i=1}^n |z_i|^2}$ .

En  $\mathbb{C}^n$  también se pueden definir las normas

$$\|\mathbf{z}\|_1 = \sum_{i=1}^n |z_i|, \quad \|\mathbf{z}\|_\infty = \max\{|z_j| : 1 \leq j \leq n\}$$

que son equivalentes a la anterior en virtud de las desigualdades

$$\|\mathbf{z}\|_1 / n \leq \|\mathbf{z}\|_2 \leq \|\mathbf{z}\|_1; \quad \|\mathbf{z}\|_\infty \leq \|\mathbf{z}\|_2 \leq \sqrt{n} \|\mathbf{z}\|_\infty$$

(En el teorema 3.18 se establecerá que en  $\mathbb{C}^n$  todas las normas sobre son equivalentes).

Finalizamos mostrando ejemplos de espacios normados complejos de funciones continuas similares a los considerados en el caso real. Sea  $C([0, 1], \mathbb{C})$  el espacio vectorial complejo formado por las funciones continuas  $f : [a, b] \rightarrow \mathbb{C}$ .

Si  $f(t) = f_1(t) + if_2(t)$  se define

$$\int_a^b f(t)dt = \int_a^b f_1(t)dt + i \int_a^b f_2(t)dt$$

y es fácil comprobar que la integral  $f \rightarrow \int_a^b f(t)dt$  es una forma lineal sobre el espacio vectorial complejo  $C([a, b], \mathbb{C})$ . La fórmula  $\langle f | g \rangle = \int_a^b f(t)\overline{g(t)}dt$  define un producto interior sobre este espacio, que lleva asociada la norma  $\|f\|_2 = \sqrt{\int_a^b |f(t)|^2 dt}$ .

En forma similar a como se ha hecho en el caso real, en  $C([0, 1], \mathbb{C})$  se pueden definir las normas  $\|\cdot\|_1$ ,  $\|\cdot\|_\infty$  y también ocurre que las topologías asociadas a las normas  $\|\cdot\|_p$ ,  $p \in \{1, 2, \infty\}$  son distintas.

# C

## Complementos al capítulo 3

### C.1. Intercambio de límites

La demostración de algunos teoremas importantes del Análisis Matemático se reduce en última instancia a la posibilidad de cambiar el orden en dos procesos sucesivos de paso al límite. El teorema 3.31 se puede interpretar como un resultado sobre intercambio de límites: Si una sucesión de funciones continuas  $\mathbf{f}_n : T \rightarrow F$  converge uniformemente hacia  $\mathbf{f} : T \rightarrow F$ , la continuidad del límite  $\mathbf{f}$  significa que para cada  $a \in T$  se cumple  $\mathbf{f}(a) = \lim_{t \rightarrow a} \mathbf{f}(t)$ , es decir,  $\lim_n \mathbf{f}_n(a) = \lim_{t \rightarrow a} (\lim_n \mathbf{f}_n(t))$ . En virtud de la continuidad de  $\mathbf{f}_n$ , podemos sustituir  $\mathbf{f}_n(a) = \lim_{t \rightarrow a} \mathbf{f}_n(t)$ , y resulta

$$\lim_n \left( \lim_{t \rightarrow a} \mathbf{f}_n(t) \right) = \lim_{t \rightarrow a} \left( \lim_n \mathbf{f}_n(t) \right)$$

En el ejercicio 3.38 se obtuvo un resultado análogo. La proposición 3.2 también proporciona un resultado sobre intercambio de límites para una función de dos variables reales  $f : M \rightarrow \mathbb{R}$  definida en  $M \supset \{(s, t) \in \mathbb{R}^2 : 0 < |s - a| < r, 0 < |t - b| < r\}$ . Si existe el límite doble  $\lim_{(s,t) \rightarrow (a,b)} f(s, t)$  y para  $0 < |s - a| < r$ ,  $0 < |t - b| < r$ , existen los límites parciales  $\lim_{t \rightarrow b} f(s, t) = \alpha(s)$ ,  $\lim_{s \rightarrow a} f(s, t) = \beta(t)$ , entonces existen los dos límites iterados y valen lo mismo que el límite doble:

$$\lim_{s \rightarrow a} \left( \lim_{t \rightarrow b} f(s, t) \right) = \lim_{(s,t) \rightarrow (a,b)} f(s, t) = \lim_{t \rightarrow b} \left( \lim_{s \rightarrow a} f(s, t) \right)$$

#### Noción general de límite

Con el fin de dar un tratamiento unificado a la cuestión de la permutabilidad de límites, conviene empezar dando la noción general de límite según una base de filtro que incluye, como casos particulares, a las distintas nociones de límite que intervienen en el Análisis Matemático: Límite de una sucesión, límite de una función en un punto, límite de sumas de Riemann, límite de sumas finitas, límite de una red, etc.

Una base de filtro  $\mathcal{B}$  en un conjunto  $T \neq \emptyset$  es una familia no vacía de partes no vacías de  $T$ , que verifica

[\* ] Para cada par  $B_1, B_2 \in \mathcal{B}$  existe  $B \in \mathcal{B}$  tal que  $B \subset B_1 \cap B_2$ .

En particular  $B_1 \cap B_2 \neq \emptyset$  para cada par  $B_1, B_2 \in \mathcal{B}$ .

**Definición C.1** Sea  $(F, \rho)$  un espacio métrico y  $\mathcal{B}$  una base de filtro en  $T$ .

Se dice que  $\mathbf{f} : T \rightarrow F$  tiene límite  $\mathbf{b}$  según la base de filtro  $\mathcal{B}$  cuando para cada  $\epsilon > 0$  existe  $B_\epsilon \in \mathcal{B}$  tal que para todo  $t \in B_\epsilon$  se cumple  $\rho(\mathbf{f}(t), \mathbf{b}) \leq \epsilon$ . En este caso se escribe  $\mathcal{B} - \lim_t \mathbf{f}(t) = \mathbf{b}$ .

Cuando se sobreentiende por el contexto la base de filtro  $\mathcal{B}$  que se considera en  $T$ , escribiremos, más brevemente,  $\lim_t \mathbf{f}(t)$  en lugar de  $\mathcal{B} - \lim_t \mathbf{f}(t)$ .

### Ejemplos C.2

a) Si en  $T = \mathbb{N}$ , se considera la base de filtro de Fréchet

$$\mathcal{F} = \{F_m : m \in \mathbb{N}\}, \quad \text{con } F_m = \{n \in \mathbb{N} : n \geq m\}$$

entonces la convergencia de una sucesión  $(\mathbf{x}_n)$  en el espacio métrico  $(F, \rho)$  es precisamente la convergencia, según la base de filtro  $\mathcal{F}$ , de la aplicación  $\mathbf{f}(n) = \mathbf{x}_n$ .

b) Si  $T = M$ , donde  $M$  es un subconjunto de un espacio métrico  $(E, d)$ , y  $a \in M'$ , la noción de límite funcional  $\lim_t \rightarrow_a \mathbf{f}(t)$  es la que resulta cuando en  $M$  se considera la base de filtro

$$\mathcal{B}_a = \{B_a(r) : r > 0\}, \quad \text{con } B_a(r) = \{t \in M : 0 < d(t, a) < r\}$$

c) Si  $T = \mathbb{R}^n$ , la noción de límite  $\lim_t \rightarrow_\infty \mathbf{f}(t)$  es la que resulta considerando la base de filtro  $\mathcal{B}_\infty = \{B_\infty(r) : r > 0\}$  donde  $B_\infty(r) = \{x \in \mathbb{R}^n : \|x\| > r\}$ .

d) Una familia infinita  $\{\mathbf{x}_j : j \in J\}$  de vectores en un espacio vectorial normado  $(E, \|\cdot\|)$  se dice que es *sumable*, con suma  $\mathbf{s}$  cuando se verifica lo siguiente:

Para cada  $\epsilon > 0$  existe un conjunto finito  $H_\epsilon \subset J$ , tal que si  $H \subset J$  es finito y  $H_\epsilon \subset H$ , entonces  $\left\| \sum_{j \in H} \mathbf{x}_j - \mathbf{s} \right\| < \epsilon$ .

La noción de familia sumable se puede formular, en términos de las sumas finitas  $\mathbf{s}_H = \sum_{j \in H} \mathbf{x}_j$ , como otro caso particular de la noción de límite según una base de filtro: Si  $H \subset J$  es finito, y  $[H]$  es la familia de los subconjuntos finitos de  $J$  que contienen a  $H$ , entonces  $\mathcal{B} = \{[H] : H \subset J, H \text{ finito}\}$  es una base de filtro en el conjunto de las partes finitas de  $J$ . Decir que  $\{\mathbf{x}_j : j \in J\}$  es sumable, con suma  $\mathbf{s}$  es lo mismo que decir que la aplicación  $H \rightarrow \mathbf{s}_H$  tiene límite  $\mathbf{s}$  según esta base de filtro. ■

El teorema 3.3 es una versión particular del siguiente resultado general

**Proposición C.3** [Condición de Cauchy] Sea  $(F, \rho)$  un espacio métrico completo y  $T$  un conjunto no vacío dotado de una base de filtro  $\mathcal{B}$ . La siguiente condición es necesaria y suficiente para que  $\mathbf{f} : T \rightarrow F$  tenga límite según  $\mathcal{B}$ :

Para cada  $\epsilon > 0$  existe  $B_\epsilon \in \mathcal{B}$  tal que  $x, y \in B_\epsilon \Rightarrow \rho(\mathbf{f}(x), \mathbf{f}(y)) < \epsilon$ .

DEM: La necesidad de la condición es una consecuencia inmediata de la definición de límite. Para demostrar que la condición es suficiente empezamos viendo que si  $t_n \in B_{1/n}$  entonces la sucesión  $(\mathbf{f}(t_n))$  es de Cauchy.

Dado  $\epsilon > 0$ , sea  $n \in \mathbb{N}$  tal que  $2/n < \epsilon$ . Si  $p > q \geq n$ , en virtud de [\*] podemos elegir  $z \in B_{1/p} \cap B_{1/q}$ . Como  $z, t_p \in B_{1/p}$ , y  $z, t_q \in B_{1/q}$ , se cumple  $\rho(\mathbf{f}(t_p), \mathbf{f}(z)) < 1/p$ ,  $\rho(\mathbf{f}(t_q), \mathbf{f}(z)) < 1/q$ , luego

$$\rho(\mathbf{f}(t_p), \mathbf{f}(t_q)) \leq \rho(\mathbf{f}(t_p), \mathbf{f}(z)) + \rho(\mathbf{f}(z), \mathbf{f}(t_q)) \leq 1/p + 1/q \leq 2/n < \epsilon$$

Como el espacio métrico  $(F, \rho)$  es completo, la sucesión de Cauchy  $(\mathbf{f}(t_n))$  es convergente y para terminar la demostración basta ver que su límite  $\mathbf{b} \in F$  es el límite de  $\mathbf{f}$  según  $\mathcal{B}$ : Dado  $\epsilon > 0$ , existe  $m \in \mathbb{N}$  verificando  $2/m < \epsilon$ , y  $\rho(\mathbf{f}(t_m), \mathbf{b}) < \epsilon/2$ . Para todo  $t \in B_{1/m}$ , se tiene  $t, t_m \in B_{1/m}$ , luego  $\rho(\mathbf{f}(t), \mathbf{f}(t_m)) < 1/m < \epsilon/2$  y se sigue que

$$\rho(\mathbf{f}(t), \mathbf{b}) \leq \rho(\mathbf{f}(t), \mathbf{f}(t_m)) + \rho(\mathbf{f}(t_m), \mathbf{b}) \leq \epsilon/2 + \epsilon/2 = \epsilon$$

■

### Relación entre los límites iterados y el límite doble.

Para dar un tratamiento general al problema de la permutabilidad de límites suponemos el lo que sigue que  $S, T$  son conjuntos no vacíos dotados, respectivamente con las bases de filtro  $\mathcal{A}, \mathcal{B}$ . Es fácil ver que

$$\mathcal{U} = \mathcal{A} \times \mathcal{B} = \{A \times B : A \in \mathcal{A}, B \in \mathcal{B}\}$$

es una base de filtro en  $S \times T$ . Dada una aplicación  $\mathbf{f} : S \times T \rightarrow F$ , se pueden considerar, si existen, los límites iterados

$$\mathcal{A} - \lim_s [\mathcal{B} - \lim_t \mathbf{f}(s, t)], \quad \mathcal{B} - \lim_t [\mathcal{A} - \lim_s \mathbf{f}(s, t)]$$

y el límite doble  $\mathcal{U} - \lim_{(s,t)} \mathbf{f}(s, t)$ , que en lo sucesivo, para simplificar la notación, escribiremos de modo más breve

$$\lim_s [\lim_t \mathbf{f}(s, t)], \quad \lim_t [\lim_s \mathbf{f}(s, t)], \quad \lim_{(s,t)} \mathbf{f}(s, t)$$

El siguiente resultado es una versión abstracta del obtenido en la proposición 3.2

**Proposición C.4** *Sea supone que  $\mathbf{f} : S \times T \rightarrow F$  verifica*

- a) *Para cada  $s \in S$  existe el límite parcial  $\lim_t \mathbf{f}(s, t) = \alpha(s)$ ;*
- b) *Existe el límite doble  $\lim_{(s,t)} \mathbf{f}(s, t) = \mathbf{b}$ ;*

*Entonces existe el límite iterado  $\lim_s \alpha(s) = \mathbf{b}$ , es decir,  $\lim_s [\lim_t \mathbf{f}(s, t)] = \lim_{(s,t)} \mathbf{f}(s, t)$ . En particular, si existe el límite doble y los dos límites iterados,  $\lim_s [\lim_t \mathbf{f}(s, t)]$ ,  $\lim_t [\lim_s \mathbf{f}(s, t)]$ , los tres límites son iguales.*

DEM: Según b), para cada  $\epsilon > 0$  existe  $U_\epsilon = A_\epsilon \times B_\epsilon \in \mathcal{A} \times \mathcal{B}$  tal que

$$(s, t) \in A_\epsilon \times B_\epsilon \Rightarrow \rho(\mathbf{f}(s, t), \mathbf{b}) < \epsilon/2$$

Según a), para cada  $s \in A_\epsilon$  existe  $B_s \in \mathcal{B}$  tal que

$$t \in B_s \Rightarrow \rho(\mathbf{f}(s, t), \alpha(s)) < \epsilon/2$$

Como  $B_s \cap B_\epsilon \neq \emptyset$ , para cada  $s \in A_\epsilon$  podemos elegir  $t_s \in B_s \cap B_\epsilon$ , y se obtiene

$$\rho(\alpha(s), \mathbf{b}) \leq \rho(\alpha(s), \mathbf{f}(s, t_s)) + \rho(\mathbf{f}(s, t_s), \mathbf{b}) \leq \epsilon/2 + \epsilon/2 = \epsilon$$

■

Las tres proposiciones siguientes proporcionan condiciones suficientes, en términos de límites iterados, para la existencia del límite doble.

**Proposición C.5** *Se supone que  $\mathbf{f} : S \times T \rightarrow F$  verifica*

- a) *Para cada  $s \in S$  existe el límite parcial  $\lim_t \mathbf{f}(s, t) = \alpha(s)$  y el límite es uniforme respecto de  $s \in S$  (es decir, para todo  $\epsilon > 0$  existe  $B_\epsilon \in \mathcal{B}$  tal que si  $t \in B_\epsilon$  entonces todo  $s \in S$  cumple  $\rho(\mathbf{f}(s, t), \alpha(s)) < \epsilon$ ).*
- b) *Existe el límite iterado  $\lim_s [\lim_t \mathbf{f}(s, t)] = \lim_s \alpha(s) = \mathbf{b}$ .*

*Entonces existe el límite doble y vale lo mismo que el límite iterado:*

$$\lim_{(s,t)} \mathbf{f}(s, t) = \lim_s [\lim_t \mathbf{f}(s, t)]$$

DEM: Dado  $\epsilon > 0$ , sea  $B_\epsilon \in \mathcal{B}$  el proporcionado por la condición a). Según b) existe  $A_\epsilon \in \mathcal{A}$  tal que  $s \in A_\epsilon \Rightarrow \rho(\alpha(s), \mathbf{b}) < \epsilon$ . Entonces  $U_\epsilon = A_\epsilon \times B_\epsilon \in \mathcal{U}$  cumple

$$(s, t) \in U_\epsilon \Rightarrow \rho(\mathbf{f}(s, t), \mathbf{b}) \leq \rho(\mathbf{f}(s, t), \alpha(s)) + \rho(\alpha(s), \mathbf{b}) \leq \epsilon + \epsilon = 2\epsilon$$

■

NOTA: Sea  $\mathbf{f}_n : M \rightarrow F$  una sucesión de funciones continuas que converge uniformemente hacia  $\mathbf{f} : M \rightarrow F$ . Dado  $a \in M$ , podemos considerar  $M$  dotado del filtro  $\mathcal{B}_a$  (véase C.2 b)), y  $T = \mathbb{N}$  dotado del filtro de Fréchet. Entonces la aplicación  $(s, n) \rightarrow \mathbf{f}_n(s)$  cumple las hipótesis de la proposición C.5 (obsérvese que existe el límite iterado  $\lim_n \lim_s \mathbf{f}_n(s) = \lim_n \mathbf{f}_n(a) = \mathbf{f}(a)$ ) y se concluye que existe el otro límite iterado y vale lo mismo, es decir  $\lim_s \lim_n \mathbf{f}_n(s) = \lim_s \mathbf{f}(s) = \mathbf{f}(a)$ , lo que significa que  $\mathbf{f}$  es continua en  $a$ . Vemos así que el teorema 3.31 se puede considerar como un caso particular de la proposición C.5.

El resultado obtenido en el ejercicio 3.39 también es otro caso particular de esta proposición. Si  $\mathbf{f}_n : M \rightarrow F$  es una sucesión de funciones equicontinuas en  $a \in M$ , que converge puntualmente hacia  $\mathbf{f} : M \rightarrow F$ , podemos considerar  $T = M$ , dotado del filtro  $\mathcal{B}_a$  y  $S = \mathbb{N}$  con el filtro de Fréchet. Ahora la aplicación  $(n, t) \rightarrow \mathbf{f}_n(t)$  cumple las hipótesis de la proposición C.5 (La equicontinuidad en  $a$  significa que el límite  $\lim_t \mathbf{f}_n(t) = \mathbf{f}_n(a)$  es uniforme respecto de la variable  $n \in \mathbb{N}$ , y la convergencia puntual de la sucesión hace que se cumpla b).

Entonces, en virtud de la proposición C.5 existen y son iguales los dos límites iterados  $\lim_n [\lim_t \mathbf{f}_n(t)] = \lim_t [\lim_n \mathbf{f}_n(t)]$  es decir,  $\mathbf{f}(a) = \lim_t \mathbf{f}(t)$ , y por lo tanto  $\mathbf{f}$  es continua en  $a$ .

**Proposición C.6** *Se supone que  $\mathbf{f} : S \times T \rightarrow F$  verifica*

- a) *Para cada  $s \in S$  existe el límite parcial  $\lim_t \mathbf{f}(s, t) = \alpha(s)$ , y el límite es uniforme respecto de  $s \in S$ .*
- b) *Existe el límite iterado  $\lim_t [\lim_s \mathbf{f}(s, t)] = \mathbf{b}$ .*

*Entonces existe el límite doble y el otro límites iterado y los tres límites son iguales:*

$$\lim_{(s,t)} \mathbf{f}(s, t) = \lim_s (\lim_t \mathbf{f}(s, t)) = \lim_t (\lim_s \mathbf{f}(s, t))$$

DEM: Basta demostrar que existe el límite doble  $\lim_{(s,y)} \mathbf{f}(s, t) = \mathbf{b}$  y aplicar luego la proposición C.4 para obtener la existencia del otro límite iterado y la igualdad de los tres límites.

Según b) la función  $\beta(t) = \lim_s \mathbf{f}(s, t)$  está definida para todo  $t \in T$ , y tiene límite  $\mathbf{b}$  según la base de filtro  $\mathcal{B}$ . Por la definición de límite, para cada  $\epsilon > 0$  existe  $B'_\epsilon \in \mathcal{B}$  tal que  $t \in B'_\epsilon \Rightarrow \rho(\beta(t), \mathbf{b}) < \epsilon/3$ .

Por otra parte, de a) se deduce que existe  $B_\epsilon \in \mathcal{B}$  tal que si  $t, t' \in B_\epsilon$  entonces todo  $s \in S$  cumple  $\rho(\mathbf{f}(s, t), \mathbf{f}(s, t')) < \epsilon/3$ . En virtud de la propiedad [\*] de las bases de filtro, existe  $B''_\epsilon \in \mathcal{B}$ ,  $B''_\epsilon \subset B_\epsilon \cap B'_\epsilon$ .

Fijando un punto  $t' \in B''_\epsilon$ , y utilizando que  $\beta(t') = \lim_s \mathbf{f}(s, t')$ , obtenemos  $A_\epsilon \in \mathcal{A}$  tal que  $s \in A_\epsilon \Rightarrow \rho(\mathbf{f}(s, t'), \beta(t')) < \epsilon/3$ . Entonces todo  $(s, t) \in A_\epsilon \times B''_\epsilon$  cumple

$$\rho(\mathbf{f}(s, t), \mathbf{b}) \leq \rho(\mathbf{f}(s, t), \mathbf{f}(s, t')) + \rho(\mathbf{f}(s, t'), \beta(t')) + \rho(\beta(t'), \mathbf{b}) \leq \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon$$

■

**Proposición C.7** *Se supone que  $(F, \rho)$  es completo y que  $\mathbf{f} : S \times T \rightarrow F$  verifica*

- a) *Para cada  $s \in S$  existe el límite parcial  $\lim_t \mathbf{f}(s, t) = \alpha(s)$ , y el límite es uniforme respecto de  $s \in S$ .*
- b) *Para todo  $t \in T$  existe  $\lim_s \mathbf{f}(s, t) = \beta(t)$ .*

*Entonces existe el límite doble y los dos límites iterados y valen lo mismo*

$$\lim_{(s,t)} \mathbf{f}(s, t) = \lim_s (\lim_t \mathbf{f}(s, t)) = \lim_t (\lim_s \mathbf{f}(s, t))$$

DEM: Para demostrar que existe el límite doble basta ver que se cumple la condición de Cauchy C.3. Según a), dado  $\epsilon > 0$  existe  $B_\epsilon \in \mathcal{B}$  tal que si  $t, t' \in B_\epsilon$  entonces todo  $s \in S$  cumple  $\rho(\mathbf{f}(s, t), \mathbf{f}(s, t')) < \epsilon/3$ . Fijado  $t'' \in B_\epsilon$ , en virtud de b), existe  $A_\epsilon \in \mathcal{A}$  tal que para cualquier par  $s, s' \in A_\epsilon$  se verifica  $\rho(\mathbf{f}(s, t''), \mathbf{f}(s', t'')) < \epsilon/3$ .

Entonces para todo par de puntos  $(s, t), (s', t') \in A_\epsilon \times B_\epsilon$  se verifica

$$\begin{aligned} \rho(\mathbf{f}(s, t), \mathbf{f}(s', t')) &\leq \rho(\mathbf{f}(s, t), \mathbf{f}(s, t'')) + \rho(\mathbf{f}(s, t''), \mathbf{f}(s', t'')) + \rho(\mathbf{f}(s', t''), \mathbf{f}(s', t')) \leq \\ &\leq \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon \end{aligned}$$



Con esto queda probado que existe el límite doble, y aplicando la proposición C.4 se termina la demostración. ■

NOTA: Sea  $\mathbf{f}_n : M \rightarrow F$  una sucesión de funciones que converge uniformemente hacia  $\mathbf{f} : M \rightarrow F$ , tal que cada  $\mathbf{f}_n$  tiene límite en  $a \in M'$ . Si tomamos  $S = M$ , dotado del filtro  $\mathcal{B}_a$  considerado en C.2 b) y  $T = \mathbb{N}$ , dotado del filtro de Fréchet, es claro que la aplicación  $(s, n) \rightarrow \mathbf{f}_n(s)$  cumple las hipótesis de la proposición anterior, lo que pone de manifiesto que 3.38 es un caso particular de esta proposición.

## C.2. Convergencia uniforme de series de funciones vectoriales

**Teorema C.8** [Weierstrass] *Si el espacio normado  $(F, \|\cdot\|)$  es completo, una condición suficiente para que una serie  $\sum_{n=1}^{\infty} \mathbf{f}_n$  de funciones  $\mathbf{f}_n : T \rightarrow F$ , sea uniformemente convergente es que sea convergente la serie numérica*

$$\sum_{n=1}^{\infty} \|\mathbf{f}_n\|_T < +\infty$$

DEM: Como  $(F, \|\cdot\|)$  es completo, basta ver que la sucesión de sumas parciales  $\mathbf{s}_n(t) = \sum_{k=1}^n \mathbf{f}_k(t)$  cumple la condición de Cauchy para la convergencia uniforme (véase 3.35). Como la serie  $\sum_{n=1}^{\infty} \|\mathbf{f}_n\|_T$  converge, dado  $\epsilon > 0$ , existe  $m \in \mathbb{N}$  tal que  $\sum_{n>m} \|\mathbf{f}_n\|_T < \epsilon$ . Si  $p > q \geq n(\epsilon)$ , para todo  $t \in T$  se cumple

$$\|\mathbf{s}_p(t) - \mathbf{s}_q(t)\| = \left\| \sum_{k=q+1}^p \mathbf{f}_k(t) \right\| \leq \sum_{k=q+1}^p \|\mathbf{f}_k(t)\| \leq \sum_{k>q} \|\mathbf{f}_k\|_T \leq \sum_{k>m} \|\mathbf{f}_k\|_T \leq \epsilon$$

■

En las aplicaciones del criterio de Weierstrass, generalmente no es preciso calcular explícitamente los valores  $\|\mathbf{f}_n\|_T$ . Basta encontrar una serie numérica convergente  $\sum_{n=1}^{\infty} M_n$  tal que, desde un valor de  $n$  en adelante, se cumpla  $\|\mathbf{f}_n(t)\| \leq M_n$  para todo  $t \in T$ .

El siguiente teorema proporciona criterios útiles para establecer convergencia uniforme de series que no son absolutamente convergentes.

**Teorema C.9** [Abel y Dirichlet] *Sea  $(F, \|\cdot\|)$  un espacio normado completo y  $\sum_{n=1}^{\infty} \mathbf{f}_n$  una serie de funciones de la forma  $\mathbf{f}_n(t) = a_n(t)\mathbf{b}_n(t)$ , donde  $a_n : T \rightarrow \mathbb{R}$ ,  $\mathbf{b}_n : T \rightarrow F$ . Cada una de las siguientes condiciones es suficiente para la convergencia uniforme de la serie.*

- La serie  $\sum_{n=1}^{\infty} \mathbf{b}_n$  converge uniformemente, la sucesión de funciones  $a_n$  es uniformemente acotada y para cada  $t \in T$  la sucesión  $a_n(t)$  es monótona.
- La serie  $\sum_{n=1}^{\infty} \mathbf{b}_n$  converge uniformemente y existe  $C > 0$  tal que, para todo  $t \in T$ , se cumple  $|a_1(t)| + \sum_{n=1}^{\infty} |a_n(t) - a_{n+1}(t)| \leq C$ .

c) La sucesión de sumas  $\sum_{n=1}^m \mathbf{b}_n$  está uniformemente acotada, la sucesión de funciones  $a_n$  converge uniformemente hacia 0 y la sucesión  $a_n(t)$  es monótona para cada  $t \in T$ .

d) La sucesión de sumas  $\sum_{n=1}^m \mathbf{b}_n$  está uniformemente acotada, la sucesión de funciones  $a_n$  converge uniformemente hacia 0, y la serie  $\sum_{n=1}^{\infty} |a_n(t) - a_{n+1}(t)|$  converge uniformemente.

DEM: La demostración se basa en la fórmula de sumación parcial de Abel:

$$F_n^m(t) = a_m(t)B_n^m(t) + \sum_{j=n}^{m-1} B_n^j(t)(a_j(t) - a_{j+1}(t)) \quad [*]$$

donde

$$F_n^m(t) = \sum_{j=n}^m \mathbf{f}_j(t), \quad B_n^m(t) = \sum_{j=n}^m \mathbf{b}_j(t)$$

Para establecerla se supone, por comodidad, que  $n = 1$ :

$$\begin{aligned} & a_m(\mathbf{b}_1 + \mathbf{b}_2 + \cdots + \mathbf{b}_m) + (a_1 - a_2)\mathbf{b}_1 + (a_2 - a_3)(\mathbf{b}_1 + \mathbf{b}_2) + \cdots + (a_{m-1} - a_m)(\mathbf{b}_1 + \mathbf{b}_2 + \cdots + \mathbf{b}_{m-1}) = \\ & = a_m(\mathbf{b}_1 + \mathbf{b}_2 + \cdots + \mathbf{b}_m) + (a_1 - a_m)\mathbf{b}_1 + (a_2 - a_m)\mathbf{b}_2 + (a_3 - a_m)\mathbf{b}_3 + \cdots + (a_{m-1} - a_m)\mathbf{b}_{m-1} = \\ & = a_1\mathbf{b}_1 + a_2\mathbf{b}_2 + \cdots + a_m\mathbf{b}_m = \mathbf{f}_1 + \mathbf{f}_2 + \cdots + \mathbf{f}_m = F_1^m \end{aligned}$$

Utilizando [\*] vamos a demostrar si se cumple b) ó d) entonces se verifica la condición de Cauchy para la convergencia uniforme. Para ello se introducen las sucesiones

$$\epsilon(n) = \sup_{m \geq n} \|B_n^m\|_T; \quad \delta(n) = \sup_{m \geq n} \|F_n^m\|_T.$$

Si se cumple b), la condición de Cauchy para la convergencia uniforme de la serie  $\sum_n \mathbf{b}_n$  se traduce en que  $\lim_n \epsilon(n) = 0$ .

Observemos en primer lugar que si  $t \in T$ , y  $m \in \mathbb{N}$  se verifica

$$|a_m(t)| \leq |a_1(t)| + |a_m(t) - a_1(t)| \leq |a_1(t)| + \sum_{i=1}^{m-1} |a_{i+1}(t) - a_i(t)| \leq C$$

Por otra parte, para cada  $j \geq n$  y cada  $t \in T$  se cumple  $|B_n^j(t)| \leq \epsilon(n)$ , y aplicando [\*] se obtiene

$$|F_n^m(t)| \leq C\epsilon(n) + \epsilon(n) \sum_{j=1}^{m-1} |a_j(t) - a_{j+1}(t)| \leq 2C\epsilon(n)$$

de donde se sigue que  $\delta(n) \leq 2C\epsilon(n)$ , y por lo tanto  $\lim_n \delta(n) = 0$ , lo que significa que la serie  $\sum_n \mathbf{f}_n$  cumple la condición de Cauchy para la convergencia uniforme.

Si se cumple d), sea  $R > 0$  tal que  $|B_1^m(t)| \leq R$  para todo  $t \in T$  y todo  $m \in \mathbb{N}$ . Entonces  $|B_n^m(t)| \leq 2R$ , y utilizando [\*] se deduce que para todo  $t \in T$  y todo  $m \geq n$  se cumple:

$$|F_n^m(t)| \leq 2R \left( \|a_m\|_T + \sum_{j=n}^{\infty} |a_j(t) - a_{j+1}(t)| \right)$$

luego  $\delta(n) \leq 2R(\alpha(n) + \beta(n))$  donde,

$$\alpha(n) = \sup_{m \geq n} \|a_m\|_T, \quad \beta(n) = \sup_{t \in T} \sum_{j=n}^{\infty} |a_j(t) - a_{j+1}(t)|$$

En virtud de las hipótesis,  $\lim_n \alpha(n) = \lim_n \beta(n) = 0$ , luego  $\lim_n \delta(n) = 0$ , y se concluye como antes que la serie  $\sum_n \mathbf{f}_n$  cumple la condición de Cauchy para la convergencia uniforme.

Para terminar la demostración del teorema basta observar que a)  $\Rightarrow$  b), y c)  $\Rightarrow$  d). Efectivamente, si se cumple a) y  $|a_n(t)| \leq M$  para todo  $n \in \mathbb{N}$  y todo  $t \in T$ , suponiendo que cada sucesión  $a_n(t)$  es decreciente, se obtiene

$$\begin{aligned} \sum_{n=1}^m |a_n(t) - a_{n+1}(t)| &= a_1(t) - a_2(t) + a_2(t) - a_3(t) + \cdots + a_m(t) - a_{m+1}(t) = \\ &= a_1(t) - a_{m+1}(t) \leq 2M \end{aligned}$$

luego se verifica b), ya que

$$|a_1(t)| + \sum_{n=1}^{\infty} |a_n(t) - a_{n+1}(t)| \leq M + 2M = 3M \text{ para todo } t \in T$$

Por otra parte, si se cumple c) y cada sucesión  $a_n(t)$  es decreciente, es claro que la sucesión de funciones  $\sum_{n=1}^m |a_n(t) - a_{n+1}(t)| = a_1(t) - a_{m+1}(t)$  converge uniformemente hacia la función  $a_1(t)$  y por lo tanto se cumple d). ■

NOTA: El apartado a) del teorema C.9 es el clásico criterio de Abel, [2, Ejer.9.13]; y el apartado b) es una ligera mejora de este. El apartado c) es el criterio de Dirichlet, [2, teo. 9.15], y el apartado d) es una versión algo más general del mismo.

# D

## Integración de funciones vectoriales

En esta sección se exponen dos alternativas para definir la integral de una función de variable real con valores en un espacio normado completo. La primera de ellas proporciona una ilustración interesante del teorema de extensión de aplicaciones uniformemente continuas 3.26 Se comienza definiendo una integral elemental sobre el espacio vectorial de las funciones escalonadas, y luego se extiende a las funciones que son límite uniforme de estas funciones. Esta es una clase de funciones bastante amplia que incluye a las continuas, y a otras muchas que no lo son, como las escalonadas y las de variación acotada.

Otra vía para definir la integral consiste en extender directamente la integral de Riemann definiéndola como límite, cuando exista, de sumas de Riemann asociadas a particiones cada vez más finas. Igual que en el caso de las funciones con valores reales, la demostración de que las funciones continuas son integrables se basa en la continuidad uniforme, pero el no poder considerar sumas superiores e inferiores hace que ahora el razonamiento sea algo más complicado.

Una vez que se ha definido la integral de una función continua con valores en un espacio normado completo se demuestra, en la forma usual, el teorema fundamental del cálculo que se necesita para obtener la fórmula integral del resto en el desarrollo de Taylor de funciones con valores en espacios normados completos.

### D.1. Integración de funciones regladas

**Definición D.1** Sea  $(F, \| \cdot \|)$  un espacio normado y  $\mathbf{h} : [a, b] \rightarrow F$ . Si existe una subdivisión  $p \in \mathcal{P}([a, b])$ ,  $p = (x_0 < x_1 < \cdots < x_m)$  tal que  $\mathbf{h}$  es constante en cada intervalo abierto  $(x_{i-1}, x_i)$  se dice que  $\mathbf{h}$  es escalonada y que  $p$  es una subdivisión admisible para  $\mathbf{h}$ . Si  $\mathbf{f} : [a, b] \rightarrow F$  es límite uniforme de una sucesión de funciones escalonadas  $\mathbf{h}_n : [a, b] \rightarrow F$  se dice que  $\mathbf{f}$  es reglada.

El conjunto de puntos de discontinuidad  $D(\mathbf{f})$  de una función reglada  $\mathbf{f}$  es numerable. En efecto, en las condiciones de la definición anterior, cada  $\mathbf{h}_n$  es continua excepto

en los puntos de un conjunto finito  $D(\mathbf{h}_n)$ , y aplicando el teorema 3.31, se obtiene que  $D(\mathbf{f}) \subset \cup_n D(\mathbf{h}_n)$  es numerable.

En lo que sigue  $\mathcal{E}([a, b], F)$  denotará el subespacio de  $l^\infty([a, b], F)$  formado por las funciones escalonadas y  $\overline{\mathcal{E}([a, b], F)}$  su clausura en  $l^\infty([a, b], F)$  para la norma

$$\|\mathbf{f}\|_\infty = \sup\{\|\mathbf{f}(t)\| : t \in [a, b]\}$$

Según la definición D.1 el conjunto de las funciones regladas  $\mathbf{f} : [a, b] \rightarrow F$  es  $\overline{\mathcal{E}([a, b], F)}$ . En virtud de la siguiente proposición, el subespacio vectorial de las funciones continuas  $C([a, b], F) \subset l^\infty([a, b], F)$  está contenido en  $\overline{\mathcal{E}([a, b], F)}$ .

**Proposición D.2** *Toda función continua  $\mathbf{f} : [a, b] \rightarrow F$  es reglada, es decir,*

$$C([a, b], F) \subset \overline{\mathcal{E}([a, b], F)}$$

DEM: Si  $\mathbf{f} : [a, b] \rightarrow F$  es continua, en virtud de 3.14,  $\mathbf{f}([a, b])$  es acotado y por lo tanto  $\mathbf{f} \in l^\infty([a, b], F)$ . Por otra parte, como  $\mathbf{f}$  es uniformemente continua, (véase 3.24) dado  $\epsilon > 0$  existe  $\delta > 0$  tal que si  $s, t \in [a, b]$  y  $|s - t| < \delta$  entonces  $\|\mathbf{f}(t) - \mathbf{f}(s)\| < \epsilon$ .

Sea  $p = (a = x_0 < x_1 < \dots < x_m = b)$  una subdivisión tal que  $\Delta(p) = \max\{x_k - x_{k-1} : 1 \leq k \leq m\} < \delta$ . La función escalonada  $\mathbf{h}_\epsilon : [a, b] \rightarrow F$  definida por

$$\mathbf{h}(a) = \mathbf{f}(a), \quad \mathbf{h}(t) = \mathbf{f}(x_k) \quad \text{si } t \in (x_{k-1}, x_k]$$

verifica  $\|\mathbf{f}(t) - \mathbf{h}_\epsilon(t)\| < \epsilon$  para todo  $t \in [a, b]$ , es decir,  $\|\mathbf{f} - \mathbf{h}_\epsilon\|_\infty \leq \epsilon$  y queda probado que  $\mathbf{f} \in l^\infty([a, b], F)$  es adherente a  $\mathcal{E}([a, b], F)$ . ■

Sea  $\mathbf{h} : [a, b] \rightarrow F$  una función escalonada y  $p = (a = x_0 < x_1 < \dots < x_m = b)$  una subdivisión admisible para  $\mathbf{h}$ . Si  $\mathbf{v}_k$  es el valor constante de  $\mathbf{h}$  en  $(x_{k-1}, x_k)$ , se define  $S(\mathbf{h}, p) = \sum_{k=1}^m (x_k - x_{k-1})\mathbf{v}_k$ . Es fácil comprobar que si  $q \in \mathcal{P}([a, b])$  es otra subdivisión admisible para  $\mathbf{h}$  entonces  $S(\mathbf{h}, p) = S(\mathbf{h}, q)$ . Este hecho permite formular la siguiente definición

**Definición D.3** *Si  $\mathbf{h} : [a, b] \rightarrow F$  es escalonada y  $p \in \mathcal{P}([a, b])$  es una subdivisión admisible para  $\mathbf{h}$ , se define*

$$\int_a^b \mathbf{h}(t)dt = S(\mathbf{h}, p) = \sum_{i=1}^m (x_k - x_{k-1})\mathbf{v}_k$$

donde  $p = (x_0 < x_1 < \dots < x_m)$  y  $\mathbf{v}_k$  es el valor constante de  $\mathbf{h}$  en  $(x_{k-1}, x_k)$ ,

Obsérvese que en la definición de la integral  $\int_a^b \mathbf{h}(t)dt$  no intervienen los valores de  $\mathbf{h}$  en los puntos  $x_k$ .

**Lema D.4** *Si en  $\mathcal{E}([a, b], F)$  se considera la norma  $\|\mathbf{h}\|_\infty = \sup\{\|\mathbf{h}(t)\| : t \in [a, b]\}$ , entonces la integral  $I : \mathcal{E}([a, b], F) \rightarrow F$ ,  $I(\mathbf{h}) = \int_a^b \mathbf{h}(t)dt$ , es una aplicación lineal continua de norma  $\|I\| \leq (b - a)$ .*

DEM: La linealidad de la integral es inmediata:

Es evidente que  $\int_a^b \alpha \mathbf{h}(t) dt = \alpha \int_a^b \mathbf{h}(t) dt$ . Por otra parte, si  $\mathbf{h}_1, \mathbf{h}_2 \in \mathcal{E}([a, b], F)$  es claro que existe una subdivisión  $p \in \mathcal{P}([a, b])$  que es admisible para  $\mathbf{h}_1$  y para  $\mathbf{h}_2$ . Entonces  $p$  es admisible para  $\mathbf{h} = \mathbf{h}_1 + \mathbf{h}_2$  y por lo tanto

$$\int_a^b \mathbf{h}(t) dt = S(\mathbf{h}, p) = S(\mathbf{h}_1, p) + S(\mathbf{h}_2, p) = \int_a^b \mathbf{h}_1(t) dt + \int_a^b \mathbf{h}_2(t) dt$$

Finalmente, si  $p \in \mathcal{P}([a, b])$  es admisible para la función escalonada  $\mathbf{h}$ , en virtud de la desigualdad triangular  $\|S(\mathbf{h}, p)\| \leq S(\|\mathbf{h}\|, p)$ , luego

$$\left\| \int_a^b \mathbf{h}(t) dt \right\| \leq \int_a^b \|\mathbf{h}(t)\| dt \leq \int_a^b \|\mathbf{h}\|_\infty dt = (b-a) \|\mathbf{h}\|_\infty$$

Es decir  $\|I(\mathbf{h})\| \leq (b-a) \|\mathbf{h}\|_\infty$  para todo  $\mathbf{h} \in \mathcal{E}([a, b], F)$ , luego  $I$  es una aplicación lineal continua de norma  $\|I\| \leq (b-a)$ . ■

**Proposición D.5** *Si el espacio normado  $(F, \|\cdot\|)$  es completo, la integral de las funciones escalonadas  $I : \mathcal{E}([a, b], F) \rightarrow F$ , se puede extender a una (única) aplicación lineal continua*

$$\hat{I} : \overline{\mathcal{E}([a, b], F)} \rightarrow F$$

*de norma  $\|\hat{I}\| \leq (b-a)$ . (El espacio  $\overline{\mathcal{E}([a, b], F)}$  se considera con la norma de la convergencia uniforme  $\|\mathbf{f}\|_\infty = \sup\{\|\mathbf{f}(t)\| : t \in [a, b]\}$ ).*

DEM: En virtud de D.4 la aplicación lineal  $I$  es continua y por lo tanto uniformemente continua. Como  $F$  es completo, aplicando el teorema de extensión de aplicaciones uniformemente continuas (véase 3.26) se obtiene una única extensión uniformemente continua  $\hat{I} : \overline{\mathcal{E}([a, b], F)} \rightarrow F$ . Según este teorema, para cada  $\mathbf{f} \in \overline{\mathcal{E}([a, b], F)}$  la extensión viene dada por

$$\hat{I}(\mathbf{f}) = \lim_n I(\mathbf{h}_n)$$

donde  $\mathbf{h}_n$  es cualquier sucesión de funciones escalonadas que converge uniformemente hacia  $\mathbf{f}$  (e.d. en la norma  $\|\cdot\|_\infty$ ). Utilizando este hecho y la linealidad de  $I$  sobre el espacio de las funciones escalonadas, se comprueba fácilmente que  $\hat{I}$  una aplicación lineal.

Por otra parte, si  $\mathbf{f} \in \overline{\mathcal{E}([a, b], F)}$  y  $\mathbf{h}_n$  es una sucesión de funciones escalonadas con  $\lim_n \|\mathbf{f} - \mathbf{h}_n\|_\infty = 0$ , se verifica  $\|\mathbf{f}\|_\infty = \lim_n \|\mathbf{h}_n\|_\infty$ . Como  $\hat{I}(\mathbf{f}) = \lim_n I(\mathbf{h}_n)$ , en virtud del lema D.4 se obtiene  $\|\hat{I}(\mathbf{f})\| = \lim_n \|I(\mathbf{h}_n)\| \leq \lim_n (b-a) \|\mathbf{h}_n\|_\infty = (b-a) \|\mathbf{f}\|_\infty$ , luego  $\hat{I}$  es una aplicación lineal continua de norma  $\|\hat{I}\| \leq (b-a)$ . ■

La proposición anterior permite definir la integral de una función reglada y en particular de una función continua

**Definición D.6** *Si  $(F, \|\cdot\|)$  es completo, y  $f \in \overline{\mathcal{E}([a, b], F)}$  es una función reglada se define  $\int_a^b \mathbf{f}(t) dt = \hat{I}(\mathbf{f})$  donde  $\hat{I}$  es la única extensión lineal continua de la integral elemental  $I(\mathbf{h}) = \int_a^b \mathbf{h}(t) dt$  definida sobre las funciones escalonadas.*

NOTA: En virtud de la definición, la integral es una aplicación lineal continua sobre  $\mathcal{E}([a, b], F)$ . Por lo tanto, si  $\mathbf{f}_n : [a, b] \rightarrow \mathbb{R}$  es una sucesión de funciones regladas que converge uniformemente hacia la función  $\mathbf{f} : [a, b] \rightarrow F$  ( que necesariamente es reglada), se cumple:

$$\int_a^b \mathbf{f}(t) dt = \lim_n \int_a^b \mathbf{f}_n(t) dt$$

**Proposición D.7** Si  $(F, \| \cdot \|)$  es completo y  $\mathbf{f} : [a, b] \rightarrow F$  es reglada se verifica

$$a) \left\| \int_a^b \mathbf{f}(t) dt \right\| \leq \int_a^b \|\mathbf{f}(t)\| dt;$$

$$b) \int_a^b \mathbf{f}(t) dt = \int_a^c \mathbf{f}(t) dt + \int_c^b \mathbf{f}(t) dt \text{ si } a \leq c \leq b.$$

DEM: a) Si  $\mathbf{h}_n$  es una sucesión de funciones escalonadas que converge uniformemente hacia  $\mathbf{f}$ , en virtud de la desigualdad

$$| \|\mathbf{f}(t)\| - \|\mathbf{h}_n(t)\| | \leq \|\mathbf{f}_n(t) - \mathbf{f}(t)\| \leq \|\mathbf{f}_n - \mathbf{f}\|_\infty$$

la sucesión de funciones escalonadas reales  $\|\mathbf{h}_n(t)\|$  converge uniformemente hacia la función real  $\|\mathbf{f}(t)\|$ , luego, en virtud de la nota que sigue a D.6 se cumple

$$\int_a^b \|\mathbf{f}(t)\| dt = \lim_n \int_a^b \|\mathbf{h}_n(t)\| dt$$

Pasando al límite en la desigualdad  $\left\| \int_a^b \mathbf{h}_n(t) dt \right\| \leq \int_a^b \|\mathbf{h}_n(t)\| dt$  válida para todo  $n \in \mathbb{N}$  se obtiene  $\left\| \int_a^b \mathbf{f}(t) dt \right\| \leq \int_a^b \|\mathbf{f}(t)\| dt$ .

b) El resultado es inmediato en el caso particular de las funciones escalonadas. El caso general se deduce de este considerando una sucesión de funciones escalonadas que converge uniformemente hacia  $\mathbf{f}$ . ■ .

## D.2. Definición general de la integral de Riemann

La integral de Riemann para funciones  $\mathbf{f} : [a, b] \rightarrow F$  con valores en un espacio normado completo  $(F, \| \cdot \|)$  se puede definir tomando como base las sumas de Riemann, pero la teoría no es tan satisfactoria como en el caso finito dimensional  $F = \mathbb{R}^n$ , pues propiedades básicas como D.14 ya no son ciertas: Pueden existir funciones integrables tales que  $\|\mathbf{f}\|$  no sea integrable o tales que el conjunto de sus puntos de discontinuidad no sea de medida nula. Sin embargo los resultados requeridos para obtener la fórmula integral del resto en el desarrollo de Taylor, son bastante modestos: Basta demostrar la integrabilidad de las funciones continuas y obtener, en el marco de estas funciones, los resultados imprescindibles para obtener el teorema fundamental del cálculo D.11.

**Definición D.8** Una función  $\mathbf{f} : [a, b] \rightarrow F$  con valores en un espacio normado  $(F, \|\cdot\|)$  se dice que es integrable Riemann si existe  $\mathbf{v} \in F$  verificando: Para cada  $\epsilon > 0$  existe  $p_\epsilon \in \mathcal{P}([a, b])$  tal que si  $p \in \mathcal{P}([a, b])$ , es más fina que  $p_\epsilon$  entonces toda suma de Riemann asociada a  $p$  verifica  $\|\Sigma(\mathbf{f}, p, \eta) - \mathbf{v}\| < \epsilon$ .

En este caso la integral  $\int_a^b \mathbf{f}(t)dt$  es el único  $\mathbf{v} \in F$  que cumple esta condición.

**Teorema D.9** Si  $(F, \|\cdot\|)$  es completo, toda función continua  $\mathbf{f} : [a, b] \rightarrow F$  es integrable Riemann.

DEM: Para cada  $\delta > 0$  se define el módulo de continuidad uniforme:

$$\omega(\delta) = \sup\{\|\mathbf{f}(s) - \mathbf{f}(t)\| : s, t \in [a, b], |s - t| < \delta\}$$

Como  $\mathbf{f}$  es uniformemente continua en el compacto  $[a, b]$ , dado  $\epsilon > 0$  existe  $\delta_\epsilon > 0$  tal que  $\|\mathbf{f}(s) - \mathbf{f}(t)\| < \epsilon$  siempre que  $s, t \in [a, b]$  con  $|s - t| < \delta_\epsilon$ . Entonces es claro que  $0 < \delta < \delta_\epsilon \Rightarrow 0 \leq \omega(\delta) \leq \omega(\delta_\epsilon)$ , es decir,  $\lim_{\delta \rightarrow 0} \omega(\delta) = 0$ .

Comenzamos demostrando:

$$[\alpha] : p, q \in \mathcal{P}([a, b]), p \subset q, \text{ y } \Delta(p) < \delta \Rightarrow \|\Sigma(\mathbf{f}, p) - \Sigma(\mathbf{f}, q)\| \leq \omega(\delta)(b - a).$$

Si  $p = (t_0 < t_1 < t_2 < \dots < t_m)$ , y  $j \in \{1, 2, \dots, m\}$ , sea  $q_j$  la subdivisión que  $q$  induce en  $[t_{j-1}, t_j]$ . Si  $s_k$  son los puntos de  $q_j$ , consideramos la suma

$$\mathbf{v}_j = (t_j - t_{j-1})\mathbf{f}(t_{j-1}) - \Sigma(\mathbf{f}, q_j) = \sum_k (s_k - s_{k-1})(\mathbf{f}(t_{j-1}) - \mathbf{f}(s_{k-1}))$$

Como  $s_{k-1} \in [t_{j-1}, t_j]$ , se cumple  $|s_{k-1} - t_{j-1}| \leq |t_j - t_{j-1}| \leq \Delta(p) \leq \delta$ , luego

$$\|\mathbf{f}(t_{j-1}) - \mathbf{f}(s_{k-1})\| \leq \omega(\delta)$$

y en virtud de la desigualdad triangular,

$$\|\mathbf{v}_j\| \leq \omega(\delta) \sum_k (s_k - s_{k-1}) = \omega(\delta)(t_j - t_{j-1})$$

Teniendo en cuenta que  $\Sigma(\mathbf{f}, q) = \Sigma_{j=1}^m \Sigma(\mathbf{f}, q_j)$ , resulta

$$\Sigma(\mathbf{f}, p) - \Sigma(\mathbf{f}, q) = \sum_{j=1}^m [(t_j - t_{j-1})\mathbf{f}(t_{j-1}) - \Sigma(\mathbf{f}, q_j)] = \sum_{j=1}^m \mathbf{v}_j$$

luego

$$\|\Sigma(\mathbf{f}, p) - \Sigma(\mathbf{f}, q)\| \leq \sum_{j=1}^m \|\mathbf{v}_j\| \leq \sum_{j=1}^m \omega(\delta)(t_j - t_{j-1}) = \omega(\delta)(b - a)$$

Finalmente, utilizamos  $[\alpha]$  para demostrar que  $\mathbf{f}$  es integrable: Sea  $p_n \in \mathcal{P}([a, b])$  una sucesión tal que  $\delta_n = \Delta(p_n)$  tiende hacia 0, y  $p_n \subset p_{n+1}$  para cada  $n \in \mathbb{N}$ .

La sucesión  $\mathbf{u}_n = \Sigma(\mathbf{f}, p_n)$  es de Cauchy pues, en virtud de  $[\alpha]$

$$m \geq n \Rightarrow \|\Sigma(\mathbf{f}, p_n) - \Sigma(\mathbf{f}, p_m)\| \leq \omega(\delta_n)(b - a)$$



y la sucesión  $\omega(\delta_n)$  tiende hacia 0. Como  $F$  es completo, la sucesión  $\mathbf{u}_n$  es convergente y sólo queda por verificar que el límite  $\mathbf{u} = \lim_n \mathbf{u}_n$  cumple los requisitos de la definición D.9. En efecto, dado  $\epsilon > 0$  existe  $n \in \mathbb{N}$  tal que

$$2(b-a)\omega(\delta_n) + \|\mathbf{u}_n - \mathbf{u}\| < \epsilon$$

Si  $p \in \mathcal{P}([a, b])$  es más fina que  $p_n$ , se cumple  $\Delta(p) \leq \Delta(\delta_n)$ , luego toda suma de Riemann  $\Sigma(\mathbf{f}, p, \eta)$  asociada a  $\mathbf{p}$  verifica

$$\|\Sigma(\mathbf{f}, p, \eta) - \Sigma(\mathbf{f}, p)\| \leq (b-a)\omega(\delta_n)$$

Por otra parte, usando otra vez  $[\alpha]$  se obtiene

$$\|\Sigma(\mathbf{f}, p) - \Sigma(\mathbf{f}, p_n)\| \leq (b-a)\omega(\delta_n)$$

luego, en virtud de la desigualdad triangular

$$\begin{aligned} \|\Sigma(\mathbf{f}, p, \eta) - \mathbf{u}\| &\leq \|\Sigma(\mathbf{f}, p, \eta) - \Sigma(\mathbf{f}, p)\| + \|\Sigma(\mathbf{f}, p) - \Sigma(\mathbf{f}, p_n)\| + \|\Sigma(\mathbf{f}, p_n) - \mathbf{u}\| \leq \\ &\leq 2(b-a)\omega(\delta_n) + \|\mathbf{u}_n - \mathbf{u}\| < \epsilon \end{aligned}$$

■

**Proposición D.10** *Si  $(F, \|\cdot\|)$  es completo y  $\mathbf{f} : [a, b] \rightarrow F$  es continua se verifica*

- a)  $\left\| \int_a^b \mathbf{f}(t) dt \right\| \leq \int_a^b \|\mathbf{f}(t)\| dt$ ;
- b)  $\int_a^b \mathbf{f}(t) dt = \int_a^c \mathbf{f}(t) dt + \int_c^b \mathbf{f}(t) dt$  si  $a \leq c \leq b$ .

DEM: Sea  $p_n \in \mathcal{P}([a, b])$  una sucesión de subdivisiones tal que  $\delta_n = \Delta(p_n)$  tiende hacia 0, y  $p_n \subset p_{n+1}$  para cada  $n \in \mathbb{N}$ . Según la demostración de D.9 se verifica

$$\int_a^b \mathbf{f}(t) dt = \lim_n \Sigma(\mathbf{f}, p_n); \quad \int_a^b \|\mathbf{f}(t)\| dt = \lim_n \Sigma(\|\mathbf{f}\|, p_n)$$

En virtud de la desigualdad triangular  $\|\Sigma(\mathbf{f}, p_n)\| \leq \Sigma(\|\mathbf{f}\|, p_n)$ , y pasando al límite se obtiene a). Para demostrar b) podemos suponer que  $c \in p_n$  para cada  $n \in \mathbb{N}$ . En este caso si  $p'_n$ , y  $p''_n$  son las subdivisiones que  $p_n$  induce en  $[a, c]$  y en  $[c, b]$  respectivamente, se verifica

$$\int_a^b \mathbf{f}(t) dt = \lim_n \Sigma(\mathbf{f}, p_n); \quad \int_a^c \mathbf{f}(t) dt = \lim_n \Sigma(\mathbf{f}, p'_n); \quad \int_c^b \mathbf{f}(t) dt = \lim_n \Sigma(\mathbf{f}, p''_n)$$

Es claro que  $\Sigma(\mathbf{f}, p_n) = \Sigma(\mathbf{f}, p'_n) + \Sigma(\mathbf{f}, p''_n)$ , y pasando al límite se obtiene el resultado.

■

**Teorema D.11** *Sea  $\mathbf{f} : [a, b] \rightarrow F$  una función continua con valores en un espacio normado completo  $(F, \|\cdot\|)$ . Entonces la función  $\mathbf{g} : [a, b] \rightarrow F$  definida por  $\mathbf{g}(x) = \int_a^x \mathbf{f}(t) dt$  es derivable en  $[a, b]$  y  $\mathbf{g}'(x) = \mathbf{f}(x)$  para todo  $x \in [a, b]$  (en  $a$  y  $b$  se consideran las derivadas laterales correspondientes).*

DEM: La demostración es análoga a la del caso de funciones con valores reales. Fijado  $x_0 \in [a, b]$ , dado  $\epsilon > 0$  existe  $\delta > 0$  tal que si  $t \in [a, b]$  y  $|t - x_0| < \delta$  se cumple  $\|\mathbf{f}(t) - \mathbf{f}(x_0)\| < \epsilon$ . Supongamos que  $a < x_0 < b$ , y sea  $x \in [a, b]$  tal que  $|x - x_0| < \delta$ , y  $x > x_0$ . En virtud de la proposición D.10 b)

$$\left\| \frac{\mathbf{g}(x) - \mathbf{g}(x_0)}{x - x_0} - \mathbf{f}(x_0) \right\| = \frac{1}{x - x_0} \left\| \int_{x_0}^x \mathbf{f}(t) dt - \int_{x_0}^x \mathbf{f}(x_0) dt \right\|$$

y aplicando la desigualdad D.10 a) se obtiene

$$\left\| \frac{\mathbf{g}(x) - \mathbf{g}(x_0)}{x - x_0} - \mathbf{f}(x_0) \right\| \leq \frac{1}{x - x_0} \int_{x_0}^x \|\mathbf{f}(t) - \mathbf{f}(x_0)\| dt \leq \epsilon$$

donde la última desigualdad se obtiene teniendo en cuenta que para todo  $t \in [x_0, x]$  se cumple  $\|\mathbf{f}(t) - \mathbf{f}(x_0)\| < \epsilon$ . Queda demostrado así que  $\mathbf{g}$  es derivable por la derecha en  $x_0$  con  $\mathbf{g}'_d(x_0) = \mathbf{f}(x_0)$ . Análogamente se demuestra que  $\mathbf{g}$  es derivable por la izquierda en  $x_0$  con  $\mathbf{g}'_i(x_0) = \mathbf{f}(x_0)$  y con ello queda demostrado el teorema. ■

**Corolario D.12** [Regla de Barrow] *Si  $(F, \|\cdot\|)$  es completo y  $\mathbf{g} : [a, b] \rightarrow F$  es derivable con derivada continua se verifica*

$$\mathbf{g}(b) - \mathbf{g}(a) = \int_a^b \mathbf{g}'(t) dt$$

DEM: Como  $\mathbf{g}'$  es continua, según el teorema D.11, la función  $\mathbf{h}(x) = \int_a^x \mathbf{g}'(t) dt$  es derivable y  $\mathbf{h}'(x) = \mathbf{g}'(x)$  para todo  $x \in [a, b]$ . En virtud del corolario 4.9 la diferencia  $\mathbf{g}(x) - \mathbf{h}(x)$  es constante. Su valor constante es  $\mathbf{g}(a) - \mathbf{h}(a) = \mathbf{g}(a)$ , luego  $\mathbf{g}(x) - \mathbf{h}(x) = \mathbf{g}(a)$  para todo  $x \in [a, b]$ . Con  $x = b$  se obtiene el resultado. ■

**Corolario D.13** [Integración por partes] *Si  $(F, \|\cdot\|)$  es completo y las funciones  $\mathbf{f} : [a, b] \rightarrow F$ ,  $\varphi : [a, b] \rightarrow \mathbb{R}$  son derivables con derivada continua, se verifica*

$$\varphi(b)\mathbf{f}(b) - \varphi(a)\mathbf{f}(a) = \int_a^b \varphi'(t)\mathbf{f}(t) dt + \int_a^b \varphi(t)\mathbf{f}'(t) dt$$

DEM: En virtud de 4.5 ii) la función  $\mathbf{g}(t) = \varphi(t)\mathbf{f}(t)$  es derivable y su derivada es  $\mathbf{g}'(t) = \varphi'(t)\mathbf{f}(t) + \varphi(t)\mathbf{f}'(t)$ . Como  $\mathbf{g}'$  es continua, con el corolario D.12 se obtiene el resultado. ■

**Proposición D.14** *Si  $\mathbf{f} : [a, b] \rightarrow \mathbb{R}^n$  es integrable, y  $\|\cdot\|$  es una norma sobre  $\mathbb{R}^n$  entonces la función  $\|\mathbf{f}\|$  también es integrable y*

$$\left\| \int_a^b \mathbf{f}(t) dt \right\| \leq \int_a^b \|\mathbf{f}(t)\| dt$$

DEM: La integrabilidad de  $\|\mathbf{f}\|$  se puede demostrar usando el teorema de Lebesgue que asegura que una función acotada  $g : [a, b] \rightarrow \mathbb{R}$  es integrable Riemann si y sólo si el conjunto de sus puntos de discontinuidad, denotado  $D(g)$ , tiene medida nula (véase 10.24). Como los conjuntos  $D(f_i)$  tienen medida nula se sigue que  $D(\|\mathbf{f}\|) \subset D(f_1) \cup D(f_2) \cup \dots \cup D(f_n)$  también tiene medida nula y por lo tanto  $\|\mathbf{f}\|$  es integrable. Con el lema 4.13 podemos conseguir una sucesión  $p_k \in \mathcal{P}([a, b])$ , donde cada  $p_k$  es más fina que  $p_{k-1}$ , tal que

$$\int_a^b \mathbf{f}(t) dt = \lim_k \Sigma(\mathbf{f}, p_k) \quad \text{y} \quad \int_a^b \|\mathbf{f}(t)\| dt = \lim_k \Sigma(\|\mathbf{f}\|, p_k)$$

En virtud de la desigualdad triangular  $\|\Sigma(\mathbf{f}, p_k)\| \leq \Sigma(\|\mathbf{f}\|, p_k)$  y usando la continuidad de la norma se obtiene

$$\left\| \int_a^b \mathbf{f}(t) dt \right\| = \lim_k \|\Sigma(\mathbf{f}, p_k)\| \leq \lim_k \Sigma(\|\mathbf{f}\|, p_k) = \int_a^b \|\mathbf{f}(t)\| dt.$$

■

**Ejercicio D.15** *Compruebe que la definición de la integral dada en D.8 es consistente con la dada en D.6 Es decir, toda función reglada es integrable según la definición D.8 y las dos definiciones de integral dan el mismo valor.*

# E

## Complementos sobre diferenciabilidad

### E.1. Caracterización de las funciones de clase $C^1$

En la siguiente proposición se caracterizan las funciones de clase  $C^1$  mediante una condición donde no intervienen las derivadas parciales. Esta condición es la que se suele utilizar para definir las funciones de clase  $C^1$  cuando  $E$  no es finito dimensional.

**Proposición E.1** *Una función  $\mathbf{f} : \Omega \rightarrow \mathbf{F}$ , definida en un abierto  $\Omega \subset E = \mathbb{R}^n$ , con valores en un espacio normado  $(F, \|\cdot\|)$  es de clase  $C^1$  si y sólo si  $\mathbf{f}$  es diferenciable en cada  $\mathbf{x} \in \Omega$  y la aplicación  $d\mathbf{f} : \Omega \rightarrow \mathcal{L}(E, F)$  es continua.*

DEM: Si  $\mathbf{x} \rightarrow d\mathbf{f}(\mathbf{x})$  es continua, para cada  $\mathbf{u} \in E$  también lo es  $\mathbf{x} \rightarrow D_{\mathbf{u}}\mathbf{f}(\mathbf{x}) = d\mathbf{f}(\mathbf{x})\mathbf{u}$ , ya que

$$\|d\mathbf{f}(\mathbf{x})\mathbf{u} - d\mathbf{f}(\mathbf{a})\mathbf{u}\| \leq \|d\mathbf{f}(\mathbf{x}) - d\mathbf{f}(\mathbf{a})\| \|\mathbf{u}\|$$

En particular, las derivadas parciales  $D_i\mathbf{f}(\mathbf{x})$ ,  $1 \leq i \leq n$  son continuas.

Recíprocamente, si las derivadas parciales  $D_i\mathbf{f}(\mathbf{x})$ ,  $1 \leq i \leq n$ , son continuas, dado  $\mathbf{a} \in \Omega$  y  $\epsilon > 0$ , existe  $\delta > 0$  tal que si  $\|\mathbf{x} - \mathbf{a}\| < \delta$ , para todo  $j \in \{1, \dots, n\}$  se cumple  $\|D_j\mathbf{f}(\mathbf{x}) - D_j\mathbf{f}(\mathbf{a})\| < \epsilon$ . Entonces, para todo  $\mathbf{u} \in \mathbb{R}^n$  con  $\|\mathbf{u}\|_2 \leq 1$  se verifica

$$\begin{aligned} \|[d\mathbf{f}(\mathbf{x}) - d\mathbf{f}(\mathbf{a})]\mathbf{u}\| &= \left\| \sum_{j=1}^n [D_j\mathbf{f}(\mathbf{x}) - D_j\mathbf{f}(\mathbf{a})]u_j \right\| \leq \\ &\leq \sum_{j=1}^n \|D_j\mathbf{f}(\mathbf{x}) - D_j\mathbf{f}(\mathbf{a})\| |u_j| \leq \epsilon \|\mathbf{u}\|_1 \leq \sqrt{n}\epsilon \|\mathbf{u}\|_2 = \sqrt{n}\epsilon \end{aligned}$$

Considerando el supremo de  $\|[d\mathbf{f}(\mathbf{x}) - d\mathbf{f}(\mathbf{a})]\mathbf{u}\|$  cuando  $\|\mathbf{u}\|_2 \leq 1$  se obtiene que  $\|d\mathbf{f}(\mathbf{x}) - d\mathbf{f}(\mathbf{a})\| \leq \sqrt{n}\epsilon$  si  $\|\mathbf{x} - \mathbf{a}\| < \delta$ . ■

## E.2. La definición general de diferencial segunda

La definición de diferencial segunda que hemos dado en 6.1, en términos de la diferenciabilidad de las derivadas parciales, sólo tiene sentido cuando  $E = \mathbb{R}^n$ . Con la siguiente proposición se pone de manifiesto que la diferencial segunda  $d^2\mathbf{f}(\mathbf{a})$  es realmente la diferencial en  $\mathbf{a}$  de la diferencial primera  $d\mathbf{f}$ . Este hecho es el que permite extender la definición de diferencial segunda al caso en que  $E$  sea un espacio normado arbitrario.

Previamente conviene recordar que el espacio vectorial  $\mathcal{L}(E, F)$ , formado por las aplicaciones lineales continuas  $L : E \rightarrow F$ , lo podemos considerar como espacio normado, con la norma

$$\|L\| = \sup\{\|L(\mathbf{x})\|_F : \|\mathbf{x}\|_E \leq 1\}$$

Para lo que sigue, conviene tener presente que para cada  $\mathbf{u} \in E$  la evaluación

$$\delta_{\mathbf{u}} : \mathcal{L}(E, F) \rightarrow F, \quad \delta_{\mathbf{u}}(L) = L(\mathbf{u})$$

es una aplicación lineal continua, pues, para cada  $L \in \mathcal{L}(E, F)$ , se cumple

$$\|\delta_{\mathbf{u}}(L)\| \leq M \|L\| \quad \text{con } M = \|\mathbf{u}\|$$

**Proposición E.2** *Si  $\mathbf{f} : V \rightarrow F$  es una aplicación diferenciable definida en un abierto  $V$  de un espacio normado  $E$ .*

- a) *Si  $\mathbf{g} = d\mathbf{f} : V \rightarrow \mathcal{L}(E, F)$  es diferenciable en  $\mathbf{a} \in V$  entonces para cada par  $(\mathbf{u}, \mathbf{v}) \in E \times E$  existen las derivadas segundas  $D_{\mathbf{u}\mathbf{v}}\mathbf{f}(\mathbf{a})$ ,  $D_{\mathbf{v}\mathbf{u}}\mathbf{f}(\mathbf{a})$  y*

$$[d\mathbf{g}(\mathbf{a})(\mathbf{u})]\mathbf{v} = D_{\mathbf{u}\mathbf{v}}\mathbf{f}(\mathbf{a}) = D_{\mathbf{v}\mathbf{u}}\mathbf{f}(\mathbf{a})$$

- b) *Cuando  $E = \mathbb{R}^n$ , una condición necesaria y suficiente para que  $\mathbf{g} = d\mathbf{f}$  sea diferenciable en  $\mathbf{a} \in V$  es que todas las derivadas parciales  $D_i\mathbf{f} : V \rightarrow F$ ,  $1 \leq i \leq n$  sean diferenciables en  $\mathbf{a}$ . En este caso*

$$d^2\mathbf{f}(\mathbf{a})(\mathbf{u}, \mathbf{v}) = [d\mathbf{g}(\mathbf{a})(\mathbf{u})](\mathbf{v})$$

DEM: a) Como  $\mathbf{f}$  es diferenciable en  $V$ , para cada  $\mathbf{v} \in E$  y cada  $\mathbf{x} \in V$  existe la derivada  $D_{\mathbf{v}}\mathbf{f}(\mathbf{x}) = d\mathbf{f}(\mathbf{x})\mathbf{v}$ . La función  $\varphi(\mathbf{x}) = D_{\mathbf{v}}\mathbf{f}(\mathbf{x})$ , es el resultado de componer  $\mathbf{g} = d\mathbf{f}$  con la evaluación  $\delta_{\mathbf{v}} : \mathcal{L}(E, F) \rightarrow F$  que es lineal continua. Como la diferencial de una aplicación lineal continua es ella misma, en virtud de la regla de la cadena,

$$\varphi(\mathbf{x}) = D_{\mathbf{v}}\mathbf{f}(\mathbf{x}) = \delta_{\mathbf{v}}(d\mathbf{f}(\mathbf{x})) = (\delta_{\mathbf{v}} \circ \mathbf{g})(\mathbf{x})$$

es diferenciable en  $\mathbf{a}$ , y  $d\varphi(\mathbf{a}) = \delta_{\mathbf{v}} \circ d\mathbf{g}(\mathbf{a})$ .

Entonces, para cada  $\mathbf{u} \in E$  existe la derivada  $D_{\mathbf{u}}\varphi(\mathbf{a}) = d\varphi(\mathbf{a})(\mathbf{u})$ , igualdad que, en términos de  $\mathbf{f}$  y  $\mathbf{g} = d\mathbf{f}$ , se escribe en la forma

$$D_{\mathbf{u}}(D_{\mathbf{v}}\mathbf{f})(\mathbf{a}) = [\delta_{\mathbf{v}} \circ d\mathbf{g}(\mathbf{a})](\mathbf{u}) = [d\mathbf{g}(\mathbf{a})(\mathbf{u})](\mathbf{v})$$

Obsérvese que la aplicación lineal continua  $B = d\mathbf{g}(\mathbf{a}) : E \rightarrow \mathcal{L}(E, F)$  asigna a cada  $\mathbf{u} \in E$  un elemento  $B(\mathbf{u}) \in \mathcal{L}(E, F)$ , luego  $[B(\mathbf{u})](\mathbf{v}) \in F$  para cada  $\mathbf{v} \in E$ .

La aplicación lineal  $B : E \rightarrow \mathcal{L}(E, F)$  se identifica con la aplicación  $E \times E \rightarrow F$  que asocia al par  $(\mathbf{u}, \mathbf{v}) \in E \times E$  el vector  $[B(\mathbf{u})](\mathbf{v}) \in F$ , y se escribe  $B(\mathbf{u}, \mathbf{v})$  en vez de  $[B(\mathbf{u})](\mathbf{v})$ . Con este convenio  $(\mathbf{u}, \mathbf{v}) \rightarrow B(\mathbf{u}, \mathbf{v})$  es una aplicación bilineal y si  $M$  es la norma de  $B$ , como aplicación lineal continua de  $E$  en  $\mathcal{L}(E, F)$ , es fácil ver que:  $\|B(\mathbf{u}, \mathbf{v})\| \leq M \|\mathbf{u}\| \|\mathbf{v}\|$  para todo  $(\mathbf{u}, \mathbf{v}) \in E \times E$ . En lo que sigue  $d\mathbf{g}(\mathbf{a}) = d(df)(\mathbf{a})$  se considera como una aplicación bilineal  $B : E \times E \rightarrow F$ .

Cuando  $E = \mathbb{R}^n$  sabemos que esta aplicación bilineal es simétrica, pero la demostración dada 6.7, basada en la simetría de las derivadas parciales, no sirve en la situación actual y debemos adaptarla al caso general:

La función de dos variables reales  $\psi(s, t) = \mathbf{f}(\mathbf{a} + s\mathbf{u} + t\mathbf{v})$  está definida y es diferenciable en un cierto entorno de  $(0, 0)$ ,  $U = \{(s, t) : |s| < r, |t| < r\}$ . Claramente

$$D_1\psi(s, t) = d\mathbf{f}(\mathbf{a} + s\mathbf{u} + t\mathbf{v})\mathbf{u}, \quad D_2\psi(s, t) = d\mathbf{f}(\mathbf{a} + s\mathbf{u} + t\mathbf{v})\mathbf{v}$$

y en virtud de la regla de la cadena  $D_1\psi$  y  $D_2\psi$  son diferenciables en  $(0, 0)$ . Teniendo en cuenta que

$$D_2\psi(s, 0) = d\mathbf{f}(\mathbf{a} + s\mathbf{u})\mathbf{v} = D_{\mathbf{v}}\mathbf{f}(\mathbf{a} + s\mathbf{u}), \quad D_1\psi(0, t) = d\mathbf{f}(\mathbf{a} + t\mathbf{v})\mathbf{u} = D_{\mathbf{u}}\mathbf{f}(\mathbf{a} + t\mathbf{v})$$

y usando la definición de derivada parcial se obtiene

$$D_{\mathbf{u}}D_{\mathbf{v}}\mathbf{f}(\mathbf{a}) = D_{12}\psi(0, 0), \quad D_{21}\psi(0, 0) = D_{\mathbf{v}}D_{\mathbf{u}}\mathbf{f}(\mathbf{a})$$

Aplicando el teorema 6.7 a la función  $\psi$  se concluye que  $D_{\mathbf{u}\mathbf{v}}\mathbf{f}(\mathbf{a}) = D_{\mathbf{v}\mathbf{u}}\mathbf{f}(\mathbf{a})$ .

b) Si  $E = \mathbb{R}^n$ , la función  $D_j\mathbf{f}(\mathbf{x})$ , es el resultado de componer  $d\mathbf{f} : V \rightarrow \mathcal{L}(\mathbb{R}^n, F)$  con la evaluación  $\delta_{\mathbf{e}_j} : \mathcal{L}(\mathbb{R}^n, F) \rightarrow F$  (que es lineal y continua), luego, según la regla de la cadena,  $D_j\mathbf{f}(\mathbf{x})$  es diferenciable en  $\mathbf{a}$  para todo  $j \in \{1 \cdots n\}$ .

Recíprocamente, supongamos que todas las derivadas parciales  $D_j\mathbf{f}(\mathbf{x})$ , son diferenciables en  $\mathbf{a}$ . Si consideremos  $F^n$  dotado de la norma

$$\|(\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_n)\|_{\infty} = \max\{\|\mathbf{y}_j\| : 1 \leq j \leq n\}$$

la aplicación  $\mathbf{x} \rightarrow (D_1\mathbf{f}(\mathbf{x}), \cdots, D_n\mathbf{f}(\mathbf{x}))$ , definida en  $V$  con valores en  $F^n$ , es diferenciable en  $\mathbf{a}$  (basta razonar como en 5.18).

Para cada  $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_n) \in F^n$  la aplicación lineal  $T_{\mathbf{y}} : \mathbb{R}^n \rightarrow F$ , definida por  $T_{\mathbf{y}}(\mathbf{x}) = \sum_{i=1}^n x_i \mathbf{y}_i$ , es continua y verifica  $\|T_{\mathbf{y}}(\mathbf{x})\| \leq \|\mathbf{y}\|_{\infty} \|\mathbf{x}\|_1$ , luego  $\|T_{\mathbf{y}}\| \leq \|\mathbf{y}\|_{\infty}$ . Esto significa que la aplicación lineal  $T : F^n \rightarrow \mathcal{L}(\mathbb{R}^n, F)$ ,  $\mathbf{y} \rightarrow T_{\mathbf{y}}$  es continua. Como  $d\mathbf{f} : U \rightarrow \mathcal{L}(\mathbb{R}^n, F)$  es el resultado de componer  $\mathbf{x} \rightarrow (D_1\mathbf{f}(\mathbf{x}), \cdots, D_n\mathbf{f}(\mathbf{x}))$  con la aplicación lineal continua  $T$ , en virtud de la regla de la cadena,  $d\mathbf{f}$  es diferenciable en  $\mathbf{a}$ . ■

### E.3. Teorema de Schwarz sobre la igualdad de las derivadas mixtas

Aunque las derivadas  $D_{11}f(a, b)$ ,  $D_{22}f(a, b)$  no intervienen en la conclusión del teorema de Young y sus corolarios 6.5, 6.6, sin embargo las hipótesis de estos resultados llevan implícita la existencia de estas derivadas. Para el caso de funciones de dos variables hay otras hipótesis, donde no intervienen las derivadas  $D_{11}f$ ,  $D_{22}f$ , que garantizan la igualdad de las derivadas mixtas  $D_{12}\mathbf{f}(\mathbf{a}) = D_{21}\mathbf{f}(\mathbf{a})$ . Uno de ellos, que se obtendrá más adelante como consecuencia del teorema de Green afirma que si en un entorno  $U$  de  $(a, b)$  existen y son continuas las derivadas parciales  $D_1f(x, y)$ ,  $D_2f(x, y)$ ,  $D_{12}f(x, y)$  y  $D_{21}f(x, y)$  entonces  $D_{21}f(x, y) = D_{12}f(x, y)$  para todo  $(x, y) \in U$ . Una mejora sustancial de este resultado y del corolario 6.6 es el siguiente teorema de Schwarz

**Teorema E.3** [Schwarz] *Sea  $\mathbf{f} : \Omega \rightarrow E$  una aplicación definida en un abierto  $\Omega \subset \mathbb{R}^2$  con valores en un espacio normado  $(E, \|\cdot\|)$  tal que en un entorno de  $(a, b) \in \Omega$  existen las derivadas parciales  $D_1\mathbf{f}(x, y)$ ,  $D_2\mathbf{f}(x, y)$ ,  $D_{21}\mathbf{f}(x, y)$ . Si  $D_{21}\mathbf{f}(x, y)$  es continua en  $(a, b)$  entonces existe  $D_{12}\mathbf{f}(a, b)$  y se cumple  $D_{12}\mathbf{f}(a, b) = D_{21}\mathbf{f}(a, b)$ .*

DEM: Basta hacer la demostración cuando  $(a, b) = (0, 0)$  y  $D_{21}\mathbf{f}(0, 0) = \mathbf{0}$  pues el caso general se reduce a este considerando  $\mathbf{g}(x, y) = \mathbf{f}(a + x, b + y) - xyD_{21}\mathbf{f}(a, b)$ . En este caso el objetivo es demostrar que existe y vale  $\mathbf{0}$  el límite

$$D_{12}\mathbf{f}(0, 0) = \lim_{h \rightarrow 0} \frac{D_2\mathbf{f}(h, 0) - D_2\mathbf{f}(0, 0)}{h}$$

Teniendo en cuenta que

$$D_2\mathbf{f}(h, 0) = \lim_{k \rightarrow 0} \frac{\mathbf{f}(h, k) - \mathbf{f}(h, 0)}{k}; \quad D_2\mathbf{f}(0, 0) = \lim_{k \rightarrow 0} \frac{\mathbf{f}(0, k) - \mathbf{f}(0, 0)}{k}$$

todo se reduce a demostrar que existe y vale  $\mathbf{0}$  el límite iterado

$$D_{12}\mathbf{f}(0, 0) = \lim_{h \rightarrow 0} \left( \lim_{k \rightarrow 0} \frac{\Delta(h, k)}{hk} \right)$$

donde

$$\Delta(h, k) = \mathbf{f}(h, k) - \mathbf{f}(h, 0) - \mathbf{f}(0, k) + \mathbf{f}(0, 0)$$

Por hipótesis, para algún  $r > 0$ , las derivadas parciales  $D_1\mathbf{f}$ ,  $D_2\mathbf{f}$ ,  $D_{21}\mathbf{f}$  existen en todo punto de  $B(r) = \{(x, y) : |x| < r, |y| < r\}$ .

Dado  $\epsilon > 0$ , en virtud de la continuidad de  $D_{21}\mathbf{f}$  en  $(0, 0)$ , existe  $\delta \in (0, r)$  tal que

$$|s| < \delta, |t| < \delta \Rightarrow \|D_{21}\mathbf{f}(s, t)\| < \epsilon$$

En lo que sigue se supone  $|h| < \delta$  y  $|k| < \delta$ . Para cada  $s \in (-\delta, \delta)$ , la función  $t \rightarrow D_1\mathbf{f}(s, t)$  es derivables en el intervalo  $(-\delta, \delta)$ , donde su derivada cumple

$\|D_{21}\mathbf{f}(s, t)\| < \epsilon$ . Aplicando a la función  $t \rightarrow D_1\mathbf{f}(s, t)$  el teorema del incremento finito en el intervalo de extremos  $0, k$ , se obtiene

$$|s| < \delta \Rightarrow \|D_1\mathbf{f}(s, k) - D_1\mathbf{f}(s, 0)\| \leq \epsilon|k|$$

Consideremos ahora la función  $s \rightarrow \mathbf{F}(s, k) = \mathbf{f}(s, k) - \mathbf{f}(s, 0)$ , que está definida para  $|s| < r$ . Cuando  $|s| < \delta$  su derivada cumple

$$\|D_1\mathbf{F}(s, k)\| = \|D_1\mathbf{f}(s, k) - D_1\mathbf{f}(s, 0)\| < \epsilon|k|$$

luego, en virtud del teorema del incremento finito  $\|\mathbf{F}(h, k) - \mathbf{F}(0, k)\| \leq \epsilon|h||k|$ . Como  $\mathbf{F}(h, k) - \mathbf{F}(0, k) = \Delta(h, k)$  queda establecido que

$$|h| < \delta, |k| < \delta \Rightarrow \|\Delta(h, k)\| \leq \epsilon|h||k|$$

lo que significa que existe el límite doble

$$\lim_{(h,k) \rightarrow (0,0)} \frac{\Delta(h, k)}{hk} = 0$$

Por otra parte, en virtud de las hipótesis, para cada  $h \in (-r, r)$  existe el límite

$$\lim_{k \rightarrow 0} \frac{\Delta(h, k)}{hk} = \frac{1}{h}(D_2\mathbf{f}(h, 0) - D_2\mathbf{f}(0, 0))$$

luego, según 3.2, (véase también C.4) debe existir el límite iterado

$$\lim_{h \rightarrow 0} \left( \lim_{k \rightarrow 0} \frac{\Delta(h, k)}{hk} \right) = 0$$

lo que significa que existe y vale  $\mathbf{0}$  la derivada

$$D_{12}\mathbf{f}(0, 0) = \lim_{h \rightarrow 0} \frac{1}{h}(D_2\mathbf{f}(h, 0) - D_2\mathbf{f}(0, 0))$$

■



# F

## Funciones convexas

### F.1. Caracterización de las funciones convexas de una variable

En este epígrafe repasamos y ampliamos los resultados básicos sobre las funciones convexas de una variable que se suelen estudiar en el curso de Análisis Matemático I. Como novedad, merece la pena destacar la caracterización integral de las funciones convexas F.5. Esta sencilla caracterización, que es una mejora sustancial de la habitual en el contexto de las funciones derivables F.7, no se suele mencionar en los textos habituales dedicados al cálculo diferencial e integral de las funciones de una variable.

Si  $-\infty \leq a < b \leq +\infty$ , en lo que sigue denotaremos por  $|a, b| \subset \mathbb{R}$  un intervalo genérico de extremos  $a, b$  (un intervalo acotado de la forma  $(a, b)$ ,  $(a, b]$ ,  $[a, b)$ ,  $[a, b]$ , con  $-\infty < a < b < +\infty$ , o un intervalo no acotado de la forma  $(a, +\infty)$ ,  $[a, +\infty)$ ,  $(-\infty, b)$ ,  $(-\infty, b]$ ,  $(-\infty, +\infty)$ ).

Las funciones convexas se suelen definir geoméricamente en la siguiente forma: *La función  $f : |a, b| \rightarrow \mathbb{R}$  es convexa cuando cada par de puntos distintos de su gráfica determinan un segmento que queda por encima de la gráfica.*

Dada una función  $f : |a, b| \rightarrow \mathbb{R}$  y un intervalo  $[u, v] \subset |a, b|$ , denotaremos por

$$m(u, v) = \frac{f(v) - f(u)}{v - u}$$

la pendiente de la recta que pasa por los puntos  $(u, f(u))$ ,  $(v, f(v))$ . Por lo tanto, la ecuación de esta recta se puede escribir en la forma  $r(x) = f(u) + m(u, v)(x - u)$ , y también en la forma  $r(x) = f(v) + m(u, v)(x - v)$ .

La definición geométrica de función convexa admite las formulaciones analíticas que recoge la siguiente proposición:

**Proposición F.1** *Una función  $f : |a, b| \rightarrow \mathbb{R}$  es convexa si y sólo si cumple alguna de las condiciones equivalentes*

[C1]:  $u < x < v \Rightarrow m(u, x) \leq m(u, v)$  en cada  $[u, v] \subset |a, b|$ .

[C2]:  $u < x < v \Rightarrow m(u, v) \leq m(x, v)$  en cada  $[u, v] \subset |a, b|$ .

[C3]:  $u < x < v \Rightarrow m(u, x) \leq m(x, v)$  en cada  $[u, v] \subset |a, b|$ .

[C4]:  $\alpha + \beta = 1, \alpha > 0, \beta > 0 \Rightarrow f(\alpha u + \beta v) \leq \alpha f(u) + \beta f(v)$ .

DEM: Si  $f$  es convexa, para cada  $[u, v] \subset |a, b|$  y cada  $x \in (u, v)$  se cumple que  $f(x) \leq r(x)$  donde  $r(x) = f(u) + m(u, v)(x - u)$  es la función afín cuya gráfica es la recta que pasa por los puntos  $(u, f(u)), (v, f(v))$ . Como  $x - u > 0$ , la desigualdad  $f(x) \leq r(x)$  adopta la forma  $m(u, x) \leq m(u, v)$  luego  $f$  es convexa si y sólo si cumple [C1].

Análogamente, usando la expresión  $r(x) = f(v) + m(u, v)(x - v)$ , como  $x - v < 0$ , se llega a que la desigualdad  $f(x) \leq r(x)$  adopta la forma  $m(u, v) \leq m(x, v)$ , luego  $f$  es convexa si y sólo si cumple [C2]. Si  $f$  es convexa, combinando [C1] con [C2], se obtiene que  $f$  cumple [C3].

Por otra parte, si  $f$  cumple [C3], también cumple [C4]: Sea  $x = \alpha u + \beta v$ , donde  $\alpha, \beta$  son números positivos que cumplen  $\alpha + \beta = 1$ . Según la hipótesis

$$\begin{aligned} 0 \leq m(x, v) - m(u, x) &= \frac{f(v) - f(x)}{v - x} - \frac{f(u) - f(x)}{u - x} = \\ &= \frac{f(v) - f(x)}{\alpha(v - u)} - \frac{f(x) - f(u)}{\beta(v - u)} = \frac{\alpha f(u) + \beta f(v) - f(x)}{\alpha\beta(v - u)} \end{aligned}$$

es decir,  $f(\alpha u + \beta v) = f(x) \leq \alpha f(u) + \beta f(v)$ .

Queda por demostrar que si  $f$  cumple [C4] entonces  $f$  es convexa: Todo  $x \in (u, v)$ , se puede representar en la forma  $x = \alpha u + \beta v$ , donde

$$\alpha = (v - x)/(v - u) > 0, \beta = (x - u)/(v - u) > 0$$

son números positivos que cumplen  $\alpha + \beta = 1$ . Según la hipótesis se cumple la desigualdad  $f(\alpha u + \beta v) \leq \alpha f(u) + \beta f(v)$  que escrita en la forma:

$$f(x) \leq f(u) + \beta(f(v) - f(u)) = f(u) + m(u, v)\beta(v - u) = f(u) + m(u, v)(x - u) = r(x)$$

pone de manifiesto que, para cada  $[u, v] \subset |a, b|$ , la recta que pasa por los puntos  $(u, f(u)), (v, f(v))$ , queda por encima de la gráfica de  $f$ , luego  $f$  es convexa. ■

OBSERVACIÓN: La condición [C4] es equivalente a la que resulta suponiendo  $\alpha \geq 0, \beta \geq 0$ , y  $\alpha + \beta = 1$ , ya que la desigualdad se convierte en una igualdad cuando  $\alpha = 0$  ó  $\beta = 0$ . Por otra parte, en la formulación de [C4] basta suponer  $u \neq v$ , aunque no hay inconveniente en admitir que  $u < v$ .

**Proposición F.2** Sea  $f : |a, b| \rightarrow \mathbb{R}$  una función convexa y  $x \in |a, b|$ .

[P1] Si  $x < b$  la función  $t \rightarrow m(x, t)$  es creciente en el intervalo  $(x, b]$ .

[P2] Si  $a < x$ , la función  $t \rightarrow m(x, t)$  es creciente en el intervalo  $|a, x)$ .

[P3] Si  $[s, t] \subset |a, b|$ , y  $s < x < t$ , se cumple  $m(s, x) \leq m(s, t) \leq m(x, t)$ .

DEM: La propiedad [P1] es consecuencia de [C1], y la propiedad [P2] es consecuencia de [C2] teniendo en cuenta que  $m(x, t) = m(t, x)$ . La propiedad [P3] es consecuencia de las propiedades [P1] y [P2] aplicadas en los intervalos  $(s, b]$ , y  $|a, t)$ , respectivamente (teniendo en cuenta que  $m(s, t) = m(t, s)$  y  $m(t, x) = m(x, t)$ ). ■.

**Proposición F.3** Si  $f : |a, b| \rightarrow \mathbb{R}$  es convexa se cumple:

- a) En cada  $x \in (a, b)$  existen las derivadas laterales,  $f'_i(x) \leq f'_d(x)$ , y las funciones  $f'_i, f'_d$  son crecientes en  $(a, b)$ .
- b) Las rectas tangentes por la izquierda y por la derecha a la gráfica de  $f$  en un punto genérico  $(x, f(x))$ , con  $x \in (a, b)$ , de ecuaciones

$$r_i(t) = f(x) + f'_i(x)(t - x), \quad r_d(t) = f(x) + f'_d(x)(t - x),$$

quedan por debajo de la gráfica de  $f$ , es decir  $r_i(t) \leq f(t)$ , y  $r_d(t) \leq f(t)$  para todo  $t \in (a, b)$ .

DEM: Según F.2, la función  $t \rightarrow m(x, t)$  es creciente en  $(x, b|$  y eligiendo un punto  $a < s < x$  obtenemos que  $m(s, x)$  es una cota inferior de  $\{m(x, t) : x < t < b\}$ , luego existe y es finito el límite

$$f'_d(x) = \lim_{t \rightarrow x^+} m(x, t) = \inf\{m(x, t) : x < t < b\} \geq m(s, x)$$

Como la desigualdad  $m(s, x) \leq f'_d(x)$  es cierta para todo  $s \in (a, x)$  y la función  $s \rightarrow m(x, s) = m(s, x)$  es creciente en  $|a, x)$ , se sigue que existe y es finito el límite

$$f'_i(x) = \lim_{s \rightarrow x^-} m(x, s) = \sup\{m(x, s) : a < s < x\} \leq f'_d(x)$$

Por otra parte, si  $a < x < y < b$ , en virtud de lo que acabamos de establecer,

$$f'_i(x) \leq f'_d(x) \leq m(x, y) = m(y, x) \leq f'_i(y) \leq f'_d(y)$$

y por lo tanto las funciones  $f'_i, f'_d$  son crecientes en  $(a, b)$ .

Las ecuaciones de las rectas tangentes, por la izquierda y por la derecha, a la gráfica de  $f$  en un punto genérico  $(x, f(x))$ , con  $a < x < b$ , se escriben así:

$$r_i(u) = f(x) + f'_i(x)(u - x); \quad r_d(u) = f(x) + f'_d(x)(u - x).$$

Si  $a < s < x < t < b$ , como  $t - x > 0$  y  $s - x < 0$ , las desigualdades

$$f'_d(x) \leq m(x, t) = (f(t) - f(x))/(t - x); \quad f'_i(t) \leq f'_d(t)$$

$$f'_i(x) \geq m(x, s) = (f(s) - f(x))/(s - x); \quad f'_i(s) \leq f'_d(s)$$

permiten afirmar que

$$t \in (x, b) \Rightarrow f(t) \geq f(x) + f'_d(x)(t - x) = r_d(t) \geq r_i(t)$$

$$s \in (a, x) \Rightarrow f(s) \geq f(x) + f'_i(x)(s - x) = r_i(s) \geq r_d(s)$$

y queda establecido que  $r_i(u) \leq f(u)$  y  $r_d(u) \leq f(u)$  para cada  $u \in (a, b)$ . ■

OBSERVACIÓN: En las condiciones de la proposición anterior si  $x$  es uno de los extremos del intervalo sólo se puede garantizar la existencia de la correspondiente

derivada lateral en sentido amplio (con valores en  $[-\infty, +\infty]$ ):

Si  $x = a \in ]a, b[$  no podemos asegurar que  $\{m(a, t) : a < t < b\}$  esté acotado inferiormente, pero debido al carácter creciente de la función  $t \rightarrow m(a, t)$ , existe el límite en la recta real ampliada

$$f'_d(a) = \lim_{t \rightarrow a^+} m(a, t) = \inf\{m(a, t) : a < t < b\} \geq -\infty$$

Análogamente, si  $x = b \in ]a, b[$ , existe la derivada en sentido amplio  $f'_i(b) \leq +\infty$ .

Según esto, haciendo intervenir derivadas laterales con valores infinitos en los extremos del intervalo que le pertenezcan, podemos afirmar que  $f'_i$  es creciente en  $(a, b]$  y  $f'_d$  es creciente en  $]a, b)$ . Además, cuando  $a$  (resp.  $b$ ) pertenece al intervalo  $]a, b[$  y la derivada lateral  $f'_d(a)$  (resp.  $f'_i(b)$ ) es finita, también se cumple que la tangente lateral en  $(a, f(a))$ , (resp.  $(b, f(b))$ ) queda por debajo de la gráfica de  $f$ .

Por consiguiente, si  $f$  es convexa y derivable en  $]a, b[$  (con derivabilidad lateral en los extremos del intervalo  $]a, b[$  que estén en el mismo) se cumple que la derivada  $f'$  es creciente en  $]a, b[$  y para todo  $x \in ]a, b[$  la gráfica de  $f$  queda por encima de su tangente en el punto  $(x, f(x))$ .

**Corolario F.4** *Toda función convexa  $f : ]a, b[ \rightarrow \mathbb{R}$  es continua en  $(a, b) \subset \mathbb{R}$ .*

DEM: La función es continua por la izquierda y por la derecha en cada  $x \in (a, b)$ , porque, según la proposición F.3, existen las derivadas laterales  $f'_i(x), f'_d(x)$  ■

En las condiciones del corolario F.4 no se puede asegurar la continuidad en un extremo del intervalo: Es inmediato que la función  $f : [a, b) \rightarrow \mathbb{R}$ , definida por  $f(a) = 1$  y  $f(t) = 0$  si  $a < t < b$ , es convexa en  $[a, b)$ , pero no es continua en  $a$ .

**Teorema F.5** *Una condición necesaria y suficiente para que  $f : (a, b) \rightarrow \mathbb{R}$  sea convexa es que exista una función creciente  $\varphi : (a, b) \rightarrow \mathbb{R}$  tal que*

$$f(y) - f(x) = \int_x^y \varphi(t) dt \quad \text{para todo intervalo } [x, y] \subset (a, b)$$

DEM: Necesidad: Demostraremos que si  $f$  es convexa la función creciente  $\varphi = f'_d$  cumple la condición del enunciado.

Si  $p = (t_0 < t_1 < \dots < t_n)$  con  $t_0 = x, t_n = y$ , es una subdivisión arbitraria del intervalo  $[x, y] \subset (a, b)$ , como  $f'_d$  es creciente, se tiene

$$s(f'_d, p) = \sum_{j=1}^n f'_d(t_{j-1})(t_j - t_{j-1}); \quad S(f'_d, p) = \sum_{j=1}^n f'_d(t_j)(t_j - t_{j-1});$$

Por otra parte, según se ha visto en la demostración de la proposición F.3

$$f'_d(t_{j-1}) \leq m(t_{j-1}, t_j) \leq f'_d(t_j)$$

Multiplicando por  $(t_j - t_{j-1}) > 0$ , y sumando para  $j = 1 \dots n$ , obtenemos

$$s(f'_d, p) \leq \sum_{j=1}^n m(t_j, t_{j-1})(t_j - t_{j-1}) \leq S(f'_d, p)$$

Teniendo en cuenta que

$$\sum_{j=1}^n m(t_j, t_{j-1})(t_j - t_{j-1}) = \sum_{j=1}^n (f(t_j) - f(t_{j-1})) = f(y) - f(x)$$

llegamos a que, para todo  $p \in \mathcal{P}([x, y])$ , se cumple

$$s(f'_d, p) \leq f(y) - f(x) \leq S(f'_d, p)$$

Como  $f'_d$  es integrable Riemann en  $[x, y]$  (por ser creciente) se sigue que

$$\int_x^y f'_d(t) dt = f(y) - f(x)$$

Queda demostrado que  $\varphi = f'_d$  cumple la condición del enunciado (con un razonamiento similar se puede ver que  $f'_i$  también la cumple).

Suficiencia: Sea  $\varphi : (a, b) \rightarrow \mathbb{R}$  una función creciente tal que para todo  $[x, y] \subset (a, b)$

$$f(y) - f(x) = \int_x^y \varphi(t) dt$$

Si  $s = \alpha x + \beta y$ , con  $\alpha > 0, \beta > 0, \alpha + \beta = 1$ , teniendo en cuenta que el máximo valor de  $\varphi$  en  $[x, s]$  es  $\varphi(s)$  y el mínimo valor de  $\varphi$  en  $[s, y]$  es  $\varphi(s)$ , se obtienen las desigualdades

$$m(x, s) = \frac{1}{s-x} \int_x^s \varphi(t) dt \leq \varphi(s) \leq \frac{1}{y-s} \int_s^y \varphi(t) dt = m(s, y)$$

Queda establecido así que en todo intervalo  $[x, y] \subset (a, b)$  cumple la condición [C3] de la proposición F.1, y por lo tanto  $f$  es convexa. ■

**Corolario F.6** Si  $f : (a, b) \rightarrow \mathbb{R}$  es convexa el conjunto de puntos donde no es derivable,  $\{x \in (a, b) : f'_i(x) < f'_d(x)\}$ , es numerable.

DEM: Es bien conocido que el conjunto de los puntos de discontinuidad  $D(\varphi)$  de una función creciente  $\varphi : (a, b) \rightarrow \mathbb{R}$  es numerable. Si  $\varphi$  es la función creciente que interviene en el teorema F.5, en virtud del teorema fundamental del cálculo podemos asegurar que la función  $f(s) = f(x) + \int_x^s \varphi(t) dt$  es derivable en cada  $s \in (a, b) \setminus D(\varphi)$ . ■

En el siguiente teorema, cuando  $a \in |a, b|$ , (resp.  $b \in |a, b|$ ) la derivabilidad de la función en  $a$  (resp. en  $b$ ) significa derivabilidad por la derecha (resp. por la izquierda) y la tangente a la gráfica en  $(a, f(a))$  (resp.  $(b, f(b))$ ) es la correspondiente tangente lateral.

**Teorema F.7** Para una función derivable  $f : |a, b| \rightarrow \mathbb{R}$  son equivalentes

- a)  $f$  es convexa.
- b) La gráfica de  $f$  queda por encima de su tangente en cualquier punto  $(x, f(x))$ , con  $x \in ]a, b[$ .
- c) La derivada  $f'$  es creciente en  $]a, b[$ .

DEM: a)  $\Rightarrow$  b) es consecuencia directa de la proposición F.3, teniendo en cuenta las observaciones que le siguen.

b  $\Rightarrow$  c): Si  $x \in ]a, b[$ , la ecuación de la recta tangente a la gráfica de  $f$  en el punto  $(x, f(x))$  es  $r(t) = f(x) + f'(x)(t - x)$ , y si suponemos que  $r(t) \leq f(t)$  para todo  $t \in ]a, b[$  se obtiene que

$$f'(x)(t - x) \leq f(t) - f(x)$$

Análogamente, cambiando los papeles de  $t$  y  $x$  podemos escribir

$$f'(t)(x - t) \leq f(x) - f(t)$$

Si  $a \leq x < t < b$  se sigue que  $f'(x)(t - x) \leq f(t) - f(x) \leq f'(t)(t - x)$ , luego  $f'(x) \leq f'(t)$ . Análogamente, si  $a < t < x \leq b$  resulta  $f'(t) \leq f'(x)$ .

c)  $\Rightarrow$  a): Suponemos ahora que la derivada  $f'$  es creciente en  $]a, b[$ . Dado un intervalo  $[u, v] \subset ]a, b[$ , si consideramos la recta que pasa por  $(u, f(u))$ , y  $(v, f(v))$ , de ecuación  $r(x) = f(u) + m(u, v)(x - u)$ , basta demostrar que  $\varphi(x) = r(x) - f(x) \geq 0$  para todo  $x \in [u, v]$ . La función  $\varphi$  es continua y derivable en  $[u, v]$  con derivada decreciente

$$\varphi'(x) = r'(x) - f'(x) = m(u, v) - f'(x)$$

Como  $\varphi(u) = \varphi(v) = 0$ , en virtud del teorema de Rolle, existe  $\eta \in (u, v)$  tal que  $\varphi'(\eta) = 0$ . Como  $\varphi'$  es decreciente podemos afirmar que  $\varphi'(s) \geq 0$  si  $u \leq s \leq \eta$ , y  $\varphi'(t) \leq 0$  si  $\eta \leq t \leq v$ , luego  $\varphi$  es creciente en  $[u, \eta]$  y decreciente en  $[\eta, v]$ , de donde se sigue que  $\varphi(x) \geq 0$  para todo  $x \in [u, v]$ .

■

OBSERVACIÓN: Para un intervalo abierto  $(a, b)$ , la implicación c)  $\Rightarrow$  a) en el teorema F.7 también se puede obtener utilizando la caracterización integral de las funciones convexas (F.5): Si  $f'$  es creciente entonces es integrable en cada intervalo  $[x, y] \subset (a, b)$  y por lo tanto

$$f(y) - f(x) = \int_x^y f'(t) dt \quad \text{para cada } [x, y] \subset (a, b)$$

**Corolario F.8** Una función dos veces derivable  $f : ]a, b[ \rightarrow \mathbb{R}$  es convexa si y sólo si  $f''(t) \geq 0$  para cada  $t \in ]a, b[$ .

DEM: Es consecuencia directa de la equivalencia a)  $\Leftrightarrow$  c) en el teorema F.7. ■

NOTA: Si en la definición de función convexa se cambian las desigualdades  $\leq$  por desigualdades estrictas  $<$  se obtiene la noción de función estrictamente convexa. En este caso las funciones  $t \rightarrow m(x, t)$  son estrictamente crecientes en los intervalos donde están definidas. Para una función derivable  $f : (a, b) \rightarrow \mathbb{R}$  son equivalentes

- a)  $f$  es estrictamente convexa.
- b)  $f'$  es estrictamente creciente.
- c) La gráfica de  $f$  queda estrictamente por encima de su tangente en un punto genérico  $(x, f(x))$  (excepto en dicho punto).

Si  $f : (a, b) \rightarrow \mathbb{R}$  es derivable dos veces en  $(a, b)$  y  $f''(x) > 0$  para todo  $x \in (a, b)$  entonces  $f$  es estrictamente convexa porque su derivada es estrictamente creciente, pero el recíproco no es cierto (la función  $f(x) = x^4$  es estrictamente convexa en  $\mathbb{R}$  pero  $f''(0) = 0$ ).

## F.2. Continuidad de las funciones convexas de varias variables

Al lector interesado le presentamos una demostración de la continuidad de las funciones convexas de varias variables. Este atractivo resultado con un enunciado tan simple no suele figurar en los textos usuales de cálculo para funciones de varias variables reales. Requiere algunos lemas preliminares de carácter técnico.

Si  $E$  es un espacio vectorial (sobre el cuerpo  $\mathbb{R}$ ), para cada par de puntos  $\mathbf{x}, \mathbf{y} \in E$  sea  $[\mathbf{x}, \mathbf{y}] = \{(1-t)\mathbf{x} + t\mathbf{y} : 0 \leq t \leq 1\}$  el segmento que los une. Recordemos que un conjunto  $A \subset E$  es convexo cuando  $[\mathbf{x}, \mathbf{y}] \subset A$  para cada par de puntos  $\mathbf{x}, \mathbf{y} \in A$ . Las siguientes propiedades se obtienen mediante comprobaciones rutinarias que se dejan al cuidado del lector:

- a) La intersección de cualquier familia (finita o no) de conjuntos convexas es un conjunto convexo.
- b) Si  $A, B \subset E$  son convexas y  $\lambda \in \mathbb{R}$  entonces también son convexas

$$A + B = \{\mathbf{a} + \mathbf{b} : \mathbf{a} \in A, \mathbf{b} \in B\}; \quad \lambda A = \{\lambda \mathbf{a} : \mathbf{a} \in A\}.$$

Si  $E$  es un espacio vectorial sobre  $\mathbb{R}$ , la envoltura convexa de  $B \subset E$ , denotada  $\text{co}(B)$  es la intersección de todos los conjuntos convexas que contienen a  $B$ . Según a),  $\text{co}(B)$  es convexo y por lo tanto es el mínimo convexo que contiene a  $B$ .

**Lema F.9** Si  $A \subset E$  es convexo y  $t_1 \geq 0, t_2 \geq 0, \dots, t_m \geq 0$  entonces

$$t_1 A + t_2 A + \dots + t_m A = \left( \sum_{i=1}^m t_i \right) A$$

DEM: Basta demostrar la inclusión  $t_1 A + t_2 A + \dots + t_m A \subset \left( \sum_{i=1}^m t_i \right) A$ , ya que la otra inclusión  $\supset$  es inmediata. Lo haremos por inducción sobre el número de

sumandos. El resultado es trivial para  $m = 1$ . Cuando  $m = 2$  podemos suponer  $t_1 + t_2 > 0$  y se obtiene la inclusión

$$t_1A + t_2A = (t_1 + t_2) \left( \frac{t_1}{t_1 + t_2}A + \frac{t_2}{t_1 + t_2}A \right) \subset (t_1 + t_2)A$$

como consecuencia de la definición de conjunto convexo, ya que

$$\frac{t_1}{t_1 + t_2} + \frac{t_2}{t_1 + t_2} = 1.$$

Si suponemos cierta la inclusión para  $m = k - 1$ , tenemos

$$t_1A + t_2A + \cdots + t_{k-1}A + t_kA \subset \left( \sum_{i=1}^{k-1} t_i \right) A + t_kA \subset \left( \sum_{i=1}^k t_i \right) A$$

donde la primera inclusión se cumple por hipótesis de inducción y la segunda por lo que ha sido demostrado en el caso  $m = 2$ . ■

El resultado que se acaba de establecer en el lema F.9 es falso cuando  $A$  no se supone convexo: Si  $E = \mathbb{R}$  el conjunto  $A = \{0, 1\}$  verifica  $2A = \{0, 2\} \not\subset \{0, 1, 2\} = A + A$ .

Cada expresión de la forma  $\sum_{i=1}^m t_i \mathbf{b}_i$  donde  $t_i \geq 0$  y  $\sum_{i=1}^m t_i = 1$  se dice que es una combinación convexa de los vectores  $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m$ . El siguiente lema se puede enunciar diciendo que la envoltura convexa de  $B \subset E$  está formada por el conjunto de las combinaciones convexas de elementos de  $B$ .

**Lema F.10** *Si  $E$  es un espacio vectorial sobre  $\mathbb{R}$  y  $B \subset E$ , se verifica*

$$co(B) = \left\{ \sum_{i=1}^m t_i \mathbf{b}_i : \mathbf{b}_i \in B, t_i \geq 0, \sum_{i=1}^m t_i = 1, m \in \mathbb{N} \right\}$$

DEM: Según F.9 cada convexo  $A \supset B$  contiene a todas las combinaciones convexas de elementos de  $B$  pues

$$\sum_{i=1}^m t_i \mathbf{b}_i \in t_1A + t_2A + \cdots + t_mA = \left( \sum_{i=1}^m t_i \right) A = A$$

Por lo tanto basta comprobar que el conjunto de las combinaciones convexas de  $B$  es convexo. Dadas dos combinaciones convexas de elementos de  $B$

$$\mathbf{x} = \sum_{i=1}^m t_i \mathbf{b}_i; \quad \mathbf{y} = \sum_{i=1}^m t'_i \mathbf{b}'_i$$

es claro que, para cada  $t \in [0, 1]$ , el vector  $(1 - t)\mathbf{x} + t\mathbf{y}$  también se puede expresar como combinación convexa de elementos de  $B$  ya que

$$(1 - t)\mathbf{x} + t\mathbf{y} = \sum_{i=1}^m (1 - t)t_i \mathbf{b}_i + \sum_{i=1}^m tt'_i \mathbf{b}'_i$$



donde

$$(1-t)t_i \geq 0; \quad tt'_i \geq 0; \quad \sum_{i=1}^m (1-t)t_i + \sum_{i=1}^m tt_i = 1$$

■

**Lema F.11** Para  $r > 0$  sea  $B_n(r) = [-r, r]^n \subset \mathbb{R}^n$  la bola cerrada de centro 0 y radio  $r > 0$  para la norma  $\|\cdot\|_\infty$  de  $\mathbb{R}^n$ . Se cumple que  $B_n(r) = \text{co}(V_n(r))$  donde

$$V_n(r) = \{(v_1, v_2, \dots, v_n) : |v_1| = r, |v_2| = r, \dots, |v_n| = r\}$$

es el conjunto de los vértices de  $B_n(r)$ .

DEM: El resultado es inmediato para  $n = 1$  y para hacer la demostración por inducción sobre la dimensión, lo suponemos cierto en  $\mathbb{R}^{n-1}$ . En lo que sigue cada vector  $\mathbf{x} \in \mathbb{R}^n = \mathbb{R}^{n-1} \times \mathbb{R}$  lo representamos en la forma  $\mathbf{x} = (\mathbf{y}, t)$  donde  $\mathbf{y} \in \mathbb{R}^{n-1}$  y  $t \in \mathbb{R}$ .

Dado  $\mathbf{x} = (\mathbf{y}, t) \in B_n(r) = B_{n-1}(r) \times [-r, r]$ , por la validez del resultado para  $n = 1$  podemos escribir  $t = \alpha(-r) + \beta r$  con  $\alpha \geq 0$ ,  $\beta \geq 0$ , y  $\alpha + \beta = 1$ , luego  $\mathbf{x} = \alpha \mathbf{p} + \beta \mathbf{q}$  donde  $\mathbf{p} = (\mathbf{y}, -r)$ ,  $\mathbf{q} = (\mathbf{y}, r)$ . Según la hipótesis de inducción  $\mathbf{y} \in \mathbb{R}^{n-1}$  se puede expresar como una combinación convexa

$$\mathbf{y} = \sum_{\mathbf{v} \in V_{n-1}(r)} t(\mathbf{v}) \mathbf{v}$$

de vectores de  $V_{n-1}(r)$ .

Con los mismos coeficientes  $t(\mathbf{v})$  podemos expresar  $\mathbf{p} = (\mathbf{y}, -r)$ , y  $\mathbf{q} = (\mathbf{y}, r)$  como combinaciones convexas

$$\mathbf{p} = \sum_{\mathbf{v} \in V_{n-1}(r)} t(\mathbf{v})(\mathbf{v}, -r); \quad \mathbf{q} = \sum_{\mathbf{v} \in V_{n-1}(r)} t(\mathbf{v})(\mathbf{v}, +r)$$

Obsérvese que  $V_n(r) = V_n(r)^+ \cup V_n(r)^-$  donde

$$V_n(r)^+ = \{(\mathbf{v}, +r) : \mathbf{v} \in V_{n-1}(r)\}, \quad V_n(r)^- = \{(\mathbf{v}, -r) : \mathbf{v} \in V_{n-1}(r)\}$$

Si para  $\mathbf{w} = (\mathbf{v}, \pm r) \in V_n(r)$  definimos  $s(\mathbf{w}) = t(\mathbf{v})$ , es claro que

$$\mathbf{x} = \alpha \mathbf{p} + \beta \mathbf{q} = \sum_{\mathbf{w} \in V_n(r)^+} \alpha s(\mathbf{w}) \mathbf{w} + \sum_{\mathbf{w} \in V_n(r)^-} \beta s(\mathbf{w}) \mathbf{w}$$

y así se obtiene  $\mathbf{x}$  como combinación convexa de vectores de  $V_n(r)$ . ■

**Teorema F.12** Toda función convexa  $f : \Omega \rightarrow \mathbb{R}$  definida en un abierto convexo  $\Omega \subset \mathbb{R}^n$  es continua.

DEM: Comencemos con una sencilla observación: Si  $\mathbf{p} = \sum_{j=1}^m t_j \mathbf{p}_j$  es una combinación convexa de puntos  $\mathbf{p}_j \in \Omega$  se demuestra fácilmente (por inducción sobre  $m$ ) que

$$f(\mathbf{p}) \leq \sum_{j=1}^m t_j f(\mathbf{p}_j)$$

Fijado un punto  $\mathbf{a} \in \Omega$ , para demostrar que  $f$  es continua en  $\mathbf{a}$  basta demostrar que es continua en 0 la función convexa  $g(\mathbf{x}) = f(\mathbf{x} + \mathbf{a}) - f(\mathbf{a})$ , que está definida en el abierto convexo  $\Omega_{\mathbf{a}} = -\mathbf{a} + \Omega$ . Puesto que  $0 \in \Omega_{\mathbf{a}}$  existe  $r > 0$  tal que  $B_n(r) = [-r, r]^n \subset \Omega_{\mathbf{a}}$ . Según el lema F.11 cada  $\mathbf{x} \in B_n(r)$  se puede escribir como combinación convexa de los vértices de  $B_n(r)$ :

$$\mathbf{x} = \sum_{\mathbf{v} \in V_n(r)} t(\mathbf{v}) \mathbf{v}$$

Como el conjunto de vértices  $V_n(r)$  es finito (con  $2^n$  elementos) podemos considerar el máximo  $M_r = \max\{g(\mathbf{v}) : \mathbf{v} \in V_n(r)\}$ , y en virtud de la observación preliminar, aplicada a la función  $g$ , resulta

$$g(\mathbf{x}) \leq \sum_{\mathbf{v} \in V_n(r)} t(\mathbf{v}) g(\mathbf{v}) \leq \sum_{\mathbf{v} \in V_n(r)} t(\mathbf{v}) M_r = M_r$$

luego  $M_r$  es una cota superior de  $g$  en  $B_n(r)$ . Si  $0 < \epsilon < 1$ , y  $\mathbf{x}/\epsilon \in B_n(r)$ , en virtud de la convexidad de  $g$  se obtiene

$$g(\mathbf{x}) \leq (1 - \epsilon)g(\mathbf{0}) + \epsilon g(\mathbf{x}/\epsilon) = \epsilon g(\mathbf{x}/\epsilon) \leq \epsilon M_r$$

Usando otra vez la convexidad de  $g$  podemos escribir

$$0 = g(\mathbf{0}) \leq (1/2)g(-\mathbf{x}) + (1/2)g(\mathbf{x})$$

luego,  $-g(\mathbf{x}) \leq g(-\mathbf{x}) \leq \epsilon M_r$  y se concluye que  $\mathbf{x}/\epsilon \in B_n(r) \Rightarrow |g(\mathbf{x})| \leq \epsilon M_r$  es decir

$$\|\mathbf{x}\|_{\infty} \leq \epsilon r \Rightarrow |g(\mathbf{x})| \leq \epsilon M_r$$

y así queda demostrado que  $g$  es continua en  $\mathbf{0}$ . ■

# G

## Funciones analíticas

### G.1. Funciones analíticas

En el caso de funciones de una sola variable es bien conocido que hay funciones de clase  $C^\infty$  cuya serie de Taylor en un cierto punto no converge hacia la función en ningún entorno del punto. Un ejemplo típico lo proporciona la función  $f(x) = e^{-1/x^2}$  si  $x \neq 0$ ,  $f(0) = 0$ , cuyas derivadas sucesivas en  $x = 0$  son todas nulas, con lo cual todos los polinomios de Taylor de  $f$  en  $x = 0$  son idénticamente nulos. Las funciones de clase  $C^\infty$  que no presentan esta patología se llaman analíticas.

La definición de función analítica de varias variables reales requiere la consideración de series de potencias en varias variables reales.

**Series de potencias.** Una serie de potencias de  $n$  variables reales  $(x_1, x_2, \dots, x_n)$ , centrada en  $\mathbf{a} = (a_1, a_2, \dots, a_n) \in \mathbb{R}^n$ , con coeficientes en un espacio normado completo  $(F, \|\cdot\|)$ , es una serie de la forma

$$\sum_{k=0}^{\infty} \mathbf{A}_k(\mathbf{x} - \mathbf{a}) = \sum_{k=0}^{\infty} \left( \sum_{|\mathbf{p}|=k} \mathbf{a}_{\mathbf{p}}(\mathbf{x} - \mathbf{a})^{\mathbf{p}} \right)$$

donde el término  $\mathbf{A}_k(\mathbf{x} - \mathbf{a}) = \sum_{|\mathbf{p}|=k} \mathbf{a}_{\mathbf{p}}(\mathbf{x} - \mathbf{a})^{\mathbf{p}}$  es un polinomio homogéneo de grado  $k$  en la variable  $\mathbf{h} = \mathbf{x} - \mathbf{a} = (x_1 - a_1, x_2 - a_2, \dots, x_n - a_n)$ , con coeficientes  $\mathbf{a}_{\mathbf{p}} \in F$ . Se suele escribir, más brevemente, en la forma  $\sum_{\mathbf{p}} \mathbf{a}_{\mathbf{p}}(\mathbf{x} - \mathbf{a})^{\mathbf{p}}$ .

En lo que sigue, para simplificar la escritura, supondremos frecuentemente que la serie de potencias están centradas en  $\mathbf{a} = \mathbf{0}$ . Esto no es restrictivo ya que, con el cambio de variable  $\mathbf{h} = \mathbf{x} - \mathbf{a}$ , la serie de potencias  $\sum_{\mathbf{p}} \mathbf{a}_{\mathbf{p}}(\mathbf{x} - \mathbf{a})^{\mathbf{p}}$  se transforma en una serie de potencias  $\sum_{\mathbf{p}} \mathbf{a}_{\mathbf{p}}\mathbf{h}^{\mathbf{p}}$  centrada en  $\mathbf{0}$ . Si  $\mathbf{r} = (r_1, r_2, \dots, r_n)$  con  $r_j > 0$  para  $1 \leq j \leq n$ , introducimos las notaciones  $B(\mathbf{r}) = \{\mathbf{x} \in \mathbb{R}^n : |x_k| < r_k, 1 \leq k \leq n\}$ ;  $K(\mathbf{r}) = \{\mathbf{x} \in \mathbb{R}^n : |x_k| \leq r_k, 1 \leq k \leq n\}$  para denotar el bloque abierto y el bloque cerrado de centro  $\mathbf{0}$  y lados  $2r_j$ .

Por otra parte, dado  $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ , el vector  $(|x_1|, |x_2|, \dots, |x_n|)$  lo designaremos con la notación abreviada  $|\mathbf{x}|$ . De acuerdo con esta notación, si  $\mathbf{p} = (p_1, p_2, \dots, p_n)$  se tiene,  $|\mathbf{x}|^{\mathbf{p}} = |x_1|^{p_1} |x_2|^{p_2} \dots |x_n|^{p_n} = |x_1^{p_1} x_2^{p_2} \dots x_n^{p_n}| = |\mathbf{x}^{\mathbf{p}}|$ .

Si la serie  $\sum_{k=0}^{\infty} \sum_{|\mathbf{p}|=k} \|\mathbf{a}_{\mathbf{p}}\| |\mathbf{x}|^{\mathbf{p}}$  es convergente se dice que la serie de potencias  $\sum_{\mathbf{p}} \mathbf{a}_{\mathbf{p}} \mathbf{x}^{\mathbf{p}}$  es absolutamente convergente en el punto  $\mathbf{x}$ . En este caso, como el espacio normado  $(F, \|\cdot\|)$  es completo, la serie  $\sum_{k=0}^{\infty} \mathbf{A}_k(\mathbf{x})$  es convergente: En efecto, aplicando la desigualdad triangular a cada suma finita  $\mathbf{A}_k(\mathbf{x}) = \sum_{|\mathbf{p}|=k} \mathbf{a}_{\mathbf{p}} \mathbf{x}^{\mathbf{p}}$  resulta  $\|\mathbf{A}_k(\mathbf{x})\| \leq \alpha_k(\mathbf{x})$  donde  $\alpha_k(\mathbf{x}) = \sum_{|\mathbf{p}|=k} \|\mathbf{a}_{\mathbf{p}}\| |\mathbf{x}|^{\mathbf{p}}$  es el término general de una serie convergente. Por el criterio de comparación  $\sum_{k=0}^{\infty} \|\mathbf{A}_k(\mathbf{x})\| < +\infty$ , y como  $(F, \|\cdot\|)$  es completo se concluye que la serie  $\sum_{k=0}^{\infty} \mathbf{A}_k(\mathbf{x})$  converge.

Más aún, si la serie converge absolutamente en un punto  $\mathbf{r} = (r_1, r_2, \dots, r_n)$  con  $r_j > 0$  para  $1 \leq j \leq n$ , entonces también converge absolutamente en cada punto del bloque compacto  $K(\mathbf{r})$ . Además la serie  $\sum_{k=0}^{\infty} \mathbf{A}_k(\mathbf{x})$  converge uniformemente sobre  $K(\mathbf{r})$  pues razonando como antes es claro que para todo  $\mathbf{x} \in K(\mathbf{r})$  se cumple  $\|\mathbf{A}_k(\mathbf{x})\| \leq \alpha_k(\mathbf{r})$  y aplicando el criterio de Weierstrass C.8 se obtiene el resultado.

Volvemos a insistir, para el lector que desee situarse en una situación más concreta, no hay inconveniente en suponer  $F = \mathbb{R}$ . Sin embargo, en este caso particular apenas se simplifica el asunto pues los resultados y razonamientos que siguen son esencialmente los mismos que intervienen en el caso de funciones con valores reales.

**Definición G.1** Una función  $\mathbf{f} : \Omega \rightarrow \mathbb{R}$ , definida en un abierto  $\Omega \subset \mathbb{R}^n$ , con valores en un espacio normado completo  $(F, \|\cdot\|)$ , se dice que es analítica en  $\mathbf{a} \in \Omega$  si existe una bola  $B(\mathbf{a}, r) \subset \Omega$  donde  $\mathbf{f}$  se puede representar mediante una serie de potencias absolutamente convergente, centrada en  $\mathbf{a}$ :

$$\mathbf{f}(\mathbf{x}) = \sum_{k=0}^{\infty} \sum_{|\mathbf{p}|=k} \mathbf{a}_{\mathbf{p}} (\mathbf{x} - \mathbf{a})^{\mathbf{p}} \text{ para todo } \mathbf{x} \in B(\mathbf{a}, r)$$

Si  $\mathbf{f}$  es analítica en cada  $\mathbf{a} \in \Omega$  se dice que es analítica en  $\Omega$ .

**Teorema G.2** Sea  $\mathbf{f}(\mathbf{x}) = \sum_{\mathbf{p}} \mathbf{a}_{\mathbf{p}} \mathbf{x}^{\mathbf{p}}$ , una serie de potencias, con coeficientes  $\mathbf{a}_{\mathbf{p}}$  en un espacio normado completo, que converge absolutamente en un punto  $\mathbf{r} = (r_1, r_2, \dots, r_n)$  con todas las coordenadas positivas. Entonces  $\mathbf{f}$  es de clase  $C^{\infty}$  en el bloque abierto  $B(\mathbf{r})$  y sus derivadas parciales sucesivas admiten desarrollos en serie de potencias que se obtienen derivando término a término la serie dada, y estas series de potencias siguen siendo absolutamente convergentes en  $\Omega_{\mathbf{z}}$ .

DEM: Para las derivadas primeras, con el fin de simplificar la escritura, consideramos el caso de la derivación respecto a la variable  $x_1$ . Para las derivadas segundas el resultado se obtendrá repitiendo el proceso con las series obtenidas para las derivadas primeras y así sucesivamente. Como pretendemos derivar respecto a la variable  $x_1$ , es conveniente escribir cada  $\mathbf{x} \in B(\mathbf{r})$  en la forma  $\mathbf{x} = (x_1, \mathbf{y})$ , donde  $\mathbf{y} = (x_2, x_3, \dots, x_n)$ , y cada multi-índice  $\mathbf{p} = (p_1, p_2, \dots, p_n)$  en la forma  $\mathbf{p} = (p_1, \mathbf{q})$  donde  $\mathbf{q} = (p_2, p_3, \dots, p_n)$ , de modo que  $\mathbf{x}^{\mathbf{p}} = x_1^{p_1} \mathbf{y}^{\mathbf{q}}$ , con  $k = p_1$ .

Si  $\mathbf{x} \in B(\mathbf{r})$ , la serie absolutamente convergente  $\sum_{\mathbf{p}} \mathbf{a}_{\mathbf{p}} \mathbf{x}^{\mathbf{p}}$  se puede sumar por paquetes organizados según las potencias de  $x_1$ :

$$\sum_{\mathbf{p}} \mathbf{a}_{\mathbf{p}} \mathbf{x}^{\mathbf{p}} = \sum_{(k, \mathbf{q})} \mathbf{a}_{(k, \mathbf{q})} x_1^k \mathbf{y}^{\mathbf{q}} = \sum_{k=0}^{\infty} \left( \sum_{\mathbf{q}} \mathbf{a}_{(k, \mathbf{q})} \mathbf{y}^{\mathbf{q}} \right) x_1^k = \sum_{k=0}^{\infty} \varphi_k(\mathbf{y}) x_1^k$$

donde las series  $\varphi_k(\mathbf{y}) = \sum_{\mathbf{q}} \mathbf{a}_{(k,\mathbf{q})} \mathbf{y}^{\mathbf{q}}$  siguen siendo absolutamente convergentes. Cuando se deriva en el punto  $\mathbf{x}$  respecto  $x_1$  el vector  $\mathbf{y} = (x_2, \dots, x_n) \in \mathbb{R}^{n-1}$  permanece fijo, y la serie

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(x_1, \mathbf{y}) = \sum_{k=0}^{\infty} \varphi_k(\mathbf{y}) x_1^k$$

hay que considerarla como una serie de potencias en la variable  $x_1$  que converge absolutamente en  $I_1 = \{x_1 : |x_1| < r_1\}$ , donde se puede derivar término a término.

Para cada  $\mathbf{p} = (k, \mathbf{q})$  con  $k \geq 1$  sea  $\mathbf{p}' = (k-1, \mathbf{q})$ . Entonces  $x_1^{k-1} \mathbf{y}^{\mathbf{q}} = \mathbf{x}^{\mathbf{p}'}$ , y la derivada de cada término de la serie adopta la forma

$$k x_1^{k-1} \varphi_k(\mathbf{y}) = \sum_{\mathbf{p}_1=k} k \mathbf{a}_{\mathbf{p}} \mathbf{x}^{\mathbf{p}'} = \sum_{\mathbf{p}_1=k} D_1(\mathbf{a}_{\mathbf{p}} \mathbf{x}^{\mathbf{p}})$$

luego

$$D_1 \mathbf{f}(\mathbf{x}) = \sum_{k=1}^{\infty} k x_1^{k-1} \varphi_k(\mathbf{y}) = \sum_{k=1}^{\infty} \sum_{\mathbf{p}_1=k} \mathbf{a}_{\mathbf{p}} D_1(\mathbf{x}^{\mathbf{p}}) = \sum_{\mathbf{p}} \mathbf{a}_{\mathbf{p}} D_1(\mathbf{x}^{\mathbf{p}})$$

Pasa obtener la última igualdad basta observar que la serie iterada

$$\sum_{k=1}^{\infty} \sum_{\mathbf{p}_1=k} \mathbf{a}_{\mathbf{p}} D_1(\mathbf{x}^{\mathbf{p}})$$

se obtiene formando paquetes en la serie  $\sum_{\mathbf{p}} \mathbf{a}_{\mathbf{p}} D_1(\mathbf{x}^{\mathbf{p}})$ , y para justificar esta suma-ción por paquetes debemos demostrar que

$$\sum_{\mathbf{p}} \|\mathbf{a}_{\mathbf{p}}\| |D_1(\mathbf{x}^{\mathbf{p}})| < +\infty$$

Para ello consideramos la serie  $g(\mathbf{s}) = \sum_{\mathbf{p}} \|\mathbf{a}_{\mathbf{p}}\| \mathbf{s}^{\mathbf{p}}$ , en un punto  $\mathbf{s} = (s_1, s_2, \dots, s_n)$  tal que  $|x_j| < s_j < r_j$ . Como todos los términos de esta serie son positivos lo mismo le ocurre a la serie que se obtiene derivando cada término respecto a la variable  $s_1$ , lo que justifica la igualdad

$$\sum_{\mathbf{p}} \|\mathbf{a}_{\mathbf{p}}\| D_1(\mathbf{s}^{\mathbf{p}}) = \sum_{k=1}^{\infty} \sum_{\mathbf{p}_1=k} p_1 \|\mathbf{a}_{\mathbf{p}}\| \mathbf{s}^{\mathbf{p}'} < +\infty$$

donde la suma de la derecha es finita (porque su suma es  $D_1 g(\mathbf{s})$ , en virtud del mismo razonamiento empleado al iniciar el cálculo de la derivada  $D_1 \mathbf{f}(\mathbf{x})$ ). Teniendo en cuenta que  $|x_j| < s_j < r_j$ ,  $1 \leq j \leq n$ , se obtiene la desigualdad

$$\sum_{\mathbf{p}} \|\mathbf{a}_{\mathbf{p}}\| |D_1(\mathbf{x}^{\mathbf{p}})| \leq \sum_{\mathbf{p}} \|\mathbf{a}_{\mathbf{p}}\| D_1(\mathbf{s}^{\mathbf{p}}) < +\infty$$

Así queda demostrado que en cada  $\mathbf{x} \in B(\mathbf{r})$  existe la derivada parcial  $D_1 \mathbf{f}(\mathbf{x})$  que coincide con la suma de la serie derivada

$$D_1 \mathbf{f}(\mathbf{x}) = \sum_{\mathbf{p}} D_1(\mathbf{a}_{\mathbf{p}} \mathbf{x}^{\mathbf{p}})$$

Para las derivadas parciales de orden superior observemos que dado un multi-índice de derivación  $\mathbf{q} = (q_1, q_2, \dots, q_n)$ , el operador  $D^{\mathbf{q}}$  sólo produce resultado no nulo en los términos  $\mathbf{x}^{\mathbf{p}}$  con  $\mathbf{p} \geq \mathbf{q}$  (e.d.  $p_j \geq q_j$ , para  $1 \leq j \leq n$ ): Como

$$\frac{\partial(x_1^{p_1} x_2^{p_2} \cdots x_n^{p_n})}{\partial x_1^{q_1} \partial x_2^{q_2} \cdots \partial x_n^{q_n}} = \frac{p_1! x_1^{p_1 - q_1}}{(p_1 - q_1)!} \frac{p_2! x_2^{p_2 - q_2}}{(p_2 - q_2)!} \cdots \frac{p_n! x_n^{p_n - q_n}}{(p_n - q_n)!}$$

el resultado  $D^{\mathbf{q}}\mathbf{x}^{\mathbf{p}}$  adopta la forma

$$D^{\mathbf{q}}\mathbf{x}^{\mathbf{p}} = \frac{\mathbf{p}!}{(\mathbf{p} - \mathbf{q})!} \mathbf{x}^{\mathbf{p} - \mathbf{q}}, \quad \text{luego} \quad D^{\mathbf{q}}\mathbf{f}(\mathbf{x}) = \sum_{\mathbf{p} \geq \mathbf{q}} \frac{\mathbf{p}!}{(\mathbf{p} - \mathbf{q})!} \mathbf{a}_{\mathbf{p}} \mathbf{x}^{\mathbf{p} - \mathbf{q}}$$

■

Una consecuencia directa del teorema G.2 es que toda función analítica es de clase  $C^\infty$  y sus derivadas parciales sucesivas siguen siendo analíticas.

### Ejemplo G.3

Si  $|x_1| < 1$  y  $|x_2| < 1$ , efectuando el producto de convolución de las dos series geométricas absolutamente convergentes

$$\frac{1}{1 - x_1} = 1 + x_1 + x_1^2 + \cdots + x_1^k + \cdots$$

$$\frac{1}{1 - x_2} = 1 + x_2 + x_2^2 + \cdots + x_2^k + \cdots$$

se obtiene un desarrollo en serie de potencias de la función de dos variables reales

$$\frac{1}{(1 - x_1)(1 - x_2)} = \sum_{k=1}^{\infty} \left[ \sum_{i+j=k} x_1^i x_2^j \right]$$

válido en el cuadrado  $U = \{(x_1, x_2) : |x_1| < 1, |x_2| < 1\}$ . En este caso  $A_0(x_1, x_2) = 1$ ,  $A_1(x_1, x_2) = x_1 + x_2$ ,  $A_2(x_1, x_2) = x_1^2 + x_1 x_2 + x_2^2$ ,  $A_3(x_1, x_2) = x_1^3 + x_1^2 x_2 + x_1 x_2^2 + x_2^3$ ; etc. Con la notación abreviada la última igualdad se escribe en la forma

$$\frac{1}{(1 - x_1)(1 - x_2)} = \sum_{\mathbf{p}} \mathbf{x}^{\mathbf{p}}$$

donde  $\mathbf{p}$  recorre los índices de la forma  $\mathbf{p} = (p_1, p_2)$  con  $p_1, p_2 \in \{0, 1, 2, \dots\}$ .

Análogamente, si  $|x_k| < 1$  para  $1 \leq k \leq n$ , resulta el desarrollo en serie de potencias de  $n$  variables

$$\frac{1}{(1 - x_1)(1 - x_2) \cdots (1 - x_n)} = \sum_{\mathbf{p}} \mathbf{x}^{\mathbf{p}} = \sum_{k=0}^{\infty} A_k(\mathbf{x})$$

Ahora  $\mathbf{x} = (x_1, x_2, \dots, x_n)$ ,  $\mathbf{p}$  recorre los índices de la forma  $\mathbf{p} = (p_1, p_2, \dots, p_n)$  con  $p_1, p_2, \dots, p_n \in \{0, 1, 2, \dots\}$  y  $A_k(\mathbf{x}) = \sum_{|\mathbf{p}|=k} x_1^{p_1} x_2^{p_2} \cdots x_n^{p_n}$ .

Con un cálculo rutinario se obtiene la derivada  $D^{\mathbf{q}}f(\mathbf{x})$ , de la función

$$f(\mathbf{x}) = \frac{1}{(1-x_1)(1-x_2)\cdots(1-x_n)}$$

$$D^{\mathbf{q}}f(\mathbf{x}) = \frac{q_1!}{(1-x_1)^{1+q_1}} \frac{q_2!}{(1-x_2)^{1+q_2}} \cdots \frac{q_n!}{(1-x_n)^{1+q_n}}$$

Si  $0 < t < 1$  el valor de esta derivada en el punto  $\mathbf{t} = (t, t, \dots, t)$  viene dado por

$$D^{\mathbf{q}}f(\mathbf{t}) = \frac{\mathbf{q}!}{(1-t)^{n+|\mathbf{q}|}}$$

Por otra parte, derivando la serie de potencias, según la regla obtenida anteriormente, y sustituyendo luego  $\mathbf{x} = \mathbf{t}$ , resulta

$$D^{\mathbf{q}}f(\mathbf{t}) = \sum_{\mathbf{p} \geq \mathbf{q}} \frac{\mathbf{p}!}{(\mathbf{p}-\mathbf{q})!} t^{|\mathbf{p}-\mathbf{q}|}$$

Se obtiene así la siguiente igualdad que será utilizada en el teorema G.4

$$\sum_{\mathbf{p} \geq \mathbf{q}} \frac{\mathbf{p}!}{(\mathbf{p}-\mathbf{q})!} t^{|\mathbf{p}-\mathbf{q}|} = \frac{\mathbf{q}!}{(1-t)^{n+|\mathbf{q}|}} \quad \text{si } 0 < t < 1.$$

El siguiente teorema proporciona una condición, bastante útil en la práctica, para justificar que una función concreta es analítica.

**Teorema G.4** Si  $\mathbf{f} : \Omega \rightarrow \mathbf{F}$  es de clase  $C^\infty$  en un abierto  $\Omega \subset \mathbb{R}^n$ , con valores en un espacio completo  $(F, \|\cdot\|)$ , son equivalentes:

- a)  $\mathbf{f}$  es analítica en  $\Omega$ .
- b) Para cada compacto  $K \subset \Omega$  existen constantes  $M > 0$  y  $r > 0$  tales que

$$\mathbf{x} \in K, \quad |\mathbf{p}| = k \Rightarrow |D^{\mathbf{p}}\mathbf{f}(\mathbf{x})| \leq Mk!R^k$$

DEM: b)  $\Rightarrow$  a): Para cada  $\mathbf{a} \in \Omega$  aplicamos la hipótesis b) a una bola compacta  $K = \overline{B(\mathbf{a}, \delta)} \subset \Omega$ , y obtenemos que se cumple la condición que interviene en el teorema 7.15, luego hay una bola  $B(\mathbf{a}, \rho) \subset \Omega$  donde  $\mathbf{f}|_{B(\mathbf{a}, \rho)}$  se puede representar mediante la suma de su serie de Taylor en  $\mathbf{a}$ .

a)  $\Rightarrow$  b): Cada  $\mathbf{a} \in \Omega$  posee un entorno  $B_\infty(\mathbf{a}, \delta) \subset \Omega$  donde  $\mathbf{f}$  se puede representar mediante una serie de potencias absolutamente convergente

$$\mathbf{f}(\mathbf{a} + \mathbf{h}) = \sum_{\mathbf{p}} \mathbf{a}_{\mathbf{p}} \mathbf{h}^{\mathbf{p}} \quad \text{si } \|\mathbf{h}\|_\infty < \delta$$

Sea  $\mathbf{r} = (r_1, r_2, \dots, r_n)$  un punto, con  $\eta = \min\{r_j : 1 \leq j \leq n\} > 0$ , tal que  $\sum_{\mathbf{p}} \mathbf{a}_{\mathbf{p}} \mathbf{r}^{\mathbf{p}}$  es absolutamente convergente. En este punto todos los términos de la serie están acotados y podemos considerar el supremo

$$C = \sup_{\mathbf{p}} \{\|\mathbf{a}_{\mathbf{p}}\| \mathbf{r}^{\mathbf{p}}\}$$

Sea  $0 < t < 1$  tal que  $\rho = t\eta < \delta$ . Si  $\|\mathbf{h}\|_\infty < \rho$ , para todo  $j \in \{1, 2, \dots, n\}$  se cumple  $|h_j| \leq tr_j$ . Con el desarrollo en serie de  $D^{\mathbf{q}}\mathbf{f}(\mathbf{a} + \mathbf{h})$  obtenido inmediatamente antes del ejemplo G.3 se obtiene

$$\|D^{\mathbf{q}}\mathbf{f}(\mathbf{a} + \mathbf{h})\| \leq \sum_{\mathbf{p} \geq \mathbf{q}} \|\mathbf{a}_{\mathbf{p}}\| \frac{\mathbf{p}!}{(\mathbf{p} - \mathbf{q})!} t^{|\mathbf{p}-\mathbf{q}|} \mathbf{r}^{\mathbf{p}-\mathbf{q}} \leq \frac{C}{\mathbf{r}^{\mathbf{q}}} \sum_{\mathbf{p} \geq \mathbf{q}} \frac{\mathbf{p}!}{(\mathbf{p} - \mathbf{q})!} t^{|\mathbf{p}-\mathbf{q}|}$$

Usando la igualdad

$$\sum_{\mathbf{p} \geq \mathbf{q}} \frac{\mathbf{p}!}{(\mathbf{p} - \mathbf{q})!} t^{|\mathbf{p}-\mathbf{q}|} = \frac{\mathbf{q}!}{(1-t)^{n+|\mathbf{q}|}}$$

establecida antes de la definición G.1, resulta

$$\|D^{\mathbf{q}}\mathbf{f}(\mathbf{a} + \mathbf{h})\| \leq \frac{C}{(1-t)^n} \frac{\mathbf{q}!}{\mathbf{r}^{\mathbf{q}}(1-t)^{|\mathbf{q}|}} \leq \frac{C}{(1-t)^n} \frac{\mathbf{q}!}{\eta^{|\mathbf{q}|}(1-t)^{|\mathbf{q}|}}$$

Tomando  $M = C/(1-t)^n$ , y  $1/R = \eta(1-t)$ , para todo índice  $\mathbf{q}$  y todo  $\mathbf{x} \in B(\mathbf{a}, \rho)$ , se cumple  $\|D^{\mathbf{q}}\mathbf{f}(\mathbf{x})\| \leq MR^k k!$ , con  $k = |\mathbf{q}|$ , (pues  $k!/|\mathbf{q}|!$  es un entero  $\geq 1$ ).

Finalmente, si  $K \subset \Omega$  es compacto, por el razonamiento anterior, para cada  $\mathbf{a} \in K$  hay una bola  $B(\mathbf{a}, \rho(\mathbf{a})) \subset \Omega$ , y constantes  $M(\mathbf{a}) > 0$ ,  $R(\mathbf{a}) > 0$ , tales que para todo  $\mathbf{x} \in B(\mathbf{a}, \rho(\mathbf{a}))$  y todo índice  $\mathbf{q}$  se cumple

$$\|D^{\mathbf{q}}\mathbf{f}(\mathbf{x})\| \leq M(\mathbf{a})R(\mathbf{a})^k k!, \quad \text{donde } k = |\mathbf{q}|.$$

Con un número finito de estas bolas  $B(\mathbf{a}_j, \rho(\mathbf{a}_j))$ ,  $1 \leq j \leq m$ , se recubre el compacto  $K$  y las constantes

$$M = \max\{M(\mathbf{a}_j) : 1 \leq j \leq m\}, \quad R = \max\{R(\mathbf{a}_j) : 1 \leq j \leq m\}$$

cumplen la condición b) del enunciado. ■

### Otros resultados.

- En las condiciones del teorema G.2 la función  $\mathbf{f}(\mathbf{x}) = \sum_{\mathbf{p}} \mathbf{a}_{\mathbf{p}} \mathbf{x}^{\mathbf{p}}$ , definida en un bloque  $B(\mathbf{r}) = \{\mathbf{x} \in \mathbb{R}^n : |x_j| < r_j\}$  por una serie de potencias es analítica. Más aún, para cada  $\mathbf{a} \in B(\mathbf{r})$  la función  $\mathbf{f}$  se puede desarrollar en serie de potencias en

$$B(\mathbf{a}, \mathbf{s}) = \{\mathbf{x} \in \mathbb{R}^n : |x_j - a_j| < s_j\}, \quad \text{donde } 0 < s_j = r_j - |a_j|, \quad 1 \leq j \leq n$$

- La función  $\mathbf{f} : \Omega \rightarrow \mathbb{R}^m$  es analítica en un abierto  $\Omega \subset \mathbb{R}^n$  si y sólo si cada componente  $f_j : \Omega \rightarrow \mathbb{R}$  lo es. En este caso, si  $\mathbf{g} : V \rightarrow F$  es analítica en un abierto  $V \subset \mathbb{R}^m$  y  $\mathbf{f}(\Omega) \subset V$ , la función compuesta  $\mathbf{g} \circ \mathbf{f}$  también es analítica en  $\Omega$ .

- *Principio de prolongación analítica:* Si  $\Omega \subset \mathbb{R}^n$  es un abierto conexo y  $\mathbf{f}, \mathbf{g} : \Omega \rightarrow \mathbb{R}$  son funciones analíticas que coinciden en un abierto no vacío  $U \subset \Omega$ , entonces  $\mathbf{f} = \mathbf{g}$ .

- En general, dada una función analítica  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , puede ocurrir que  $f$  no admita una representación global mediante una serie de potencias convergente en todo  $\mathbb{R}^n$ : Ya en el caso  $n = 1$  hay funciones como  $f(x) = 1/(1+x^2)$ , que son analíticas en todo  $\mathbb{R}$  y sin embargo su desarrollo en serie de potencias alrededor de un punto  $a \in \mathbb{R}$  nunca converge en todo  $\mathbb{R}$  (esta afirmación resultará evidente para el lector que conozca la teoría de las funciones analíticas de variable compleja).



# H

## Dependencia funcional. Subvariedades diferenciables

### H.1. Dependencia e independencia funcional

**Definición H.1** Sean  $f, f_1, \dots, f_p : \Omega \rightarrow \mathbb{R}$  funciones de clase  $C^k(\Omega)$  definidas en un abierto  $\Omega \subset \mathbb{R}^n$ , donde  $1 \leq p \leq n$ , y  $k \geq 1$ . Se dice que  $f$  depende funcionalmente de  $f_1, \dots, f_p$  en el punto  $\mathbf{a} \in \Omega$  si existen un entorno abierto de  $\mathbf{a}$ ,  $V \subset \Omega$ , un entorno abierto de  $\mathbf{b} = (f_1(\mathbf{a}), \dots, f_p(\mathbf{a}))$ ,  $U \subset \mathbb{R}^p$  y una función  $F : U \rightarrow \mathbb{R}$ , de clase  $C^k(U)$ , tales que  $U = \{(f_1(\mathbf{x}), \dots, f_p(\mathbf{x})) : \mathbf{x} \in V\}$  y

$$f(\mathbf{x}) = F(f_1(\mathbf{x}), \dots, f_p(\mathbf{x})), \text{ para todo } \mathbf{x} \in V.$$

**Teorema H.2** Sean  $f, f_1, \dots, f_p : \Omega \rightarrow \mathbb{R}$  funciones de clase  $C^k(\Omega)$  definidas en un abierto  $\Omega \subset \mathbb{R}^n$ , donde  $1 \leq p \leq n$ , y  $k \geq 1$ , y sea  $\mathbf{a} \in \Omega$  un punto tal que las formas lineales  $df_1(\mathbf{a}), \dots, df_p(\mathbf{a})$  son linealmente independientes.

Si las formas lineales  $df(\mathbf{x}), df_1(\mathbf{x}), \dots, df_p(\mathbf{x})$  son linealmente dependientes en todos los puntos  $\mathbf{x}$  de algún entorno  $\Omega_{\mathbf{a}}$  de  $\mathbf{a}$ , entonces  $f$  depende funcionalmente de  $f_1, \dots, f_p$  en el punto  $\mathbf{a}$ .

DEM: Puesto que los vectores  $\nabla f_1(\mathbf{a}), \dots, \nabla f_p(\mathbf{a})$  son linealmente independientes, la matriz cuyas filas son estos vectores es de rango  $p$ , y reordenando las variables si es necesario, podemos suponer que no es nulo el determinante

$$\frac{\partial(f_1, \dots, f_p)}{\partial(x_1, \dots, x_p)}(\mathbf{a}) \neq 0$$

Sea  $\mathbf{b} = (f_1(\mathbf{a}), \dots, f_p(\mathbf{a}))$ . Consideremos la función  $\Phi : \Omega \rightarrow \mathbb{R}^n$  definida por:

$$\Phi(x_1, \dots, x_n) = (f_1(\mathbf{x}), f_2(x), \dots, f_p(\mathbf{x}), x_{p+1}, \dots, x_n)$$

Es claro que en todo  $\mathbf{x} \in \Omega$  se cumple

$$\frac{\partial(\Phi_1, \dots, \Phi_n)}{\partial(x_1, \dots, x_n)}(\mathbf{x}) = \frac{\partial(f_1, \dots, f_p)}{\partial(x_1, \dots, x_p)}(\mathbf{x})$$

Como el determinante anterior no es nulo cuando  $\mathbf{x} = \mathbf{a}$ , según el teorema de la función inversa, existe  $V \subset \Omega_{\mathbf{a}}$ , entorno abierto de  $\mathbf{a}$  y  $B'$  entorno abierto de  $\mathbf{b}' = (\mathbf{b}, a_{p+1}, \dots, a_n)$ , tales que  $\Phi|_V : V \rightarrow B'$  es un  $C^k$ -difeomorfismo. No hay inconveniente en suponer que  $B' = U \times W$ , donde  $U \subset \mathbb{R}^p$  es un entorno de  $\mathbf{b}$  y  $W = \{(y_{p+1}, \dots, y_n) : |y_j - a_j| < \varepsilon, p < j \leq n\}$ . Tampoco es restrictivo suponer que  $\frac{\partial(\Phi_1, \dots, \Phi_n)}{\partial(x_1, \dots, x_n)}(\mathbf{x}) \neq 0$ , para todo  $\mathbf{x} \in V$ , lo que significa que las formas lineales  $d\Phi_1(\mathbf{x}), \dots, d\Phi_n(\mathbf{x})$  son linealmente independientes para todo  $\mathbf{x} \in V$ . Según la definición de las componentes de  $\Phi$ , esto significa que las formas lineales  $df_1(\mathbf{x}), \dots, df_p(\mathbf{x}), dx_{p+1}, \dots, dx_n$  son linealmente independientes para todo  $\mathbf{x} \in V$ .

La función  $F : U \times W \rightarrow \mathbb{R}$ , definida por  $F = f \circ (\Phi|_V)^{-1}$ , verifica  $f|_V = F \circ \Phi|_V$  luego, según la regla de la cadena, para cada  $\mathbf{x} \in V$  se cumple:

$$df(\mathbf{x}) = \sum_{j=1}^n \frac{\partial F}{\partial x_j}(\Phi(\mathbf{x})) d\Phi_j(\mathbf{x}) = \sum_{j=1}^p \frac{\partial F}{\partial x_j}(\Phi(\mathbf{x})) df_j(\mathbf{x}) + \sum_{j=p+1}^n \frac{\partial F}{\partial x_j}(\Phi(\mathbf{x})) dx_j.$$

Por hipótesis, las formas lineales  $df(\mathbf{x}), df_1(\mathbf{x}), \dots, df_p(\mathbf{x})$  son linealmente dependientes, mientras que las  $n$  formas  $df_1(\mathbf{x}), \dots, df_p(\mathbf{x}), dx_{p+1}, \dots, dx_n$  son linealmente independientes, luego la igualdad anterior implica que

$$\frac{\partial F}{\partial x_j}(\Phi(\mathbf{x})) = 0, \quad \forall \mathbf{x} \in V, \quad \forall j \in \{p+1, \dots, n\}$$

es decir:

$$\frac{\partial F}{\partial x_j}(\mathbf{y}) = 0, \quad \forall \mathbf{y} \in U \times W, \quad \forall j \in \{p+1, \dots, n\}.$$

Teniendo en cuenta que  $W$  es un paralelepípedo, la anulación de las derivadas parciales significa que  $F$  no depende de las variables  $y_{p+1}, \dots, y_n$  en  $U \times W$ , o bien, que  $F$  sólo depende de  $y_1, \dots, y_p$ , por lo que podemos considerarla definida en  $U \subset \mathbb{R}^p$ .

Volviendo ahora a la relación  $f|_V = F \circ \Phi|_V$ , ésta significa que si  $\mathbf{x} \in V$ , entonces  $f(\mathbf{x}) = F(f_1(\mathbf{x}), \dots, f_p(\mathbf{x}))$ . Por otra parte, si  $\mathbf{u} \in U$ , entonces  $(\mathbf{u}, a_{p+1}, \dots, a_n) \in U \times W$ , por lo que existe  $\mathbf{x} \in V$  tal que  $\Phi(\mathbf{x}) = (\mathbf{u}, a_{p+1}, \dots, a_n)$ , es decir  $\mathbf{u} = (f_1(\mathbf{x}), \dots, f_p(\mathbf{x}))$ . ■

**Definición H.3** Sean  $f_1, \dots, f_p : \Omega \rightarrow \mathbb{R}$  funciones de clase  $C^1(\Omega)$  definidas en un abierto  $\Omega \subset \mathbb{R}^n$ . Se dice que  $f_1, \dots, f_p$  son funcionalmente independientes en  $\mathbf{a} \in \Omega$  si cada función real continua  $F$ , definida en un entorno de  $\mathbf{b} = (f_1(\mathbf{a}), \dots, f_p(\mathbf{a}))$ , que verifique  $F(f_1(\mathbf{x}), \dots, f_p(\mathbf{x})) = 0$  en los puntos  $\mathbf{x}$  de un entorno de  $\mathbf{a}$ , debe ser idénticamente nula en algún entorno de  $\mathbf{b}$ .

**Teorema H.4** Sean  $f_1, \dots, f_p : \Omega \rightarrow \mathbb{R}$  funciones de clase  $C^1(\Omega)$  en un abierto  $\Omega \subset \mathbb{R}^n$ , con  $p \leq n$ . Si en el punto  $\mathbf{a} \in \Omega$  el rango de la matriz  $(D_i f_j(\mathbf{a}))_{1 \leq i \leq n, 1 \leq j \leq p}$  es  $p$  entonces las funciones  $f_1, \dots, f_p$  son funcionalmente independientes en  $\mathbf{a}$ .

DEM: Es consecuencia directa del teorema de la aplicación abierta aplicado a la función  $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_p(\mathbf{x}))$ . En virtud de la hipótesis, existe un entorno abierto

$U$  de  $\mathbf{a}$  tal que  $\mathbf{f}|_U$  es abierta. En tal caso si  $F$  es una función continua definida en un entorno  $V$  de  $\mathbf{b}$ , si existe un entorno  $U_1$  de  $\mathbf{a}$  tal que

$$F(f_1(\mathbf{x}), \dots, f_m(\mathbf{x})) = 0 \quad \forall \mathbf{x} \in U_1$$

debe cumplirse que  $f(U_1) \subset V$ , luego  $W = \mathbf{f}(U \cap U_1)$  es un entorno de  $\mathbf{b}$  tal que  $F(\mathbf{y}) = 0$ , para todo  $\mathbf{y} \in W$ . ■

Un resultado análogo al anterior, en el caso de rango no máximo (pero constante) proporciona una condición suficiente de dependencia funcional.

**Teorema H.5** Sean  $f_1, \dots, f_m : \Omega \rightarrow \mathbb{R}$  funciones de clase  $C^1(\Omega)$  en un abierto  $\Omega \subset \mathbb{R}^n$ . Si existe un entorno de  $\mathbf{a} \in \Omega$ ,  $V \subset \Omega$ , tal que para todo  $x \in V$ , el rango de la matriz  $(D_i f_j(\mathbf{x}))_{1 \leq i \leq n, 1 \leq j \leq m}$  es  $p < m$ , entonces  $m - p$  de estas funciones dependen funcionalmente en  $\mathbf{a}$  de las restantes.

DEM: Sea  $\mathbf{f} : \Omega \rightarrow \mathbb{R}^m$ , definida por  $f(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))$ . Podemos suponer que  $\nabla f_1(\mathbf{a}), \dots, \nabla f_p(\mathbf{a})$  son linealmente independientes (reordenando las funciones si es necesario). Por hipótesis, para cada  $j \in \{p+1, \dots, m\}$   $\mathbf{x} \in V$  el rango de  $\mathbf{f}'(\mathbf{x})$  es siempre  $p$ , luego, para cada  $j$  tenemos que las formas lineales  $df_1(\mathbf{x}), \dots, df_p(\mathbf{x}), df_j(\mathbf{x})$ , son linealmente dependientes y aplicando el teorema H.1 se obtiene el resultado. ■

Por ejemplo, si en el teorema anterior, suponemos  $p = m-1$  y  $df_1(\mathbf{a}), \dots, df_{m-1}(\mathbf{a})$  son linealmente independientes, entonces existe una función  $F$  de clase  $C^1$  en un entorno de  $\mathbf{a}$  tal que:

$$f_m(\mathbf{x}) = F(f_1(\mathbf{x}), \dots, f_{m-1}(\mathbf{x}))$$

para  $\mathbf{x}$  en un cierto entorno de  $\mathbf{a}$ .

Si  $p = m - 2$  y suponemos  $df_1(\mathbf{a}), \dots, df_{m-2}(\mathbf{a})$  linealmente independientes, existen funciones  $F_1, F_2$  de clase  $C^1$  en un entorno de  $\mathbf{a}$  tales que:

$$\begin{aligned} f_{m-1}(\mathbf{x}) &= F_1(f_1(\mathbf{x}), \dots, f_{m-2}(\mathbf{x})) \\ f_m(\mathbf{x}) &= F_2(f_1(\mathbf{x}), \dots, f_{m-2}(\mathbf{x})) \end{aligned}$$

para  $\mathbf{x}$  en un cierto entorno de  $\mathbf{a}$ .

### Ejemplo H.6

Sean  $f_1, f_2, f_3 : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  dadas por:

$$\begin{aligned} f_1(r, \theta) &= r \cos \theta \\ f_2(r, \theta) &= r \operatorname{sen} \theta \\ f_3(r, \theta) &= r \end{aligned}$$

En un entorno de  $(0, \theta_0)$  el rango de  $df(r, \theta)$  no permanece constante, aunque siempre es menor que 3. Se satisface la relación funcional

$$f_1^2 + f_2^2 - f_3^2 = 0$$

pero no es posible expresar una de las funciones como una función de clase  $C^1$  de las otras dos: Basta observar que el punto  $(0, 0, 0)$  es el vértice del cono  $x^2 + y^2 - z^2 = 0$ .

Este ejemplo muestra que en el teorema anterior la hipótesis «rango constante» (naturalmente también menor que  $m$ ) es esencial. ■

## H.2. Parametrizaciones regulares

Frecuentemente se asume sin demostración que la parametrización habitual de un trozo de esfera usando la longitud y la latitud como parámetros conduce a una parametrización regular. En el siguiente ejemplo se puede ver una demostración detallada de este hecho.

### Ejemplo H.7 *Parametrización regular de un trozo de esfera*

Sea  $U_{\alpha\beta} = \{(s, t) \in \mathbb{R}^2 : |s| < \beta, 0 < t < \beta\}$ , donde  $0 < \alpha \leq \pi/2$ ,  $0 < \beta \leq 2\pi$ . Si suponemos que  $s$  (resp.  $t$ ) representa la latitud (resp. longitud) de un punto de la esfera  $S = \{(x, y, z) : x^2 + y^2 + z^2 = R^2\}$ , es fácil visualizar geoméricamente el trozo de esfera  $S_{\alpha\beta} = \varphi(U_{\alpha\beta})$  obtenido como imagen de  $U_{\alpha\beta}$  mediante la aplicación

$$\varphi(s, t) = (R \cos s \cos t, R \cos s \sin t, R \sin s)$$

Seguidamente vemos con detalle que la parametrización  $\varphi : U_{\alpha\beta} \rightarrow \mathbb{R}^3$  es regular. Observemos en primer lugar que para todo  $(s, t) \in U_{\alpha\beta}$  los vectores no nulos

$$D_1\varphi(s, t) = (-R \sin s \cos t, -R \sin s \sin t, R \cos s)$$

$$D_2\varphi(s, t) = (-R \cos s \sin t, R \cos s \cos t, 0)$$

son linealmente independientes por ser ortogonales. También es fácil ver que  $\varphi$  es inyectiva en el abierto  $U_{\alpha\beta}$ : Sean  $(s, t), (s', t') \in U_{\alpha\beta}$  tales que  $\varphi(s, t) = \varphi(s', t')$ , lo que significa que se cumplen las tres igualdades.

$$\cos s \cos t = \cos s' \cos t'; \quad \cos s \sin t = \cos s' \sin t', \quad \sin s = \sin s'$$

Elevando al cuadrado las dos primeras y sumando resulta  $\cos^2 s = \cos^2 s'$ , y teniendo en cuenta que  $\cos s > 0$  y  $\cos s' > 0$  (porque  $s, s' \in (\pi/2, \pi/2)$ ) resulta  $\cos s = \cos s'$ . Esta igualdad, combinada con la última,  $\sin s = \sin s'$ , conduce a que  $s - s'$  es un múltiplo entero de  $2\pi$ , y teniendo en cuenta que  $s, s' \in (0, 2\pi)$ , se obtiene  $s = s'$ . Ahora, utilizando las dos primeras igualdades y teniendo en cuenta que  $\cos s = \cos s' > 0$  se obtiene que  $\cos t = \cos t'$ ,  $\sin t = \sin t'$ , y con un argumento similar se concluye que  $t = t'$ .

Para terminar debemos demostrar que la inversa de la biyección  $\varphi : U_{\alpha\beta} \rightarrow \varphi(U_{\alpha\beta})$  es continua. Lo haremos viendo que si  $\mathbf{p}_j = \varphi(s_j, t_j)$  es una sucesión en  $\varphi(U_{\alpha\beta})$  que converge hacia  $\mathbf{p} = \varphi(s, t) \in \varphi(U_{\alpha\beta})$  entonces la sucesión  $(s_j, t_j) \in U_{\alpha\beta}$  converge hacia  $(s, t) \in U_{\alpha\beta}$  (véase el corolario 2.8). Como la sucesión  $(s_j, t_j)$  está contenida en el compacto  $K = [-\pi/2, \pi/2] \times [0, 2\pi]$ , bastará que ver que cualquier subsucesión convergente  $(s_{j_k}, t_{j_k})$  converge hacia  $(s, t) \in K$ . En lo que sigue consideramos que  $\varphi$  está definida, por las mismas fórmulas, en el compacto  $K = [-\pi/2, \pi/2] \times [0, 2\pi]$ . Si la subsucesión  $(s_{j_k}, t_{j_k})$  converge hacia  $(s', t') \in K$ , por continuidad se debe cumplir

$$\varphi(s', t') = \lim_k \varphi(s_{j_k}, t_{j_k}) = \lim_k \mathbf{p}_{j_k} = \mathbf{p} = \varphi(s, t)$$

Como  $|s| < \pi/2$ ,  $|s'| \leq \pi/2$ ,  $t \in (0, 2\pi)$ ,  $t' \in [0, 2\pi]$  con el razonamiento realizado para demostrar que  $\varphi$  es inyectiva sobre  $U_{\alpha\beta}$  se concluye que  $(s, t) = (s', t')$ . ■

La observación que sigue a la demostración del teorema 9.4 tiene una consecuencia interesante formulada en la proposición H.9, en términos de la siguiente definición:

**Definición H.8** *Dos parametrizaciones,  $\varphi_j : U_j \rightarrow \mathbb{R}^n$ ,  $j = 1, 2$ , de clase  $C^m$  y dimensión  $k$ , se dice que son  $C^m$ -equivalentes cuando existe un  $C^m$ -difeomorfismo  $\mathbf{g} : U_1 \rightarrow U_2$ , tal que  $\varphi_1 = \varphi_2 \circ \mathbf{g}$ .*

**Proposición H.9** *Dos parametrizaciones regulares  $\varphi_j : U_j \rightarrow \mathbb{R}^n$ ,  $j = 1, 2$ , de clase  $C^m$  y dimensión  $k$ , con la misma imagen son  $C^m$ -equivalentes.*

DEM: Si  $\varphi_2^{-1} : S \rightarrow U_2$  denota la inversa del homeomorfismo  $\varphi_2 : U_2 \rightarrow S$  entonces  $\mathbf{g} = \varphi_2^{-1} \circ \varphi_1 : U_1 \rightarrow U_2$  es un homeomorfismo que verifica  $\varphi_1 = \varphi_2 \circ \mathbf{g}$ , y basta demostrar que  $\mathbf{g}$  es de clase  $C^m$  (pues el mismo razonamiento, cambiando los papeles de los subíndices, asegurará que su inversa también es de clase  $C^m$ ).

Esto lo haremos viendo que cada  $\mathbf{a} \in U_1$  posee un entorno  $O_{\mathbf{a}} \subset U_1$  donde  $\mathbf{g}|_{O_{\mathbf{a}}}$  es de clase  $C^m$ . Según la observación que sigue al teorema 9.4 el punto  $\mathbf{p} = \varphi_1(\mathbf{a})$  de la subvariedad  $S = \varphi_2(U_2)$  posee un entorno abierto  $W_{\mathbf{p}} \subset \mathbb{R}^n$ , en el que hay definida una función  $\Psi : W_{\mathbf{p}} \rightarrow U_2$  de clase  $C^m$ , que verifica

$$\Psi(\mathbf{x}) = \varphi_2^{-1}(\mathbf{x}) \quad \text{para cada } \mathbf{x} \in S \cap W_{\mathbf{p}}$$

luego  $O_{\mathbf{a}} = \varphi_1^{-1}(W_{\mathbf{p}}) \subset U_1$  es un entorno abierto de  $\mathbf{a}$  tal que  $\mathbf{g}|_{O_{\mathbf{a}}} = \Psi \circ (\varphi_1|_{O_{\mathbf{a}}})$ . Como la composición de funciones de clase  $C^m$  es de clase  $C^m$ , se concluye que  $\mathbf{g}|_{O_{\mathbf{a}}}$  es de clase  $C^m$ . ■

En lo que sigue a las subvariedades diferenciables de  $\mathbb{R}^n$  consideradas en 9.6 a), que se pueden describir como imagen de una parametrización regular, las llamaremos *k-hipersuperficies paramétricas regulares* (*superficies paramétricas regulares* en el caso  $n = 3$ ,  $k = 2$ ). Con esta terminología la proposición H.9 dice que si  $M$  es una *k-hipersuperficie paramétrica regular* de clase  $C^m$ , entonces todas sus representaciones paramétricas regulares de clase  $C^m$  son  $C^m$ -equivalentes.

**Espacio tangente a una parametrización.** Si  $\varphi : U \rightarrow \mathbb{R}^n$  es una parametrización, de clase  $C^m$  y dimensión  $k$ , a cada  $\mathbf{u} \in U$ , le podemos asociar el espacio vectorial  $E(\varphi, \mathbf{u}) := d\varphi(\mathbf{u})(\mathbb{R}^k) \subset \mathbb{R}^n$ , generado por los vectores  $D_j\varphi(\mathbf{u})$ ,  $1 \leq j \leq k$ .

Sabemos que los vectores de  $E(\varphi, \mathbf{u})$ , son tangentes a  $M = \varphi(U)$  en el punto  $\mathbf{p} = \varphi(\mathbf{u})$ , es decir  $E(\varphi, \mathbf{u}) \subset T_{\mathbf{p}}(M)$ , y por ello se suele decir que  $E(\varphi, \mathbf{u})$  es el espacio vectorial tangente a la parametrización  $\varphi$ , en el punto  $\mathbf{p} = \varphi(\mathbf{u})$ , para el valor  $\mathbf{u}$  del parámetro (se puede prescindir de la última frase si  $\varphi$  es inyectiva).

**Proposición H.10** *Sean  $\varphi_i : U_i \rightarrow \mathbb{R}^n$ , ( $i = 1, 2$ ), parametrizaciones  $C^m$ -equivalentes y  $\mathbf{g} : U_1 \rightarrow U_2$ , un  $C^m$  difeomorfismo con  $\varphi_1 = \varphi_2 \circ \mathbf{g}$ . Si  $\mathbf{v} = \mathbf{g}(\mathbf{u}) \in U_2$ , se cumple*

$$E(\varphi_2, \mathbf{v}) = E(\varphi_1, \mathbf{u})$$

DEM: En virtud de la regla de la cadena  $d\varphi_2(\mathbf{v}) \circ d\mathbf{g}(\mathbf{u}) = d\varphi_1(\mathbf{u})$ , y teniendo en cuenta que la aplicación lineal  $d\mathbf{g}(\mathbf{u}) : \mathbb{R}^k \rightarrow \mathbb{R}^k$  es sobreyectiva resulta

$$E(\varphi_1, \mathbf{u}) = d\varphi_1(\mathbf{u})(\mathbb{R}^k) = d\varphi_2(\mathbf{v})(d\mathbf{g}(\mathbf{u})(\mathbb{R}^k)) = d\varphi_2(\mathbf{v})(\mathbb{R}^k) = E(\varphi_2, \mathbf{v})$$

■

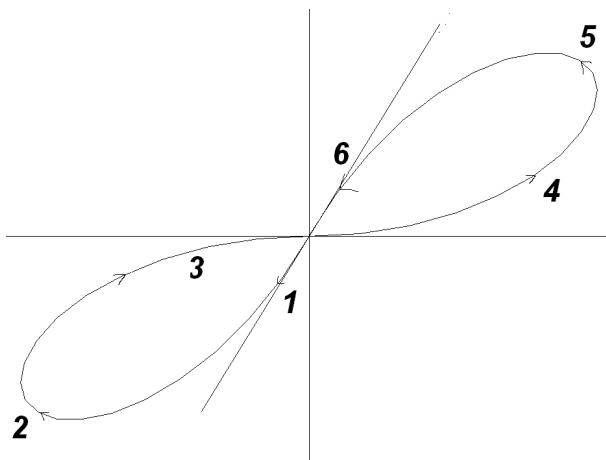
OBSERVACIÓN: Cuando la parametrización  $\varphi$  es regular sabemos que  $M = \varphi(U)$  es una subvariedad diferenciable y por lo tanto, para cada  $\mathbf{p} = \varphi(\mathbf{u}) \in M$ , se cumple la igualdad  $E(\varphi, \mathbf{u}) = T_{\mathbf{p}}(M)$ , lo que significa que  $E(\varphi, \mathbf{u})$  sólo depende de la imagen  $M = \varphi(U)$  y del punto  $\mathbf{p} \in M$ . Pero conviene advertir que en general, para una parametrización no regular, el espacio tangente  $E(\varphi, \mathbf{u})$  depende de  $\varphi$ , y de  $\mathbf{u}$ . Además, puede ocurrir que siendo  $\varphi$  inyectiva, y  $E(\varphi, \mathbf{u})$  un espacio vectorial de dimensión  $k$ , sin embargo  $T_{\mathbf{p}}(M)$  no sea espacio vectorial y  $E(\varphi, \mathbf{u}) \subsetneq T_{\mathbf{p}}(M)$  (es claro que entonces  $M = \varphi(U)$  no es subvariedad diferenciable de  $\mathbb{R}^n$ ). Esto se pone de manifiesto con el ejemplo H.11, referente al caso  $n = 2, k = 1$ .

**Ejemplo H.11** *Una parametrización inyectiva no regular*

Consideremos la parametrización  $\varphi : U \rightarrow \mathbb{R}^2$ , definida en el intervalo  $U = (-\pi/3, \pi/3)$  mediante las ecuaciones

$$\varphi(t) = (\sin 3t \cos t, \sin 3t |\sin t|)$$

Su imagen es una curva en forma de 8 recorrida en el sentido indicado en la figura. Obsérvese que para  $t \in [0, \pi/3)$  es  $\varphi(t) = (\sin 3t \cos t, \sin 3t \sin t)$  luego  $\varphi([0, \pi/3))$  está en el primer cuadrante y coincide con el arco de curva cuya ecuación en coordenadas polares es  $r = \sin 3t$ ,  $0 \leq t \leq \pi/3$ , luego  $r$  crece desde 0 hasta 1 en el intervalo  $0 \leq t \leq \pi/6$ , y decrece desde 1 hasta 0 (sin llegar a valer 0) en el intervalo  $\pi/6 \leq t < \pi/3$ . Es fácil comprobar que para  $t \in (-\pi/3, t]$  es  $\varphi(t) = -\varphi(-t)$ . Con esta información se aprecia que cuando  $t$  recorre  $U$  en sentido creciente, su imagen  $\varphi(t)$  recorre la curva  $S = \varphi(U)$  en el sentido que muestra la figura:



$\varphi(t)$  arranca en el tercer cuadrante muy cerca de  $(0, 0)$ , según la dirección de la recta  $y = \sqrt{3}x$ , pasa por  $(0, 0)$  en el instante  $t = 0$ , con tangente horizontal, y entra en el primer cuadrante, con tangente horizontal, para finalizar su recorrido acercándose a  $(0, 0)$  según la dirección de la recta  $y = \sqrt{3}x$ .

Es fácil ver que  $\varphi$  es inyectiva y de clase  $C^1$ , con  $\varphi'(t) \neq 0$  para todo  $t \in U$  (obsérvese que  $\lim_{t \rightarrow 0^+} \varphi'(t) = \lim_{t \rightarrow 0^-} \varphi'(t) = (3, 0)$ ), es decir  $\varphi$  es una parametrización inyectiva de clase  $C^1$  y dimensión 1, cuyo espacio tangente, para  $t = 0$ , es la recta  $E(\varphi, 0) = \{(x, y) : y = 0\}$ . Es geoméricamente evidente que el conjunto de vectores tangentes en  $\mathbf{0} = (0, 0)$  al conjunto  $S = \varphi(U)$  no es espacio vectorial:

$$T_{\mathbf{0}}(S) = \{(x, y) : y = \sqrt{3}x\} \cup \{(x, y) : y = 0\}$$

Obsérvese que la parametrización  $\varphi$  no es regular porque  $\varphi : U \rightarrow S$  no es homeomorfismo: Hay puntos de  $S$  tan próximos a  $(0, 0) = \varphi(0)$  como queramos que son imágenes de puntos  $t$  con  $1 < |t| < \pi/3$ . ■

### H.3. Subvariedades orientables

Sea  $M \subset \mathbb{R}^n$  una subvariedad diferenciable de clase  $C^m$  y dimensión  $k \leq n$  y para cada  $\mathbf{p} \in M$  sea  $O_{\mathbf{p}}$  una orientación del espacio tangente  $T_{\mathbf{p}}(M)$ . Se dice que  $\{O_{\mathbf{p}} : \mathbf{p} \in M\}$  es un *sistema continuo de orientaciones* de  $M$  cuando cada  $\mathbf{p} \in M$  posee un entorno abierto  $G_{\mathbf{p}}$  tal que en  $G_{\mathbf{p}} \cap M$  se puede definir una función continua  $\beta : G_{\mathbf{p}} \cap M \rightarrow (\mathbb{R}^n)^k$  tal que para cada  $\mathbf{y} \in G_{\mathbf{p}} \cap M$ ,  $\beta(\mathbf{y}) = (\beta_1(\mathbf{y}), \beta_2(\mathbf{y}), \dots, \beta_k(\mathbf{y}))$  es una base de  $T_{\mathbf{p}}(M)$  positiva para la orientación  $O_{\mathbf{y}}$ .

**Definición H.12** *Una subvariedad diferenciable  $M \subset \mathbb{R}^n$ , de clase  $C^m$  y dimensión  $k \leq n$ , se dice que es orientable cuando admite un sistema continuo de orientaciones. En ese caso, una vez que se ha fijado en  $M$  un sistema continuo de orientaciones se dice que  $M$  está orientada.*

Si  $M \subset \mathbb{R}^n$  es una subvariedad orientada mediante el sistema continuo de orientaciones  $\{O_{\mathbf{p}} : \mathbf{p} \in M\}$ , dado un abierto  $\Omega \subset \mathbb{R}^n$  con  $M \cap \Omega = M_0 \neq \emptyset$  es inmediato que  $M_0$  es una subvariedad orientable con la orientación inducida, que es la definida por el sistema continuo de orientaciones  $\{O_{\mathbf{p}} : \mathbf{p} \in M_0\}$ .

Un ejemplo trivial de subvariedad orientable lo proporciona cualquier subespacio vectorial  $M \subset \mathbb{R}^n$ . En este caso  $T_{\mathbf{p}}(M) = M$  para todo  $\mathbf{p} \in M$ , y si orientamos  $M$  como espacio vectorial eligiendo una base  $\beta = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k)$  de  $M$ , es evidente que la aplicación constante  $\beta(\mathbf{y}) = \beta$  define en  $M$  un sistema continuo de orientaciones, por lo que  $M$  queda orientado como subvariedad diferenciable.

**Proposición H.13** *Si  $M = \varphi(U)$  es la imagen de una parametrización regular  $\varphi : U \rightarrow \mathbb{R}^n$  de clase  $C^m$  ( $m \geq 1$ ) definida en un abierto  $U \subset \mathbb{R}^k$ , entonces  $M$  es una subvariedad orientable. Si  $O_{\mathbf{p}}$  es la orientación de  $T_{\mathbf{p}}(M)$  definida por la base*

$$(D_1\varphi(\varphi^{-1}(\mathbf{p})), D_2\varphi(\varphi^{-1}(\mathbf{p})), \dots, D_k\varphi(\varphi^{-1}(\mathbf{p})))$$

*entonces  $\{O_{\mathbf{p}} : \mathbf{p} \in M\}$  es un sistema continuo de orientaciones en  $M$ .*

DEM: Basta observar que al ser  $\varphi : U \rightarrow M$  un homeomorfismo entonces

$$(D_1\varphi(\varphi^{-1}(\mathbf{p})), D_2\varphi(\varphi^{-1}(\mathbf{p})), \dots, D_k\varphi(\varphi^{-1}(\mathbf{p})))$$

es una base de  $T_{\mathbf{p}}(M)$  que depende continuamente de  $\mathbf{p} \in M$ , y por lo tanto define un sistema continuo de orientaciones en  $M$ . ■

En las condiciones de la última proposición la subvariedad  $M$  queda orientada eligiendo una de sus parametrizaciones regulares. La orientación que la parametrización regular  $\varphi$  define en  $M$  es la indicada en el enunciado de esta proposición.

En lo que sigue  $\mathbb{R}^n$  siempre se supondrá orientado con la orientación usual para la que la base canónica es positiva. Un subespacio vectorial  $T \subset \mathbb{R}^n$  de dimensión  $k$  se puede orientar eligiendo un sistema de  $n - k$  vectores linealmente independientes  $(\mathbf{z}_1, \dots, \mathbf{z}_{n-k})$  tales que  $\mathbf{z}_j \notin T$  para todo  $j \in [1, \dots, n - k]$ . Se comprueba fácilmente que todas las bases  $(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$  de  $T$  tales que  $(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k, \mathbf{z}_1, \dots, \mathbf{z}_{n-k})$  es una base positiva de  $\mathbb{R}^n$  tienen la misma orientación. Diremos que ésta es la orientación de  $T$  definida por el sistema de vectores linealmente independientes  $(\mathbf{z}_1, \dots, \mathbf{z}_{n-k})$ . En particular, un hiperplano  $T \subset \mathbb{R}^n$  queda orientado mediante un vector  $\mathbf{z} \notin T$ . Las subvariedades diferenciables de  $\mathbb{R}^n$  se pueden orientar usando un procedimiento análogo:

**Proposición H.14** *Sea  $M \subset \mathbb{R}^n$  una subvariedad diferenciable de clase  $C^m$  y dimensión  $k \leq n$  y para cada  $\mathbf{p} \in M$  sea  $(\mathbf{z}_1(\mathbf{p}), \dots, \mathbf{z}_{n-k}(\mathbf{p}))$  un sistema de vectores linealmente independientes que dependen continuamente de  $\mathbf{p}$ , y tal que  $\mathbf{z}_j(\mathbf{p}) \notin T_{\mathbf{p}}(M)$  para todo  $\mathbf{p} \in M$  y cada  $j \in [1, 2, \dots, n - k]$ . Si  $O_{\mathbf{p}}$  es la orientación de  $T_{\mathbf{p}}(M)$  definida por el sistema de vectores linealmente independientes  $(\mathbf{z}_1(\mathbf{p}), \dots, \mathbf{z}_{n-k}(\mathbf{p}))$ , entonces  $\{O_{\mathbf{p}} : \mathbf{p} \in M\}$  es un sistema continuo de orientaciones en  $M$ .*

DEM: Cada  $\mathbf{p} \in M$  tiene un entorno abierto  $\Omega_{\mathbf{p}} \subset \mathbb{R}^n$  tal que  $M \cap \Omega_{\mathbf{p}} = \varphi(U)$  donde  $\varphi : U \rightarrow \mathbb{R}^n$  es una parametrización regular de clase  $C^m$ , que podemos suponer definida en un abierto conexo  $U \subset \mathbb{R}^k$ , para tener garantizado que su imagen  $M \cap \Omega_{\mathbf{p}}$  también es conexa. Para cada  $\mathbf{y} \in M \cap \Omega_{\mathbf{p}}$  y cada  $j \in [1, \dots, k]$  sea  $\mathbf{v}_j(\mathbf{y}) = D_j\varphi(\varphi^{-1}(\mathbf{y}))$ . Entonces  $(\mathbf{v}_1(\mathbf{y}), \dots, \mathbf{v}_k(\mathbf{y}))$  es una base de  $T_{\mathbf{y}}(M)$ , que depende continuamente de  $\mathbf{y} \in M \cap \Omega_{\mathbf{p}}$  luego, en virtud de las hipótesis,

$$(\mathbf{v}_1(\mathbf{y}), \dots, \mathbf{v}_k(\mathbf{y}), \mathbf{z}_1(\mathbf{y}), \dots, \mathbf{z}_{n-k}(\mathbf{y}))$$

es una base de  $\mathbb{R}^n$  que depende continuamente de  $\mathbf{y} \in M \cap \Omega_{\mathbf{p}}$ . Su determinante  $\Delta(\mathbf{y})$  es una función continua que no se anula en el conjunto conexo  $M \cap \Omega_{\mathbf{p}}$ , y por lo tanto conserva un signo constante. Podemos suponer, (cambiando  $\mathbf{v}_k(\mathbf{y})$  por  $-\mathbf{v}_k(\mathbf{y})$  si es preciso) que  $\Delta(\mathbf{y}) > 0$  para todo  $\mathbf{y} \in M \cap \Omega_{\mathbf{p}}$ , luego la base  $(\mathbf{v}_1(\mathbf{y}), \dots, \mathbf{v}_k(\mathbf{y}))$  de  $T_{\mathbf{y}}(M)$  es positiva para la orientación  $O_{\mathbf{y}}$ . Así queda establecido que  $\{O_{\mathbf{p}} : \mathbf{p} \in M\}$  es un sistema continuo de orientaciones en  $M$ . ■



**Corolario H.15** Sea  $M = \{\mathbf{x} \in \Omega : g_1(\mathbf{x}) = g_2(\mathbf{x}) = \cdots = g_{n-k}(\mathbf{x}) = 0\}$  donde  $g_1, g_2, \cdots, g_{n-k} : \Omega \rightarrow \mathbb{R}^{n-k}$  son funciones de clase  $C^m$  ( $m \geq 1$ ) definidas en un abierto  $\Omega \subset \mathbb{R}^n$ , tales que para todo  $\mathbf{p} \in M$  los vectores  $(\nabla g_1(\mathbf{p}), \nabla g_2(\mathbf{p}), \cdots, \nabla g_{n-k}(\mathbf{p}))$  son linealmente independientes. Entonces  $M$  es una subvariedad orientable.

Una orientación de  $M$  es la definida por el sistema de vectores linealmente independientes  $(\nabla g_1(\mathbf{p}), \nabla g_2(\mathbf{p}), \cdots, \nabla g_{n-k}(\mathbf{p}))$ .

DEM: Es una consecuencia directa de la proposición H.14 ■

# I

## Extremos y formas cuadráticas

### I.1. Extremos y formas cuadráticas

En este apéndice se consideran algunas aplicaciones interesantes de la teoría de extremos condicionados al estudio de las formas cuadráticas. Se recomienda comenzar con los ejercicios 9.17, 9.18 que contienen casos particulares de los resultados generales que se exponen aquí. El siguiente teorema es un resultado bien conocido del álgebra lineal del que ofrecemos una demostración alternativa basada en optimizaciones sucesivas de una forma cuadrática sobre la intersección de la esfera unidad con una sucesión decreciente de subespacios vectoriales. La idea clave es que el mayor y el menor autovalor de una matriz simétrica real proporcionan el máximo y el mínimo absoluto, sobre la esfera unidad, de la forma cuadrática asociada a la matriz.

**Teorema I.1** Sea  $A = (\alpha_{ij})_{1 \leq i, j \leq n}$  una matriz simétrica,  $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$  la aplicación lineal asociada y  $Q : \mathbb{R}^n \rightarrow \mathbb{R}$  la forma cuadrática asociada:

$$L(\mathbf{x}) = (L_1(\mathbf{x}), L_2(\mathbf{x}), \dots, L_n(\mathbf{x})), \text{ donde } L_k(\mathbf{x}) = \sum_{j=1}^n \alpha_{kj} x_j$$

$$Q(\mathbf{x}) = \langle L(\mathbf{x}) \mid \mathbf{x} \rangle = \sum_{i,j=1}^n \alpha_{ij} x_i x_j$$

Entonces se verifica: Todos los autovalores de  $A$  son reales. Si  $\mu_1 \geq \mu_2 \geq \dots \geq \mu_n$  son los autovalores de  $A$  se cumple

$$\mu_1 = \max\{Q(\mathbf{x}) : \|\mathbf{x}\|_2 = 1\}, \quad \mu_n = \min\{Q(\mathbf{x}) : \|\mathbf{x}\|_2 = 1\}$$

Existe una base ortonormal de  $\mathbb{R}^n$   $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$  formada por vectores propios

$$L(\mathbf{u}_j) = \mu_j \mathbf{u}_j, \quad i \leq j \leq n$$

y respecto a esta base la matriz de  $Q$  es diagonal: Si  $\mathbf{x} = \sum_{i=1}^n t_i \mathbf{u}_i$  entonces

$$Q(\mathbf{x}) = \mu_1 t_1^2 + \mu_2 t_2^2 + \dots + \mu_n t_n^2$$

DEM: La función continua  $Q(\mathbf{x})$  alcanza un máximo y un mínimo absoluto sobre la esfera compacta  $S = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_2 = 1\}$ . Sea  $\mathbf{y} \in S$  tal que

$$Q(\mathbf{y}) = \max\{Q(\mathbf{x}) : \mathbf{x} \in S\}$$

$S = \{\mathbf{x} \in \mathbb{R}^n : g(\mathbf{x}) = 0\}$  donde  $g(\mathbf{x}) = x_1^2 + x_2^2 + \cdots + x_n^2 - 1$  con  $\nabla g(\mathbf{x}) = 2\mathbf{x} \neq \mathbf{0}$  para cada  $\mathbf{x} \in S$  luego  $\mathbf{y} \in S$  es un punto estacionario para  $Q|_S$ , es decir

$$\nabla Q(\mathbf{y}) = \lambda \nabla g(\mathbf{y})$$

para algún  $\lambda \in \mathbb{R}$ . Como  $D_k Q(\mathbf{x}) = 2 \sum_{j=1}^n \alpha_{kj} x_j = 2L_k(\mathbf{x})$ ,  $1 \leq k \leq n$ , resulta  $2L(\mathbf{y}) = 2\lambda \mathbf{y}$ , luego  $\mathbf{y}$  es un vector propio de  $L$  y  $\lambda$  es el autovalor asociado. Nótese que

$$Q(\mathbf{y}) = \langle L(\mathbf{y}) | \mathbf{y} \rangle = \lambda \|\mathbf{y}\|^2 = \lambda$$

Empezamos la construcción con  $\mathbf{u}_1 = \mathbf{y}$ , y  $\mu_1 = \lambda$ , que cumplen  $Q(\mathbf{u}_1) = \mu_1$ . Tenemos demostrado así que  $L$  tiene un autovalor real. (De hecho hemos detectado el mayor autovalor, pues si  $\nu$  es otro autovalor,  $L(\mathbf{v}) = \nu \mathbf{v}$  para algún  $\mathbf{v} \in S$ , luego

$$\nu = \langle \nu \mathbf{v} | \mathbf{v} \rangle = \langle L(\mathbf{v}) | \mathbf{v} \rangle = Q(\mathbf{v}) \leq Q(\mathbf{y}) = \mu_1).$$

Continuamos la construcción considerando

$$S_1 = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_2 = 1, \langle \mathbf{u}_1 | \mathbf{x} \rangle = 0\}$$

y un punto  $\mathbf{z} \in S_1$  tal que

$$Q(\mathbf{z}) = \max\{Q(\mathbf{x}) : \mathbf{x} \in S_1\} \leq Q(\mathbf{y}) = \mu_1$$

$S_1 = \{\mathbf{x} \in \mathbb{R}^n : g(\mathbf{x}) = 0, g_1(\mathbf{x}) = 0\}$  donde  $g_1(\mathbf{x}) = \langle \mathbf{u}_1 | \mathbf{x} \rangle$ . Como los vectores  $\nabla g(\mathbf{z}) = 2\mathbf{z}$ ,  $\nabla g_1(\mathbf{z}) = \mathbf{u}_1$  son independientes (por ser ortogonales) podemos asegurar que  $\mathbf{z} \in S_1$  es estacionario para  $Q|_{S_1}$ , luego existen  $\lambda', \lambda'' \in \mathbb{R}$  tales que

$$\nabla Q(\mathbf{z}) = \lambda' \nabla g(\mathbf{z}) + \lambda'' \nabla g_1(\mathbf{z})$$

es decir

$$2L(\mathbf{z}) = 2\lambda' 2\mathbf{z} + \lambda'' \mathbf{u}_1$$

Como  $\mathbf{u}_1$  y  $\mathbf{z}$  son ortogonales, usando la simetría de  $A$  se obtiene que  $\mathbf{u}_1$  y  $L(\mathbf{z})$  también lo son:

$$\langle L(\mathbf{z}) | \mathbf{u}_1 \rangle = \langle \mathbf{z} | L(\mathbf{u}_1) \rangle = \langle \mathbf{z} | \mu_1 \mathbf{u}_1 \rangle = 0.$$

Se sigue de esto que  $\lambda'' = 0$ , que  $\mathbf{z}$  es un vector propio y que  $\lambda'$  es el autovalor correspondiente. Si tomamos  $\mu_2 = \lambda'$  y  $\mathbf{u}_2 = \mathbf{z}$  es claro que

$$\langle \mathbf{u}_1 | \mathbf{u}_2 \rangle = 0; \quad \mu_2 = \langle L(\mathbf{u}_2) | \mathbf{u}_2 \rangle = Q(\mathbf{u}_2) \leq \mu_1.$$

La construcción continúa ahora considerando el compacto  $S_2 \subset S_1$  definido por

$$S_2 = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_2 = 1, \langle \mathbf{u}_1 | \mathbf{x} \rangle = 0, \langle \mathbf{u}_2 | \mathbf{x} \rangle = 0\}$$

y un punto  $\mathbf{w} \in S_2$  tal que

$$Q(\mathbf{w}) = \text{máx}\{Q(\mathbf{x}) : \mathbf{x} \in S_2\}$$

Ahora  $S_2 = \{\mathbf{x} \in \mathbb{R}^n : g(\mathbf{x}) = 0, g_1(\mathbf{x}) = 0, g_2(\mathbf{x}) = 0\}$  donde  $g_2(\mathbf{x}) = \langle \mathbf{u}_2 | \mathbf{x} \rangle$ .

Como los vectores  $\nabla g(\mathbf{w}) = 2\mathbf{w}$  y  $\nabla g_1(\mathbf{w}) = \mathbf{u}_1$ ,  $\nabla g_2(\mathbf{w}) = \mathbf{u}_2$  son independientes (por ser ortogonales) el punto  $\mathbf{w} \in S_2$  es estacionario para  $Q|_{S_2}$  y existen  $\mu', \mu'', \mu''' \in \mathbb{R}$  tales que

$$\nabla Q(\mathbf{w}) = \mu' \nabla g(\mathbf{w}) + \mu'' \nabla g_1(\mathbf{w}) + \mu''' \nabla g_2(\mathbf{w})$$

es decir

$$2L(\mathbf{w}) = 2\mu' \mathbf{w} + \mu'' \mathbf{u}_1 + \mu''' \mathbf{u}_2$$

Como  $\mathbf{w}$  es ortogonal a  $\mathbf{u}_1$ , y  $\mathbf{u}_2$ , usando la simetría de  $A$  se obtiene que  $L(\mathbf{w})$  también lo es:

$$\langle L(\mathbf{w}) | \mathbf{u}_i \rangle = \langle \mathbf{w} | L(\mathbf{u}_i) \rangle = \langle \mathbf{w} | \mu_i \mathbf{u}_i \rangle = 0$$

Se sigue de esto que  $\mu'' = \mu''' = 0$ , que  $\mathbf{w}$  es un vector propio, y que  $\mu'$  es el autovalor asociado. Si tomamos  $\mu_3 = \mu'$  y  $\mathbf{u}_3 = \mathbf{w}$  es claro que se cumple

$$\langle \mathbf{u}_1 | \mathbf{u}_3 \rangle = 0; \quad \langle \mathbf{u}_2 | \mathbf{u}_3 \rangle = 0; \quad \mu_3 = \langle L(\mathbf{u}_3) | \mathbf{u}_3 \rangle = Q(\mathbf{u}_3) \leq \mu_2.$$

La construcción sigue de esta forma hasta que acaba en un número finito de pasos. ■

A la hora de aplicar la condición suficiente de extremo condicionado dada en el apartado b) del teorema 9.11 se plantea el problema de saber cuando la restricción a un subespacio  $T \subset \mathbb{R}^n$  de la forma cuadrática  $Q(\mathbf{u}) = \sum_{ij=1}^n \alpha_{ij} u_i u_j$ , asociada a una matriz simétrica, es definida positiva o definida negativa.

**Proposición I.2** *Sea  $Q(\mathbf{u}) = \sum_{ij=1}^n \alpha_{ij} u_i u_j$  la forma cuadrática asociada a una matriz simétrica y  $T \subset \mathbb{R}^n$  un subespacio vectorial  $k$ -dimensional de ecuaciones implícitas*

$$\sum_{j=1}^n \beta_{ij} u_j = 0, \quad 1 \leq i \leq m.$$

donde  $m = n - k$  y los vectores  $(\beta_{i1}, \beta_{i2}, \dots, \beta_{in})$ ,  $1 \leq i \leq m$  son linealmente independientes. Si todas las raíces del polinomio

$$\Delta(\sigma) = \begin{vmatrix} \alpha_{11} - \sigma & \alpha_{12} & \cdots & \alpha_{1n} & \beta_{11} & \cdots & \beta_{m1} \\ \alpha_{21} & \alpha_{22} - \sigma & \cdots & \alpha_{2n} & \beta_{12} & \cdots & \beta_{m2} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} - \sigma & \beta_{1n} & \cdots & \beta_{mn} \\ \beta_{11} & \beta_{12} & \cdots & \beta_{1n} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \beta_{2n} & 0 & \cdots & 0 \\ \beta_{m1} & \beta_{m2} & \cdots & \beta_{mn} & 0 & \cdots & 0 \end{vmatrix}$$

son positivas entonces la restricción de  $Q$  al subespacio  $T \subset \mathbb{R}^n$  es definida positiva.

DEM: Razonando como en el lema 6.14 es claro que  $Q(\mathbf{u}) > 0$  para todo  $\mathbf{u} \in T \setminus \{\mathbf{0}\}$  si y sólo si el mínimo absoluto de la función continua  $Q$  sobre el compacto  $S \cap T = \{\mathbf{x} \in T : \|\mathbf{x}\|_2 = 1\}$ , que se alcanza en algún  $\mathbf{h} \in S \cap T$ , es positivo:

$$\min\{Q(\mathbf{u}) : \mathbf{u} \in T, \|\mathbf{u}\|_2 = 1\} = Q(\mathbf{h}) > 0$$

$S \cap T$  está definido mediante las  $m + 1$  condiciones de ligadura:

$$\begin{aligned} \beta_{11}x_1 + \beta_{12}x_2 + \cdots + \beta_{1n}x_n &= 0 \\ \beta_{21}x_1 + \beta_{22}x_2 + \cdots + \beta_{2n}x_n &= 0 \\ \dots \dots \dots \dots \dots \dots & \\ \beta_{m1}x_1 + \beta_{m2}x_2 + \cdots + \beta_{mn}x_n &= 0 \\ x_1^2 + x_2^2 + \cdots + x_n^2 - 1 &= 0 \end{aligned}$$

Para cada  $\mathbf{x} \in S \cap T$  los gradientes de las condiciones de ligadura

$$(\beta_{11}, \beta_{12}, \dots, \beta_{1n}), (\beta_{21}, \beta_{22}, \dots, \beta_{2n}), (\beta_{m1}, \beta_{m2}, \dots, \beta_{mn}), (2x_1, 2x_2, \dots, 2x_n)$$

son independientes (los  $m$  primeros vectores son independientes y el último es ortogonal a todos ellos) luego, en virtud del teorema 9.10,  $\mathbf{h}$  es un punto estacionario de  $Q$  sobre  $S \cap T$ , es decir, existen coeficientes  $\sigma, \lambda_1, \lambda_2, \dots, \lambda_m$  tales que

$$D_i Q(\mathbf{h}) = 2\sigma h_i + 2\lambda_1 \beta_{1i} + \cdots + 2\lambda_m \beta_{mi} = 0, \quad 1 \leq i \leq n.$$

es decir

$$\sum_{j=1}^n \alpha_{ij} h_j = \sigma h_i + \sum_{r=1}^m \lambda_r \beta_{ri} = 0, \quad 1 \leq i \leq n.$$

Multiplicando la ecuación  $i$ -ésima por  $h_i$ , sumando, y utilizando que las componentes de  $\mathbf{h} = (h_1, h_2, \dots, h_n) \in T \cap S$  satisfacen las ecuaciones

$$\beta_{i1}h_1 + \beta_{i2}h_2 + \cdots + \beta_{in}h_n = 0, \quad h_1^2 + h_2^2 + \cdots + h_n^2 = 1$$

se concluye que

$$Q(\mathbf{h}) = \sum_{ij} \alpha_{ij} h_i h_j = \sigma$$

Hemos demostrado así que si el mínimo absoluto de  $Q$  sobre  $S \cap T$  se alcanza en  $\mathbf{h} \in S \cap T$  entonces el mínimo absoluto  $Q(\mathbf{h})$  es el valor del multiplicador  $\sigma$  asociado a la condición de ligadura  $x_1^2 + x_2^2 + \cdots + x_n^2 = 1$ .

Sabemos que  $h_1, h_2, \dots, h_n, -\lambda_1, -\lambda_2, \dots, -\lambda_m$  son soluciones del sistema homogéneo de  $n + m$  ecuaciones con  $n + m$  incógnitas

$$\begin{aligned} (\alpha_{11} - \sigma)h_1 + \alpha_{12}h_2 + \cdots + \alpha_{1n}h_n + \beta_{11}\lambda_1 + \beta_{21}\lambda_2 + \cdots + \beta_{m1}\lambda_m &= 0 \\ \alpha_{21}h_1 + (\alpha_{22} - \sigma)h_2 + \cdots + \alpha_{2n}h_n + \beta_{12}\lambda_1 + \beta_{22}\lambda_2 + \cdots + \beta_{m2}\lambda_m &= 0 \end{aligned}$$

$$\begin{array}{l} \dots\dots\dots \qquad \dots\dots\dots \qquad \dots\dots\dots \\ \alpha_{n1}h_1 + \alpha_{n2}h_2 + \dots + (\alpha_{nn} - \sigma)h_n + \beta_{1n}\lambda_1 + \beta_{2n}\lambda_2 + \dots + \beta_{mn}\lambda_m = 0 \\ \beta_{11}h_1 + \beta_{12}h_2 + \dots + \beta_{1n}h_n = 0 \\ \dots\dots\dots \qquad \dots\dots\dots \\ \beta_{m1}h_1 + \beta_{m2}h_2 + \dots + \beta_{mn}h_n = 0 \end{array}$$

Como este sistema admite soluciones no triviales, su determinante  $\Delta(\sigma)$  se anula, es decir  $\sigma$  es solución de la ecuación  $\Delta(\sigma) = 0$ . Podemos afirmar entonces que si el polinomio  $\Delta(\sigma)$  tiene todas sus raíces positivas entonces la forma cuadrática  $Q|_T$  es definida positiva. ■

# J

## Cambio de variable en la integral de Riemann

### J.1. Preliminares

En esta sección se recogen algunos resultados preliminares que intervienen en la demostración del teorema del cambio de variable, que tienen interés por sí mismos.

Si todos los lados de un intervalo cerrado  $Q = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n] \subset \mathbb{R}^n$  tienen la misma longitud  $b_j - a_j = l$ , diremos que  $Q$  es un cubo cerrado de lado  $l$  y centro  $\mathbf{c} = (c_1, c_2, \dots, c_n)$ , con  $c_j = (a_j + b_j)/2$ . Análogamente se define el cubo abierto de centro  $\mathbf{c}$  y lado  $l$ . Para el estudio de las cuestiones de cálculo integral que se abordan en este capítulo es conveniente utilizar en  $\mathbb{R}^n$  la norma  $\|\cdot\|_\infty$  porque con ella una bola cerrada (resp. abierta) de centro  $\mathbf{c}$  y radio  $r > 0$ , no es otra cosa que el cubo cerrado (resp. abierto) de centro  $\mathbf{c}$  y lado  $l = 2r$ .

**Lema J.1** *Para un conjunto  $H \subset \mathbb{R}^n$  son equivalentes*

- i)  $H$  tiene medida nula (resp. contenido nulo).*
- ii) Para cada  $\epsilon > 0$  existe una sucesión infinita (resp. finita) de cubos cerrados  $(Q_j)$  tal que  $H \subset \bigcup_j Q_j$ , y  $\sum_j v(Q_j) < \epsilon$ .*
- iii) Para cada  $\epsilon > 0$  existe una sucesión infinita (resp. finita) de cubos abiertos  $(U_j)$  tal que  $H \subset \bigcup_j U_j$ , y  $\sum_j v(U_j) < \epsilon$ .*

DEM: Observación preliminar: Si  $R = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$  es un intervalo cerrado con lados de longitud racional  $b_k - a_k = r_k \in \mathbb{Q}$ ,  $1 \leq k \leq n$ , entonces existe  $p \in \mathcal{P}(R)$  que descompone a  $R$  en cubos cerrados que no se solapan

$$R = \bigcup \{S : S \in \Delta(p)\}, \quad \text{y} \quad v(R) = \sum_{S \in \Delta(p)} v(S)$$

Basta escribir los números racionales  $r_k = n_k/m$  como fracciones con un denominador común  $m$  y descomponer cada intervalo  $[a_k, b_k]$  en  $n_k$  intervalos de la misma longitud  $1/m$  para conseguir una subdivisión  $p$  con la propiedad requerida.

i)  $\Rightarrow$  ii) Si  $H$  tiene medida nula, para cada  $\epsilon > 0$  existe una sucesión de rectángulos cerrados  $\{E_k : k \in \mathbb{N}\}$  tal que  $H \subset \bigcup_{k=1}^{\infty} E_k$ , y  $\sum_{k=1}^{\infty} v(E_k) < \epsilon/2$ .

Para cada  $k \in \mathbb{N}$  existe un rectángulo cerrado  $R_k \supset E_k$  con todos sus lados de longitud racional, tal que  $v(R_k) \leq v(E_k) + \epsilon/2^{k+1}$  (basta tener en cuenta que  $\mathbb{Q}$  es denso en  $\mathbb{R}$  y que el volumen de un rectángulo acotado depende continuamente de las longitudes de sus lados). Según la observación previa, cada  $R_k$  se descompone en una cantidad finita de cubos cerrados  $\{Q_j : j \in M_k\}$  tal que  $\sum_{j \in M_k} v(Q_j) = v(R_k)$ , donde no hay inconveniente en suponer que los conjuntos  $M_k \subset \mathbb{N}$  son disjuntos. Si  $M = \bigcup_{k=1}^{\infty} M_k$ , es claro que  $\{Q_j : j \in M\}$  es una familia numerable de cubos cerrados que recubre  $H$  y verifica

$$\sum_{j \in M} v(Q_j) = \sum_{k=1}^{\infty} \sum_{j \in M_k} v(Q_j) = \sum_{k=1}^{\infty} v(R_k) \leq \sum_{k=1}^{\infty} (v(E_k) + \epsilon/2^{k+1}) \leq \epsilon/2 + \epsilon/2 = \epsilon$$

ii)  $\Rightarrow$  iii) Sea  $(Q_j)$  una sucesión de cubos cerrados que cumple ii). Para cada  $j$  existe un cubo abierto  $U_j \supset Q_j$ , tal que  $v(U_j) < \delta/2^j$ , donde  $\delta = \epsilon - \sum_j v(Q_j) > 0$ , y es claro que la sucesión de cubos abiertos  $(U_j)$  cumple iii).

iii)  $\Rightarrow$  i) Es inmediato.

La caracterización alternativa de los conjuntos de contenido nulo se deja al cuidado del lector.  $\blacksquare$

**Proposición J.2** *Sea  $H \subset \mathbb{R}^n$  un conjunto de medida nula (resp. contenido nulo). Si la aplicación  $\mathbf{g} : H \rightarrow \mathbb{R}^n$  es lipschitziana entonces  $\mathbf{g}(H)$  tiene medida nula (resp. contenido nulo).*

DEM: En  $\mathbb{R}^n$  consideramos la norma  $\|\cdot\|_{\infty}$  que tiene la propiedad de que sus bolas son cubos. Como todas las normas de  $\mathbb{R}^n$  son equivalentes es fácil ver que  $\mathbf{g}$  sigue siendo lipschitziana para esta norma, luego existe  $C > 0$  tal que

$$\|\mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{y})\|_{\infty} \leq C \|\mathbf{x} - \mathbf{y}\|_{\infty} \quad \text{para todo par } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$$

Si  $Q \subset \mathbb{R}^n$  es un cubo abierto de lado  $l$  tal que  $H \cap Q \neq \emptyset$ , entonces  $\mathbf{g}(H \cap Q)$  está contenido en un cubo  $U \subset \mathbb{R}^n$  de lado  $2Cl$ . (En efecto, si  $\mathbf{a} \in H \cap Q$ , para cada  $\mathbf{x} \in H \cap Q$  se verifica  $\|\mathbf{x} - \mathbf{a}\|_{\infty} < l$ , y así,  $\|\mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{a})\|_{\infty} < Cl$ , luego  $U = B_{\infty}(\mathbf{g}(\mathbf{a}), Cl)$  es un cubo abierto de lado  $2Cl$  que contiene a  $\mathbf{g}(H \cap Q)$ ).

Si  $H$  tiene medida nula, dado  $\epsilon > 0$ , en virtud de J.1 existe una sucesión de cubos cerrados  $(Q_j)$  que cubre  $H$  y verifica  $\sum_j v(Q_j) < (2C)^{-n}\epsilon$ .

Si  $l_j$  es el lado del cubo  $Q_j$  existe un cubo  $U_j$  con lado  $2Cl_j$  que contiene a  $\mathbf{g}(H \cap Q_j)$ . La sucesión de cubos  $(U_j)$  cubre a  $\mathbf{g}(H) = \bigcup_j \mathbf{g}(H \cap Q_j)$ , y verifica:

$$\sum_j v(U_j) = \sum_j (2Cl_j)^n = (2C)^n \sum_j v(Q_j) < \epsilon$$

La demostración del resultado alternativo, para el caso de un conjunto  $H$  de contenido nulo, se deja al cuidado del lector.  $\blacksquare$



**Proposición J.3** Sea  $\mathbf{g} : \Omega \rightarrow \mathbb{R}^n$  de clase  $C^1$  en un abierto  $\Omega \subset \mathbb{R}^n$ .

- a) Si  $H$  tiene contenido nulo y  $\overline{H} \subset \Omega$ , entonces  $\mathbf{g}(H)$  tiene contenido nulo.  
 b) Si  $H \subset \Omega$  tiene medida nula, entonces  $\mathbf{g}(H)$  tiene medida nula.

DEM: a) Si  $H$  tiene contenido nulo entonces  $\overline{H}$  es compacto (porque es cerrado y acotado). La hipótesis  $\overline{H} \subset \Omega$  permite asegurar que la familia de los cubos abiertos  $Q$  con  $\overline{Q} \subset \Omega$  es un cubrimiento abierto de  $\overline{H}$  del que se puede extraer un subrecubrimiento finito  $Q_1, Q_2, \dots, Q_r$ . Como  $\mathbf{g}$  es de clase  $C^1(\Omega)$ , la función  $\mathbf{x} \rightarrow \|d\mathbf{g}(\mathbf{x})\|$  es continua en  $\Omega$ , y por lo tanto está acotada sobre cada cubo compacto  $\overline{Q}_j$ , ( $1 \leq j \leq r$ ). Según el teorema del incremento finito 5.22  $\mathbf{g}|_{Q_j}$  es lipschitziana y la proposición J.2 nos dice que cada  $\mathbf{g}(H \cap Q_j)$  tiene contenido nulo, luego  $\mathbf{g}(H) = \cup_{j=1}^r \mathbf{g}(H \cap Q_j)$  tiene contenido nulo.

b) Sea  $(C_k)$  una sucesión de compactos cuya unión es  $\Omega$ , por ejemplo

$$C_k = \{\mathbf{x} \in \Omega : d(\mathbf{x}, \Omega^c) \geq 1/k\} \cap \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_\infty \leq k\}$$

(donde  $d(\mathbf{x}, \Omega^c) = \inf\{\|\mathbf{x} - \mathbf{y}\|_\infty : \mathbf{y} \in \Omega^c\}$ ). Cada conjunto  $H_k = H \cap C_k$  tiene medida nula y  $\overline{H_k} \subset C_k \subset \Omega$  es compacto. Razonando como en el apartado a), pero manejando series en vez de sumas finitas, se obtiene que cada  $\mathbf{g}(H_k)$  tiene medida nula, de donde se sigue que  $\mathbf{g}(H) = \cup_{k=1}^\infty \mathbf{g}(H_k)$  tiene medida nula. ■

NOTA: La hipótesis  $\overline{H} \subset \Omega$  en el apartado a) de la proposición J.3 es esencial: La función  $g(x) = \text{Arc tg } x$ , definida en  $\Omega = (-\pi/2, \pi/2)$  transforma el conjunto  $H = \{\pi/2 - 1/n : n \in \mathbb{N}\}$ , que tiene contenido nulo, en un conjunto que no tiene contenido nulo porque no es acotado.

**Ejercicio J.4** Sea  $\mathbf{g} : S \rightarrow \mathbb{R}^n$  una aplicación lipschitziana en un rectángulo cerrado  $S \subset \mathbb{R}^n$  con constante de Lipschitz  $\mu$ , es decir:

$$\|\mathbf{g}(\mathbf{y}) - \mathbf{g}(\mathbf{x})\|_\infty \leq \mu \|\mathbf{y} - \mathbf{x}\|_\infty \quad \text{para todo } \mathbf{x}, \mathbf{y} \in S$$

Entonces  $\mathbf{g}(S)$  está contenido en un conjunto medible de contenido  $\leq 2\mu^n v(S)$ .

DEM: Si  $S$  es un cubo, de centro  $\mathbf{a}$  y lado  $l = 2r$  el resultado es inmediato:

$$\mathbf{x} \in S \Rightarrow \|\mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{a})\|_\infty \leq \mu \|\mathbf{x} - \mathbf{a}\|_\infty \leq \mu r$$

luego  $\mathbf{g}(S)$  está contenido en el cubo de centro  $\mathbf{g}(\mathbf{a})$  y lado  $2\mu r$ , cuyo volumen es  $\mu^n (2r)^n = \mu^n v(S)$ . De aquí se sigue el resultado para el caso de un rectángulo cerrado  $S$  con lados de longitud racional, porque, mediante una partición apropiada, lo podemos descomponer en cubos. Un rectángulo cerrado arbitrario  $S$  lo podemos cubrir con un rectángulo cerrado  $R$  de lados racionales y volumen  $v(R) \leq 2v(S)$ , y según lo que acabamos de ver  $\mathbf{g}(R)$  se puede cubrir con un conjunto medible Jordan de contenido  $\mu^n v(R) \leq 2\mu^n v(S)$ . ■

**Ejercicio J.5** Sea  $H \subset \mathbb{R}^n$  y  $\mathbf{g} : H \rightarrow \mathbb{R}^m$  una aplicación lipschitziana. Utilice el ejercicio 10.5.10 para demostrar las siguientes afirmaciones:

- Si  $n < m$ ,  $\mathbf{g}(H)$  tiene medida nula.
- Si  $n < m$  y  $H$  es acotado,  $\mathbf{g}(H)$  tiene contenido nulo.

**Ejercicio J.6** Sea  $\Omega \subset \mathbb{R}^n$  abierto y  $\mathbf{g} : \Omega \rightarrow \mathbb{R}^m$  una aplicación de clase  $C^1(\Omega)$ . Justifique las siguientes afirmaciones:

- Si  $n < m$  entonces  $\mathbf{g}(\Omega)$  tiene medida nula.
- Si  $n < m$ ,  $H$  es acotado, y  $\overline{H} \subset \Omega$ , entonces  $\mathbf{g}(H)$  tiene contenido nulo.

Muestre un ejemplo que ponga de manifiesto que no se cumple d) cuando la condición  $\overline{H} \subset \Omega$  se sustituye por  $H \subset \Omega$ .

### Transformaciones lineales de conjuntos medibles.

**Lema J.7** Toda aplicación lineal no singular  $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , se puede expresar como composición de aplicaciones lineales elementales  $E$  de los siguientes tipos

- $E(\mathbf{x}) = \mathbf{y}$ , donde hay una coordenada  $i \in \{1, 2, \dots, n\}$ , y un número real  $\lambda \neq 0$ , tales que  $y_i = \lambda x_i$ ,  $y_j = x_j$  para todo  $j \neq i$ .
- $E(\mathbf{x}) = \mathbf{y}$ , donde hay un par de coordenadas  $i, k \in \{1, 2, \dots, n\}$  tales que  $y_k = x_i$ ,  $y_i = x_k$ ,  $y_j = x_j$  para cada  $j \notin \{i, k\}$ .
- $E(\mathbf{x}) = \mathbf{y}$ , donde hay un par de coordenadas  $i, k \in \{1, 2, \dots, n\}$  tales que  $y_k = x_k + x_i$ ,  $y_j = x_j$  para cada  $j \neq k$ .

DEM: Sea  $\mathcal{D}$  la clase de las aplicaciones lineales  $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$  que se pueden expresar como composición de aplicaciones lineales elementales de los tipos considerados en el enunciado. Es fácil ver que pertenecen a  $\mathcal{D}$  las aplicaciones lineales  $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$  que son de la forma

- $L(\mathbf{x}) = \mathbf{y}$ , donde hay un par de coordenadas  $i, k \in \{1, 2, \dots, n\}$  y un número real  $\lambda \neq 0$  tales que  $y_k = x_k + \lambda x_i$ ,  $y_j = x_j$ , para cada  $j \neq k$ .

Una aplicación lineal no singular  $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$  queda determinada mediante los vectores linealmente independientes  $\mathbf{v}_j = T^{-1}(\mathbf{e}_j)$ ,  $1 \leq j \leq n$ . Como el vector  $\mathbf{v}_1 = (v_{11}, v_{12}, \dots, v_{1n})$  no es nulo, para algún  $k \in \{1, 2, \dots, n\}$  es  $v_{1k} \neq 0$ . Entonces, con aplicaciones lineales de los tipos  $\alpha$  y  $\beta$  el vector  $\mathbf{v}_1$  se puede transformar en un vector de la forma  $\mathbf{v}'_1 = (1, v'_{12}, \dots, v'_{1n})$  el cual, con aplicaciones lineales de tipo  $\delta$  se transforma en  $\mathbf{e}_1$ . Queda establecido así que existe  $A_1 \in \mathcal{D}$  tal que  $A_1(\mathbf{v}_1) = \mathbf{e}_1$ .

El vector  $A_1(\mathbf{v}_2) = \mathbf{u}_2 = (u_{21}, u_{22}, \dots, u_{2n})$  tiene alguna componente no nula  $u_{2j} \neq 0$ , con  $j \geq 2$  (en caso contrario  $A_1(\mathbf{v}_1) = \mathbf{e}_1$  y  $A_1(\mathbf{v}_2) = t\mathbf{e}_1$  lo que es imposible porque  $A_1$  es no singular y  $\mathbf{v}_1, \mathbf{v}_2$  son linealmente independientes). Procediendo como antes, con aplicaciones lineales de los tipos  $\alpha$  y  $\beta$  que dejen fija la primera coordenada podemos transformar  $\mathbf{u}_2$  en un vector de la forma  $\mathbf{u}'_2 = (u'_{21}, 1, u'_{23}, \dots, u'_{2n})$ , el cual, con aplicaciones lineales de tipo  $\delta$  que dejen fija la primera coordenada, se puede

transformar en  $\mathbf{e}_2$ . Puesto que las transformaciones lineales usadas en esta etapa dejan fijo el vector  $\mathbf{e}_1$  queda justificado que existe  $A_2 \in \mathcal{D}$  tal que  $A_2 \circ A_1(\mathbf{v}_1) = \mathbf{e}_1$ , y  $A_2 \circ A_1(\mathbf{v}_2) = \mathbf{e}_2$ . Continuando con el proceso se obtienen  $A_3, \dots, A_n \in \mathcal{D}$  tales que para todo  $j \in \{1, 2, \dots, n\}$  es  $A_n \circ A_{n-1} \circ \dots \circ A_2 \circ A_1(\mathbf{v}_j) = \mathbf{e}_j = T(\mathbf{v}_j)$ , luego  $T = A_n \circ \dots \circ A_2 \circ A_1 \in \mathcal{D}$ . ■

**Teorema J.8** *Si  $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$  es lineal y  $M \subset \mathbb{R}^n$  es medible Jordan entonces  $T(M)$  es medible Jordan y  $c_n(T(M)) = |\det T| c_n(M)$ .*

DEM: i) Comenzamos suponiendo  $\det T = 0$ . En este caso  $T(\mathbb{R}^n)$  tiene medida nula porque está contenido en un hiperplano  $H$  (esto es consecuencia directa del ejercicio 10.5.11). Si  $M \subset \mathbb{R}^n$  es medible Jordan,  $\overline{M}$  es compacto y  $T(\overline{M}) \subset H$  es un compacto de medida nula, luego tiene contenido nulo. Se sigue que  $T(M) \subset T(\overline{M})$  tiene contenido nulo, luego es medible Jordan y se cumple la igualdad  $c_n(T(M)) = 0 = |\det T| c_n(M)$ .

ii) En el caso  $\det T \neq 0$  para ver que  $T$  transforma conjuntos medibles Jordan en conjuntos medibles Jordan utilizaremos que  $T$  es lipschitziana,

$$\|T(\mathbf{x}) - T(\mathbf{y})\| \leq \|T\| \|\mathbf{x} - \mathbf{y}\| \quad \text{para cada } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$$

y la proposición J.2 que nos dice que  $T$  transforma conjuntos de contenido nulo en conjuntos de contenido nulo. Como  $T$  y su inversa  $T^{-1}$  son continuas (por ser lineales)  $T$  es un homeomorfismo que establece una biyección entre la frontera de  $M \subset \mathbb{R}^n$  y la frontera de su imagen, es decir  $\partial T(M) = T(\partial M)$ . Si  $\partial M$  tiene contenido nulo también tiene contenido nulo su imagen  $\partial T(M) = \mathbf{T}(\partial M)$ , luego, en virtud del teorema 10.26,  $T(M)$  es medible Jordan si  $M$  lo es.

Estableceremos la igualdad  $c_n(T(M)) = |\det T| c_n(M)$  en varias etapas:

*Primera etapa:* Si  $S \subset \mathbb{R}^n$  es un rectángulo cerrado y  $E : \mathbb{R}^n \rightarrow \mathbb{R}^n$  es una aplicación lineal elemental de las consideradas en el lema J.7 se cumple la igualdad

$$c_n(E(S)) = |\det E| c_n(S)$$

i) Si  $E$  es de tipo  $(\alpha)$ , de la forma  $E(x_1, x_2, \dots, x_i, \dots, x_n) = (x_1, x_2, \dots, \lambda x_i, \dots, x_n)$ , entonces  $|\det E| = |\lambda|$ , y es inmediato que  $c_n(E(S)) = |\lambda| c_n(S) = |\det E| c_n(S)$ .

ii) Si  $E$  es de tipo  $(\beta)$ , es claro que  $|\det E| = 1$ , y  $c_n(E(S)) = c_n(S)$ .

iii) Finalmente, cuando  $E$  es de tipo  $(\gamma)$ , de la forma

$$E(x_1, x_2, \dots, x_k, \dots, x_n) = (x_1, x_2, \dots, x_k + x_i, \dots, x_n)$$

es fácil ver que  $|\det E| = 1$ . El contenido de  $W = E(S)$  se puede calcular con el teorema de Fubini en  $\mathbb{R}^n = \mathbb{R}^{n-1} \times \mathbb{R}$  suponiendo, para simplificar la escritura, que  $k = n$ : Si  $W^y = \{t \in \mathbb{R} : (\mathbf{y}, t) \in W\}$ , con  $\mathbf{y} = (x_1, x_2, \dots, x_{n-1}) \in \mathbb{R}^{n-1}$ , se tiene

$$c_n(E(S)) = \int c_1(W^y) d\mathbf{y}$$

Teniendo en cuenta que

$$W^{\mathbf{y}} = \{t \in \mathbb{R} : (\mathbf{y}, t - x_n) \in S\} = \{t \in \mathbb{R} : t - x_n \in S^{\mathbf{y}}\} = x_n + S^{\mathbf{y}}$$

se obtiene que  $c_1(W^{\mathbf{y}}) = c_1(S^{\mathbf{y}})$ , luego

$$c_n(E(S)) = \int c_1(W^{\mathbf{y}}) d\mathbf{y} = \int c_1(S^{\mathbf{y}}) d\mathbf{y} = c_n(S) = |\det E| c_n(S)$$

(Obsérvese que  $S^{\mathbf{y}}$  es un segmento y por lo tanto  $W^{\mathbf{y}} = x_n + S^{\mathbf{y}}$  también lo es).

b) Si  $E : \mathbb{R}^n \rightarrow \mathbb{R}^n$  es una aplicación lineal elemental (de las considerados en el lema J.7) entonces para todo conjunto medible Jordan  $M \subset \mathbb{R}^n$  se cumple

$$c_n(E(M)) \leq |\det E| c_n(M)$$

Fijamos un rectángulo cerrado  $A \supset M$ . Para cada  $p \in \mathcal{P}(A)$  se tiene:

$$E(M) \subset \bigcup_{S \in \Delta'(p)} E(S), \quad \text{donde } \Delta'(p) = \{S \in \Delta(p) : S \cap M \neq \emptyset\}$$

y según 10.8 c) se cumple la desigualdad  $c_n(E(M)) \leq \sum_{S \in \Delta'(p)} c_n(E(S))$ . Según lo demostrado en la etapa a),  $c_n(E(S)) = |\det E| c_n(S)$ , y llegamos a la desigualdad

$$c_n(E(M)) \leq |\det E| \sum_{S \in \Delta'(p)} c_n(S) = |\det E| S(\chi_M, p)$$

válida para cada  $p \in \mathcal{P}(A)$ , luego  $c_n(E(M)) \leq |\det E| \inf\{S(\chi_M, p) : p \in \mathcal{P}(A)\}$ , es decir  $c_n(E(M)) \leq |\det E| c_n(M)$ .

c) Como consecuencia del lema J.7 y de lo establecido en la etapa b), para todo conjunto medible Jordan  $M \subset \mathbb{R}^n$  y toda aplicación lineal no singular  $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$  se cumple la desigualdad  $c_n(T(M)) \leq |\det T| c_n(M)$ .

d) Terminamos viendo que para todo conjunto medible Jordan  $M \subset \mathbb{R}^n$  y toda aplicación lineal no singular  $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$  se verifica  $c_n(T(M)) = |\det T| c_n(M)$ : Aplicando la desigualdad c) al conjunto medible Jordan  $T(M)$  y a la aplicación lineal no singular  $T^{-1}$ , resulta  $c_n(M) \leq |\det T^{-1}| c_n(T(M))$ , y combinando esta desigualdad con la obtenida en c) obtenemos la desigualdad opuesta

$$c_n(T(M)) \leq |\det T| c_n(M) \leq |\det T| |\det T^{-1}| c_n(T(M)) = c_n(T(M))$$

luego  $c_n(T(M)) = |\det T| c_n(M)$ . ■

**Corolario J.9** Si  $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$  es una aplicación lineal que conserva el producto escalar entonces  $|\det T| = 1$ .

DEM:  $T$  es una isometría para la norma euclídea  $\|\cdot\|_2$  luego  $T$  transforma la bola  $B_1 = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| \leq 1\}$  en sí misma y aplicando el teorema J.8 con  $M = B_1$  se obtiene el resultado. ■

## J.2. La demostración del teorema de cambio de variable

**Proposición J.10** *Sea  $\mathbf{g} : \Omega \rightarrow \mathbb{R}^n$  una función de clase  $C^1$  en un abierto  $\Omega \subset \mathbb{R}^n$  y  $M \subset \Omega$  un conjunto medible Jordan tal que  $\overline{M} \subset \Omega$ , y  $\det \mathbf{g}'(\mathbf{x}) \neq 0$  para todo  $\mathbf{x} \in M^\circ$ . Entonces  $\mathbf{g}(M)$  es medible Jordan.*

DEM: Si  $M$  es medible Jordan, su frontera  $\partial M \subset \overline{M} \subset \Omega$  es un conjunto cerrado de contenido nulo y en virtud de la proposición J.3 a), su imagen  $\mathbf{g}(\partial M)$  tiene contenido nulo. Para obtener que  $\mathbf{g}(M)$  es medible Jordan basta ver que su frontera  $\partial \mathbf{g}(M)$  tiene contenido nulo. Esto es consecuencia de la inclusión  $\partial \mathbf{g}(M) \subset \mathbf{g}(\partial M)$  que demostramos a continuación: Según el teorema 8.8 la restricción de  $\mathbf{g}$  al interior de  $M$  es abierta y por lo tanto  $\mathbf{g}(M^\circ)$  es un conjunto abierto, de modo que  $\mathbf{g}(M^\circ) \subset \mathbf{g}(M)^\circ$ . Por otra parte, como  $\overline{M} \subset \Omega$  es compacto también lo es  $\mathbf{g}(\overline{M})$  de donde se sigue que  $\overline{\mathbf{g}(M)} \subset \mathbf{g}(\overline{M})$ . Por continuidad se cumple  $\mathbf{g}(\overline{M}) \subset \overline{\mathbf{g}(M)}$  y queda establecida la igualdad  $\mathbf{g}(\overline{M}) = \overline{\mathbf{g}(M)}$ , con la que se obtiene la inclusión

$$\partial \mathbf{g}(M) = \overline{\mathbf{g}(M)} \setminus \mathbf{g}(M)^\circ = \mathbf{g}(\overline{M}) \setminus \mathbf{g}(M)^\circ \subset \mathbf{g}(\overline{M}) \setminus \mathbf{g}(M^\circ) \subset \mathbf{g}(\partial M)$$

■

**Lema J.11** *Sea  $\mathbf{g} : \Omega \rightarrow \mathbb{R}^n$  una aplicación de clase  $C^1$  en un abierto  $\Omega \subset \mathbb{R}^n$  tal que  $\det \mathbf{g}'(\mathbf{x}) \neq 0$  para todo  $\mathbf{x} \in \Omega$ . Entonces las funciones  $h, k : \Omega \times \Omega \rightarrow \mathbb{R}$ , definidas por  $h(\mathbf{x}, \mathbf{y}) = \|\mathbf{d}\mathbf{g}(\mathbf{y})^{-1} \circ \mathbf{d}\mathbf{g}(\mathbf{x})\|$ ,  $k(\mathbf{x}, \mathbf{y}) = |\det \mathbf{g}'(\mathbf{y})|^{-1} |\det \mathbf{g}'(\mathbf{x})|$ , son continuas en  $\Omega \times \Omega$ .*

(En  $\mathcal{L}(\mathbb{R}^n)$  se considera la norma  $\|L\| = \sup\{\|L(\mathbf{x})\|_\infty : \|\mathbf{x}\|_\infty \leq 1\}$ .)

DEM: Demostramos la continuidad de  $h$ , y dejamos al cuidado del lector el caso más sencillo de la continuidad de la función  $k$ .

En los espacios normados de dimensión finita todas las normas son equivalentes, de donde se sigue que todo isomorfismo algebraico entre dos espacios normados finito dimensionales es un homeomorfismo topológico, luego el espacio normado finito dimensional  $(\mathcal{L}(\mathbb{R}^n), \|\cdot\|)$  se identifica algebraicamente y topológicamente con el espacio  $\mathcal{M}$  formado por las matrices cuadradas  $n \times n$  de números reales, dotado de la topología usual de  $\mathbb{R}^{n^2}$ . En la demostración del teorema 8.12 ya hemos visto que la aplicación  $\det : \mathcal{M} \rightarrow \mathbb{R}$  es continua, luego el conjunto de las matrices invertibles  $\{M \in \mathcal{M} : \det(M) \neq 0\}$  es un subconjunto abierto en  $\mathcal{M}$ , y se sigue de esto que  $\Gamma = \{L \in \mathcal{L}(\mathbb{R}^n) : \det L \neq 0\}$  es un subconjunto abierto de  $\mathcal{L}(\mathbb{R}^n)$  y que la aplicación  $\text{Inv} : \Gamma \rightarrow \Gamma$ ,  $\text{Inv}(L) = L^{-1}$ , es continua.

En virtud de las hipótesis,  $\mathbf{x} \rightarrow \mathbf{d}\mathbf{g}(\mathbf{x})$  es una aplicación continua definida en  $\Omega$  con valores en  $\Gamma$  (véase la proposición E.1), y según la regla de la cadena también es continua la aplicación  $\mathbf{y} \rightarrow \mathbf{d}\mathbf{g}(\mathbf{y})^{-1} = [\text{Inv} \circ \mathbf{d}\mathbf{g}](\mathbf{y})$ . Queda establecida así la continuidad de  $\Psi : \Omega \times \Omega \rightarrow \mathcal{L}(\mathbb{R}^n) \times \mathcal{L}(\mathbb{R}^n)$ , definida por  $\Psi(\mathbf{x}, \mathbf{y}) = (\mathbf{d}\mathbf{g}(\mathbf{y})^{-1}, \mathbf{d}\mathbf{g}(\mathbf{x}))$ .

Por otra parte, es fácil comprobar que si  $L_1, L_2 \in \mathcal{L}(\mathbb{R}^n)$ , y  $L = L_1 \circ L_2$ , entonces  $\|L\| \leq \|L_1\| \|L_2\|$ , de donde se sigue, con un razonamiento estandar, que la operación

de composición  $C : \mathcal{L}(\mathbb{R}^n) \times \mathcal{L}(\mathbb{R}^n) \rightarrow \mathcal{L}(\mathbb{R}^n)$ ,  $C(L_1, L_2) = L_1 \circ L_2$  es continua. Invocando otra vez la regla de la cadena obtenemos la continuidad de la función  $h(\mathbf{x}, \mathbf{y}) = \|C \circ \Psi(\mathbf{x}, \mathbf{y})\|$ . ■

**Lema J.12** Sea  $\mathbf{g} : \Omega \rightarrow \mathbb{R}^n$  una aplicación inyectiva de clase  $C^1$  en un abierto  $\Omega \subset \mathbb{R}^n$  tal que  $\det \mathbf{g}'(\mathbf{x}) \neq 0$  para todo  $\mathbf{x} \in \Omega$ . Si  $A \subset \Omega$  es un rectángulo cerrado entonces  $\mathbf{g}(A)$  es medible Jordan y

$$c_n(\mathbf{g}(A)) \leq \int_A |\det \mathbf{g}'(\mathbf{x})| \, d\mathbf{x}$$

DEM: Según la proposición J.10 el conjunto  $\mathbf{g}(A)$  es medible Jordan, y la demostración de la desigualdad del enunciado la desglosamos en dos etapas.

a) En la primera etapa suponemos que  $A$  es un cubo cerrado. En este caso, dado  $\epsilon > 0$  existe  $p \in \mathcal{P}(A)$  que descompone al cubo  $A$  en cubos cerrados  $\{S_1, S_2, \dots, S_m\}$ , del mismo lado, tales que si  $\mathbf{a}_j$  es el centro de  $S_j$ , y  $T_j = d\mathbf{g}(\mathbf{a}_j)$ , se verifica

- i)  $c_n(\mathbf{g}(S_j)) \leq (1 + \epsilon)^n |\det T_j| c_n(S_j)$
- ii)  $(1 - \epsilon) |\det T_j| c_n(S_j) \leq \int_{S_j} |\det \mathbf{g}'(\mathbf{x})| \, d\mathbf{x}$ .

Para demostrar esto consideramos las funciones

$$h(\mathbf{x}, \mathbf{y}) = \|d\mathbf{g}(\mathbf{y})^{-1} \circ d\mathbf{g}(\mathbf{x})\|, \quad k(\mathbf{x}, \mathbf{y}) = |\det \mathbf{g}'(\mathbf{y})|^{-1} |\det \mathbf{g}'(\mathbf{x})|$$

que según el lema J.11 son continuas en  $\Omega \times \Omega$ , y por lo tanto uniformemente continuas sobre el compacto  $A \times A \subset \mathbb{R}^n \times \mathbb{R}^n$ . Por lo tanto existe  $\delta > 0$  tal que

$$\left\{ \begin{array}{l} \|\mathbf{x} - \mathbf{a}\|_\infty < \delta \quad \mathbf{x}, \mathbf{a} \in A \\ \|\mathbf{y} - \mathbf{b}\|_\infty < \delta \quad \mathbf{y}, \mathbf{b} \in A \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} |h(\mathbf{x}, \mathbf{y}) - h(\mathbf{a}, \mathbf{b})| < \epsilon \\ |k(\mathbf{x}, \mathbf{y}) - k(\mathbf{a}, \mathbf{b})| < \epsilon \end{array} \right\}$$

En particular, cuando  $\mathbf{a} = \mathbf{b} = \mathbf{y} \in A$ ,  $\mathbf{x} \in A$ , resulta

$$\|\mathbf{x} - \mathbf{a}\|_\infty < \delta \Rightarrow |h(\mathbf{x}, \mathbf{a}) - 1| < \epsilon, \quad |k(\mathbf{x}, \mathbf{a}) - 1| < \epsilon,$$

luego, si  $\mathbf{x}, \mathbf{a} \in A$  con  $\|\mathbf{x} - \mathbf{a}\|_\infty < \delta$  se cumple  $h(\mathbf{x}, \mathbf{a}) < 1 + \epsilon$ , y  $k(\mathbf{x}, \mathbf{a}) > 1 - \epsilon$ .

Sea  $p \in \mathcal{P}(A)$  una subdivisión del cubo  $A$  en cubos  $\{S_1, S_2, \dots, S_m\}$  del mismo lado  $2r < \delta$ , y centros  $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m\}$ . Si  $\mathbf{x} \in S_j$  se cumple  $\|\mathbf{x} - \mathbf{a}_j\|_\infty < \delta$ , luego

- i')  $\mathbf{x} \in S_j \Rightarrow \|T_j^{-1} \circ d\mathbf{g}(\mathbf{x})\| = h(\mathbf{x}, \mathbf{a}_j) < 1 + \epsilon$ .
- ii')  $\mathbf{x} \in S_j \Rightarrow |\det \mathbf{g}'(\mathbf{a}_j)|^{-1} |\det \mathbf{g}'(\mathbf{x})| = k(\mathbf{x}, \mathbf{a}_j) > 1 - \epsilon$

En lo que sigue escribimos  $\mathbf{g}_j = T_j^{-1} \circ \mathbf{g}$ . De acuerdo con la regla de la cadena,  $d\mathbf{g}_j(\mathbf{x}) = T_j^{-1} \circ d\mathbf{g}(\mathbf{x})$ , y según i') para todo  $\mathbf{x} \in S_j$  se cumple  $\|d\mathbf{g}_j(\mathbf{x})\| < 1 + \epsilon$ . Entonces, con el teorema del incremento finito se obtiene que para todo  $\mathbf{x} \in S_j$  vale la desigualdad  $\|\mathbf{g}_j(\mathbf{x}) - \mathbf{g}_j(\mathbf{a}_j)\| \leq (1 + \epsilon) \|\mathbf{x} - \mathbf{a}_j\|$  que nos dice que  $\mathbf{g}_j(S_j)$

está contenido en un cubo de centro  $\mathbf{g}_j(\mathbf{a}_j)$ , con lado de longitud  $\leq (1 + \epsilon)$  lado( $S_j$ ), luego

$$c_n(\mathbf{g}_j(S_j)) \leq (1 + \epsilon)^n c_n(S_j)$$

Según el teorema J.8  $c_n(\mathbf{g}_j(S_j)) = c_n(T_j^{-1}(\mathbf{g}(S_j))) = |\det T_j|^{-1} c_n(\mathbf{g}(S_j))$ , y con la última desigualdad se obtiene i):

$$c_n(\mathbf{g}(S_j)) \leq |\det T_j| (1 + \epsilon)^n c_n(S_j)$$

Por otra parte, según i") para todo  $\mathbf{x} \in S_j$  se cumple  $|\det \mathbf{g}'(\mathbf{x})| \geq (1 - \epsilon) |\det T_j|$ , e integrando sobre  $S_j$  se obtiene ii).

Con las desigualdades i) y ii) se llega a la desigualdad

$$c_n(\mathbf{g}(S_j)) \leq \frac{(1 + \epsilon)^n}{1 - \epsilon} \int_{S_j} |\det \mathbf{g}'(\mathbf{x})| d\mathbf{x}$$

y teniendo en cuenta que  $\mathbf{g}(A) = \cup_{j=1}^m \mathbf{g}(S_j)$  resulta

$$\begin{aligned} c_n(\mathbf{g}(A)) &\leq \sum_{j=1}^m c_n(\mathbf{g}(S_j)) \leq \frac{(1 + \epsilon)^n}{1 - \epsilon} \sum_{j=1}^m \int_{S_j} |\det \mathbf{g}'(\mathbf{x})| d\mathbf{x} = \\ &= \frac{(1 + \epsilon)^n}{1 - \epsilon} \int_A |\det \mathbf{g}'(\mathbf{x})| d\mathbf{x} \end{aligned}$$

Si en esta desigualdad pasamos al límite cuando  $\epsilon \rightarrow 0$  se obtiene la desigualdad

$$c_n(\mathbf{g}(A)) \leq \int_A |\det \mathbf{g}'(\mathbf{x})| d\mathbf{x}$$

b) En la segunda etapa suponemos que  $A$  es un rectángulo cerrado. Si las longitudes de todos sus lados son racionales, existe una partición  $q \in \mathcal{P}(A)$  tal que cada  $S \in \Delta(q)$  es un cubo. Utilizando lo demostrado en la primera etapa se obtiene

$$c_n(\mathbf{g}(A)) \leq \sum_{S \in \Delta(q)} c_n(\mathbf{g}(S)) \leq \sum_{S \in \Delta(q)} \int_S |\det \mathbf{g}'(\mathbf{x})| d\mathbf{x} = \int_A |\det \mathbf{g}'(\mathbf{x})| d\mathbf{x}$$

Si  $A \subset \Omega$  es un rectángulo cerrado arbitrario, es fácil ver que existe una sucesión decreciente de rectángulos cerrados  $\Omega \supset A_i \supset A_{i+1} \cdots \supset A$  tal que los lados de cada  $A_i$  son de longitud racional, y  $\lim_i c_n(A_i \setminus A) = \lim_i [c_n(A_i) - c_n(A)] \rightarrow 0$ .

Utilizando que la función continua  $\mathbf{x} \rightarrow |\det \mathbf{g}'(\mathbf{x})|$  está acotada en el compacto  $A_1$  se obtiene fácilmente que  $\lim_i \int_{A_i} |\det \mathbf{g}'(\mathbf{x})| d\mathbf{x} = \int_A |\det \mathbf{g}'(\mathbf{x})| d\mathbf{x}$ . Por lo que ya hemos demostrado sabemos que para cada  $i \in \mathbb{N}$  se cumple la desigualdad

$$c_n(\mathbf{g}(A)) \leq c_n(\mathbf{g}(A_i)) \leq \int_{A_i} |\det \mathbf{g}'(\mathbf{x})| d\mathbf{x}$$

y pasando al límite obtenemos la desigualdad del enunciado para el caso de un rectángulo cerrado arbitrario  $A \subset \Omega$ . ■

**Teorema J.13** [Primera versión] *Sea  $\mathbf{g} : \Omega \rightarrow \mathbb{R}^n$  una aplicación inyectiva de clase  $C^1$  en un abierto  $\Omega \subset \mathbb{R}^n$ , tal que  $\det \mathbf{g}'(\mathbf{x}) \neq 0$  para cada  $\mathbf{x} \in \Omega$ . Si  $\overline{M} \subset \Omega$  y  $M$  es medible Jordan entonces  $\mathbf{g}(M)$  también es medible Jordan. Además  $f : \mathbf{g}(M) \rightarrow \mathbb{R}$  es integrable Riemann sobre  $\mathbf{g}(M)$  si y sólo si  $(f \circ \mathbf{g})|\det \mathbf{g}'| : M \rightarrow \mathbb{R}$  es integrable Riemann sobre  $M$  y en ese caso*

$$\int_{\mathbf{g}(M)} f(\mathbf{y}) d\mathbf{y} = \int_M f(\mathbf{g}(\mathbf{x})) |\det \mathbf{g}'(\mathbf{x})| d\mathbf{x}$$

DEM: En virtud del teorema 8.8  $U = \mathbf{g}(\Omega)$  es abierto y según el corolario 8.14 la inversa  $\mathbf{g}^{-1} : U \rightarrow \Omega$  es de clase  $C^1$  (es decir,  $\mathbf{g}$  establece un  $C^1$ -difeomorfismo entre  $\Omega$  y su imagen  $U$ ). Ya hemos demostrado en la proposición J.10 que  $E = \mathbf{g}(M)$  es medible Jordan. Es claro que los conjuntos

$$D(f, E) = \{\mathbf{y} \in E : f|_E \text{ es discontinua en } \mathbf{y}\}$$

$$D(f \circ \mathbf{g}, M) = \{\mathbf{x} \in M : (f \circ \mathbf{g})|_M \text{ es discontinua en } \mathbf{x}\}$$

se corresponden mediante  $\mathbf{g} : \Omega \rightarrow U$ , es decir  $\mathbf{g}^{-1}(D(f, E)) = D(f \circ \mathbf{g}, M)$ .

Entonces, según la proposición J.3,  $D(f \circ \mathbf{g}, M)$  tiene medida nula si y sólo si  $D(f, E)$  tiene medida nula y con el teorema 10.27 se obtiene que  $f$  es integrable sobre  $E$  si y sólo si  $(f \circ \mathbf{g})$  es integrable sobre  $M$ . Como  $|\det \mathbf{g}'|$  es una función continua que no se anula sobre  $M$ , y el producto de funciones integrables es integrable se sigue que  $(f \circ \mathbf{g})|\det \mathbf{g}'|$  es integrable sobre  $M$  si y sólo si  $(f \circ \mathbf{g})$  es integrable sobre  $M$ .

Observemos en primer lugar que la fórmula del cambio de variable basta establecerla para el caso de una función  $f \geq 0$ , y esto es lo que supondremos en lo que sigue. Desglosamos la demostración en tres etapas:

*Primera etapa:* Si  $A \subset \Omega$  es un rectángulo cerrado entonces

$$\int_{\mathbf{g}(A)} f(\mathbf{y}) d\mathbf{y} \leq \int_A f(\mathbf{g}(\mathbf{x})) |\det \mathbf{g}'(\mathbf{x})| d\mathbf{x}$$

Dada una partición  $p \in \mathcal{P}(A)$ , como  $\mathbf{g}(A) = \bigcup_{S \in \Delta(p)} \mathbf{g}(S)$ , (unión que en general no es disjunta) obtenemos que  $f\chi_{\mathbf{g}(A)} \leq \sum_{S \in \Delta(p)} f\chi_{\mathbf{g}(S)}$ , luego

$$\int_{\mathbf{g}(A)} f(\mathbf{y}) d\mathbf{y} \leq \sum_{S \in \Delta(p)} \int_{\mathbf{g}(S)} f(\mathbf{y}) d\mathbf{y} \leq \sum_{S \in \Delta(p)} M(f \circ \mathbf{g}, S) c_n(\mathbf{g}(S))$$

donde  $M(f \circ \mathbf{g}, S) = \sup\{f(\mathbf{g}(\mathbf{x})) : \mathbf{x} \in S\}$ . Utilizando la desigualdad establecida en el lema J.12 obtenemos

$$(D1) \quad \int_{\mathbf{g}(A)} f(\mathbf{y}) d\mathbf{y} \leq \sum_{S \in \Delta(p)} \int_S M(f \circ \mathbf{g}, S) |\det \mathbf{g}'(\mathbf{x})| d\mathbf{x}$$

Como la función  $\varphi := (f \circ \mathbf{g})|\det \mathbf{g}'|$  es integrable sobre  $A$ , dado  $\epsilon > 0$  podemos encontrar  $p \in \mathcal{P}(A)$  tal que



$$(D2) \quad S(\varphi, p) = \sum_{S \in \Delta(p)} M(\varphi, S)v(S) \leq \int_A \varphi(\mathbf{x}) \, d\mathbf{x} + \epsilon$$

Por otra parte, como la función continua  $\mathbf{x} \rightarrow \det \mathbf{g}'(\mathbf{x})$  es uniformemente continua sobre el compacto  $A$ , existe  $\delta > 0$  tal que

$$\mathbf{x}, \mathbf{t} \in A, \|\mathbf{x} - \mathbf{t}\|_\infty < \delta \Rightarrow |\det \mathbf{g}'(\mathbf{x}) - \det \mathbf{g}'(\mathbf{t})| < \epsilon/K$$

donde  $K > 0$  es una cota superior del conjunto  $f(\mathbf{g}(A))$ . Podemos suponer que la partición  $p \in \mathcal{P}(A)$  ha sido elegida de modo que  $\text{diam}(S) < \delta$  para cada  $S \in \Delta(p)$ . Entonces, si  $\mathbf{x}, \mathbf{t} \in S \in \Delta(p)$  se cumple  $\|\mathbf{x} - \mathbf{t}\|_\infty < \delta$  lo que lleva consigo la desigualdad

$$|\det \mathbf{g}'(\mathbf{x})| \leq |\det \mathbf{g}'(\mathbf{t})| + \epsilon/K$$

y multiplicando por  $f(\mathbf{g}(\mathbf{t})) \geq 0$  se obtiene que para todo para  $\mathbf{x}, \mathbf{t} \in S$  se cumple

$$f(\mathbf{g}(\mathbf{t}))|\det \mathbf{g}'(\mathbf{x})| \leq f(\mathbf{g}(\mathbf{t}))|\det \mathbf{g}'(\mathbf{t})| + \epsilon \leq M(\varphi, S) + \epsilon$$

Dejando  $\mathbf{x} \in S$  fijo y tomando supremos en  $\mathbf{t} \in S$  se llega a la desigualdad

$$M(f \circ \mathbf{g}, S)|\det \mathbf{g}'(\mathbf{x})| \leq M(\varphi, S) + \epsilon$$

Incorporando esta desigualdad a (D1) y teniendo en cuenta (D2) se llega a

$$\int_{\mathbf{g}(A)} f(\mathbf{y}) \, d\mathbf{y} \leq \sum_{S \in \Delta(p)} (M(\varphi, S) + \epsilon)v(S) \leq \int_A \varphi(\mathbf{x}) \, d\mathbf{x} + \epsilon + \epsilon v(A)$$

Como  $\epsilon > 0$  es arbitrario se obtiene que  $\int_{\mathbf{g}(A)} f(\mathbf{y}) \, d\mathbf{y} \leq \int_A \varphi(\mathbf{x}) \, d\mathbf{x}$ .

*Segunda etapa:* Demostramos ahora que si  $\overline{M} \subset \Omega$  y  $M$  es medible Jordan, entonces

$$\int_{\mathbf{g}(M)} f(\mathbf{y}) \, d\mathbf{y} \leq \int_M f(\mathbf{g}(\mathbf{x}))|\det \mathbf{g}'(\mathbf{x})| \, d\mathbf{x}$$

Fijado un rectángulo cerrado  $A \supset M$ , utilizando que  $\chi_M$  es integrable Riemann sobre  $A$  podemos encontrar, para cada  $k \in \mathbb{N}$ , una subdivisión  $p_k \in \mathcal{P}(A)$  tal que

$$S(\chi_M, p_k) - s(\chi_M, p_k) < 1/k$$

No hay inconveniente en suponer que, de modo recurrente, hemos elegido  $p_{k+1}$  más fina que  $p_k$ . Entonces, para cada  $k \in \mathbb{N}$ , las figuras elementales

$$E_k = \bigcup \{S \in \Delta(p_k) : S \subset M\}, \quad F_k = \bigcup \{S \in \Delta(p_k) : S \cap M \neq \emptyset\}$$

verifican  $E_k \subset E_{k+1} \subset M \subset F_{k+1} \subset F_k$ , y

$$c_n(F_k \setminus E_k) = c_n(F_k) - c_n(E_k) = S(\chi_M, p_k) - s(\chi_M, p_k) < 1/k$$

Como  $\overline{M} \subset \Omega$  se cumple que  $\eta := \text{dist}(\overline{M}, \Omega^c) > 0$  y podemos suponer que la partición  $p_1$  ha sido elegida verificando  $\max\{\text{diam}(S) : S \in \Delta(p_1)\} < \eta$ , y así tenemos

asegurado que  $F_1$  es un subconjunto compacto de  $\Omega$  sobre el que la función continua  $|\det \mathbf{g}'|$  está definida y es acotada. En lo que sigue

$$\alpha = \max\{|\det \mathbf{g}'(\mathbf{x})| : \mathbf{x} \in F_1\} < +\infty; \quad \beta = \sup\{f(\mathbf{y}) : \mathbf{y} \in \mathbf{g}(M)\}$$

$$\Delta'_k = \{S \in \Delta(p_k) : S \subset M\}, \quad \Delta''_k = \{S \in \Delta(p_k) : S \cap M \neq \emptyset\}$$

Como  $f \geq 0$ , y  $\mathbf{g}(E_k) = \cup\{\mathbf{g}(S) : S \in \Delta'_k\}$ , teniendo en cuenta lo demostrado en la primera etapa y la propiedad aditiva de la integral respecto al intervalo resulta

$$\begin{aligned} \int_{\mathbf{g}(E_k)} f(\mathbf{y}) \, d\mathbf{y} &\leq \sum_{S \in \Delta'_k} \int_{g(S)} f(\mathbf{y}) \, d\mathbf{y} \leq \sum_{S \in \Delta'_k} \int_S f(\mathbf{g}(\mathbf{x})) |\det \mathbf{g}'(\mathbf{x})| \, d\mathbf{x} = \\ &= \int_{E_k} f(\mathbf{g}(\mathbf{x})) |\det \mathbf{g}'(\mathbf{x})| \, d\mathbf{x} \leq \int_M f(\mathbf{g}(\mathbf{x})) |\det \mathbf{g}'(\mathbf{x})| \, d\mathbf{x} \end{aligned}$$

La conclusión se obtendrá demostrando que  $\lim_k \int_{\mathbf{g}(E_k)} f(\mathbf{y}) \, d\mathbf{y} = \int_{\mathbf{g}(M)} f(\mathbf{y}) \, d\mathbf{y}$ , y pasando al límite en la desigualdad que acabamos de establecer

$$\int_{\mathbf{g}(E_k)} f(\mathbf{y}) \, d\mathbf{y} \leq \int_M f(\mathbf{g}(\mathbf{x})) |\det \mathbf{g}'(\mathbf{x})| \, d\mathbf{x}$$

Teniendo en cuenta la desigualdad

$$\left| \int_{\mathbf{g}(M)} f - \int_{\mathbf{g}(E_k)} f \right| = \int_{\mathbf{g}(M) \setminus \mathbf{g}(E_k)} f \leq \beta \, c_n(\mathbf{g}(M \setminus E_k))$$

basta demostrar que  $\lim_k c_n(\mathbf{g}(M \setminus E_k)) = 0$ . En virtud de la inclusión

$$M \setminus E_k \subset F_k \setminus E_k \subset \bigcup \{S : S \in \Delta''_k \setminus \Delta'_k\}$$

se cumple  $\mathbf{g}(M \setminus E_k) \subset D_k$ , con  $D_k = \bigcup_{S \in \Delta''_k \setminus \Delta'_k} \mathbf{g}(S)$ , luego

$$c_n(\mathbf{g}(M \setminus E_k)) \leq c_n(D_k) \leq \sum_{S \in \Delta''_k \setminus \Delta'_k} c_n(\mathbf{g}(S))$$

y con el lema J.12 se obtiene

$$\begin{aligned} \sum_{S \in \Delta''_k \setminus \Delta'_k} c_n(\mathbf{g}(S)) &\leq \sum_{S \in \Delta''_k \setminus \Delta'_k} \int_S |\det \mathbf{g}'| = \sum_{S \in \Delta''_k} \int_S |\det \mathbf{g}'| - \sum_{S \in \Delta'_k} \int_S |\det \mathbf{g}'| = \\ &= \int_{F_k} |\det \mathbf{g}'| - \int_{E_k} |\det \mathbf{g}'| = \int_{F_k \setminus E_k} |\det \mathbf{g}'| \leq \alpha c_n(F_k \setminus E_k) \leq \alpha/k \end{aligned}$$

Queda establecida así la desigualdad  $c_n(\mathbf{g}(M \setminus E_k)) \leq \alpha/k$ , con la que se obtiene el resultado deseado:  $\lim_k c_n(\mathbf{g}(M \setminus E_k)) = 0$ .

*Tercera etapa:* Ya hemos visto que en virtud del teorema 8.8 y el corolario 8.14,  $U = \mathbf{g}(\Omega)$  es abierto  $g$  es un  $C^1$ -difeomorfismo entre  $\Omega$  y su imagen  $U$ , luego la

inversa  $\mathbf{g}_1 = \mathbf{g}^{-1} : U \rightarrow \Omega$  cumple que  $\det \mathbf{g}'_1(\mathbf{y}) \neq 0$  para cada  $\mathbf{y} \in U$ . También sabemos que el conjunto  $M_1 = \mathbf{g}(M) \subset U$  es medible Jordan y que  $\overline{M}_1 = \mathbf{g}(\overline{M}) \subset \mathbf{g}(\Omega) = U$ . Aplicando lo demostrado en la segunda etapa, a la función  $f_1(\mathbf{x}) = f(\mathbf{g}(\mathbf{x}))|\det \mathbf{g}'(\mathbf{x})|$ , y al conjunto medible Jordan  $M_1$ , con el cambio de variable  $\mathbf{g}_1$ , se obtiene la desigualdad

$$\int_{\mathbf{g}_1(M_1)} f_1(\mathbf{x}) \, d\mathbf{x} \leq \int_{M_1} f_1(\mathbf{g}_1(\mathbf{y}))|\det \mathbf{g}'_1(\mathbf{y})| \, d\mathbf{y}$$

Como  $f_1(\mathbf{g}_1(\mathbf{y})) = f(\mathbf{y})|\det \mathbf{g}'(\mathbf{g}_1(\mathbf{y}))|$  resulta

$$\int_M f(\mathbf{g}(\mathbf{x}))|\det \mathbf{g}'(\mathbf{x})| \, d\mathbf{x} \leq \int_{\mathbf{g}(M)} f(\mathbf{y})|\det \mathbf{g}'(\mathbf{g}_1(\mathbf{y}))|\det \mathbf{g}'_1(\mathbf{y})| \, d\mathbf{y}$$

Si  $\mathbf{y} = \mathbf{g}(\mathbf{x})$  entonces  $\mathbf{g}_1(\mathbf{y}) = \mathbf{x}$ , y según el teorema 8.11 la matriz  $\mathbf{g}'_1(\mathbf{y})$  es inversa de la matriz  $\mathbf{g}'(\mathbf{x})$ , luego  $\det \mathbf{g}'(\mathbf{g}_1(\mathbf{y}))\det \mathbf{g}'_1(\mathbf{y}) = 1$ , y se obtiene así la desigualdad opuesta a la obtenida en la segunda etapa:

$$\int_M f(\mathbf{g}(\mathbf{x}))|\det \mathbf{g}'(\mathbf{x})| \, d\mathbf{x} \leq \int_{\mathbf{g}(M)} f(\mathbf{y}) \, d\mathbf{y}$$

■

**Teorema J.14** [Segunda versión] *Sea  $\mathbf{g} : \Omega \rightarrow \mathbb{R}^n$  una transformación de clase  $C^1$  definida en un abierto  $\Omega \subset \mathbb{R}^n$ . Sea  $M \subset \mathbb{R}^n$  un conjunto medible Jordan tal que*

- a)  $\overline{M} \subset \Omega$
- b)  $\mathbf{g}$  es inyectiva sobre  $M^\circ$
- c)  $\det \mathbf{g}'(\mathbf{x}) \neq 0$  para cada  $\mathbf{x} \in M^\circ$

Entonces  $\mathbf{g}(M)$  es medible Jordan, y si  $f : \mathbf{g}(M) \rightarrow \mathbb{R}$  es integrable Riemann sobre  $\mathbf{g}(M)$ , la función  $(f \circ \mathbf{g})|\det \mathbf{g}'|$  es integrable Riemann sobre  $M$  y vale la igualdad.

$$\int_{\mathbf{g}(M)} f(\mathbf{y}) \, d\mathbf{y} = \int_M f(\mathbf{g}(\mathbf{x}))|\det \mathbf{g}'(\mathbf{x})| \, d\mathbf{x}$$

DEM: Según la proposición J.10 el conjunto  $\mathbf{g}(M)$  es medible Jordan. Fijado un rectángulo cerrado  $A \supset M$ , para cada  $\epsilon > 0$  existe una partición  $p \in \mathcal{P}(A)$  tal que

$$E = \{S \in \Delta(p) : S \subset M^\circ\}, \quad F = \{S \in \Delta(p) : S \cap \overline{M} \neq \emptyset\}$$

son figuras elementales que verifican  $c_n(F \setminus E) < \epsilon$ , y  $E \subset M^\circ \subset \overline{M} \subset F$  (recuérdese que al ser  $M$  medible Jordan es  $c_n(\partial M) = 0$ , de donde se sigue que  $M^\circ$  y  $\overline{M}$  son medibles Jordan, y  $c_n(M^\circ) = c_n(M) = c_n(\overline{M})$ ).

Aplicando el teorema J.13 al cambio de variable  $\mathbf{g}|_{M^\circ}$  y al conjunto cerrado medible Jordan  $E \subset M^\circ$  se obtiene la igualdad

$$i) \int_{\mathbf{g}(E)} f(\mathbf{y}) \, d\mathbf{y} = \int_E f(\mathbf{g}(\mathbf{x})) |\det \mathbf{g}'(\mathbf{x})| \, d\mathbf{x}$$

Como  $F$  es compacto, y la función  $\mathbf{x} \rightarrow \|\mathbf{d}\mathbf{g}(\mathbf{x})\|$  es continua podemos asegurar que  $\mu = \max\{\|\mathbf{d}\mathbf{g}(\mathbf{x})\| : \mathbf{x} \in F\} < +\infty$ . En virtud del teorema del incremento finito, para cada rectángulo cerrado  $S \in \Delta(p)$  con  $S \subset F$ , se cumple

$$\mathbf{x}, \mathbf{y} \in S \Rightarrow \|\mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{y})\|_\infty \leq \mu \|\mathbf{x} - \mathbf{y}\|_\infty$$

luego, según el ejercicio J.4,  $\mathbf{g}(S)$  está contenido en un conjunto medible Jordan de contenido  $\leq 2\mu^n v(S)$ . Como  $F \setminus E$  está incluido en una unión finita de tales rectángulos  $S$  con suma de volúmenes  $< \epsilon$ , se sigue que  $c_n(\mathbf{g}(F \setminus E)) \leq 2\mu^n \epsilon$ .

Si  $\alpha = \max\{|\det \mathbf{g}'(\mathbf{x})| : \mathbf{x} \in F\}$ , y  $\beta = \sup\{|f(\mathbf{x})| : \mathbf{y} \in \mathbf{g}(M)\}$ , se verifica

$$ii) \left| \int_{\mathbf{g}(M)} f(\mathbf{y}) \, d\mathbf{y} - \int_{\mathbf{g}(E)} f(\mathbf{y}) \, d\mathbf{y} \right| \leq \beta c_n(\mathbf{g}(M \setminus E)) \leq \beta c_n(\mathbf{g}(F \setminus E)) \leq \beta 2\mu^n \epsilon.$$

$$iii) \left| \int_M (f \circ \mathbf{g}) |\det \mathbf{g}'| - \int_E (f \circ \mathbf{g}) |\det \mathbf{g}'| \right| \leq \alpha \beta c_n(M \setminus E) \leq \alpha \beta c_n(F \setminus E) \leq \alpha \beta \epsilon.$$

Combinando i), ii) y iii) y teniendo en cuenta que  $\epsilon > 0$  es arbitrario se obtiene

$$\int_{\mathbf{g}(M)} f(\mathbf{y}) \, d\mathbf{y} = \int_M f(\mathbf{g}(\mathbf{x})) |\det \mathbf{g}'(\mathbf{x})| \, d\mathbf{x}$$

■

Para terminar vemos como la fórmula del cambio de variable se extiende al caso de integrales de Riemann impropias (de funciones absolutamente integrables, y de funciones no negativas localmente integrables).

**Teorema J.15** [Tercera versión] *Sea  $\mathbf{g} : \Omega \rightarrow U$  un  $C^1$ -difeomorfismo entre dos abiertos  $\Omega, U \subset \mathbb{R}^n$ , y  $f : U \rightarrow \mathbb{R}$  una función localmente integrable Riemann sobre  $U$ . Entonces  $(f \circ \mathbf{g}) |\det \mathbf{g}'| : \Omega \rightarrow \mathbb{R}$  es localmente integrable Riemann sobre  $\Omega$  y si  $f \geq 0$  se cumple la igualdad (eventualmente con valor  $= +\infty$ )*

$$\int_U f(\mathbf{u}) \, d\mathbf{u} = \int_\Omega f(\mathbf{g}(\mathbf{x})) |\det \mathbf{g}'(\mathbf{x})| \, d\mathbf{x}$$

*Si la función  $f : U \rightarrow \mathbb{R}$  es absolutamente integrable Riemann sobre  $U$  entonces  $(f \circ \mathbf{g}) |\det \mathbf{g}'|$  es absolutamente integrable Riemann sobre  $\Omega$  y la igualdad anterior se cumple con valor finito.*

DEM: El homeomorfismo  $\mathbf{g} : \Omega \rightarrow U$  establece una biyección  $K \rightarrow \mathbf{g}(K)$  entre los subconjuntos compactos de  $\Omega$  y los subconjuntos compactos de  $U$ . En virtud de la proposición J.10, en esta biyección se corresponden los compactos medibles Jordan, es decir (con la notación del capítulo 12)  $K \in \mathcal{K}_\Omega$  si y sólo si  $H = \mathbf{g}(K) \in \mathcal{K}_U$ . Más aún, es fácil comprobar que si  $K_j \in \mathcal{K}_\Omega$  es una sucesión expansiva en  $\Omega$  entonces  $H_j = \mathbf{g}(K_j) \in \mathcal{K}_U$  es una sucesión expansiva en  $U$ .

Si  $f : U \rightarrow \mathbb{R}$  es localmente integrable sobre  $U$ , para cada  $K \in \mathcal{K}_\Omega$ ,  $f$  es integrable sobre  $H = \mathbf{g}(K) \in \mathcal{K}_U$ , y con el teorema J.13 se obtiene que  $(f \circ \mathbf{g}) |\det \mathbf{g}'|$  es integrable sobre  $K \in \mathcal{K}_\Omega$ , cumpliéndose la igualdad

$$\int_H f(\mathbf{u}) \, d\mathbf{u} = \int_K f(\mathbf{g}(\mathbf{x})) |\det \mathbf{g}'(\mathbf{x})| \, d\mathbf{x}$$

Así queda demostrado que  $(f \circ \mathbf{g})|\det \mathbf{g}'|$  es localmente integrable sobre  $\Omega$ . Cuando  $f \geq 0$ , tomando supremos en la última igualdad se obtiene que

$$\sup_{H \in \mathcal{K}_U} \int_U f(\mathbf{u}) \, d\mathbf{u} = \sup_{K \in \mathcal{K}_\Omega} \int_K f(\mathbf{g}(\mathbf{x}))|\det \mathbf{g}'(\mathbf{x})| \, d\mathbf{x} \leq +\infty$$

es decir,  $\int_U f(\mathbf{u}) \, d\mathbf{u} = \int_\Omega f(\mathbf{g}(\mathbf{x}))|\det \mathbf{g}'(\mathbf{x})| \, d\mathbf{x} \leq +\infty$ .

Lo que acabamos de demostrar, aplicado a la función  $|f|$ , nos dice que si  $f : U \rightarrow \mathbb{R}$  es absolutamente integrable sobre  $U$ , entonces  $(f \circ \mathbf{g})|\det \mathbf{g}'|$  es absolutamente integrable sobre  $\Omega$ . En este caso la igualdad del enunciado se obtiene pasando al límite en la igualdad

$$\int_{H_j} f(\mathbf{u}) \, d\mathbf{u} = \int_{K_j} f(\mathbf{g}(\mathbf{x}))|\det \mathbf{g}'(\mathbf{x})| \, d\mathbf{x}$$

donde  $K_j \in \mathcal{K}_\Omega$ , es una sucesión expansiva en  $\Omega$  y  $H_j = \mathbf{g}(K_j) \in \mathcal{K}_U$  es la correspondiente sucesión expansiva en  $U$ . ■

# K

## Formas diferenciales

### K.1. Producto mixto y producto vectorial

En lo que sigue  $E$  denotará un espacio euclídeo de dimensión  $k$  (habitualmente será un subespacio vectorial  $k$ -dimensional de  $\mathbb{R}^n$ ,  $1 \leq k \leq n$ , con el producto escalar inducido por el producto escalar usual de  $\mathbb{R}^n$ ). Fijada una base ortonormal  $\beta = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$  en  $E$ , este espacio euclídeo queda identificado con  $\mathbb{R}^k$ , mediante la aplicación lineal

$$T_\beta : \mathbb{R}^k \rightarrow E, \quad T_\beta(\mathbf{x}) = \sum_{j=1}^k x_j \mathbf{u}_j$$

Obsérvese que  $T_\beta$  conserva el producto escalar y por lo tanto es una isometría lineal.

Un conjunto  $M \subset E$  se dice que es medible Jordan en  $E$  cuando  $M_\beta = T_\beta^{-1}(M)$  es medible Jordan en  $\mathbb{R}^k$ , y en ese caso se define  $c_E(M) = c_k(M_\beta)$ . Esta definición no depende de la base ortonormal  $\beta$  fijada en  $E$ : Si  $\beta' = \{\mathbf{u}'_1, \mathbf{u}'_2, \dots, \mathbf{u}'_k\}$  es otra base ortonormal de  $E$ , la aplicación lineal  $T = T_\beta^{-1} \circ T_{\beta'} : \mathbb{R}^k \rightarrow \mathbb{R}^k$  es una isometría que conserva el producto escalar, luego  $|\det T| = 1$  (esto, que es un resultado bien conocido de geometría euclídea, ha sido establecido en el corolario J.9). Como  $T(M_{\beta'}) = M_\beta$ , en virtud del teorema J.8 se cumple que  $M_\beta$  es medible Jordan si  $M_{\beta'}$  es medible Jordan, y en ese caso  $c_k(M_\beta) = |\det T| c_k(M_{\beta'}) = c_k(M_{\beta'})$ .

En lo que sigue  $\mathcal{M}_E$  será la familia de los conjuntos  $M \subset E$  que son medibles Jordan, y  $c_E : \mathcal{M}_E \rightarrow [0, +\infty)$  el contenido de Jordan en  $E$  que se acaba de definir.

Los resultados recogidos en el siguiente ejercicio, que se obtienen reformulando con las nuevas definiciones resultados conocidos, se dejan al cuidado del lector.

**Ejercicio K.1** Sean  $E, F$  espacios euclídeos de dimensión  $k$ .

- Si  $G \subset E$  es un subespacio propio y  $M \subset G$  es acotado entonces  $M \in \mathcal{M}_E$  y  $c_E(M) = 0$
- Si  $T : E \rightarrow F$  es una aplicación lineal y  $M \in \mathcal{M}_E$  entonces  $T(M) \in \mathcal{M}_F$  y  $c_F(T(M)) = |\det T| c_E(M)$ , donde  $\det T$  es el determinante de  $T$  respecto a una base ortonormal en  $E$  y una base ortonormal en  $F$ .

El paralelepípedo definido por  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k \in E$  es el conjunto

$$P(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) = L(Q)$$

donde  $Q = [0, 1]^k$ , y  $L: \mathbb{R}^k \rightarrow E$  es la aplicación lineal  $L(\mathbf{x}) = \sum_{j=1}^k x_j \mathbf{v}_j$ .

Si  $\beta = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$  es una base ortonormal de  $E$ , y  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k \in E$ , utilizaremos la notación  $\det_\beta(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$  para designar el determinante de la matriz cuadrada  $(\alpha_{ij})_{1 \leq i, j \leq k}$ , formada con las coordenadas de los vectores  $\mathbf{v}_i$  respecto a la base  $\beta$ , es decir  $\mathbf{v}_i = \sum_{j=1}^k \alpha_{ij} \mathbf{u}_j$ , ( $1 \leq i \leq k$ ).

**Proposición K.2** *Si  $E$  es un espacio euclídeo  $k$ -dimensional, y  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k \in E$  entonces  $P(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$  es medible Jordan en  $E$ , y su contenido vale*

$$c_E(P(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)) = |\det_\beta(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)| = \sqrt{|\det(\langle \mathbf{v}_i | \mathbf{v}_j \rangle)_{1 \leq i, j \leq k}|}$$

donde  $\beta = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k)$  es una base ortonormal de  $E$ .

DEM: Supongamos en primer lugar que los vectores  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$  son linealmente independientes. Consideremos las aplicaciones lineales  $L, T_\beta: \mathbb{R}^k \rightarrow E$  definidas por

$$L(\mathbf{x}) = \sum_{j=1}^k x_j \mathbf{v}_j; \quad T_\beta(\mathbf{x}) = \sum_{j=1}^k x_j \mathbf{u}_j$$

Para justificar que  $P := P(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) = L(Q)$  es medible Jordan en  $E$  debemos comprobar que  $T_\beta^{-1}(P)$  es medible Jordan en  $\mathbb{R}^k$ . Según la proposición J.8 esto ocurre porque  $T_\beta^{-1}(P)$  es la imagen de  $Q = [0, 1]^k$  mediante la aplicación lineal  $T = T_\beta^{-1} \circ L$ . Además, teniendo en cuenta la definición de  $c_E$ , según esta proposición

$$c_E(P) = c_k(T_\beta^{-1}(P)) = c_k(T(Q)) = |\det T| c_k(Q) = |\det T|$$

Para terminar debemos ver que  $|\det T|$  vale lo que figura en el enunciado. Según la definición,  $\det_\beta(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$  es el determinante de la matriz  $A = (\alpha_{ij})_{1 \leq i, j \leq k}$ , donde  $\mathbf{v}_i = \sum_{j=1}^k \alpha_{ij} \mathbf{u}_j$ . Obsérvese que  $T_\beta(\sum_{i=1}^k \alpha_{ij} \mathbf{e}_i) = \sum_{i=1}^k \alpha_{ij} \mathbf{u}_j = \mathbf{v}_j$ , luego

$$T(\mathbf{e}_i) = T_\beta^{-1} L(\mathbf{e}_i) = T_\beta^{-1}(\mathbf{v}_i) = (\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{ik}) \in \mathbb{R}^k$$

lo que significa que la matriz de la aplicación lineal  $T = T_\beta^{-1} \circ L: \mathbb{R}^k \rightarrow \mathbb{R}^k$  (respecto a la base canónica de  $\mathbb{R}^k$ ) tiene como columnas las filas de la matriz  $A = (\alpha_{ij})_{1 \leq i, j \leq k}$ , y por lo tanto  $\det T = \det A = \det_\beta(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$ .

Obsérvese que si los vectores  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$  son linealmente dependientes el paralelepípedo  $P(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$  es un conjunto acotado contenido en un subespacio propio de  $E$  y por lo tanto, según el ejercicio K.1, tiene contenido nulo, luego la primera igualdad del enunciado es evidente porque  $\det_\beta(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) = 0$ .

Finalmente, para establecer la segunda igualdad del enunciado basta observar que  $\langle \mathbf{v}_i | \mathbf{v}_j \rangle = \sum_{p=1}^k \alpha_{ip} \alpha_{jp}$ , lo que significa que la matriz  $B = (\langle \mathbf{v}_i | \mathbf{v}_j \rangle)_{1 \leq i, j \leq k}$  coincide

con el producto  $AA^t$  (donde  $A^t$  es la traspuesta de  $A$ ). Se sigue de esto que  $\det B = \det A \det A^t = (\det A)^2$ , y queda establecido que

$$|\det_{\beta}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)| = \sqrt{|\det(\langle \mathbf{v}_i | \mathbf{v}_j \rangle)_{1 \leq i, j \leq k}|}$$

■

**Orientación de un espacio vectorial.** Sea  $E$  un espacio vectorial de dimensión  $k$ , y  $\beta = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k)$ ,  $\beta' = (\mathbf{u}'_1, \mathbf{u}'_2, \dots, \mathbf{u}'_k)$ , bases ordenadas de  $E$ . Sea  $A = (\alpha_{ij})_{1 \leq i, j \leq k}$  la matriz de la aplicación lineal que transforma la base  $\beta$  en la base  $\beta'$ :

$$\mathbf{u}'_i = \sum_{j=1}^k \alpha_{ij} \mathbf{u}_j, \quad 1 \leq i \leq k.$$

Si  $\det A > 0$  se dice que las dos bases tienen la misma orientación. Así queda definida una relación de equivalencia en la familia de las bases ordenadas de  $E$  con la que esta familia queda descompuesta en dos clases de equivalencia. Se dice que el espacio vectorial  $E$  está orientado cuando se ha elegido una de las dos clases de equivalencia que se declara como clase positiva. A la otra clase de equivalencia se le llama clase negativa y define en  $E$  la orientación opuesta. En la práctica, un espacio vectorial se orienta eligiendo una de las dos orientaciones posibles mediante uno de sus representantes, es decir eligiendo una base ordenada  $\beta = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k)$ , como base positiva. La orientación canónica del espacio  $\mathbb{R}^n$  es la definida con la base canónica ordenada en la forma habitual,  $\beta = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ .

Aunque  $\mathbb{R}^n$  tiene una orientación canónica, sin embargo para un subespacio  $k$ -dimensional  $E \subset \mathbb{R}^n$  no hay definida de forma natural una orientación canónica y para algunas de las cuestiones que se estudian más adelante convendrá elegir de forma adecuada las orientaciones de los subespacios que intervienen.

Dado un hiperplano  $E = \{\mathbf{x} \in \mathbb{R}^n : \langle \mathbf{x} | \mathbf{z} \rangle = 0\}$  determinado por un vector normal  $\mathbf{z} \neq 0$ , es fácil comprobar que dos bases  $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{n-1}\}$ ,  $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{n-1}\}$  de  $E$  tienen la misma orientación si y sólo si  $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{n-1}, \mathbf{z}\}$  y  $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{n-1}, \mathbf{z}\}$  son bases de  $\mathbb{R}^n$  con la misma orientación. Esto permite dar la siguiente definición: *La orientación inducida en  $E$  por el vector normal  $\mathbf{z}$*  es la determinada por una base  $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{n-1}\}$  de  $E$  tal que  $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{n-1}, \mathbf{z}\}$  es una base positiva para la orientación canónica de  $\mathbb{R}^n$ . Es claro que para cada  $t > 0$  (resp.  $t < 0$ ) los vectores  $\mathbf{z}$  y  $t\mathbf{z}$  inducen la misma orientación (resp. orientaciones opuestas) en  $E$ .

**Producto mixto.** Sea  $E$  un espacio euclídeo orientado de dimensión  $k$  y  $\beta = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k)$  una base ortonormal positiva. El *producto mixto* de  $k$  vectores ordenados  $(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) \in E^k$ , denotado  $[\mathbf{v}_1 \cdot \mathbf{v}_2 \cdots \mathbf{v}_k]$ , se define como el valor del determinante

$$\det_{\beta}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) = \det(v_{ij})_{1 \leq i, j \leq k}, \quad \text{donde} \quad \sum_{j=1}^k v_{ij} \mathbf{u}_j = \mathbf{v}_i, \quad 1 \leq i \leq k$$

El producto mixto  $[\mathbf{v}_1 \cdot \mathbf{v}_2 \cdots \mathbf{v}_k]$  es no nulo si y sólo si los vectores  $\{\mathbf{v}_i : 1 \leq i \leq k\}$  son linealmente independientes, y en este caso, de acuerdo con la proposición K.2,



si  $(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$  es una base positiva para la orientación de  $E$ , el producto mixto  $[\mathbf{v}_1 \cdot \mathbf{v}_2 \cdots \mathbf{v}_k]$  coincide con el valor del contenido de Jordan en  $E$  del paralelepípedo  $P(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$ , y con el valor opuesto cuando  $(\mathbf{v}_1, \dots, \mathbf{v}_k)$  es una base negativa. Esto pone de manifiesto que el valor de  $[\mathbf{v}_1 \cdot \mathbf{v}_2 \cdots \mathbf{v}_k] = \det_{\beta}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$  es independiente de la base ortonormal positiva  $\beta$  que se ha elegido.

La aplicación  $\Lambda : E^k \rightarrow \mathbb{R}$ ,  $\Lambda(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) = [\mathbf{v}_1 \cdot \mathbf{v}_2 \cdots \mathbf{v}_k]$  recibe el nombre de  $k$ -forma fundamental del espacio euclídeo orientado  $E$ . Aunque para el cálculo explícito de  $[\mathbf{v}_1 \cdot \mathbf{v}_2 \cdots \mathbf{v}_k]$  hay que considerar una base ortonormal positiva de  $E$  sin embargo esta aplicación está definida de modo intrínseco, ya que sólo depende de la estructura euclídea y de la orientación de  $E$ . Conviene hacer notar que no es imprescindible usar siempre la misma base ortonormal positiva de  $E$ , por lo que desde el punto de vista práctico, para el cálculo de un valor concreto  $[\mathbf{v}_1 \cdot \mathbf{v}_2 \cdots \mathbf{v}_k]$ , puede resultar cómodo utilizar una base ortonormal positiva que dependa de  $(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$ .

**Producto vectorial.** Sea  $E$  un espacio euclídeo orientado de dimensión  $k$  y  $\beta = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k)$  una base ortonormal positiva de  $E$ . Dados  $(k-1)$  vectores ordenados  $(\mathbf{v}_2, \mathbf{v}_2, \dots, \mathbf{v}_k)$  de  $E$  consideremos la aplicación  $L : E \rightarrow \mathbb{R}$ , definida por

$$L(\mathbf{x}) = [\mathbf{x} \cdot \mathbf{v}_1 \cdots \mathbf{v}_k] = \det_{\beta}(\mathbf{x}, \mathbf{v}_2, \dots, \mathbf{v}_k)$$

Obsérvese que esta aplicación no depende de la base ortonormal positiva  $\beta$  que hayamos elegido y que, en virtud de las propiedades de los determinantes,  $L$  es lineal. Usando la identificación canónica entre vectores de un espacio euclídeo y formas lineales sobre el mismo (proposición B.8), existe un único vector  $\mathbf{z} \in E$  tal que para todo  $\mathbf{x} \in E$  se cumple  $L(\mathbf{x}) = \langle \mathbf{x} | \mathbf{z} \rangle$ , es decir

$$\langle \mathbf{x} | \mathbf{z} \rangle = [\mathbf{x} \cdot \mathbf{v}_2 \cdots \mathbf{v}_k] = \det_{\beta}(\mathbf{x}, \mathbf{v}_2, \dots, \mathbf{v}_k)$$

Este vector, denotado  $\mathbf{z} = \mathbf{v}_2 \times \mathbf{v}_2 \times \cdots \times \mathbf{v}_k$ , recibe el nombre de *producto vectorial* los  $(k-1)$  vectores ordenados  $(\mathbf{v}_2, \mathbf{v}_2, \dots, \mathbf{v}_k)$ . De la definición se deduce que  $\mathbf{z}$  es ortogonal a los vectores  $\{\mathbf{v}_j : 2 \leq j \leq k\}$ , y que  $\mathbf{z} \neq 0$  si y sólo si estos vectores son linealmente independientes. En este caso, sea  $F \subset E$  el hiperplano de  $E$  generado por los vectores  $\{\mathbf{v}_j : 2 \leq j \leq k\}$ , y  $\mathbf{n}$  el vector unitario ortogonal a  $F$  para el cual  $(\mathbf{n}, \mathbf{v}_2, \dots, \mathbf{v}_k)$  es una base positiva de  $E$ . Como  $\langle \mathbf{n} | \mathbf{z} \rangle = \det_{\beta}(\mathbf{n}, \mathbf{v}_2, \dots, \mathbf{v}_k) > 0$ , se sigue que  $\mathbf{z}$  tiene la dirección y el sentido de  $\mathbf{n}$ , luego  $(\mathbf{z}, \mathbf{v}_2, \dots, \mathbf{v}_k)$  es una base positiva de  $E$ , y de acuerdo con la proposición K.2 y el ejercicio K.3

$$\|\mathbf{z}\|_2 = \langle \mathbf{n} | \mathbf{z} \rangle = \det_{\beta}(\mathbf{n}, \mathbf{v}_2, \dots, \mathbf{v}_k) = c_E(P(\mathbf{n}, \mathbf{v}_2, \dots, \mathbf{v}_k)) = c_F(P(\mathbf{v}_2, \dots, \mathbf{v}_k))$$

es decir, la norma euclídea del producto vectorial  $\mathbf{z} = \mathbf{v}_2 \times \cdots \times \mathbf{v}_k$  es el volumen  $(k-1)$ -dimensional del paralelepípedo generado por los vectores  $(\mathbf{v}_2, \dots, \mathbf{v}_k)$  luego, según la proposición K.2, también se cumple que

$$\|\mathbf{z}\|_2 = \sqrt{|\det(\langle \mathbf{v}_i | \mathbf{v}_j \rangle)_{2 \leq i, j \leq k}|}$$

Si se conocen las coordenadas de los vectores  $\mathbf{v}_j$ ,  $2 \leq j \leq k$  respecto a una base ortonormal positiva  $\beta = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k)$ ,  $\mathbf{v}_i = \sum_{j=1}^k v_{ij} \mathbf{u}_j$ ,  $2 \leq i \leq k$ , para

obtener las coordenadas del producto vectorial  $\mathbf{z} = \mathbf{v}_2 \times \cdots \times \mathbf{v}_k$  respecto a esta base  $\mathbf{z} = \sum_{j=1}^k z_j \mathbf{u}_j$ , basta calcular los productos escalares

$$z_j = \langle \mathbf{u}_j | \mathbf{z} \rangle = [\mathbf{u}_j \cdot \mathbf{v}_2 \cdots \mathbf{v}_k] = \det_{\beta}(\mathbf{u}_j, \mathbf{v}_2, \cdots, \mathbf{v}_k)$$

luego

$$z_j = \begin{vmatrix} 0 & 0 & \cdots & 1 & \cdots & 0 \\ v_{21} & v_{22} & \cdots & v_{2j} & \cdots & v_{2k} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ v_{k1} & v_{k2} & \cdots & v_{kj} & \cdots & v_{kk} \end{vmatrix}$$

es decir, el producto vectorial  $\mathbf{v}_2 \times \cdots \times \mathbf{v}_k$  es el vector  $\mathbf{z}$  que resulta cuando se desarrolla formalmente el determinante

$$\mathbf{z} = \begin{vmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_j & \cdots & \mathbf{u}_k \\ v_{21} & v_{22} & \cdots & v_{2j} & \cdots & v_{2k} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ v_{k1} & v_{k2} & \cdots & v_{kj} & \cdots & v_{kk} \end{vmatrix}$$

A continuación vemos el significado geométrico del valor absoluto de las coordenadas del producto vectorial  $\mathbf{z} = \mathbf{v}_2 \times \cdots \times \mathbf{v}_k$  (respecto a una base ortonormal positiva  $\beta = (\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k)$  de  $E$ ). Por comodidad en la notación razonamos, sin pérdida de generalidad, con la primera coordenada de  $\mathbf{z}$ .

Con la fórmula  $z_1 = \det_{\beta}(\mathbf{u}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k)$  y la proposición K.2 obtenemos que  $|z_1|$  es el contenido en  $E$  del paralelepípedo  $P(\mathbf{u}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k)$ :

$$|z_1| = |\det_{\beta}(\mathbf{u}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k)| = c_E(P(\mathbf{u}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k))$$

Desarrollando el determinante anterior por la primera fila se obtiene que

$$|z_1| = |\det_{\beta'}(\mathbf{v}'_2, \cdots, \mathbf{v}'_k)|$$

donde  $\beta' = (\mathbf{u}_2, \cdots, \mathbf{u}_k)$  y los vectores  $\mathbf{v}'_i = \sum_{j=2}^k v_{ij} \mathbf{u}_j$  son las proyecciones ortogonal de los vectores  $\mathbf{v}_i$  sobre el subespacio  $F_1 = \{\mathbf{x} \in E : \langle \mathbf{x} | \mathbf{u}_1 \rangle = 0\}$ .

Como  $\beta'$  es una base ortonormal de  $F_1$  resulta que  $|z_1|$  también es el contenido en  $F_1$  del paralelepípedo  $P(\mathbf{v}'_2, \cdots, \mathbf{v}'_k) \subset F_1$ , es decir

$$|z_1| = c_{F_1}(P(\mathbf{v}'_2, \cdots, \mathbf{v}'_k))$$

Por otra parte  $z_1 = \langle \mathbf{u}_1 | \mathbf{z} \rangle = \|\mathbf{z}\|_2 \cos \theta$ , donde  $\theta$  es el ángulo agudo formado por las rectas que generan  $\mathbf{u}_1$  y  $\mathbf{z}$ . Obsérvese que esto está de acuerdo con el resultado del ejercicio K.4 según el cual  $c_{F_1}(P(\mathbf{v}'_2, \cdots, \mathbf{v}'_k)) = c_F(P(\mathbf{v}_2, \cdots, \mathbf{v}_k)) \cos \theta$ , ya que  $P(\mathbf{v}'_2, \cdots, \mathbf{v}'_k)$  es la proyección ortogonal sobre  $F_1$  del paralelepípedo  $P(\mathbf{v}_2, \cdots, \mathbf{v}_k)$  y  $\theta$  es el ángulo agudo que forman los hiperplanos  $F$  y  $F_1$ .

Análogamente,  $|z_j|$  es el contenido, en el subespacio  $F_j = \{\mathbf{x} \in E : \langle \mathbf{x} | \int \mathbf{u}_j \rangle = 0\}$ , del la proyección ortogonal de  $P(\mathbf{v}_2, \cdots, \mathbf{v}_k)$  sobre este subespacio.

**Ejercicio K.3** Sean  $\{\mathbf{v}_i : 1 \leq i \leq k\}$  vectores linealmente independientes de un espacio euclídeo  $k$ -dimensional  $E$ , y sea  $E_i \subset E$ , el subespacio  $(k-1)$ -dimensional generado por los vectores  $\{\mathbf{v}_1, \dots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_k\}$ . Se llama base  $i$ -ésima del paralelepípedo  $P(\mathbf{v}_1, \dots, \mathbf{v}_k) \subset E$ , al paralelepípedo de dimensión  $(k-1)$

$$B_i = P(\mathbf{v}_1, \dots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_k) \subset E_i$$

La longitud de la altura correspondiente a esta base es  $h_i = |\langle \mathbf{v}_i | \mathbf{n}_i \rangle|$ , donde  $\mathbf{n}_i \in E$  es un vector unitario ortogonal a  $E_i$ . Demuestre que

$$c_E(P(\mathbf{v}_1, \dots, \mathbf{v}_k)) = c_{E_i}(B_i)h_i$$

(así, para  $k=3$ , el volumen de un paralelepípedo es igual al producto del área de una de sus bases por la longitud de la correspondiente altura).

SOLUCIÓN

Por comodidad en la notación suponemos  $i=1$ . Sea  $\beta_1 = (\mathbf{u}_2, \mathbf{u}_3, \dots, \mathbf{u}_k)$  una base ortonormal de  $E_1$ , y  $\mathbf{u}_1 \in E$  un vector unitario ortogonal a  $E_1$ , de modo que  $\beta = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k)$  es una base ortonormal de  $E$ . Cambiando el signo de  $\mathbf{u}_1$  si es preciso, podemos suponer que  $h_1 = \langle \mathbf{v}_1 | \mathbf{u}_1 \rangle > 0$ , de modo que es  $h_1$  es la longitud de la altura que corresponde a la base  $B_1 = P(\mathbf{v}_2, \dots, \mathbf{v}_k)$ .

Descomponiendo  $\mathbf{v}_1 = \mathbf{y}_1 + \mathbf{z}_1$ , con  $\mathbf{y}_1 = \langle \mathbf{v}_1 | \mathbf{u}_1 \rangle \mathbf{u}_1$ ,  $\mathbf{z}_1 = \mathbf{v}_1 - \mathbf{y}_1$ , es claro que  $\langle \mathbf{z}_1 | \mathbf{u}_1 \rangle = 0$ , luego  $\mathbf{z}_1 \in E_1$ . Como  $E_1$  está generado por los vectores  $\{\mathbf{v}_2, \dots, \mathbf{v}_k\}$ , se sigue que  $\det_\beta(\mathbf{z}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) = 0$ , luego

$$\begin{aligned} \det_\beta(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) &= \det_\beta(\mathbf{y}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) + \det_\beta(\mathbf{z}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) = \\ &= \det_\beta(\mathbf{y}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) \end{aligned}$$

La primera fila del último determinante, formada por las coordenadas de  $\mathbf{y}_1$  respecto a la base  $\beta$ , es  $(h_1, 0, \dots, 0)$ , donde  $h_1 = \langle \mathbf{y}_1 | \mathbf{u}_1 \rangle = \langle \mathbf{v}_1 | \mathbf{u}_1 \rangle > 0$  es la longitud de la altura que corresponde a la base  $B_1$ . Para  $i > 1$  la fila  $i$ -ésima de este determinante, formada por las coordenadas de  $\mathbf{v}_i$  respecto a la base  $\beta$  de  $E$ , es de la forma  $(0, \alpha_{i2}, \dots, \alpha_{ik})$ , donde  $(\alpha_{i2}, \dots, \alpha_{ik})$  son las coordenadas respecto a la base  $\beta_1$  de  $E_1$ . Desarrollando el determinante  $\det_\beta(\mathbf{y}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$  por la primera fila se obtiene que su valor es  $h_1 \det_{\beta_1}(\mathbf{v}_2, \dots, \mathbf{v}_k)$  y así queda demostrado que

$$c_E(P(\mathbf{v}_1, \dots, \mathbf{v}_k)) = h_1 c_{E_1}(P(\mathbf{v}_2, \dots, \mathbf{v}_k))$$

**Ejercicio K.4** Sean  $F, G \subset E$  hiperplanos distintos del espacio euclídeo  $k$ -dimensional  $E$ , y  $\theta$  el ángulo agudo que forman estos hiperplanos. Si  $\pi : E \rightarrow G$  es la proyección ortogonal de  $E$  sobre  $G$  y  $M \subset F$  es medible Jordan en  $F$ , entonces  $\pi(M) \subset G$  es medible Jordan en  $G$  y  $c_G(\pi(M)) = c_F(M) \cos \theta$ .

SOLUCIÓN

$F \cap G \subset E$  es un subespacio de dimensión  $(k-2)$ , en el que podemos fijar una base ortonormal  $(\mathbf{u}_3, \dots, \mathbf{u}_k)$ . Existen vectores unitarios  $\mathbf{v} \in F$ ,  $\mathbf{w} \in G$  tales que

$$\beta_F = (\mathbf{v}, \mathbf{u}_3, \dots, \mathbf{u}_k), \quad \beta_G = (\mathbf{w}, \mathbf{u}_3, \dots, \mathbf{u}_k)$$

son bases ortonormales de  $F$  y  $G$  respectivamente. Además, cambiando si es preciso el signo de  $\mathbf{w}$ , podemos suponer que  $0 < \langle \mathbf{v} | \mathbf{w} \rangle = \cos \theta$ , donde  $\theta$  es el ángulo agudo entre los vectores  $\mathbf{v}$ ,  $\mathbf{w}$ .

Sea  $p : F \rightarrow G$  la restricción de la proyección ortogonal  $\pi : E \rightarrow G$ , que viene dada por  $\pi(\mathbf{x}) = \mathbf{x} - \langle \mathbf{x} | \mathbf{n} \rangle \mathbf{n}$ , donde  $\mathbf{n} \in E$  es un vector unitario ortogonal a  $G$ .

La matriz de la aplicación lineal  $p : F \rightarrow G$  respecto a las bases  $\beta_F, \beta_G$  tiene como columnas las coordenadas de  $p(\mathbf{v}), p(\mathbf{u}_3), \dots, p(\mathbf{u}_k)$  respecto a la base  $\beta_G = (\mathbf{w}, \mathbf{u}_3, \dots, \mathbf{u}_k)$ . Las coordenadas de  $p(\mathbf{v})$  vienen dadas por los productos escalares

$$\langle p(\mathbf{v}) | \mathbf{w} \rangle = \langle \mathbf{v} | \mathbf{w} \rangle = \cos \theta; \quad \langle p(\mathbf{v}) | \mathbf{u}_j \rangle = \langle \mathbf{v} | \mathbf{u}_j \rangle = 0, \quad \text{si } j \geq 3$$

Por otra parte, para  $j \geq 3$  es  $p(\mathbf{u}_j) = \mathbf{u}_j$ , de modo que la matriz de la aplicación lineal  $p$  respecto a las bases indicadas es diagonal, con diagonal  $(\cos \theta, 1, 1, \dots, 1)$ . Si  $M \subset F$  es medible Jordan en  $F$ , según el ejercicio K.1 b),  $\pi(M) = p(M) \subset G$  es medible Jordan en  $G$  y se cumple  $c_G(\pi(M)) = |\det p| c_F(M) = \cos \theta c_F(M)$ . ■

## K.2. Formas multilineales alternadas

Sea  $G_k$  el grupo de las permutaciones de  $\{1, 2, \dots, k\}$ . Para cada  $\sigma \in G_k$  sea  $\nu(\sigma)$  el número de parejas  $(i, j)$  tales que  $1 \leq i < j \leq k$  y  $\sigma(i) > \sigma(j)$  (número de inversiones de  $\sigma$ ). La aplicación signatura  $\varepsilon : G_k \rightarrow \{-1, 1\}$ , definida por  $\varepsilon(\sigma) = (-1)^{\nu(\sigma)}$  tiene las siguientes propiedades que la caracterizan:

- i)  $\varepsilon(\sigma) = 1$  si  $\sigma$  es la identidad;
- ii)  $\varepsilon(\sigma) = -1$  si  $\sigma$  es una transposición;
- iii)  $\varepsilon(\sigma\tau) = \varepsilon(\sigma)\varepsilon(\tau)$  para cada  $\sigma, \tau \in G_k$ .

En lo que sigue  $E$  es un espacio vectorial real de dimensión  $n \geq k$ .

**Definición K.5** Una aplicación multilineal  $T : E^k \rightarrow \mathbb{R}$  se dice que es *alternada* (o *antisimétrica*) cuando  $T(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k) = 0$  siempre que  $\mathbf{x}_i = \mathbf{x}_j$  para algún par  $(i, j)$  con  $1 \leq i < j \leq k$ .

Para una aplicación multilineal  $T : E^k \rightarrow \mathbb{R}$  son equivalentes

- a)  $T$  es alternada;
- b)  $T(\mathbf{x}_1, \dots, \mathbf{x}_k) = \varepsilon(\sigma)T(\mathbf{x}_{\sigma(1)}, \dots, \mathbf{x}_{\sigma(k)})$  para cada  $\sigma \in G_k$  y cada  $(\mathbf{x}_1, \dots, \mathbf{x}_k) \in E^k$

Es inmediato que b)  $\Rightarrow$  a). La demostración de a)  $\Rightarrow$  b) basta hacerla cuando  $\sigma$  es una transposición: Dada una pareja de índices  $1 \leq i < j \leq k$ , utilizando a) y el carácter multilineal de  $T$  se obtiene el resultado: Si  $\mathbf{x}'_i = \mathbf{x}_j$ , y  $\mathbf{x}'_j = \mathbf{x}_i$ , resulta

$$0 = T(\mathbf{x}_1, \dots, \mathbf{x}_{i-1}, \mathbf{x}_i + \mathbf{x}_j, \mathbf{x}_{i+1}, \dots, \mathbf{x}_{j-1}, \mathbf{x}_j + \mathbf{x}_i, \mathbf{x}_{j+1}, \dots, \mathbf{x}_k) = \\ T(\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_j, \dots, \mathbf{x}_k) + T(\mathbf{x}_1, \dots, \mathbf{x}'_i, \dots, \mathbf{x}'_j, \dots, \mathbf{x}_k)$$

Se llama *alternada o antisimetrizada* de la aplicación multilineal  $B : E^k \rightarrow \mathbb{R}$  a la aplicación multilineal  $B_A : E^k \rightarrow \mathbb{R}$  definida por

$$B_A(\mathbf{x}_1, \dots, \mathbf{x}_k) = \sum_{\sigma \in G_k} \varepsilon(\sigma) B(\mathbf{x}_{\sigma(1)}, \dots, \mathbf{x}_{\sigma(k)})$$

Se deja al cuidado del lector la comprobación de que  $B_{\mathcal{A}}$  es multilineal alternada, y que  $B$  es alternada si y sólo si  $B_{\mathcal{A}} = k!B$ .

En lo que sigue  $\Gamma_k(E)$  denotará el espacio vectorial formado por todas las aplicaciones multilineales alternadas  $T : E^k \rightarrow \mathbb{R}$ . En particular  $\Gamma_1(E) = E^*$  es el espacio dual, y conviene introducir el convenio  $\Gamma_0(E) = \mathbb{R}$ . A los elementos de  $\Gamma_k(E)$  se les suele llamar  $k$ -formas exteriores sobre  $E$ ,  $k$ -formas de grado  $k$ , y también  $k$ -covectores.

El producto tensorial  $f_1 \otimes f_2 \otimes \cdots \otimes f_k$  de  $k$  formas lineales  $f_1, f_2, \dots, f_k \in E^*$ , es la aplicación multilineal  $B : E^k \rightarrow \mathbb{R}$  definida por

$$B(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k) = f_1(\mathbf{x}_1)f_2(\mathbf{x}_2) \cdots f_k(\mathbf{x}_k)$$

y su producto exterior, denotado  $f_1 \wedge f_2 \wedge \cdots \wedge f_k$ , es la aplicación multilineal alternada  $B_{\mathcal{A}}$  asociada al producto tensorial, es decir

$$(f_1 \wedge f_2 \wedge \cdots \wedge f_k)(\mathbf{x}_1, \dots, \mathbf{x}_k) = \sum_{\sigma \in G_k} \varepsilon(\sigma) f_1(\mathbf{x}_{\sigma(1)}) \cdots f_k(\mathbf{x}_{\sigma(k)}) = \det(f_i(\mathbf{x}_j))_{1 \leq i, j \leq k}$$

Es fácil comprobar que el producto exterior  $(f_1, f_2, \dots, f_k) \rightarrow f_1 \wedge \cdots \wedge f_k$  es una aplicación multilineal alternada, definida en  $(E^*)^k$ , y con valores en  $\Gamma_k(E)$ , es decir, para cada permutación  $\sigma \in G_k$  se cumple

$$f_{\sigma(1)} \wedge f_{\sigma(2)} \wedge \cdots \wedge f_{\sigma(k)} = \varepsilon(\sigma) f_1 \wedge f_2 \wedge \cdots \wedge f_k$$

Se deduce de esto que  $f_1 \wedge f_2 \wedge \cdots \wedge f_k = 0$  cuando hay dos factores iguales o proporcionales y también cuando  $f_1, f_2, \dots, f_k$  son linealmente dependientes. También se puede demostrar la validez del recíproco, de modo que  $f_1 \wedge f_2 \wedge \cdots \wedge f_k \neq 0$  si y sólo si  $f_1, f_2, \dots, f_k$  son linealmente independientes.

Nuestro siguiente objetivo es describir una base de  $\Gamma_k(E)$  asociada a una base  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  de  $E$ . Para ello conviene introducir alguna notación preliminar que abrevie la escritura. Sea  $\mathcal{J}_k$  la familia formada con las sucesiones finitas crecientes de  $k$  números naturales  $\{j_1 < j_2 < \cdots < j_k\}$ , con  $1 \leq j_1 < j_k \leq n$ . Dada una matriz  $A = (a_{ij})_{1 \leq i \leq k, 1 \leq j \leq n}$  con  $k$  filas y  $n$  columnas, si  $J = \{j_1 < j_2 < \cdots < j_k\}$  es un elemento de  $\mathcal{J}_k$  escribiremos  $\Delta_J(A)$  para designar el valor del determinante de la matriz obtenida extrayendo de  $A$  las  $k$  columnas indicadas por los subíndices  $j_1 < j_2 < \cdots < j_k$ ,

$$\Delta_J(A) = \sum_{\sigma \in G_k} \varepsilon(\sigma) a_{1j_{\sigma(1)}} a_{2j_{\sigma(2)}} \cdots a_{kj_{\sigma(k)}}$$

Asociada a una base  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  de  $E$  consideramos en  $E^*$  la base dual

$$\{dx_1, dx_2, \dots, dx_n\}$$

donde  $dx_j$  es la forma lineal  $dx_j(\mathbf{x}) = x_j$  que asignan al vector  $\mathbf{x} = \sum_{j=1}^n x_j \mathbf{e}_j \in E$  su coordenada  $x_j$ . Análogamente, para cada  $J = \{j_1 < j_2 < \cdots < j_k\} \in \mathcal{J}_k$  escribimos  $dx_J$  como una abreviatura del producto exterior  $dx_{j_1} \wedge dx_{j_2} \wedge \cdots \wedge dx_{j_k}$ .

**Teorema K.6** Si  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  es una base de  $E$  y  $\{dx_1, dx_2, \dots, dx_n\}$ , es su base dual en  $E^*$  entonces  $\{dx_J : J \in \mathcal{J}_k\}$  es una base de  $\Gamma_k(E)$

Sea  $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k) \in E^k$  y  $A = (a_{ij})_{1 \leq i \leq k, 1 \leq j \leq n}$  la matriz formada con las coordenadas de los vectores  $\mathbf{x}_i$  respecto a la base  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ , es decir  $\mathbf{x}_i = \sum_{j=1}^n a_{ij} \mathbf{e}_j$ .

Entonces, con la notación abreviada que hemos introducido, para cada  $J \in \mathcal{J}_k$  se verifica  $dx_J(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k) = \Delta_J(A)$ . Dada  $T \in \Gamma_k(E)$  su carácter multilineal nos permite escribir

$$T(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k) = \sum_{1 \leq i_1, i_2, \dots, i_k \leq n} a_{1i_1} a_{2i_2} \cdots a_{ki_k} T(\mathbf{e}_{i_1}, \mathbf{e}_{i_2}, \dots, \mathbf{e}_{i_k})$$

Como  $T$  es alternada, son nulos los sumandos donde intervienen subíndices repetidos, de modo que la suma la podemos suponer extendida a todos los subconjuntos  $\{i_1, i_2, \dots, i_k\} \subset \{1, 2, \dots, n\}$  de  $k$  elementos distintos. Utilizando que  $T$  es alternada podemos asociar todos los sumandos que corresponden a permutaciones  $\sigma$  de un sistema creciente  $J = \{j_1 < j_2 < \dots < j_k\} \in \mathcal{J}_k$ , cuya suma vale

$$\begin{aligned} & \sum_{\sigma \in G_k} a_{1j_{\sigma(1)}} a_{2j_{\sigma(2)}} \cdots a_{kj_{\sigma(k)}} T(\mathbf{e}_{j_{\sigma(1)}}, \mathbf{e}_{j_{\sigma(2)}}, \dots, \mathbf{e}_{j_{\sigma(k)}}) = \\ & = T(\mathbf{e}_{j_1}, \mathbf{e}_{j_2}, \dots, \mathbf{e}_{j_k}) \sum_{\sigma \in G_k} \varepsilon(\sigma) a_{1j_{\sigma(1)}} a_{2j_{\sigma(2)}} \cdots a_{kj_{\sigma(k)}} = \\ & = T(\mathbf{e}_{j_1}, \mathbf{e}_{j_2}, \dots, \mathbf{e}_{j_k}) \Delta_J(A) = T(\mathbf{e}_{j_1}, \mathbf{e}_{j_2}, \dots, \mathbf{e}_{j_k}) dx_J(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k) \end{aligned}$$

Sumando todos estos términos la suma inicial se escribe en la forma

$$T(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k) = \left[ \sum_{J \in \mathcal{J}_k} \alpha_J dx_J \right] (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k)$$

donde  $\alpha_J = T(\mathbf{e}_{j_1}, \mathbf{e}_{j_2}, \dots, \mathbf{e}_{j_k})(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k)$  cuando  $J = \{j_1 < j_2 < \dots < j_k\}$ .

Queda demostrado que  $T$  es una combinación lineal de las formas multilineales  $\{dx_J : J \in \mathcal{J}_k\}$  y para terminar basta verificar que estas formas son linealmente independientes: Si  $0 = \sum_{J \in \mathcal{J}_k} \alpha_J dx_J$  lo hacemos actuar sobre  $(e_{j_1}, e_{j_2}, \dots, e_{j_k})$  con  $(j_1 < j_2 < \dots < j_k) = J \in \mathcal{J}_k$  se obtiene que es nulo el coeficiente  $\alpha_J$ . ■

**Corolario K.7** Si  $E$  es un espacio vectorial real de dimensión  $n$  y  $k \leq n$  entonces  $\Gamma_k(E)$  es un espacio vectorial de dimensión  $\binom{n}{k}$ . (Si  $k > n$  entonces  $\Gamma_k(E) = \{0\}$ ).

DEM: Inmediato ■

Sea  $\beta = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  una base ordenada de  $E$  y  $\{dx_1, dx_2, \dots, dx_n\}$  su base dual en  $E^*$ . Si  $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) \in E^n$  sea  $\det_\beta(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$  el determinante de la matriz  $A = (a_{ij})_{1 \leq i, j \leq n}$  donde  $\mathbf{x}_i = \sum_{j=1}^n a_{ij} \mathbf{e}_j$ ,  $1 \leq i \leq n$ . Entonces

$$dx_1 \wedge dx_2 \wedge \cdots \wedge dx_n(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) = \det_\beta(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$$

Como  $\Gamma_n(E)$  es de dimensión 1, con base  $\{dx_1 \wedge dx_2 \wedge \cdots \wedge dx_n\}$  se sigue que cualquier  $n$ -forma exterior  $T \in \Gamma_n(E)$  es de la forma  $T = \mu dx_1 \wedge dx_2 \wedge \cdots \wedge dx_n$ , con  $\mu \in \mathbb{R}$ , luego  $(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n) \rightarrow \det_\beta(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n)$  es la única forma multilinear alternada de grado  $n$  que toma el valor 1 sobre  $(\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n)$ .

Análogamente a como se define el producto exterior de formas lineales se define el producto exterior de formas multilineales alternadas:

**Definición K.8** Dadas  $S \in \Gamma_p(E)$ ,  $T \in \Gamma_q(E)$ , con  $1 \leq p, q \leq n$ , su producto exterior  $S \wedge T \in \Gamma_{p+q}(E)$  es la forma multilinear alternada asociada a la forma multilinear  $(\mathbf{x}_1, \cdots, \mathbf{x}_p, \mathbf{x}_{p+1}, \cdots, \mathbf{x}_{p+q}) \rightarrow \frac{1}{p!q!} S(\mathbf{x}_1, \cdots, \mathbf{x}_p) T(\mathbf{x}_{p+1}, \cdots, \mathbf{x}_{p+q})$ , es decir

$$(S \wedge T)(\mathbf{x}_1, \cdots, \mathbf{x}_{p+q}) = \frac{1}{p!q!} \sum_{\sigma \in G_{p+q}} \varepsilon(\sigma) S(\mathbf{x}_{\sigma(1)}, \cdots, \mathbf{x}_{\sigma(p)}) T(\mathbf{x}_{\sigma(p+1)}, \cdots, \mathbf{x}_{\sigma(p+q)})$$

Es fácil comprobar que el producto exterior de formas multilineales es lineal respecto a cada factor.

Una forma multilinear alternada  $T \in \Gamma_k(E)$  se dice que es descomponible si se puede expresar como producto exterior de  $k$  formas lineales, es decir, si es de la forma  $T = f_1 \wedge f_2 \wedge \cdots \wedge f_k$  donde  $f_i \in E^*$ . En virtud del teorema K.6 toda forma exterior de grado  $k$  se puede expresar como suma de formas exteriores descomponibles de grado  $k$ .

**Lema K.9** Si  $S = f_1 \wedge f_2 \wedge \cdots \wedge f_p \in \Gamma_p(E)$ ,  $T = g_1 \wedge \cdots \wedge g_2 \wedge \cdots \wedge g_q \in \Gamma_q(E)$  son formas descomponibles, entonces  $S \wedge T = f_1 \wedge f_2 \wedge \cdots \wedge f_p \wedge g_1 \wedge \cdots \wedge g_2 \wedge \cdots \wedge g_q$ .

DEM: Sea  $W = f_1 \wedge f_2 \wedge \cdots \wedge f_p \wedge g_1 \wedge \cdots \wedge g_2 \wedge \cdots \wedge g_q$ .

Sea  $G'$  (resp.  $G''$ ) es subgrupo de  $G_{p+q}$  formado por las permutaciones que sólo permutan los  $p$  primeros (resp.  $q$  últimos) elementos, dejando fijos los restantes. En  $G_{p+q}$  se considera la siguiente relación de equivalencia:  $\sigma \sim \tau$  si  $\sigma = \tau\sigma'\sigma''$  con  $\sigma' \in G'$ ,  $\sigma'' \in G''$ . Cada clase de equivalencia contiene  $p!q!$  elementos, luego hay  $\binom{p+q}{p}$  clases de equivalencia. Si  $D \subset G_{p+q}$  es un conjunto formado con un elemento de cada clase de equivalencia, se verifica:

$$\begin{aligned} W(\mathbf{x}_1, \cdots, \mathbf{x}_{p+q}) &= \sum_{\sigma \in G_{p+q}} \varepsilon(\sigma) f_1(\mathbf{x}_{\sigma(1)}) \cdots f_p(\mathbf{x}_{\sigma(p)}) g_1(\mathbf{x}_{\sigma(p+1)}) \cdots g_q(\mathbf{x}_{\sigma(p+q)}) = \\ &= \sum_{\tau \in D} \varepsilon(\tau) S(\mathbf{x}_{\tau(1)}, \cdots, \mathbf{x}_{\tau(p)}) T(\mathbf{x}_{\tau(p+1)}, \cdots, \mathbf{x}_{\tau(p+q)}) \end{aligned}$$

donde

$$\begin{aligned} S(\mathbf{x}_{\tau(1)}, \cdots, \mathbf{x}_{\tau(p)}) &= \sum_{\sigma' \in G'} \varepsilon(\sigma') f_1(\mathbf{x}_{\tau\sigma'(1)}) \cdots f_p(\mathbf{x}_{\tau\sigma'(p)}) \\ T(\mathbf{x}_{\tau(p+1)}, \cdots, \mathbf{x}_{\tau(p+q)}) &= \sum_{\sigma'' \in G''} \varepsilon(\sigma'') g_1(\mathbf{x}_{\tau\sigma''(1)}) \cdots g_q(\mathbf{x}_{\tau\sigma''(p)}) \end{aligned}$$

Usando que  $p!S = S_{\mathcal{A}}$ ,  $q!T = T_{\mathcal{A}}$  se obtiene que

$$W(\mathbf{x}_1, \cdots, \mathbf{x}_{p+q}) =$$

$$\begin{aligned}
 &= \frac{1}{p!q!} \sum_{\tau \in D} \left[ \sum_{\sigma' \in G'} \varepsilon(\sigma') S(\mathbf{x}_{\tau\sigma'(1)}, \dots, \mathbf{x}_{\tau\sigma'(p)}) \right] \left[ \sum_{\sigma'' \in G''} \varepsilon(\sigma'') T(\mathbf{x}_{\tau\sigma''(p+1)}, \dots, \mathbf{x}_{\tau\sigma''(p+q)}) \right] = \\
 &= \frac{1}{p!q!} \sum_{\sigma \in G_{p+q}} \varepsilon(\sigma) S(\mathbf{x}_{\sigma(1)}, \dots, \mathbf{x}_{\sigma(p)}) T(\mathbf{x}_{\sigma(p+1)}, \dots, \mathbf{x}_{\sigma(p+q)}) = (S \wedge T)(\mathbf{x}_1, \dots, \mathbf{x}_{p+q})
 \end{aligned}$$

■

**Proposición K.10** *El producto exterior de formas exteriores es asociativo y anticommutativo: Si  $R \in \Gamma_k(E)$ ,  $S \in \Gamma_p(E)$ ,  $T \in \Gamma_q(E)$ , entonces*

i)  $R \wedge (S \wedge T) = (R \wedge S) \wedge T.$

ii)  $S \wedge T = (-1)^{pq} T \wedge S.$

DEM: En virtud de la bilinealidad del producto exterior de formas exteriores, basta demostrar las propiedades i) y ii) para formas exteriores descomponibles, y en este caso el resultado es consecuencia del lema K.9. ■

**Correspondencias entre sistemas de vectores y formas exteriores.** En lo que sigue se supone que  $E$  es un espacio euclídeo orientado de dimensión  $n$  y que  $\beta = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$  es una base ortonormal positiva de  $E$ . Recordemos que el producto mixto  $[\mathbf{v}_1 \cdot \mathbf{v}_2 \cdot \dots \cdot \mathbf{v}_n]$  de  $n$  vectores ordenados  $(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n) \in E^n$ , viene dado por  $\det_{\beta}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n) = \det(v_{ij})_{1 \leq i, j \leq n}$ , donde  $\sum_{j=1}^n v_{ij} \mathbf{e}_j = \mathbf{v}_i$ , ( $1 \leq i \leq n$ ). Su valor absoluto proporciona el contenido de Jordan en  $E$  del paralelepípedo  $P(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n)$ . Aunque para calcular el producto mixto hay que fijar en  $E$  una base ortonormal positiva sin embargo el producto mixto es una noción intrínseca que sólo depende de la estructura euclídea y de la orientación de  $E$ .

La  $n$ -forma fundamental del espacio euclídeo orientado  $E$  es la aplicación multilinear alternada  $\Lambda : E^n \rightarrow \mathbb{R}$ ,  $\Lambda(\mathbf{v}_1, \dots, \mathbf{v}_n) = [\mathbf{v}_1 \cdot \mathbf{v}_2 \cdot \dots \cdot \mathbf{v}_n]$ , y con ella podemos asociar a cada sistema  $(\mathbf{u}_1, \dots, \mathbf{u}_p) \in E^p$ ,  $1 \leq p < n$ , la forma exterior  $\xi_p(\mathbf{u}_1, \dots, \mathbf{u}_p) \in \Gamma_k(E)$  de grado  $k = n - p$ , mediante la aplicación  $\xi_p : E^p \rightarrow \Gamma_k(E)$ , definida por

$$\xi_p(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_p)(\mathbf{x}_1, \dots, \mathbf{x}_k) = \Lambda(\mathbf{x}_1, \dots, \mathbf{x}_k, \mathbf{u}_1, \dots, \mathbf{u}_p)$$

Cuando  $p = 0$ , y  $p = n$ , conviene usar los convenios  $E^0 = \mathbb{R}$ ,  $\Gamma_0(E) = \mathbb{R}$ , y denotar por  $\xi_0 : E^0 \rightarrow \Gamma_n(E)$ ,  $\xi_n : E^n \rightarrow \Gamma_0(E)$ , las aplicaciones  $\xi_0(t) = t\Lambda$ , y  $\xi_n = \Lambda$ .

Es inmediato que para  $1 < p \leq n$  la aplicación  $\xi_p$  es multilinear y antisimétrica.

Merece atención especial la aplicación  $\xi_1 : E \rightarrow \Gamma_{n-1}(E)$ , que es un isomorfismo porque los dos espacios vectoriales  $E$  y  $\Gamma_{n-1}(E)$  tienen dimensión  $n$ . La base de  $\Gamma_{n-1}(E)$  es  $\{\mu_i : 1 \leq i \leq n\}$  donde  $\mu_1 = dx_2 \wedge \dots \wedge dx_n$ ,  $\mu_n = dx_1 \wedge \dots \wedge dx_{n-1}$ , y para  $1 < i < n$ ,  $\mu_i = dx_1 \wedge \dots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \dots \wedge dx_n$ .

Según la demostración del teorema K.6, las coordenadas de  $\xi_1(\mathbf{x}) = \sum_{i=1}^n a_i \mu_i$  respecto a esta base vienen dadas por

$$a_i = \xi_1(\mathbf{x})(\mathbf{e}_1, \dots, \mathbf{e}_{i-1}, \mathbf{e}_{i+1}, \dots, \mathbf{e}_n) = \det_{\beta}(\mathbf{e}_1, \dots, \mathbf{e}_{i-1}, \mathbf{e}_{i+1}, \dots, \mathbf{e}_n, \mathbf{x}) =$$



$$(-1)^{n-i} \det_{\beta}(\mathbf{e}_1, \dots, \mathbf{e}_{i-1}, \mathbf{x}, \mathbf{e}_{i+1}, \dots, \mathbf{e}_n) = (-1)^{n-i} x_i$$

Por otra parte sabemos que, en virtud de la estructura euclídea de  $E$ , hay una identificación canónica entre  $E$  y  $E^*$  que se obtiene asociando a cada  $\mathbf{x} \in E$  la forma lineal  $f_{\mathbf{x}} : E \rightarrow \mathbb{R}$  definida por  $f_{\mathbf{x}}(\mathbf{y}) = \langle \mathbf{x} \mid \mathbf{y} \rangle$ . Según su definición, el producto vectorial  $\mathbf{z} = \mathbf{x}_2 \times \mathbf{x}_3 \times \dots \times \mathbf{x}_n$  es el vector que corresponde a la forma lineal  $\mathbf{x} \rightarrow \Lambda(\mathbf{x}, \mathbf{x}_2, \dots, \mathbf{x}_n) = \xi_{n-1}(\mathbf{x}_2, \dots, \mathbf{x}_n)(\mathbf{x})$  luego  $\xi_{n-1}(\mathbf{x}_2, \dots, \mathbf{x}_n) = f_{\mathbf{z}}$ .

**Ejercicio K.11** En las condiciones anteriores, si  $\mathbf{z} = \mathbf{x}_2 \times \dots \times \mathbf{x}_n$ , se verifica

$$\xi_1(\mathbf{z}) = (-1)^{n-1} f_{\mathbf{x}_2} \wedge \dots \wedge f_{\mathbf{x}_n}$$

SOLUCIÓN

Si  $\xi_1(\mathbf{z}) = \sum_{j=1}^n a_j \mu_j$ , hemos visto antes que  $a_i = (-1)^{n-i} z_i$ . Calculemos ahora las coordenadas de  $T := f_{\mathbf{x}_2} \wedge \dots \wedge f_{\mathbf{x}_n} = \sum_{i=1}^n t_i \mu_i$  respecto a la misma base. Según la demostración de K.6 estas coordenadas  $\{t_i : 1 \leq i \leq n\}$  vienen dadas por

$$t_i = T(\mathbf{e}_1, \dots, \mathbf{e}_{i-1}, \mathbf{e}_{i+1}, \dots, \mathbf{e}_n) = \det(f_{\mathbf{x}_p}(e_q))_{p \neq 1; q \neq i}$$

Si  $\mathbf{x}_p = \sum_{j=1}^n x_{pj} \mathbf{e}_j$ , el último determinante se escribe explícitamente así

$$t_i = \begin{vmatrix} x_{21} & \cdots & x_{2,i-1} & x_{2,i+1} & \cdots & x_{2n} \\ x_{31} & \cdots & x_{3,i-1} & x_{3,i+1} & \cdots & x_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ x_{n1} & \cdots & x_{n,i-1} & x_{n,i+1} & \cdots & x_{nn} \end{vmatrix}$$

luego

$$(-1)^{i-1} t_i = \begin{vmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ x_{21} & \cdots & x_{2,i-1} & x_{2i} & x_{2,i+1} & \cdots & x_{2n} \\ x_{31} & \cdots & x_{3,i-1} & x_{3i} & x_{3,i+1} & \cdots & x_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ x_{n1} & \cdots & x_{n,i-1} & x_{ni} & x_{n,i+1} & \cdots & x_{nn} \end{vmatrix}$$

Según hemos visto en el apéndice K.1 el último determinante proporciona la coordenada  $z_i$  del producto vectorial  $\mathbf{z}$ , luego  $a_i = (-1)^{n-i} z_i = (-1)^{n-i} (-1)^{i-1} t_i = (-1)^{n-1} t_i$ , y así queda establecido que  $\xi_1(\mathbf{z}) = (-1)^{n-1} T$ . ■

Cuando  $n = 3$  y  $\mathbf{z} = \mathbf{x}_2 \times \mathbf{x}_3$ , en virtud del ejercicio anterior  $\xi_1(\mathbf{z}) = f_{\mathbf{x}_2} \wedge f_{\mathbf{x}_3}$ , y según la definición del producto vectorial también se cumple que  $\xi_2(\mathbf{x}_2 \times \mathbf{x}_3) = f_{\mathbf{z}}$ .

*Interpretación geométrica de la forma multilineal  $\xi_1(\mathbf{u})$*  Comenzamos con la interpretación geométrica de la forma multilineal  $\xi_1(\mathbf{u})$  cuando  $E = \mathbb{R}^n$  y  $\mathbf{u} \in \mathbb{R}^n$ , es un vector unitario para la norma euclídea  $\|\mathbf{u}\|_2 = 1$ . Para ello consideramos el hiperplano  $H = \{\mathbf{x} \in \mathbb{R}^n : \langle \mathbf{x} \mid \mathbf{u} \rangle = 0\}$  con la orientación inducida por el vector normal  $\mathbf{u}$ , es decir, la orientación definida por una base  $(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{n-1})$  de  $H$  tal que  $(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{n-1}, \mathbf{u})$  es una base positiva para la orientación canónica de  $\mathbb{R}^n$ .

Restringiendo la forma multilineal  $\xi_1(\mathbf{u})$  a vectores de  $H$  obtenemos que si  $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n-1}) \in H^{n-1}$ , son linealmente independientes entonces

$$\xi_1(\mathbf{u})(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n-1}) = \pm c_n(P(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n-1}, \mathbf{u})) = \pm c_H(P(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n-1}))$$

con signo  $+$  (resp.  $-$ ) cuando  $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n-1})$  es una base positiva (resp. negativa) para la orientación considerada en  $H$ . Es decir,  $\xi_1(\mathbf{u})$  restringida al hiperplano  $H$  ortogonal a  $\mathbf{u}$ , orientado mediante el vector normal  $\mathbf{u}$ , actúa sobre los vectores  $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n-1}) \in H^{n-1}$  proporcionando el volumen  $(n-1)$ -dimensional, con signo, del paralelepípedo generado por ellos.

En general, cuando los vectores  $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n-1})$  no están en el hiperplano ortogonal a  $\mathbf{u}$ , si son linealmente independientes, estarán contenidos en un hiperplano  $F = \{\mathbf{x} \in \mathbb{R}^n : \langle \mathbf{x} | \mathbf{n} \rangle = 0\}$ , donde  $\mathbf{n}$  es un vector unitario elegido con la condición de que  $\langle \mathbf{n} | \mathbf{u} \rangle = \cos \theta > 0$ . Entonces el vector  $\mathbf{y} = \mathbf{u} - \langle \mathbf{u} | \mathbf{n} \rangle \mathbf{n}$  es ortogonal a  $\mathbf{n}$  y por lo tanto

$$\xi_1(\mathbf{u})(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n-1}) = \Lambda(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n-1}, \mathbf{u}) = \Lambda(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n-1}, \mathbf{n}) \cos \theta$$

luego  $\xi_1(\mathbf{u})(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n-1})$  es el volumen del paralelepípedo que se obtiene al proyectar  $P(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n-1})$  sobre  $H$ , con signo  $+1$ , (resp.  $-1$ ) si  $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n-1}, \mathbf{n})$  es una base positiva (resp. negativa) para la orientación canónica de  $\mathbb{R}^n$ .

Cuando  $\mathbf{u}$  no es unitario, la interpretación es análoga, sólo que ahora los volúmenes considerados aparecen multiplicados por  $\|\mathbf{u}\|_2$  ya que para  $\mathbf{v} = \mathbf{u} / \|\mathbf{u}\|_2$  se cumple  $\Lambda(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n-1}, \mathbf{u}) = \|\mathbf{u}\|_2 \Lambda(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n-1}, \mathbf{v})$ .

### K.3. Formas diferenciales

La teoría de las formas diferenciables permite dar un tratamiento unificado a diversos resultados del Análisis Vectorial. Las formas diferenciales, que tienen un comportamiento cómodo y flexible frente a los cambios de variable, son objetos matemáticos para los que de forma natural y mecánica se puede definir la integral respecto a una parametrización, independientemente del sistema de coordenadas curvilíneas empleado (siempre que se conserven la orientación). La teoría de las formas diferenciales, además de establecer los fundamentos rigurosos de cierto tipo de cálculos formales que intervienen en los problemas de cambio de variable clarifica y proporciona un tratamiento unificado de los teoremas clásicos del Análisis Vectorial. Por una parte, diversas nociones de origen físico, como el trabajo de un campo de fuerzas, el flujo de un campo de vectores a través de una superficie, son casos particulares de la noción de integral de una forma diferencial respecto a una aplicación. Por otra parte, los operadores diferenciales clásicos como el rotacional, la divergencia y el gradiente son casos particulares de la noción de diferencial exterior, un concepto que se puede definir de modo intrínseco (con independencia del sistema de coordenadas curvilíneas utilizado) usando las identificaciones canónicas entre campos de vectores y formas diferenciales. Un primer resultado que justifica la noción de diferencial exterior es el clásico Lema de Poincaré, que se particulariza

en diversos resultados básicos del Análisis Vectorial, como el que establece las condiciones necesarias y suficientes para que un campo vectorial sea un gradiente. Por último, podemos anunciar que las formas diferenciales son la herramienta idónea para desarrollar el cálculo integral sobre dominios curvos (subvariedades de  $\mathbb{R}^n$ ) y especialmente para establecer la versión general del teorema de Stokes donde interviene de manera decisiva la noción de diferencial exterior. Esta versión general del teorema de Stokes incluye como casos particulares distintos resultados centrales del Análisis Vectorial clásico como los teoremas de Green, Stokes, Gauss etc.

En este capítulo seguimos denotando por  $E$  un espacio euclídeo  $n$ -dimensional en el que se ha fijado una base ordenada  $\beta = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ . Su base dual en  $E^*$  la denotaremos con la notación  $\{dx_1, dx_2, \dots, dx_n\}$  habitual en el cálculo diferencial. La utilidad de esta notación se pondrá de manifiesto más adelante al estudiar los problemas de cambio de variable. Suponemos al lector familiarizado con la teoría de las formas multilineales alternadas que se expone en el apéndice K.2, donde se introducen las notaciones y los resultados básicos requeridos para este capítulo. Así, denotaremos por  $\Gamma_k(E)$  el espacio vectorial formado por las aplicaciones multilineales alternadas  $T : E^k \rightarrow \mathbb{R}$ . En particular  $\Gamma_1(E) = E^*$  es el espacio dual, y por convenio  $\Gamma_0(E) = \mathbb{R}$ . Una base de  $\Gamma_k(E)$  asociada a la base  $\beta = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  de  $E$  es  $\{dx_J : J \in \mathcal{J}_k\}$ , donde  $\mathcal{J}_k$  la familia formada con las sucesiones finitas crecientes de  $k$  números naturales  $\{j_1 < j_2 < \dots < j_k\}$ , con  $1 \leq j_1 < j_k \leq n$  y  $dx_J$  es una abreviatura del producto exterior  $dx_{j_1} \wedge dx_{j_2} \wedge \dots \wedge dx_{j_k}$ .

Si  $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k) \in E^k$  y consideramos la matriz  $A = (a_{ij})_{1 \leq i \leq k, 1 \leq j \leq n}$  formada con las coordenadas de los vectores  $\mathbf{x}_i$  respecto a la base  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ , es decir,  $\mathbf{x}_i = \sum_{j=1}^n a_{ij} \mathbf{e}_j$ , entonces

$$dx_J(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k) = \sum_{\sigma \in G_k} \varepsilon(\sigma) a_{1j_{\sigma(1)}} a_{2j_{\sigma(2)}} \dots a_{kj_{\sigma(k)}}$$

es el determinante de la matriz obtenida extrayendo de  $A$  las  $k$  columnas indicadas por los subíndices  $j_1 < j_2 < \dots < j_k$ ,

**Definición K.12** *Una forma diferencial de grado  $k$ ,  $1 \leq k \leq n$ , (brevemente,  $k$ -forma diferencial) en un abierto  $\Omega \subset E$  es un campo de formas multilineales alternadas de grado  $k$  definido en  $\Omega$ , es decir, es una aplicación  $\omega : \Omega \rightarrow \Gamma_k(E)$ .*

Una forma de grado 0 es una función  $f : \Omega \rightarrow \mathbb{R}$ , y una forma de grado 1 es un campo de formas lineales  $\omega : \Omega \rightarrow E^*$ . Toda forma de grado  $k$  se puede escribir, respecto a la base  $\{dx_J : J \in \mathcal{J}_k\}$  de  $\Gamma_k(E)$  en la forma canónica

$$\omega(\mathbf{x}) = \sum_{1 \leq j_1 < j_2 < \dots < j_k \leq n} \omega_{j_1 < j_2 < \dots < j_k}(\mathbf{x}) dx_{j_1} \wedge dx_{j_2} \wedge \dots \wedge dx_{j_k}$$

que escribiremos más brevemente en la forma

$$\omega(\mathbf{x}) = \sum_{J \in \mathcal{J}_k} \omega_J(\mathbf{x}) dx_J$$

donde  $J = (j_1 < j_2 < \dots < j_k)$ ,  $dx_J = dx_{j_1} \wedge dx_{j_2} \dots \wedge dx_{j_k}$ . Obsérvese que en el caso  $k = n - 1$ , el espacio vectorial  $\Gamma_{n-1}(E)$  es de dimensión 1 por lo que toda forma diferencial de grado  $n - 1$  se escribe en la forma

$$\omega(\mathbf{x}) = \omega_{1,2,\dots,n}(\mathbf{x}) dx_1 \wedge dx_2 \wedge \dots \wedge dx_{n-1}$$

y por lo tanto se puede identificar con la función  $\omega_{1,2,\dots,n} : \Omega \rightarrow \mathbb{R}$ .

Se dice que la forma diferencial  $\omega(\mathbf{x}) = \sum_{J \in \mathcal{J}_k} \omega_J(\mathbf{x}) dx_J$  es de clase  $C^m$  en  $\Omega$  cuando todas las funciones coordenadas  $\omega_J : \Omega \rightarrow \mathbb{R}$  son de clase  $C^m$  en  $\Omega$ . Es fácil comprobar que esta definición no depende de la base  $\beta$  que se ha elegido en  $E$ .

En el caso de ser  $q = 0$ ,  $\omega' = f$  es una función, y se conviene en que  $\omega \wedge f = f \wedge \omega = f\omega$  es la forma diferencial de grado  $p$  definida por  $(f\omega)(\mathbf{x}) := f(\mathbf{x}) \wedge \omega(\mathbf{x})$ .

### Ejemplos K.13

a) Cuando  $n = 1$  la noción de las forma diferenciales carece de interés, pues las formas diferenciales de grado 0 son funciones y las de grado 1, que son de la forma  $f(x)dx$  también se identifican con funciones.

b) Cuando  $n = 2$ , las formas diferenciales de grado 0 son funciones de dos variables, las de grado 1 se suelen escribir usando la notación  $P(x, y)dx + Q(x, y)dy$ , y las de grado 2,  $\omega(x, y) = f(x, y) dx \wedge dy$ , se identifican con funciones de dos variables  $f : \Omega \rightarrow \mathbb{R}$ .

c) Cuando  $n = 3$ , las formas diferenciales de grado 0 son funciones de tres variables, las de grado 1 se escriben en la forma  $F_1(x, y, z)dx + F_2(x, y, z)dy + F_3(x, y, z)dz$ , las de grado 2, se suelen escribir en la forma

$$A(x, y, z) dy \wedge dz + B(x, y, z) dz \wedge dx + C(x, y, z) dx \wedge dy$$

y las de grado 3,  $\omega(x, y, z) = f(x, y, z)dx \wedge dy \wedge dz$ , se identifican con funciones de tres variables  $f : \Omega \rightarrow \mathbb{R}$ .

d) Si  $f : \Omega \rightarrow \mathbb{R}$  es diferenciables en un abierto  $\Omega \subset \mathbb{R}^n$ , su diferencial  $df(\mathbf{x}) = \sum_{j=1}^n D_j f(\mathbf{x}) dx_j$  es una forma diferencial de grado 1 que será de clase  $C^{m-1}$  si  $f$  es de clase  $C^m$ .

El conjunto de las formas diferenciales de grado  $k$  en  $\Omega$  forma un espacio vectorial real, donde está definido el producto exterior de formas diferenciales: Si  $\omega, \omega'$  son formas diferenciales de grados  $p$  y  $q$  respectivamente, entonces su producto exterior  $(\omega \wedge \omega')(\mathbf{x}) : \omega(\mathbf{x}) \wedge \omega'(\mathbf{x})$  es una forma exterior de grado  $k = p + q$ . La multiplicación exterior de formas diferenciales es una operación asociativa y anticonmutativa que verifica la ley de anticonmutatividad  $\omega \wedge \omega' = (-1)^{pq} \omega' \wedge \omega$ .

**Ejemplos K.14**

a) Para  $n = 3$  el producto exterior de las formas diferenciales

$$\omega = F_1 dx + F_2 dy + F_3 dz, \quad \omega' = G_1 dx + G_2 dy + G_3 dz$$

es la forma diferencial

$$\omega \wedge \omega' = (F_2 G_3 - F_3 G_2) dy \wedge dz + (F_3 G_1 - F_1 G_3) dz \wedge dx + (F_1 G_2 - F_2 G_1) dx \wedge dy$$

b) El producto exterior de las forma diferencial de grado 1,

$$\omega = A dy \wedge dz + B dz \wedge dx + C dx \wedge dy$$

por la forma diferencial de grado 2

$$\omega' = A' dx + B' dy + C' dz$$

es la forma diferencial de grado 3,

$$\omega \wedge \omega' = (AA' + BB' + CC') dx \wedge dy \wedge dz$$

c) El producto exterior  $\omega = \omega^1 \wedge \omega^2 \wedge \omega^3$  de las formas de grado 1,

$$\omega^i = F_1^i dx + F_2^i dy + F_3^i dz, \quad 1 \leq i \leq 3$$

es la forma diferencial de grado 3,

$$\omega = \det[F_j^i]_{1 \leq i, j \leq 3} dx \wedge dy \wedge dz$$

c) Si  $f_1, f_2 : \Omega \rightarrow \mathbb{R}$  son funciones diferenciables en un abierto  $\Omega \subset \mathbb{R}^n$ , el producto exterior  $df_1 \wedge df_2 : \Omega \rightarrow \Gamma_2(\mathbb{R}^n)$  es la forma diferencial de grado 2, cuya expresión en forma canónica es

$$(df_1 \wedge df_2)(\mathbf{x}) = \sum_{i < j} \frac{D(f_1, f_2)}{D(x_i, x_j)}(\mathbf{x}) dx_i \wedge dx_j$$

**Proposición K.15** Sean  $f_i : \Omega \rightarrow \mathbb{R}$ ,  $1 \leq i \leq k$  funciones diferenciables en un abierto  $\Omega \subset \mathbb{R}^n$ . Entonces su producto exterior  $df_1 \wedge df_2 \wedge \cdots \wedge df_k : \Omega \rightarrow \Gamma_k(\mathbb{R}^n)$  es la forma diferencial de grado  $k$ , cuya expresión en forma canónica es

$$(df_1 \wedge df_2 \wedge \cdots \wedge df_k)(\mathbf{x}) = \sum_{j_1 < j_2 < \cdots < j_k} \frac{D(f_1, f_2, \cdots, f_k)}{D(x_{j_1}, x_{j_2}, \cdots, x_{j_k})}(\mathbf{x}) dx_{j_1} \wedge dx_{j_2} \wedge \cdots \wedge dx_{j_k}$$

DEM: Basta recordar que si  $J = (j_1 < j_2 < \dots < j_k)$  entonces la coordenada  $\omega_J(\mathbf{x})$  de  $\omega(\mathbf{x}) = (df_1 \wedge df_2 \wedge \dots \wedge df_k)(\mathbf{x})$  respecto a la base canónica  $\{dx_J : J \in \mathcal{J}_k\}$  de  $\Gamma_k(E)$  viene dada por

$$\begin{aligned} \omega_J(\mathbf{x}) &= (df_1 \wedge df_2 \wedge \dots \wedge df_k)(\mathbf{x})(\mathbf{e}_{j_1}, \mathbf{e}_{j_2}, \dots, \mathbf{e}_{j_k}) = \\ &= \det[df_p(\mathbf{x})\mathbf{e}_{j_q}]_{1 \leq p, q \leq k} = \det[D_{j_q}f_p(\mathbf{x})]_{1 \leq p, q \leq k} = \frac{D(f_1, f_2, \dots, f_k)}{D(x_{j_1}, x_{j_2}, \dots, x_{j_k})}(\mathbf{x}) \end{aligned}$$

(véase el apéndice K.2) ■

**Imagen recíproca de una forma diferencial.** Sean  $E, F$  espacios vectoriales euclídeos de dimensiones  $n$  y  $m$  respectivamente y  $\omega : \Omega \rightarrow \Gamma_k(E)$ , una forma diferencial de grado  $k$  definida en un abierto  $\Omega \subset E$ . Si  $\mathbf{g} : U \rightarrow \Omega$  es una aplicación diferenciable en un abierto  $U \subset F$ , la fórmula

$$\mathbf{g}^*(\omega)(\mathbf{u}) : (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) \longrightarrow \omega(\mathbf{g}(\mathbf{u}))(d\mathbf{g}(\mathbf{u})\mathbf{v}_1, \dots, d\mathbf{g}(\mathbf{u})\mathbf{v}_k)$$

define en  $U$  una forma diferencial  $\mathbf{g}^*(\omega) : U \rightarrow \Gamma_k(F)$  de grado  $k$  (obsérvese que los vectores  $\mathbf{v}_j$  pertenecen a  $F$  y sus imágenes  $d\mathbf{g}(\mathbf{u})\mathbf{v}_j$  pertenecen a  $E$ ). Se dice que  $\mathbf{g}^*(\omega)$  es la imagen recíproca de  $\omega$  por el cambio de variable  $\mathbf{g}$ , o que  $\mathbf{g}^*(\omega)$  se deduce de  $\omega$  mediante el cambio de variable  $\mathbf{x} = \mathbf{g}(\mathbf{u})$  (¡Atención  $\mathbf{u}$  y  $\mathbf{x}$ , en general son vectores de dimensiones distintas  $m, n$ ).

Cuando  $\omega = f$  es una función definida en  $\Omega$  (forma diferencial de grado 0) es conveniente definir  $\mathbf{g}^*(f) = f \circ \mathbf{g}$ .

**Proposición K.16** *Sea  $\mathbf{g} : U \rightarrow \Omega$  una aplicación diferenciable en un abierto  $U$  del espacio euclídeo  $m$  dimensional  $F$ , con valores en un abierto  $\Omega$  del espacio euclídeo  $n$ -dimensional  $E$ . Se verifica*

- i)  $\mathbf{g}^*$  es lineal (sobre el espacio vectorial de las  $k$ -formas diferenciales en  $\Omega$ ).
- ii)  $\mathbf{g}^*(df) = d(\mathbf{g}^*f)$  para cada función diferenciable  $f : \Omega \rightarrow \mathbb{R}$ .
- iii)  $\mathbf{g}^*(\omega \wedge \omega') = \mathbf{g}^*(\omega) \wedge \mathbf{g}^*(\omega')$  cuando  $\omega : \Omega \rightarrow \Gamma_p(E)$  y  $\omega' : \Omega \rightarrow \Gamma_q(E)$  son formas diferenciales en  $\Omega$ , de grados  $p$  y  $q$  respectivamente, con  $p + q \leq n$ .

DEM: La propiedad i) es inmediata. y ii) es consecuencia directa de las definiciones:

$$\begin{aligned} \mathbf{g}^*(df)(\mathbf{u})\mathbf{v} &= df(\mathbf{g}(\mathbf{u}))[d\mathbf{g}(\mathbf{u})\mathbf{v}] = \\ &= [df(\mathbf{g}(\mathbf{u})) \circ d\mathbf{g}(\mathbf{u})]\mathbf{v} = d(f \circ \mathbf{g})(\mathbf{u})\mathbf{v} = d(\mathbf{g}^*f)(\mathbf{u})\mathbf{v} \end{aligned}$$

Para demostrar iii) consideremos un punto concreto  $\mathbf{x} = \mathbf{g}(\mathbf{u}) \in \Omega$ , donde  $\mathbf{u} \in U$ , un sistema de vectores  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k \in F$  y sus imágenes  $\mathbf{w}_j = d\mathbf{g}(\mathbf{u})\mathbf{v}_j$ ,  $1 \leq j \leq k$ . Utilizando la definición del producto exterior de formas multilineales (véase K.2) resulta:

$$\mathbf{g}^*(\omega \wedge \omega')(\mathbf{u})(\mathbf{v}_1, \dots, \mathbf{v}_k) = (\omega \wedge \omega')(\mathbf{x})(\mathbf{w}_1, \dots, \mathbf{w}_k) =$$

$$\begin{aligned}
 &= \frac{1}{p!q!} \sum_{\sigma \in G_k} [\omega(\mathbf{x})(\mathbf{w}_{\sigma(1)}, \dots, \mathbf{w}_{\sigma(p)})][\omega'(\mathbf{x})(\mathbf{w}_{\sigma(p+1)}, \dots, \mathbf{w}_{\sigma(p+q)})] = \\
 &= \frac{1}{p!q!} \sum_{\sigma \in G_k} [\mathbf{g}^*(\omega)(\mathbf{u})(\mathbf{v}_{\sigma(1)}, \dots, \mathbf{v}_{\sigma(p)})][\mathbf{g}^*(\omega')(\mathbf{u})(\mathbf{v}_{\sigma(p+1)}, \dots, \mathbf{v}_{\sigma(p+q)})] = \\
 &= [\mathbf{g}^*(\omega) \wedge \mathbf{g}^*(\omega')](\mathbf{u})(\mathbf{v}_1, \dots, \mathbf{v}_k)
 \end{aligned}$$

■

Con el fin de ejercitarse en el cálculo con formas diferenciales, merece la pena comprobar la propiedad K.16 iii) en algunas situaciones concretas

### Ejemplos K.17

a) Comenzamos con una situación particular que es consecuencia inmediata de las definiciones: Si  $f, f_1, \Omega \rightarrow \mathbb{R}$  son funciones, y  $\omega : \Omega \rightarrow \Gamma_k(E)$  una forma diferencial

$$\mathbf{g}^*(f\omega) = \mathbf{g}^*(f)\mathbf{g}^*(\omega); \quad \mathbf{g}^*(ff_1) = \mathbf{g}^*(f)\mathbf{g}^*(f_1);$$

b)  $\mathbf{g}^*(dx_i \wedge dx_j) = dg_i \wedge dg_j$ . Efectivamente, como  $dx_i$  es la aplicación lineal que asigna al  $d\mathbf{g}(\mathbf{u})\mathbf{v} = (dg_1(\mathbf{u})\mathbf{v}, \dots, dg_k(\mathbf{u})\mathbf{v})$  su componente  $dg_i(\mathbf{u})\mathbf{v}$ , resulta

$$\begin{aligned}
 &\mathbf{g}^*(dx_i \wedge dx_j)(\mathbf{u})(\mathbf{v}_1, \mathbf{v}_2) = (dx_i \wedge dx_j)(d\mathbf{g}(\mathbf{u})\mathbf{v}_1, d\mathbf{g}(\mathbf{u})\mathbf{v}_2) = \\
 &= \begin{vmatrix} dx_i(d\mathbf{g}(\mathbf{u})\mathbf{v}_1) & dx_i(d\mathbf{g}(\mathbf{u})\mathbf{v}_2) \\ dx_j(d\mathbf{g}(\mathbf{u})\mathbf{v}_1) & dx_j(d\mathbf{g}(\mathbf{u})\mathbf{v}_2) \end{vmatrix} = \begin{vmatrix} dg_i(\mathbf{u})\mathbf{v}_1 & dg_i(\mathbf{u})\mathbf{v}_2 \\ dg_j(\mathbf{u})\mathbf{v}_1 & dg_j(\mathbf{u})\mathbf{v}_2 \end{vmatrix} = \\
 &= dg_i(\mathbf{u})\mathbf{v}_1 dg_j(\mathbf{u})\mathbf{v}_2 - dg_i(\mathbf{u})\mathbf{v}_2 dg_j(\mathbf{u})\mathbf{v}_1 = [dg_i(\mathbf{u}) \wedge dg_j(\mathbf{u})](\mathbf{v}_1, \mathbf{v}_2)
 \end{aligned}$$

c)  $\mathbf{g}^*(df_1 \wedge df_2) = \mathbf{g}^*(df_1) \wedge \mathbf{g}^*(df_2) = d(\mathbf{g}^*f_1) \wedge d(\mathbf{g}^*f_2)$ .

En efecto, usando la linealidad de  $\mathbf{g}^*$  y lo obtenido en b) se obtiene que la imagen recíproca de la forma diferencial

$$(df_1 \wedge df_2)(\mathbf{x}) = \sum_{i < j} \frac{D(f_1, f_2)}{D(x_i, x_j)}(\mathbf{x}) dx_i \wedge dx_j$$

viene dada por

$$\mathbf{g}^*(df_1 \wedge df_2)(\mathbf{u}) = \sum_{i < j} \frac{D(f_1, f_2)}{D(x_i, x_j)}(\mathbf{g}(\mathbf{u})) dg_i(\mathbf{u}) \wedge dg_j(\mathbf{u})$$

Por otra parte, según K.16 ii), se verifica

$$\mathbf{g}^*(df_1) \wedge \mathbf{g}^*(df_2) = d(\mathbf{g}^*f_1) \wedge d(\mathbf{g}^*f_2)$$

En virtud de la regla de la cadena, para  $i = 1, 2$ , se tiene:

$$d(\mathbf{g}^*f_i)(\mathbf{u}) = d(f_i \circ \mathbf{g})(\mathbf{u}) = df_i \circ d\mathbf{g} = \sum_{j=1}^n D_j f_i(\mathbf{g}(\mathbf{u})) dg_j(\mathbf{u})$$

Sustituyendo arriba esta expresión para  $d(\mathbf{g}^* f_i)$ ,  $i = 1, 2$ , y usando las reglas del álgebra exterior se llega a la igualdad

$$[d(\mathbf{g}^* f_1) \wedge d(\mathbf{g}^* f_2)](\mathbf{u}) = \sum_{i < j} \frac{D(f_1, f_2)}{D(x_1, x_2)}(\mathbf{g}(\mathbf{u})) dg_i(\mathbf{u}) \wedge dg_j(\mathbf{u}) = \mathbf{g}^*(df_1 \wedge df_2)(\mathbf{u})$$

En las condiciones de la proposición K.17, sea  $\{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m\}$  una base ordenada del espacio euclídeo  $F$ , y  $\{du_1, du_2, \dots, du_m\}$  su base dual en  $F^*$ . Dada una forma diferencial  $\omega : \Omega \rightarrow \Gamma_k(E)$ , en forma canónica

$$\omega(\mathbf{x}) = \sum_{1 \leq j_1 < j_2 < \dots < j_k \leq n} \omega_{j_1 j_2 \dots j_k}(\mathbf{x}) dx_{j_1} \wedge dx_{j_2} \wedge \dots \wedge dx_{j_k}$$

para obtener la forma canónica de  $\mathbf{g}^*(\omega)$ , en términos de la base  $\{du_1, du_2, \dots, du_m\}$  basta sustituir

$$dg_{j_1}(\mathbf{u}) \wedge dg_{j_2}(\mathbf{u}) \wedge \dots \wedge dg_{j_k}(\mathbf{u}) = \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq m} \frac{D(g_{j_1}, \dots, g_{j_k})}{D(u_{i_1}, \dots, u_{i_k})}(\mathbf{u}) du_{i_1} \wedge \dots \wedge du_{i_k}$$

en la expresión

$$\mathbf{g}^*(\omega)(\mathbf{u}) = \sum_{1 \leq j_1 < \dots < j_k \leq n} \omega_{j_1 j_2 \dots j_k}(\mathbf{g}(\mathbf{u})) dg_{j_1}(\mathbf{u}) \wedge dg_{j_2}(\mathbf{u}) \wedge \dots \wedge dg_{j_k}(\mathbf{u})$$

con lo que se llega a

$$\mathbf{g}^*(\omega) = \sum_{1 \leq i_1 < \dots < i_k \leq m} \left( \sum_{1 \leq j_1 < \dots < j_k \leq n} \omega_{j_1 j_2 \dots j_k}(\mathbf{g}(\mathbf{u})) \frac{D(g_{j_1}, \dots, g_{j_k})}{D(u_{i_1}, \dots, u_{i_k})}(\mathbf{u}) \right) du_{i_1} \wedge \dots \wedge du_{i_k}$$

Aunque la última fórmula tiene un aspecto aparentemente complicado, en la práctica el cambio de variable se reduce a cálculos mecánicos muy sencillos siguiendo el siguiente esquema: Para realizar el cambio de variable  $\mathbf{x} = \mathbf{g}(\mathbf{u})$  en la forma diferencial de grado  $k$ ,

$$\omega(\mathbf{x}) = \sum_{1 \leq j_1 < \dots < j_k \leq n} \omega_{j_1 j_2 \dots j_k}(\mathbf{x}) dx_{j_1} \wedge dx_{j_2} \wedge \dots \wedge dx_{j_k}$$

se efectúan las sustituciones formales,  $x_j = g_j(u_1, u_2, \dots, u_m)$ ,  $dx_j = dg_j$  y se siguen las reglas formales del álgebra exterior. Esta regla permite dar una doble interpretación a la fórmula canónica para representar una forma diferencial: Si en vez considerar que  $\mathbf{x}$  es una variable independiente, se considera como función de  $\mathbf{u}$ , mediante el cambio de variable  $\mathbf{x} = \mathbf{g}(\mathbf{u})$ , entonces  $dx_j$  se debe interpretar como la diferencial de la función  $x_j = g_j(\mathbf{u})$ ,  $1 \leq j \leq n$ . Debido a esta flexibilidad de la notación, las formas diferenciales proporcionan algoritmos formales de cálculo muy adecuados para los problemas de cambio de variable.

**Ejercicio K.18** En las condiciones de K.16, si  $\mathbf{h} : V \rightarrow U$  es otra aplicación diferenciable definida en un abierto  $V$  de un espacio euclídeo  $r$  dimensional, se verifica

$$(\mathbf{g} \circ \mathbf{h})^* \omega = \mathbf{h}^*(\mathbf{g}^* \omega)$$



**Ejemplos K.19**

a) En el caso  $E = F = \mathbb{R}^2$ , con el cambio de variable a coordenadas polares,  $(x, y) = \mathbf{g}(r, \theta)$ , dado por  $x = r \cos \theta$ ,  $y = r \sin \theta$ , la forma diferencial  $\omega = dx \wedge dy$  se transforma en la que se obtiene con la sustitución formal  $dx = \cos \theta dr - r \sin \theta d\theta$ ,  $dy = \sin \theta dr + r \cos \theta d\theta$ , con la que se obtiene

$$\mathbf{g}^*(\omega)(r, \theta) = (\cos \theta dr - r \sin \theta d\theta) \wedge (\sin \theta dr + r \cos \theta d\theta)$$

Con las reglas de cálculo  $dr \wedge dr = d\theta \wedge d\theta = 0$ ,  $d\theta r \wedge dr = -dr \wedge d\theta$ , se obtiene

$$\mathbf{g}^*(\omega)(r, \theta) = r dr \wedge d\theta$$

b) Consideremos ahora el caso  $E = F = \mathbb{R}^3$ , y el cambio de variable a coordenadas esféricas  $(x, y, z) = \mathbf{g}(\rho, \theta, \varphi)$ , dado por

$$x = \rho \cos \varphi \cos \theta, \quad y = \rho \cos \varphi \sin \theta, \quad z = \rho \sin \varphi$$

Dada la forma diferencial  $\omega = dx \wedge dy \wedge dz$ , para calcular  $\mathbf{g}^*(\omega)(\rho, \theta, \varphi)$  basta efectuar las sustitución formal

$$\begin{aligned} dx &= \cos \varphi \cos \theta d\rho - \rho \cos \varphi \sin \theta d\theta - \rho \sin \varphi \cos \theta d\varphi \\ dy &= \cos \varphi \sin \theta d\rho + \rho \cos \varphi \cos \theta d\theta - \rho \sin \varphi \sin \theta d\varphi \\ dz &= \sin \varphi d\rho + \rho \cos \varphi d\varphi \end{aligned}$$

En una primera etapa, con las reglas del álgebra exterior, obtenemos

$$dx \wedge dy = (\rho \cos^2 \varphi d\rho \wedge d\theta + \rho^2 \sin \varphi \cos \varphi d\theta \wedge d\varphi)$$

luego

$$dx \wedge dy \wedge dz = (\rho \cos^2 \varphi d\rho \wedge d\theta + \rho^2 \sin \varphi \cos \varphi d\theta \wedge d\varphi) \wedge (\sin \varphi d\rho + \rho \cos \varphi d\varphi)$$

y efectuando las operaciones se llega al resultado

$$\mathbf{g}^*(\omega)(\rho, \theta, \varphi) = \rho^2 \cos \varphi d\rho \wedge d\theta \wedge d\varphi$$

c) Consideremos la parametrización usual de la esfera de centro  $(0, 0, 0)$  y radio  $R > 0$ ,  $(x, y, z) = \mathbf{g}(\theta, \varphi)$ , donde

$$x = R \cos \varphi \cos \theta, \quad y = R \cos \varphi \sin \theta, \quad z = R \sin \varphi$$

y una forma diferencial de grado 2

$$\omega(x, y, z) = F_1(x, y, z) dy \wedge dz + F_2(x, y, z) dz \wedge dx + F_3(x, y, z) dx \wedge dy$$

Para calcular  $\mathbf{g}^*(\omega)(\theta, \varphi)$  basta efectuar las sustituciones formales

$$\begin{aligned} dx &= -R \cos \varphi \sin \theta d\theta - R \sin \varphi \cos \theta d\varphi \\ dy &= R \cos \varphi \cos \theta d\theta - R \sin \varphi \sin \theta d\varphi \\ dz &= R \cos \varphi d\varphi \end{aligned}$$

con las que se obtiene  $\mathbf{g}^*(\omega)(\theta, \varphi) = R^2 f(\theta, \varphi) d\theta \wedge d\varphi$ , donde

$$f(\theta, \varphi) = F_1^*(\theta, \varphi) \cos^2 \varphi \cos \theta + F_2^*(\theta, \varphi) \cos^2 \varphi \sin \theta + F_3^*(\theta, \varphi) \sin \varphi \cos \varphi$$

con  $F_j^* = F_j \circ \mathbf{g}$ ,  $1 \leq j \leq 3$ .

# Bibliografía

- [1] J. A. Abia, J. García, and C. Marijuán, *Cálculo Diferencial en  $R^n$* , Los autores, 1998.
- [2] T. A. Apostol, *Análisis Matemático*, seg. ed., Reverté, Barcelona, 1976.
- [3] ———, *Calculus*, seg. ed., Reverté, Barcelona, 1986.
- [4] F. Bombal, L. R. Marín, and G. Vera, *Problemas de Análisis Matemático. Cálculo Diferencial*, vol. 2, AC, 1990.
- [5] ———, *Problemas de Análisis Matemático. Cálculo Integral*, vol. 3, AC, 1990.
- [6] ———, *Problemas de Análisis Matemático. Espacios Métricos y Normados*, vol. 1, AC, 1990.
- [7] G. Choquet, *Topology*, Toray-Masson, 1971.
- [8] G. H. Edwards, Jr, *Advanced Calculus of Several Variables*, Academic Press, 1973.
- [9] J. A. Facenda and F. J. Freniche, *Integración de funciones de varias variables*, Pirámide. Madrid, 2002.
- [10] C. Fernández, Fco. J. Vázquez, and Vegas J. M., *Cálculo Diferencial de Varias Variables*, Thomson, 2002.
- [11] J. A. Fernández Viña, *Análisis Matemático II. (Integración y cálculo exterior)*, Tecnos, 1992.
- [12] ———, *Análisis Matemático II (Topología y Cálculo Diferencial)*, Tecnos, 1992.
- [13] W. H. Fleming, *Functions of Several Variables*, Springer-Verlag, 1997, En CE-SA se publicó una traducción al castellano de la primera edición.
- [14] E. H. Hewitt and K. Stromberg, *Real and Abstract Analysis*, tercera ed., vol. 25, Springer-Verlag. Graduate Text in Math., 1975.
- [15] J. E. Marsden and A. J. Tromba, *Cálculo Vectorial*, Pearson, 1998.
- [16] Descombes R., *Cours d'Analyse pour the certificat de Mathematiques I*, Vuibert, Paris, 1962.

- [17] Jean Saint Raymond, *Local inversion for differentiable functions and the Darboux property*, *Mathematika* **49** (2002), no. 1-2, 141–158 (2004). MR2059050 (2005c:26015)
- [18] L. Schwartz, *Cours d'Analyse I y II*, Hermann, 1967.
- [19] G. Vera, *Temas de Análisis Matemático*, [webs.um.es/gvb/](http://webs.um.es/gvb/), 2008.

[1, 2, 3, 6, 4, 5, 7, 16, 8, 9, 12, 11, 10, 13, 14, 15, 17, 18, 19]