McGRAW-HILL

CÁLCULO INFINITESIMAL DE VARIAS VARIABLES

Juan de Burgos

McGraw Hill

CÁLCULO INFINITESIMAL DE VARIAS VARIABLES

CÁLCULO INFINITESIMAL DE VARIAS VARIABLES

Juan de Burgos Román

Catedrático de Matemática Aplicada
Escuela Superior de Ingenieros Aeronáuticos
Universidad Politécnica de Madrid

McGraw-Hill

MADRID • BUENOS AIRES • CARACAS • GUATEMALA • LISBOA • MEXICO • NUEVA YORK
PANAMA • SAN JUAN • SANTAFE DE BOGOTA • SANTIAGO • SAO PAULO
AUCKLAND • HAMBURGO • LONDRES • MILAN • MONTREAL • NUEVA DELHI • PARIS
SAN FRANCISCO • SIDNEY • SINGAPUR • ST. LOUIS • TOKIO • TORONTO

CÁLCULO INFINITESIMAL DE VARIAS VARIABLES

McGRAW-HILL/INTERAMERICANA DE ESPAÑA, S. A.
Edificio Valrealty, 1.ª planta
Basauri, 17
28023 Aravaca (Madrid)

ISBN: 84-481-1621-6
Depósito legal: M. 31.655-1995

Editora: Isabel Capella
Diseño cubierta: Estudio F. Piñuela
Dibujos artísticos: Raúl Arias
Compuesto en MonoComp, S. A.
Impreso en Lavel, S. A.

IMPRESO EN ESPAÑA - PRINTED IN SPAIN

Contenido

Prólogo

Supone aquí el autor que las gentes nos comportamos como cabría esperar que lo hiciéramos y, consecuente con ello, imagina que, al tomar un libro, lo primero que hará el lector será enterarse de lo que de éste se dice en el prólogo, donde esperará encontrar información buena y fiable acerca de la obra. El autor, creyéndolo así u obligado a admitir dicho supuesto, procura que este primer contacto del lector con el libro incite, a aquél, a inclinarse en favor de éste y, quizá exagerando, pero sin faltar a la verdad, canta las bondades de su obra, por la que, obviamente, tiene la natural inclinación, hablando de aquello que, a su entender, más y mejor contribuya al conocimiento de lo que a continuación se ofrece. La prudencia y el buen juicio aconsejan que un prólogo no sea excesivamente dilatado; por ello y como el autor, muy a su pesar, tiene más verbosidad de lo que es razonable, se ha visto él en la precisión de glosar sólo algunos de los aspectos de la obra, que espera haber seleccionado con acierto.

Para empezar, el autor entiende que debe hablar de algo que, aunque evidente, no ha de ser por ello omitido: ya sabe él que no descubre el Mediterráneo si dice que un curso de Cálculo Infinitesimal de varias variables, como el presente, es continuación lógica del Cáculo Infinitesimal de una variable, pero debe decirlo y señalar, así, que al lector de este texto se le ha de suponer que conoce bien los resultados más sobresalientes del cálculo de una variable. Para aliviar los inconvenientes que todo ello pudiera acarrear, cada vez que, a lo largo de este texto, se acude a alguno de los teoremas o resultados que se suponen ya conocidos, el tal teorema se concreta y resume (en una nota aparte), expresándole del modo que mejor cuadra al uso que de él se va a hacer.

El lector, sin más que echar un simple vistazo al título de este libro, tendrá una información nada despreciable acerca de su contenido. A poco que conozca él del Cálculo Infinitesimal, casi le bastará con haber cursado las matemáticas del Bachillerato, el lector se barrunta que aquí se habla sobre la convergencia (mejor, sobre la topología) en los espacios ordinarios de dimensión mayor que uno (con lo que se manejará más de una variable), sobre los límites y la continuidad de funciones entre estos espacios, sobre la derivación y la diferenciación de tales funciones, sobre las aplicaciones de dichas derivadas al propio cálculo, sobre las integrales múltiples y sobre temas afines a ellas. Y no se equivocará el lector, pues es de ello de lo que trata el presente libro, como se puede comprobar echando una ojeada al índice, con lo que se obtendrá una visión bastante precisa de su contenido y de la intensidad con la que se aborda cada uno de los temas que en él se tratan.

Si bien es verdad que al autor, como a todos los autores, le gustaría tener cuantos más lectores mejor, él sabe que lo que en este texto se dice no interesa a todo el mundo, sino sólo a unas pocas personas, y Dios quiera que éstas no falten. Así, el autor, cuando escribe, se dirige a un tipo muy concreto de lector; el autor escribe para un lector imaginario que, poseyendo ya ciertos conocimientos de Cálculo Infinitesimal, está interesado en acrecentar éstos, avanzando así en el estudio de esta disciplina. Este lector imaginario ha seguido con aprovechamiento un curso de Cálculo Infinitesimal de una variable y tiene ahora precisión de extender tales conocimientos al caso de varias variables. Lo que él aprenda sobre esto, será base en la que poder sostener su posterior formación, cuando avance en sus estudios de Matemáticas, Física, Ingeniería o Economía, en los que está muy interesado.

El autor, desde que empezó a escribir el libro, estuvo obviamente preocupado en conseguir que, lo que en él se fuera a decir, se dijera cuanto más claro mejor, procurando evitar toda ambigüedad. Ello le ha llevado a presentar lo que es básico, esto es, los conceptos, las definiciones, las propiedades y los teoremas, del modo más preciso que le ha sido posible y libres de cualquier aditamento. De entre estos asuntos, aquellos de mayor importancia, es decir, los que constituyen el núcleo del texto, se han destacado del resto de lo escrito, presentándoles de modo compendiado y, para subrayar su interés, encerrados en recuadros y sombreados. Al lector le será fácil, así, distinguir lo más importante de lo que lo es menos y podrá dirigir mejor sus pasos y ganar en eficacia cuando estudie. A cada uno de los anteriores textos recuadrados se le acompaña de los necesarios complementos: comentarios, generalizaciones, notas, ampliaciones, ejemplos y ejercicios. Se ha procurado que estos últimos, los ejemplos y ejercicios, no falten nunca, pues se estima que son imprescindibles para llegar a entender, de verdad, los conceptos que se abordan y las propiedades que se demuestran.

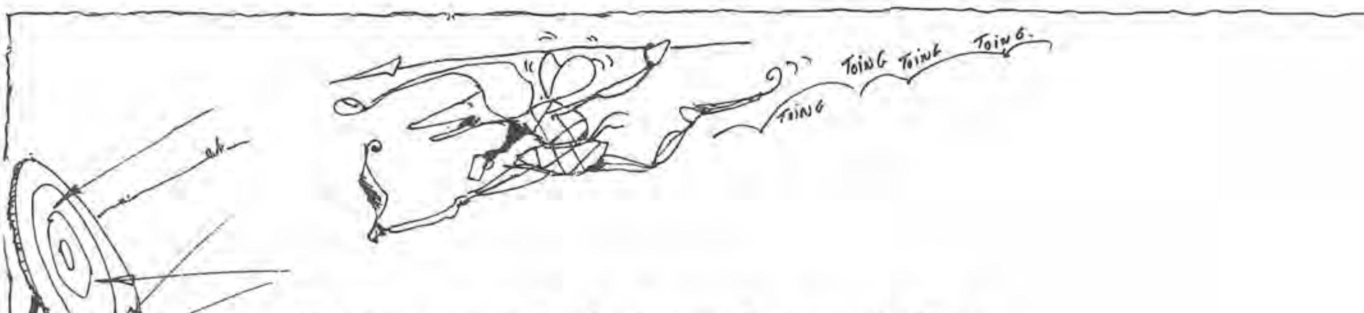
Hay una cuestión sobre la que el autor tiene especial interés en hablar; se trata de las demostraciones. El autor está convencido de que a un estudiante, cuando se las ha de ver con el teorema que le acaba de presentar su profesor o ha visto enunciado en su libro, no se le va a pasar por la cabeza el dudar de su veracidad; el alumno no va a poner en tela de juicio un resultado que viene siendo admitido, de antiguo, por todos los matemáticos, pues pensará, con buen juicio, que éstos lo habrán verificado sobradamente. El alumno se inclina a admitir la veracidad de los teoremas del mismo modo que cree a su profesor de historia cuando le dice que Cristóbal Colón descubrió el Nuevo Mundo el 12 de octubre de 1492. Así pues, las demostraciones no se les realizan a los alumnos para convencerles de la veracidad de lo demostrado. En un libro de texto o en las explicaciones del profesor, con las demostraciones no sólo se prueba que lo que se está afirmando es cierto, sino que también, y ello es mucho más importante para la formación de los alumnos, se está enseñando a razonar, a como se avanza en matemáticas, a resolver los problemas que se presentarán mañana, a manejar los conceptos, a utilizar las propiedades que ya se conocen. Los caminos que se siguen en las demostraciones y las ideas que en ellas se manejan serán las que, más tarde, habrá que utilizar para abordar con éxito las aplicaciones; mal podrá resolver las nuevas cuestiones aquel que carece de la experiencia que se adquiere al realizar las demostraciones. Consecuente con lo que piensa, el autor no ha omitido las demostraciones (salvo las obvias y, ya hacia el final, una especialmente difícil y alambicada); con ello no se desea sugerir al lector que deba él realizarlas todas, sólo se le quiere facilitar el que pueda hacerlo. Es más, el autor se permite aconsejar, al que se inicia en el estudio de esta disciplina, que realice las demostraciones sencillas, que aborde bastantes de las medianamente complicadas y que se ocupe de muy pocas de las que son más intrincadas o artificiosas de lo razonable.

Aun cuando el prólogo se está alargando más de lo que el autor tenía pensado, él no resiste la tentación de decir aquí lo que entiende por «dominar» algo en matemáticas; ello vendrá a explicar el porqué de la estructura, la organización, la hechura que se le ha dado al texto. Antes que nada, hay que decir que dominar las matemáticas no es ser especialmente habilidoso realizando cálculos más o menos rutinarios o aplicando algoritmos, tampoco es buscar la solución a problemas similares a los que, previamente, se ha visto resolver a otro. Para dominar las matemáticas hay que saber lo que significa cada uno de los objetos que se manejen y apreciar las cualidades de los mismos, hay que saber conceptuar y analizar los problemas, hay que tomar decisiones, que luego habrán de ser confirmadas o rechazadas, hay que formular debidamente las cuestiones que se planteen, expresándolas en términos precisos y libres de toda ganga, hay que llegar al fondo de los asuntos que se estudien, detectándoles el meollo, hay que saber, en cada momento, para qué se hace lo que se hace y el porqué de lo que se hace, hay que tener curiosidad y espíritu crítico.

El autor no quiere acabar el prólogo antes de manifestar que se siente gratamente obligado hacia todos los que, de uno u otro modo, le han ayudado en su tarea. El mayor apoyo le ha venido de su mujer y de sus hijos; gracias por vuestra cooperación animosa. Han sido muchos los colegas que, con sus valiosas sugerencias e inestimables consejos, han contribuido eficazmente en las elaboración de esta obra; el autor se ve en la precisión de citar aquí, reconocido, a José Manuel Vega, a Ignacio Parra, a José Antonio Nicolás, a Manuel Martínez, a Damián Rivas, a Eduardo Ahedo, a Carlos Vázquez, a Julio Ramírez, a Hermenegildo García, a Francisco Mancebo, a Carlos Martel, a Carlos Álvarez, a María Higuera, a Ignacio Delgado, a Ángel Velázquez. No puede el autor olvidarse de Eva Villacisneros, a quien le agradece su cuidadoso y excelente trabajo componiendo textos de matemáticas. El autor también da, gustoso, las gracias a todos los que en McGraw-Hill han colaborado en la confección de este libro y, en en particular, a Antonio García-Maroto, su director, y a Isabel Capella, editora.

En la villa de Madrid,
a 15 de mayo de 1995

CÁLCULO INFINITESIMAL DE VARIAS VARIABLES



El autor, que acaba de desentrañar unos insólitos manuscritos, no resiste la tentación de ofrecer, aquí, uno de ellos al lector. Bien es cierto que, lo que en aquel se dice, le cuadra a este primer capítulo; no obstante, también es verdad que tropieza con el tono de gravedad y circunspección que, de siempre, impera en los textos de matemáticas. Empero, el autor cree contar con la venia del lector y se anima y le inserta a continuación.

APRENDIENDO A CLAVAR LA LANZA CON TINO

Unos años atrás, no muchos, trabé gran amistad con un provecto caballero, antiguo profesor de matemáticas, ya apartado de la tarea de explicar esta ciencia a los mozalbetes de su pequeña ciudad. El día que conocí a Ocol-Nilep, nuestro profesor, pensé que no estaba en sus cabales: a poco de iniciada la conversación, empezó a hablarme de cosas tan extravagantes que me llevaron a suponer que desbarraba. Más tarde, sus desvaríos empezaron a tomar sentido para mí y, con el correr de los días, me llegó a parecer normal lo que decía. A estas alturas, no se si Ocol-Nilep estaba cuerdo, y me convenció de su cordura, o si, disparatando él, sus argumentos hicieron mella en mí y empecé a ver normal lo que era puro dislate. Juzgue el amable lector si fue lo uno o fue lo otro.

Decía Ocol-Nilep que, en sus años mozos, tropezó con un extraño texto, tallado en las paredes de una gruta(*), que guardó celosamente lo que allí ponía, que durante muchos años se dedicó, sin éxito, a interpretar el escrito, pero que, al fin, pudo dar con lo que en él se decía, lo cual resultó ser cosa, además de cierta, admirable: se trataba de un eficaz conjuro que permitía penetrar en el mundo de los anú yodón y dialogar con ellos. Los tales anú yodón constituyen una rara especie de gnomos voladores, que no sosiegan, vuelan incansablemente, dirigiéndose siempre, con obstinación, una vez tras otra, a un mismo lugar al que apuntan y, con no mucho tino, arrojan allí una lanza que siempre llevan consigo, intentando clavarla en él.

Hoy, que ya ha fallecido Ocol-Nilep, creo que ha llegado el momento de desvelar su secreto, para lo que me dio autorización, pues ha de interesar a muchos conocer la vida y milagros de los anú yodón. Para el conjuro, las cosas hay que hacerlas como aquí digo:

En primer lugar, se toma una estaca, no más larga que largo es el que hace el conjuro, y se clava en medio de una gran planicie. En la parte soterrada de la estaca se tallará la palabra «límite», que es la clave del conjuro, y en su parte visible se escribirá «rarraga ed-eh-et radnor ohcum-ed seupsed». Después, el conjurante se situará a gran distancia de la estaca

Ya allí, él emprenderá una alocada carrera, con mil cambios de rumbo, llena de vacilaciones, de idas y de venidas, que le irá acercando, dando vueltas a su alrededor, a su destino, a la inscripción que dice «límite» en la parte enterrada de la estaca. Durante todo este recorrido zigzagueante, entonará reiteradamente con monotonía, como lo hacían los indios americanos cuando imploraban la lluvia a Manitú, con voz monocorde, un canto que diga «ollip et-ek, otidlam, oparta et-ek».

Ya cerca de su meta, cuando le separen sólo unos codos de ella, si el conjurante ha procedido como aquí se dice y tiene confianza en conseguir su objetivo, en-

(*) Este texto estaba escrito al revés, de derecha a izquierda. Por ello, no sabemos si los nombres de las gentes y los conjuros que aquí se citan deben ser leídos como están escritos o al revés, de derecha a izquierda.

1

trará en tránsito, percibirá sensaciones extrañas, irá disminuyendo su consciencia, sentirá que se acerca sin cesar a su destino y que lo hace cada vez más y más rápidamente, a velocidad de vértigo, se le nublará la vista y perderá el conocimiento. Cuando vuelva en sí, se encontrará de pie, abrazado a la estaca; en una palabra, ha llegado al «límite», ha concluido el conjuro. Si mira a su alrededor, verá que por allí pululan los anú yodón, que le contemplarán con admiración y arrobo.

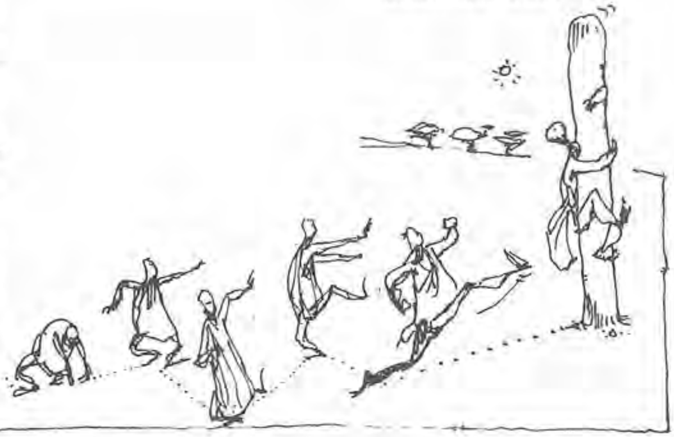
El embeleso de los anú yodón está plenamente justificado: se emboban contemplando a aquel que ha logrado, con limpieza, lo que ellos, tras mucho intentarlo, sólo consiguen de muy tarde en tarde; es decir, admiran a quien acaba de «alcanzar el límite», cosa esta que a ellos les tiene obsesionados y a la que se dedican, incesantemente, con escasísimo éxito.

Hice yo el conjuro del modo que dijo Ocol-Nilep y las cosas acontecieron como él anunció, de suerte que, tras ello, me vi, en el límite, rodeado de la admiración de los anú yodón que, quejándose con amargura de su poco tino, me pedían ayuda para «llegar al límite». Entonces recordé que, antes de hacer el conjuro y por si pudiera serme de utilidad, me había echado al bolsillo un pequeño manual sobre los límites, que Ocol-Nilep utilizaba con frecuencia; pensé que podría ser una buena ayuda para intentar un conjuro que resolviera el problema de los anú yodón, pues, al no entender yo nada de lo que allí ponía, supuse que se trataba de un libro cabalístico, lleno de invocaciones mágicas dirigidas a las fuerzas de lo oculto. ¿Qué podía ser, sino, aquella extraña colección de frases y de signos desconocidos, como $\forall$, $\exists$, $\mathbb{N}$ o $\in$, que no había manera de entender?

Así pues, me puse a buscar, en el manual, el conjuro más idóneo. Al cabo de un rato creí dar con él allí donde, bajo el epígrafe «límite de una sucesión», decía que $l = \lim x_n$ significaba que «$\forall\, \varepsilon > 0, \exists\, \nu \in \mathbb{N} / n \geqslant \nu \Rightarrow \|x_n - l\| < \varepsilon$». Aquello era, de fijo, un conjuro, ¿qué otra cosa podía ser? Lo leímos a coro, lo escribimos en todas partes, lo cantamos a varias voces, lo acompañamos de frases mágicas, danzamos mientras lo recitábamos, pero no pasó nada. Ya exhaustos, oímos la voz de Ocol-Nilep que, procedente de ultratumba, decía: no, así no, las cosas de los límites, como otra muchas, para que nos funcionen hay que entenderlas, no basta con memorizarlas y repetirlas, no, hay que ahondar en ellas, hay que penetrar en su significación más profunda.

No sé qué ocurrió en mí, quizá fue el buen influjo de Ocol-Nilep, pero el caso es que me propuse enterarme yo, primero, y explicar a los anú yodón, después, todo lo que sobre los límites se decía en el manual. Al principio las cosas resultaron harto difíciles, pero, al poco, cambió nuestra actitud y empezamos a «encontrarle el gusto» a aquello; en cuanto nos hubimos enterado, de verdad, de lo que era un límite y supimos algo acerca de cómo funcionaban, los anú yodón mejoraron muy notablemente su puntería, clavaban atinadamente sus lanzas; se produjo en ellos tal cambio que, en lugar de anú yodón, desde entonces se les conoce como los omirra emay. Y es que no hay nada como saber algo para hacer bien ese algo; no olvidemos que «saber» no es, sólo, «conocer», no, «saber» es «saber hacer».

Gudor Ben Jusá



CAPÍTULO 1

Topología, límites y continuidad

1.1. NOCIONES SOBRE LA TOPOLOGÍA DE $\mathbb{R}^p$

A lo largo de este curso vamos a precisar de pocos conocimientos acerca de la topología de $\mathbb{R}^p$, por lo que no nos excederemos aquí, hablando de este asunto; sólo nos ocuparemos de sus cuestiones más sobresalientes, de aquellas de las que luego precisaremos en este y en los próximos capítulos. Ello no obstante, se harán algunos comentarios y observaciones que amplíen y generalicen lo que aquí se diga, intentando dar una visión algo más completa de la topología usual del espacio euclídeo $\mathbb{R}^p$, que habrá de interesar a aquellos que hayan de realizar estudios posteriores de Análisis Matemático.

❑ EL ESPACIO EUCLÍDEO $\mathbb{R}^p$

Aun cuando lo que aquí se va a decir es, a buen seguro, conocido de cuantos han seguido un primer curso de Álgebra Lineal, no está de más que ahora lo recordemos. Se trata de un resumen acerca de $\mathbb{R}^p$, de su estructura de espacio vectorial euclídeo, con el producto escalar usual, destacando de entre ello. lo relativo a la distancia que se obtiene del referido producto escalar.

[01]

Consideremos el conjunto $\mathbb{R}^p$ (donde $p \in \mathbb{N}$ es dado), formado por todas las sucesiones $\boldsymbol{x} = (x_1, x_2, ..., x_p)$ de p números reales, que es un espacio vectorial con las operaciones usuales (suma y producto por un número):

$$(x_1, ..., x_p) + (y_1, ..., y_p) = (x_1 + y_1, ..., x_p + y_p), \qquad \lambda(x_1, ..., x_p) = (\lambda x_1, ..., \lambda x_p)$$

Consideremos también el *espacio euclídeo* $\mathbb{R}^p$ (de dimensión p) que se obtiene al dotar al espacio vectorial $\mathbb{R}^p$ del producto escalar ordinario, entre vectores de $\mathbb{R}^p$, que está definido mediante:

$$(x_1, ..., x_p) \cdot (y_1, ..., y_p) = x_1 y_1 + \cdots + x_p y_p$$

La *norma euclídea* de un vector $\boldsymbol{x} \in \mathbb{R}^p$ (o sea, la norma deducida del anterior producto escalar) es, pues, el siguiente número $\|\boldsymbol{x}\|$:

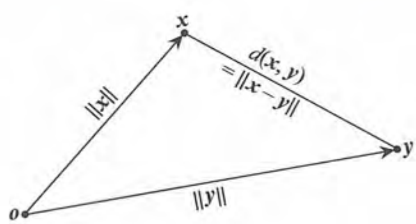
$$\|\boldsymbol{x}\| = \sqrt{\boldsymbol{x} \cdot \boldsymbol{x}} = \sqrt{(x_1)^2 + \cdots + (x_p)^2}$$

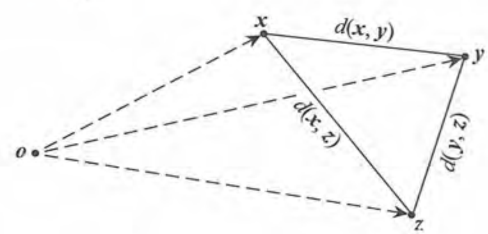
A partir de esta norma, se define la distancia euclídea de $x = (x_i)$ a $\boldsymbol{y} = (y_i)$ como el número $d(\boldsymbol{x}, \boldsymbol{y})$ definido por:

$$d(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{y} - \boldsymbol{x}\| = \sqrt{(y_1 - x_1)^2 + \cdots + (y_p - x_p)^2}$$

Propiedades características de las distancias. Se verifica que, para $\boldsymbol{x}, \boldsymbol{y}, z \in \mathbb{R}^p$:

1. $d(\boldsymbol{x}, \boldsymbol{y}) > 0$ si $\boldsymbol{x} \neq \boldsymbol{y}$; además $d(\boldsymbol{x}, \boldsymbol{x}) = 0$
2. $d(\boldsymbol{x}, \boldsymbol{y}) = d(\boldsymbol{y}, \boldsymbol{x})$
3. $d(\boldsymbol{x}, z) \leqslant d(\boldsymbol{x}, \boldsymbol{y}) + d(\boldsymbol{y}, z)$ (propiedad triangular)





Demostración

Recuérdese que para la norma euclídea (como para cualquier otra norma) se verifica que, si $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^p$ y $\lambda \in \mathbb{R}$, es:

- $\|\boldsymbol{x}\| > 0$ para $\boldsymbol{x} \neq \boldsymbol{o}$; además $\|\boldsymbol{o}\| = 0$
- $\|\lambda \boldsymbol{x}\| = |\lambda| \, \|\boldsymbol{x}\|$
- $\|\boldsymbol{x} + \boldsymbol{y}\| \leqslant \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$

Apoyándonos en estos resultados, podemos razonar como sigue:

1. $d(x, y) = \|y - x\| \begin{cases} > 0, & \text{si } x \neq y \quad (\text{pues } y - x \neq o) \\ = 0, & \text{si } x = y \quad (\text{pues } y - x = o) \end{cases}$
2. $d(x, y) = \|y - x\| = \|-(x - y)\| = |-1|\,\|x - y\| = \|x - y\| = d(y, x)$
3. $d(x, z) = \|z - x\| = \|(z - y) + (y - x)\| \leqslant \|z - y\| + \|y - x\| = d(y, z) + d(x, y)$

$[01]_1$ Observación (espacio métrico)

Esta distancia, la euclídea, no es la única posible, sino que hay otras muchas definiciones de distancia, de las que nosotros o nos ocuparemos; a todas ellas se las exige que cumplan las tres condiciones 1, 2 y 3 de [01]: tales distancias responden todas a la siguiente definición:

Se llama *distancia* entre puntos de un conjunto $C \neq \varnothing$ a toda aplicación $d: C \times C \to \mathbb{R}$ tal que, para cualesquiera $a, b, c \in C$, es:

- $d(a, b) > 0$, si $a \neq b$; $d(a, a) = 0$
- $d(a, b) = d(b, a)$
- $d(a, c) \leqslant d(a, b) + d(b, c)$

Un conjunto $C \neq \varnothing$ junto con una distancia d, definida en él, recibe el nombre de espacio métrico, al que se denota poniendo (C, d).

$[01]_2$ Observaciones (referentes a la notación)

Al espacio vectorial $\mathbb{R}^p$ se le puede considerar también como un espacio de puntos o espacio afín. Cuando así se haga, en lugar de decir que $x = (x_1, ..., x_p)$ es un vector de $\mathbb{R}^p$, se podrá hablar del punto x o X, cuyo sistema de coordenadas es $(x_1, ..., x_p)$; estas coordenadas lo son respecto de la referencia que tiene su origen en el elemento nulo de $\mathbb{R}^p$ (vector o o punto origen O) y su base de vectores es la base canónica de $\mathbb{R}^p$, esto es, la formada por los p vectores

$$(1, 0, ..., 0), (0, 1, ..., 0), ..., (0, 0, ..., 1)$$

La correspondencia entre el punto $X\,(x_i)$ y el vector $x = (x_i)$ se expresa de una cualquiera de las dos formas:

$$x = \overrightarrow{OX} \qquad X = O + x$$

Acudiendo a los puntos $X(x_i)$ e $Y(y_i)$, en lugar de utilizar los vectores $x = (x_i)$ e $y = (y_i)$, la distancia se puede expresar poniendo

$$d(X, Y) = \|\overrightarrow{XY}\|$$

[01]$_3$ Ejercicio

Compruébese que, para cualesquiera que sean $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^p$, se verifica que

$$d(\boldsymbol{x}, \boldsymbol{y}) \geqslant |d(\boldsymbol{x}, \boldsymbol{z}) - d(\boldsymbol{y}, \boldsymbol{z})|$$

Resolución

La relación a demostrar equivale al par de relaciones

$$d(\boldsymbol{x}, \boldsymbol{y}) \geqslant \begin{cases} d(\boldsymbol{x}, \boldsymbol{z}) - d(\boldsymbol{y}, \boldsymbol{z}) \\ -[d(\boldsymbol{x}, \boldsymbol{z}) - d(\boldsymbol{y}, \boldsymbol{z})] \end{cases}$$

las cuales se pueden escribir poniendo

$$d(\boldsymbol{x}, \boldsymbol{y}) + d(\boldsymbol{y}, \boldsymbol{z}) \geqslant d(\boldsymbol{x}, \boldsymbol{z}) \qquad \text{y} \qquad d(\boldsymbol{x}, \boldsymbol{y}) + d(\boldsymbol{x}, \boldsymbol{z}) \geqslant d(\boldsymbol{y}, \boldsymbol{z})$$

Como estas dos relaciones son ciertas, pues expresan la propiedad triangular, resulta que también lo es la del enunciado.

BOLAS Y ENTORNOS

[02]

En el espacio euclídeo $\mathbb{R}^p$, se llama *bola abierta* (o simplemente bola) de centro $\boldsymbol{a} \in \mathbb{R}^p$ y de radio $\rho > 0$ al conjunto

$$B(\boldsymbol{a}, \rho) = \{\boldsymbol{x} \in \mathbb{R}^p \ / \ d(\boldsymbol{a}, \boldsymbol{x}) < \rho\}$$

Se llama *bola cerrada* de centro $\boldsymbol{a} \in \mathbb{R}^p$ y de radio $\rho > 0$ al conjunto

$$\bar{B}(\boldsymbol{a}, \rho) = \{\boldsymbol{x} \in \mathbb{R}^p \ / \ d(\boldsymbol{a}, \boldsymbol{x}) \leqslant \rho\}$$

Se llama *bola reducida* de centro $\boldsymbol{a} \in \mathbb{R}^p$ y radio $\rho > 0$ al conjunto

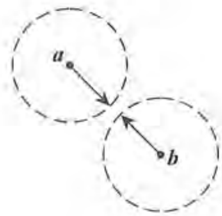
$$B^*(\boldsymbol{a}, \rho) = B(\boldsymbol{a}, \rho) - \{\boldsymbol{a}\}$$

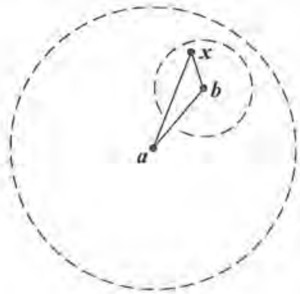
Se llaman *entornos* de un punto $\boldsymbol{a} \in \mathbb{R}^p$, en primer lugar, a las bolas abiertas de centro en $\boldsymbol{a}$ (con radio $\rho > 0$ cualquiera) y, también, a todo conjunto que incluya a una de tales bolas.

Propiedades:

1.ª Dados dos puntos distintos $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^p$, existen unos ciertos $\rho > 0$ y $\sigma > 0$ tales que $B(\boldsymbol{a}, \rho) \cap B(\boldsymbol{b}, \sigma) = \varnothing$.

2.ª Si $\boldsymbol{b} \in \mathbb{R}^p$ pertenece a una bola $B(\boldsymbol{a}, \rho)$, entonces existe un cierto $\sigma > 0$ tal que $B(\boldsymbol{b}, \sigma) \subset B(\boldsymbol{a}, \rho)$.





Demostración

1.° Por ser $\boldsymbol{a} \neq \boldsymbol{b}$, sabemos que $d(\boldsymbol{a}, \boldsymbol{b}) > 0$; tomando $\rho > 0$ y $\sigma > 0$ de manera que sea $\rho + \sigma \leqslant d(\boldsymbol{a}, \boldsymbol{b})$, por ejemplo $\rho = \sigma = \frac{1}{2}d(\boldsymbol{a}, \boldsymbol{b})$, las bolas $B(\boldsymbol{a}, \rho)$ y $B(\boldsymbol{b}, \sigma)$ son disjuntas. En efecto: si hubiera algún punto $\boldsymbol{x} \in \mathbb{R}^p$ que perteneciese a ambas bolas, sería

$$d(\boldsymbol{a}, \boldsymbol{b}) \leqslant d(\boldsymbol{a}, \boldsymbol{x}) + d(\boldsymbol{b}, \boldsymbol{x}) < \rho + \sigma \leqslant d(\boldsymbol{a}, \boldsymbol{b})$$

lo que obviamente es falso.

2.° Como $\boldsymbol{b}$ pertenece a la bola $B(\boldsymbol{a}, \rho)$, es $d(\boldsymbol{a}, \boldsymbol{b}) < \rho$, con lo que el número $\sigma = \rho - d(\boldsymbol{a}, \boldsymbol{b})$ es positivo; veamos que la bola $B(\boldsymbol{b}, \sigma)$ está incluida en la $B(\boldsymbol{a}, \rho)$. Así ocurre, pues:

$$\boldsymbol{x} \in B(\boldsymbol{b}, \sigma) \quad \Rightarrow \quad d(\boldsymbol{b}, \boldsymbol{x}) < \sigma \quad \Rightarrow \quad d(\boldsymbol{a}, \boldsymbol{x}) \leqslant d(\boldsymbol{a}, \boldsymbol{b}) + d(\boldsymbol{b}, \boldsymbol{x}) <$$
$$< d(\boldsymbol{a}, \boldsymbol{b}) + \sigma = \rho \quad \Rightarrow \quad \boldsymbol{x} \in B(\boldsymbol{a}, \rho)$$

[02]$_1$ Observaciones

- La bola abierta $B(\boldsymbol{a}, \varepsilon)$, de centro $\boldsymbol{a} \in \mathbb{R}^p$ y radio $\varepsilon > 0$, se suele llamar ε-entorno de $\boldsymbol{a}$. También se dice que $B(\boldsymbol{a}, \varepsilon)$ es el «entorno esférico» o «entorno circular», de radio $\varepsilon > 0$ del punto $\boldsymbol{a}$. Esta última denominación es debida a que en los casos $p = 3$ y $p = 2$ dicho entorno es la parte interior de la esfera y la circunferencia, respectivamente, de centro $\boldsymbol{a}$ y radio ε.
- Se llama *esfera* de centro $\boldsymbol{a} \in \mathbb{R}^p$ y radio $\rho > 0$ al conjunto

$$S(\boldsymbol{a}, \rho) = \{\boldsymbol{x} \in \mathbb{R}^p \;/\; d(\boldsymbol{a}, \boldsymbol{x}) = \rho\}$$

Nótese que

$$B(\boldsymbol{a}, \rho) \cup S(\boldsymbol{a}, \rho) = \bar{B}(\boldsymbol{a}, \rho)$$

[02]$_2$ Ejercicio

Se llaman «entornos rectangulares» del punto $\boldsymbol{a} = (a_1, a_2, \ldots, a_p)$ de $\mathbb{R}^p$ a los conjuntos del siguiente tipo, donde $\varepsilon_1 > 0$, $\varepsilon_2 > 0$, ..., $\varepsilon_p > 0$:

$$U = \{(x_1, x_2, \ldots, x_p) \in \mathbb{R}^p \;/\; |x_i - a_i| < \varepsilon_i \text{ para } i = 1, 2, \ldots, p\}$$

Compruébese que tal entorno incluye a un entorno circular y recíprocamente.

Resolución

- Llamando $\varepsilon = \text{mín}\ \{\varepsilon_1, \varepsilon_2, ..., \varepsilon_p\}$, la bola $B(\boldsymbol{a}, \varepsilon)$ está incluido en U, puesto que

$$\boldsymbol{x} \in B(\boldsymbol{a}, \varepsilon) \quad \Rightarrow \quad |x_i - a_i| \leqslant \|\boldsymbol{x} - \boldsymbol{a}\| < \varepsilon \leqslant \varepsilon_i \ (i = 1, 2, ..., p) \quad \Rightarrow \quad \boldsymbol{x} \in U$$

- Dada una bola $B(\boldsymbol{a}, \varepsilon)$, si se toma $\varepsilon_i = \varepsilon/\sqrt{p}$, el conjunto U está incluido en la bola, puesto que

$$\boldsymbol{x} \in U \quad \Rightarrow \quad |x_i - a_i| < \varepsilon/\sqrt{p}\ (i = 1, 2, ..., p) \quad \Rightarrow \quad \|\boldsymbol{x} - \boldsymbol{a}\| = \sqrt{\sum (x_i - a_i)^2} <$$
$$< \sqrt{p(\varepsilon/\sqrt{p})^2} = \varepsilon \quad \Rightarrow \quad \boldsymbol{x} \in B(\boldsymbol{a}, \varepsilon)$$

SUCESIONES CONVERGENTES

[03]

Se dice qe una sucesión $(\boldsymbol{x}_n)$, de puntos de $\mathbb{R}^p$, tiene por *límite* a $\boldsymbol{l} \in \mathbb{R}^p$ o que converge hacia $\boldsymbol{l}$, y se pone

$$\lim_{n \to \infty} \boldsymbol{x}_n = \boldsymbol{l} \qquad \text{o abreviadamente} \qquad \lim \boldsymbol{x}_n = \boldsymbol{l} \quad \text{o} \quad \boldsymbol{x}_n \to \boldsymbol{l}$$

si se verifica una cualquiera de las siguientes condiciones, equivalentes entre sí:

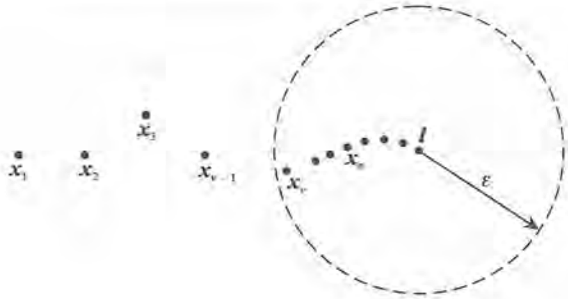
1.ª Para cada número real $\varepsilon > 0$, existe un $\nu \in \mathbb{N}$ tal que $\|\boldsymbol{x}_n - \boldsymbol{l}\| < \varepsilon$ para todo $n \geqslant \nu$ (a esta condición se la llama «$\varepsilon : \nu$»).
2.ª Dado un entorno cualquiera de $\boldsymbol{l}$, «fuera» de él hay, a lo sumo, un número finito de elementos de la sucesión.
3.ª La sucesión $(\|\boldsymbol{x}_n - \boldsymbol{l}\|)$, de números reales, tiene límite 0.

Propiedades: Sea $(\boldsymbol{x}_n)$ una sucesión de puntos de $\mathbb{R}^p$; se verifica que:

1. Si $\boldsymbol{x}_n = (x_n^1, x_n^2, ..., x_n^p)$ para $n \in \mathbb{N}$ y si $\boldsymbol{l} = (l^1, l^2, ..., l^p) \in \mathbb{R}^p$, entonces:

$$\left[\lim_{n \to \infty} \boldsymbol{x}_n = \boldsymbol{l}\right] \Leftrightarrow \left[\lim_{n \to \infty} x_n^i = l_n^i, \text{ para } i = 1, 2, ..., p\right]$$

2. Si $(\boldsymbol{x}_n)$ es convergente, entonces $(\boldsymbol{x}_n)$ tiene un solo límite.
3. Si $(\boldsymbol{x}_n)$ es convergente, entonces $(\boldsymbol{x}_n)$ está acotada (es decir, está acotada la sucesión $(\|\boldsymbol{x}_n\|)$, de números reales).
4. Si $(\boldsymbol{x}_n)$ converge hacia $\boldsymbol{l}$, entonces también converge hacia $\boldsymbol{l}$ toda subsucesión de $(\boldsymbol{x}_n)$.



Demostración

Las tres formas de definir el límite que se acaban de dar son equivalentes entre sí, pues son distintos modos de expresar un mismo concepto. Así, por ejemplo, decir que «fuera de $B(\boldsymbol{l}, \varepsilon)$ hay un número finito de elementos de $(\boldsymbol{x}_n)$» es lo mismo que afirmar que a partir de uno de ellos (el $\boldsymbol{x}_\nu$) están todos en dicho entorno, es decir, que existe $\nu \in \mathbb{N}$ tal que $\|\boldsymbol{x}_n - \boldsymbol{l}\| < \varepsilon$ para todo $n \geqslant \nu$.

1. Supongamos primero que $\boldsymbol{x}_n \to \boldsymbol{l}$. Entonces, dado $\varepsilon > 0$ (cualquiera), existe $\nu \in \mathbb{N}$ tal que $\|\boldsymbol{x}_n - \boldsymbol{l}\| < \varepsilon$ para todo $n \geqslant \nu$, luego para $i = 1, 2, ..., p$ es:

$$n \geqslant \nu \quad \Rightarrow \quad |x_n^i - l^i| = \sqrt{(x_n^i - l^i)^2} \leqslant \sqrt{\sum_{j=1}^{p} (x_n^j - l^j)^2} = \|\boldsymbol{x} - \boldsymbol{l}\| < \varepsilon$$

lo que prueba que $x_n^i \to l^i$ para $i = 1, 2, ..., p$. Recíprocamente, si $x_n^i \to l^i$ para $i = 1, 2, ..., p$, dado $\varepsilon > 0$ (cualquiera) existe $\nu_i \in \mathbb{N}$ tal que $|x_n^i - l^i| < \varepsilon/\sqrt{p}$ para $i = 1, 2, ..., p$ y, por tanto, llamando $\nu = \text{máx}\,\{\nu_1, \nu_2, ..., \nu_p\}$, se verifica entonces que

$$n \geqslant \nu \quad \Rightarrow \quad [|x_n^i - l^i| < \varepsilon/\sqrt{p} \qquad \text{para } i = 1, 2, ..., p] \quad \Rightarrow$$
$$\Rightarrow \quad \|\boldsymbol{x}_n - \boldsymbol{l}\| = \sqrt{\sum_{j=1}^{p} (x_n^j - l^j)^2} < \sqrt{p(\varepsilon/\sqrt{p})^2} = \varepsilon$$

luego $(\boldsymbol{x}_n)$ tiene límite $\boldsymbol{l}$, como había que comprobar.

2 a 4. Estas propiedades, que se cumplen cuando se consideran sucesiones reales (en lugar de sucesiones de puntos de $\mathbb{R}^p$), lo que se supone conocido, se demuestran aquí, mutatis mutandis, como en el caso real. Dichas propiedades se pueden comprobar, también, acudiendo a la anterior propiedad 1, con lo que ellas se reducen al caso real, en el que son ciertas:

- La propiedad 2 se comprueba por reducción al absurdo: suponiendo que hubiera dos límites $\boldsymbol{l}$ y $\boldsymbol{l}'$ distintos y comprobando que, entonces, la condición «$\varepsilon : \nu$» no se puede cumplir para $\varepsilon \leqslant \frac{1}{2}\|\boldsymbol{l} - \boldsymbol{l}'\|$.
- Para verificar la propiedad 3 basta con tener en cuenta que, si $\boldsymbol{l}$ es el límite de $(\boldsymbol{x}_n)$, en la bola $B(\boldsymbol{l}, 1)$ están todos los $\boldsymbol{x}_n$ salvo un número finito de ellos.
- La propiedad 4 se verifica ya que fuera de cualquier entorno de $\boldsymbol{l}$ hay un número finito de puntos de $(\boldsymbol{x}_n)$, luego lo mismo ocurre con cualquier subsucesión de $(\boldsymbol{x}_n)$.

[03]$_1$ Observación (sucesión que no tiene límite *l*)

Conviene resaltar que la negación de «la sucesión (x_n) tiene límite $l \in \mathbb{R}^p$» puede expresarse de cualquiera de las formas siguientes:

— La sucesión (x_n) no converge hacia l.
— Existe un número real $\varepsilon_0 > 0$ tal que para cada $v \in \mathbb{N}$ existe algún $n \geqslant v$ tal que $\|x_n - l\| \geqslant \varepsilon_0$.
— Existe un número real $\varepsilon_0 > 0$ y una subsucesión de (x_n) tales que para todo elemento x de la subsucesión es $\|x - l\| \geqslant \varepsilon_0$.

[03]$_2$ Ejercicio

Sean (x_n) e (y_n) dos sucesiones, de puntos de $\mathbb{R}^p$. Pruébese que

$$\left.\begin{array}{l} \lim x_n = l \in \mathbb{R}^p \\ \lim \|x_n - y_n\| = 0 \end{array}\right\} \Rightarrow \lim y_n = l$$

Resolución

Dado $\varepsilon > 0$ (cualquiera), tomando $\varepsilon/2$ (en lugar de ε) en la condición «$\varepsilon : v$» de cada uno de los dos límites de la hipótesis, se concluye que:

$$[\lim x_n = l] \Rightarrow \exists\, v_1 \in \mathbb{N} \,/\, n \geqslant v_1 \Rightarrow \|x_n - l\| < \varepsilon/2$$

$$[\lim \|x_n - y_n\| = 0] \Rightarrow \exists\, v_2 \in \mathbb{N} \,/\, n \geqslant v_2 \Rightarrow \|x_n - y_n\| < \varepsilon/2$$

Por tanto, llamando $v = \max\{v_1, v_2\}$, resulta que

$$n \geqslant v \Rightarrow \|y_n - l\| \leqslant \|y_n - x_n\| + \|x_n - l\| < \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon$$

así pues, l es el límite de (y_n).

[03]$_3$ Propiedades aritméticas de los límites

Si (x_n) e (y_n) son dos sucesiones convergentes, de puntos de $\mathbb{R}^p$, entonces también son convergentes las sucesiones (λx_n), $(x_n + y_n)$ y $(\|x_n\|)$, donde $\lambda \in \mathbb{R}$, y se verifica que:

1. $\lim\limits_{n\to\infty} (\lambda x_n) = \lambda\left(\lim\limits_{n\to\infty} x_n\right)$
2. $\lim\limits_{n\to\infty} (x_n + y_n) = \left(\lim\limits_{n\to\infty} x_n\right) + \left(\lim\limits_{n\to\infty} y_n\right)$
3. $\lim\limits_{n\to\infty} \|x_n\| = \|\lim\limits_{n\to\infty} x_n\|$

Demostración

Sean $\boldsymbol{l} = \lim \boldsymbol{x}_n$ y $\boldsymbol{m} = \lim \boldsymbol{y}_n$. Dado $\varepsilon > 0$ (cualquiera), se tiene:

1. Como $\boldsymbol{x}_n \to \boldsymbol{l}$, existe un $\nu \in \mathbb{N}$ tal que $n \geqslant \nu \;\Rightarrow\; \|\boldsymbol{x}_n - \boldsymbol{l}\| < \varepsilon_1$ (donde $\varepsilon_1 > 0$ es cualquiera, fijo), luego
$$n \geqslant \nu \quad\Rightarrow\quad \|\lambda\boldsymbol{x}_n - \lambda\boldsymbol{l}\| = |\lambda|\,\|\boldsymbol{x}_n - \boldsymbol{l}\| < |\lambda|\varepsilon_1$$
por tanto $\lambda\boldsymbol{x}_n \to \lambda\boldsymbol{l}$, pues se cumple la condición «$\varepsilon : \nu$» (tómese ε_1 de modo que $|\lambda|\varepsilon_1 < \varepsilon$).
2. Como $\boldsymbol{x}_n \to \boldsymbol{l}$ e $\boldsymbol{y}_n \to \boldsymbol{m}$, existen $\nu_1, \nu_2 \in \mathbb{N}$ tales que $n \geqslant \nu_1 \;\Rightarrow\; \|\boldsymbol{x}_n - \boldsymbol{l}\| < \varepsilon_1$ y $n \geqslant \nu_2 \;\Rightarrow\; \|\boldsymbol{y}_n - \boldsymbol{m}\| < \varepsilon_2$ (donde $\varepsilon_1 > 0$ y $\varepsilon_2 > 0$ son cualesquiera, fijos), luego si es $\nu = \max\{\nu_1, \nu_2\}$, será
$$n \geqslant \nu \quad\Rightarrow\quad \|(\boldsymbol{x}_n + \boldsymbol{y}_n) - (\boldsymbol{l} + \boldsymbol{m})\| \leqslant \|\boldsymbol{x}_n - \boldsymbol{l}\| + \|\boldsymbol{y}_n - \boldsymbol{m}\| < \varepsilon_1 + \varepsilon_2$$
por tanto, $\boldsymbol{x}_n + \boldsymbol{y}_n \to \boldsymbol{l} + \boldsymbol{m}$, pues se cumple la condición «$\varepsilon : \nu$» (tómese $\varepsilon_1 = \varepsilon_2 = \varepsilon/2$).
3. Como $\boldsymbol{x}_n \to \boldsymbol{l}$, existe $\nu \in \mathbb{N}$ tal que $n \geqslant \nu \;\Rightarrow\; \|\boldsymbol{x}_n - \boldsymbol{l}\| < \varepsilon$, luego
$$n \geqslant \nu \quad\Rightarrow\quad \big|\,\|\boldsymbol{x}_n\| - \|\boldsymbol{l}\|\,\big| \leqslant \|\boldsymbol{x}_n - \boldsymbol{l}\| < \varepsilon$$
por tanto, $\|\boldsymbol{x}_n\| \to \|\boldsymbol{l}\|$, pues se cumple la condición «$\varepsilon : \nu$».

CONDICIÓN DE CONVERGENCIA DE CAUCHY

[04]

Se dice que la sucesión $(\boldsymbol{x}_n)$, de puntos de $\mathbb{R}^p$, es una *sucesión fundamental o de Cauchy* si se verifica el siguiente requisito, que se llama *condición de Cauchy*:

- Para cada número real $\varepsilon > 0$, existe un índice $\nu \in \mathbb{N}$ tal que, para cualesquiera que sean $p \geqslant \nu$ y $q \geqslant \nu$, es $|\boldsymbol{x}_p - \boldsymbol{x}_q| < \varepsilon$.

Si $\boldsymbol{x}_n = (x_n^1, x_n^2, \ldots, x_n^p)$, la sucesión $(\boldsymbol{x}_n)$ es de Cauchy si, y sólo si, son de Cauchy las p sucesiones $(x_n^1), (x_n^2), \ldots, (x_n^p)$, llamadas componentes de $(\boldsymbol{x}_n)$.

CRITERIO GENERAL DE CONVERGENCIA DE CAUCHY. Una sucesión $(\boldsymbol{x}_n)$, de puntos de $\mathbb{R}^p$, es convergente si, y sólo si, cumple la condición de Cauchy (es decir, es una sucesión fundamental).

Demostración

1. Para comprobar que $(\boldsymbol{x}_n)$ es de Cauchy si, y sólo si, lo son las p sucesiones (x_n^i), se puede seguir el mismo camino que en la demostración de la anterior propiedad [03], 1 (aquella que dice que $(\boldsymbol{x}_n) \to \boldsymbol{l}$ equivale a $(x_n^i) \to l^i$ para $i = 1, 2, \ldots, p$). Las demostraciones de una y otra son muy similares; para obtener la de ahora, basta con sustituir, en la de antes, $\boldsymbol{x}_n - \boldsymbol{l}$ y $x_n^i - l^i$ por $\boldsymbol{x}_p - \boldsymbol{x}_q$ y $x_p^i - x_q^i$, respectivamente (además de poner $p \geqslant \nu$ y $q \geqslant \nu$ en lugar de $n \geqslant \nu$).

2. Recurriendo a que, según debemos saber, una sucesión de números reales es de Cauchy si, y sólo si, es convergente, se puede poner:

$$[(x_n) \text{ de Cauchy}] \Leftrightarrow [(x_n^i) \text{ de Cauchy, para } i = 1, 2, ..., p] \Leftrightarrow$$
$$\Leftrightarrow [(x_n^i) \text{ convergentes, para } i = 1, 2, ..., n] \Leftrightarrow [(x_n) \text{ convergente}]$$

$[04]_1$ Observación (espacios completos)

En un espacio métrico cualquiera (esto es, en un conjunto en el que se ha definido una distancia) se puede definir también, de igual modo que aquí (en [03] y [04]), los conceptos de sucesión convergente y sucesión de Cauchy. Se comprueba trivialmente que, en cualquier espacio métrico, toda sucesión convergente es de Cauchy. Sin embargo, el recíproco no se verifica en general: hay espacios métricos en los que existen sucesiones de Cauchy que no son convergentes.

Pues bien, un espacio métrico se dice que es *completo* si toda sucesión de Cauchy, de punto de él, es convergente, hacia un cierto punto del referido espacio. Según [04], el espacio métrico $\mathbb{R}^p$ es completo.

$[04]_2$ Ejercicio (sucesiones contractivas)

Se dice que una sucesión (x_n), de puntos de $\mathbb{R}^p$, es *contractiva* de constante $k \in \mathbb{R}$, con $0 < k < 1$, si para todo $n \in \mathbb{N}$ se verifica:

$$\|x_{n+1} - x_n\| \leqslant k\|x_n - x_{n-1}\|$$

Pruébese que toda sucesión contractiva de $\mathbb{R}^p$ es de Cauchy y, por tanto, convergente en $\mathbb{R}^p$.

Resolución

Aplicando reiteradamente la desigualdad del enunciado, se tiene:

$$\|x_{n+1} - x_n\| \leqslant k\|x_n - x_{n-1}\| \leqslant k^2\|x_{n-1} - x_{n-2}\| \leqslant \cdots \leqslant k^{n-1}\|x_2 - x_1\|$$

Acudiendo a la desigualdad triangular y llamando $d = \|x_2 - x_1\|$, se puede poner (se supone $p > q$):

$$\|x_p - x_q\| = \|(x_p - x_{p-1}) + (x_{p-1} - x_{p-2}) + \cdots + (x_{q+1} - x_q)\| \leqslant$$
$$\leqslant \|x_p - x_{p-1}\| + \|x_{p-1} - x_{p-2}\| + \cdots + \|x_{q+1} - x_q\| \leqslant$$
$$\leqslant k^{p-2}d + k^{p-3}d + \cdots + k^{q-1}d = \frac{k^{q-1}(1-k^{p-q})}{1-k}\, d < \frac{k^{q-1}}{1-k}\, d$$

Ahora bien, como es $0 < k < 1$, se sabe que $k^n \to 0$ y, por tanto, esta última expresión es menor que un $\varepsilon > 0$ (dado arbitrariamente) con tal de tomar q mayor que cierto $\nu \in \mathbb{N}$. Así pues,

$$p \geqslant q \geqslant \nu \Rightarrow \|x_p - x_q\| < \frac{k^{q-1}}{1-k}\, d < \varepsilon$$

luego (x_n) es una sucesión de Cauchy, como había que probar.

LÍMITES DE OSCILACIÓN. TEOREMA DE BOLZANO-WEIERSTRASS

[05]

Se dice que un punto $\alpha \in \mathbb{R}^p$ es un límite de oscilación de la sucesión (x_n), de puntos de $\mathbb{R}^p$, si se verifica una de las siguientes condiciones, que son equivalentes entre sí:

a) α es el límite de alguna subsucesión de (x_n).
b) En todo entorno de α hay infinitos elementos de (x_n).

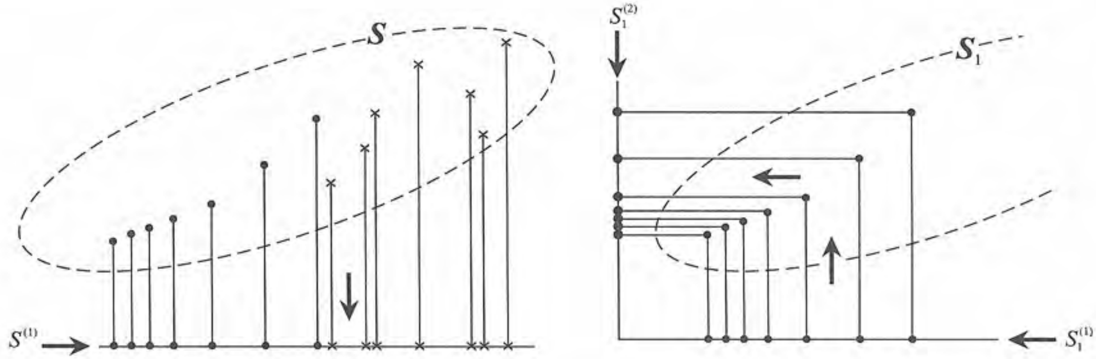
TEOREMA DE BOLZANO-WEIERSTRASS. Toda sucesión acotada(*), de puntos de $\mathbb{R}^p$, tiene algún límite de oscilación, en $\mathbb{R}^p$ (o sea, tiene alguna subsucesión convergente, hacia un punto de $\mathbb{R}^p$).

(*) Una sucesión (x_n) está acotada si existe un $k > 0$ tal que $\|x_n\| \leqslant k$ para todo $n \in \mathbb{N}$.

Demostración

1. Comprobemos que las condiciones *a*) y *b*) son equivalentes:
   - $[a \Rightarrow b]$ Si hay una subsucesión de (x_n) que converge hacia α y dado un entorno cualquiera de α, en dicho entorno hay infinitos elementos de la subsucesión, luego en él hay infinitos elementos de (x_n).
   - $[b \Rightarrow a]$ Por hipótesis, en la bola $B(\alpha, 1)$ hay algún elemento x_{i_1} de (x_n); en la bola $B(\alpha, 1/2)$ hay infinitos elementos de (x_n), luego hay uno, x_{i_2}, *con* $i_2 > i_1$; siguiendo de este modo, se llega a que, para cada $n \in \mathbb{N}$, en la bola $B(\alpha, 1/n)$ hay un elemento x_{i_n} con $i_n > i_{n-1}$. La subsucesión $x_{i_1}, x_{i_2}, ..., x_{i_n}, ...$, así obtenidos, tiene límite α, como es fácil comprobar.

2. Teorema de Bolzano-Weierstrass. Para demostrar este teorema acudiremos a que, según debemos saber, él se cumple en el caso real ($p = 1$): Sea S una sucesión acotada de puntos de $\mathbb{R}^p$; se quiere probar que S tiene alguna subsucesión convergente en $\mathbb{R}^p$. Llamemos $S^{(i)}$ a la proyección i-ésima (para $i = 1, 2, ..., p$) de S; como $S^{(1)}$ es una sucesión acotada de números reales, ella admite una subsucesión convergente, a la que llamaremos $S_1^{(1)}$. Llamemos S_1 a la subsucesión de S formada por los elementos, de S, cuyas proyecciones forman $S_1^{(1)}$. Sea $S_1^{(2)}$ la segunda proyección de S_1; como $S_1^{(2)}$ es una sucesión acotada de números reales, ella admite una subsucesión convergente, a la que llamaremos $S_{12}^{(2)}$. Acudamos a la correspondiente sucesión S_{12}, que es una subsucesión de S que tiene sus dos primeras proyecciones, $S_{12}^{(1)}$ y $S_{12}^{(2)}$, convergentes, y sea $S_{12}^{(3)}$ la tercera proyección de S_{12}; como $S_{12}^{(3)}$ es una sucesión acotada de números reales, ella admite una subsucesión convergente, a la que llamaremos $S_{123}^{(3)}$. Acudamos a la correspondiente sucesión S_{123}, que es una subsucesión de S que tiene sus tres primeras proyecciones convergentes. Continuando de este modo, al cabo de p etapas, se llega a una sucesión $S' = S_{12\cdots p}$, subsucesión de S, que tiene sus p proyecciones convergentes. Acudiendo ahora a la propiedad [03],1, se concluye que S' es convergente, hacia cierto punto de $\mathbb{R}^p$, lo que prueba el teorema.



[05]$_1$ Ejercicio

Hallar los límites de oscilación de la sucesión (x_n), de puntos de $\mathbb{R}^2$, cuyo elemento n-ésimo es

$$x_n = = \left(\frac{2n}{n+1} \operatorname{sen} \frac{n\pi}{2}, \frac{n+5}{3n} \cos \frac{n\pi}{2}\right)$$

Resolución

Consideremos las cuatro subsucesiones:

$$\text{si} \quad n = \dot{4}, \qquad x_n = \left(0, \frac{n+5}{3n}\right) \to (0, 1/3)$$

$$\text{si} \quad n = \dot{4} + 1, \quad x_n = \left(\frac{2n}{n+1}, 0\right) \to (2, 0)$$

$$\text{si} \quad n = \dot{4} + 2, \quad x_n = \left(0, -\frac{n+5}{3n}\right) \to (0, -1/3)$$

$$\text{si} \quad n = \dot{4} + 3, \quad x_n = \left(-\frac{2n}{n+1}, 0\right) \to (-2, 0)$$

A la vista de ello, es obvio que los límites de oscilación de la sucesión (x_n) son los cuatro puntos $(0, 1/3)$, $(2, 0)$, $(0, -1/3)$ y $(-2, 0)$.

[05]$_2$ Caso de límite de oscilación único

Una sucesión (x_n), de puntos de $\mathbb{R}^p$, converge hacia un punto $\alpha \in \mathbb{R}^p$ si, y sólo si, (x_n) es una sucesión acotada que tiene a α como su único límite de oscilación.

Demostración

1. Supongamos primero que (x_n) converge hacia α. Sabemos (véase [03],3) que (x_n) está entonces acotada. El punto α que, evidentemente, es límite de oscilación de (x_n), es su único límite de oscilación; en efecto, si $\beta \in \mathbb{R}^p$ es un punto distinto de α, tomando entornos de α y β que se excluyan, como fuera del de α hay sólo un número finito de elementos de (x_n), resulta que en el entorno de β no puede haber infinitos elementos de (x_n), luego β no es límite de oscilación de (x_n).
2. Supongamos ahora que (x_n) está acotada y tiene a α como su único límite de oscilación. Entonces α es el límite de (x_n) ya que, dado un entorno cualquiera de α, fuera de dicho entorno sólo hay un número finito de elementos de (x_n); en efecto: si fuera de alguno de estos entornos hubiera infinitos elementos de (x_n), entonces la subsucesión de (x_n) que forman estos infinitos elementos, como está acotada y según el teorema de Bolzano-Weierstrass (véase [05]), tendría algún límite de oscilación, que no puede ser α, y que sería, entonces, otro límite de oscilación de (x_n), lo que no es posible, por hipótesis.

LÍMITE INFINITO

[06]

- Se dice que una sucesión (x_n), de puntos de $\mathbb{R}^p$, tiene *límite infinito* si para cada número real $K > 0$ existe $v \in \mathbb{N}$ tal que $\|x_n\| > K$ para todo $n \geqslant v$. En tal caso, se pone $\lim x_n = \infty$ o $x_n \to \infty$.
- Se dice que ∞ es límite de oscilación de la sucesión (x_n) si ésta no está acotada, es decir, si para cualquiera que sea $K > 0$ la relación $\|x_n\| > K$ se verifica para infinitos índices $n \in \mathbb{N}$.

PROPIEDAD (existencia de límites de oscilación). Toda sucesión, de puntos de $\mathbb{R}^p$, tiene algún límite de oscilación (en $\mathbb{R}^p$ o ∞).

Demostración

1. Obsérvese primeramente (en relación con el concepto de límite de oscilación infinito) que una sucesión (x_n), de puntos de $\mathbb{R}^p$, no está acotada si, para cualquiera que sea $K > 0$, existe algún $n_1 \in \mathbb{N}$ tal que $\|x_{n_1}\| > K$. De ser así, la relación $\|x_n\| > K$ se verifica, no sólo para uno, sino para infinitos valores de n. En efecto: llamando $K_1 = \|x_{n_1}\|$, existe $n_2 \in \mathbb{N}$ tal que $\|x_{n_2}\| > K_1$; llamando $K_2 = \|x_{n_2}\|$, existe $n_3 \in \mathbb{N}$ tal que $\|x_{n_3}\| > K_2$; continuando así, indefinidamente, se encuentran infinitos elementos, x_{n_1}, x_{n_2}, ..., para los que $\|x_{n_i}\| > K$.
2. Comprobemos ahora que toda sucesión (x_n), de puntos de $\mathbb{R}^p$, tiene algún límite de oscilación (finito o infinito). Si (x_n) no está acotada, entonces tiene a ∞ como límite de oscilación. Si (x_n) está acotada, el teorema de Bolzano-Weierstrass (véase [05]) asegura entonces que (x_n) tiene algún límite de oscilación (en $\mathbb{R}^p$). Así pues, (x_n) tiene (esté o no acotada) un límite de oscilación, al menos.

[06]$_1$ Caso de límite de oscilación único

Una sucesión (x_n), de puntos de $\mathbb{R}^p$, tiene límite l (de $\mathbb{R}^p$ o ∞) si, y sólo si, l es su único límite de oscilación (en $\mathbb{R}^p \cup \{\infty\}$).

Demostración

Si (x_n) está acotada, esta propiedad se verifica, según ya se comprobó en $[05]_2$. Supondremos, pues, que (x_n) no está acotada, es decir, que tiene a ∞ como límite de oscilación.

- Si (x_n) tiene límite l, entonces ha de ser $l = \infty$, pues de lo contrario (x_n) estaría acotada y no lo está. Considérese un $\alpha \in \mathbb{R}^p$, cualquiera; α no puede ser límite de oscilación de (x_n). En efecto: tomando $K > \|\alpha\|$, como (x_n) tiene límite ∞, se sabe que es $\|x_n\| > K$ para todo n mayor que cierto $\nu \in \mathbb{N}$ y, por tanto, en la bola $B(\alpha, \rho)$, con $\rho = K - \|\alpha\|$, no puede haber infinitos elementos de (x_n).
- Si (x_n) tiene un único límite de oscilación (el cual es, pues, ∞), entonces (x_n) tiene límite ∞, esto es, dado $K > 0$ (cualquiera), la relación $\|x_n\| \leqslant K$ se verifica sólo para un número finito de elementos de (x_n). En efecto: si ello fuese falso, en la bola $B(o, K)$ habría infinitos elementos de (x_n), luego (x_n) tendría a o como límite de oscilación (no infinito), lo que va contra la hipótesis.

[06]$_2$ Observaciones

- Si a $\mathbb{R}^p$ se le añade el símbolo ∞, se dice que se ha ampliado $\mathbb{R}^p$; se denota por $\overline{\mathbb{R}}^p$ a esta ampliación, esto es, $\overline{\mathbb{R}}^p = \mathbb{R}^p \cup \{\infty\}$.
- Se llama entorno (circular) de ∞ de radio $K > 0$ al conjunto

$$\{x \in \mathbb{R}^p \;/\; \|x\| > K\} = \mathbb{R}^p - \overline{B}(o, K)$$

- En la definición de límite ∞, de una sucesión (x_n), poco importa la posible dirección según la que x_n tiende a infinito (se aleja del origen). Se considera un solo «punto del infinito» al que «se llega» por cualquier dirección (alejándose del origen). Nótese que en Geometría, contrariamente a lo que aquí se hace, se introduce un punto del infinito por cada dirección del espacio $\mathbb{R}^p$.

[06]$_3$ Ejercicio (sucesiones que no tienen límite *l*)

Sea (x_n) una sucesión de puntos de $\mathbb{R}^p$. Pruébese que la sucesión (x_n) no tiene límite l (de $\mathbb{R}^p$ o ∞) si, y sólo si, existe una subsucesión de (x_n) que tiene límite α (de $\mathbb{R}^p$ o ∞) distinto de l.

Resolución

Según ya sabemos (véase $[06]_1$), la sucesión (x_n) no tiene límite l si, y sólo si (x_n) tiene algún límite de oscilación α ($\mathbb{R}^p$ o ∞) distinto de l. Esto último equivale a decir que hay una subsucesión de (x_n) que tiene límite α.

CONJUNTOS ABIERTOS Y CERRADOS. FRONTERA

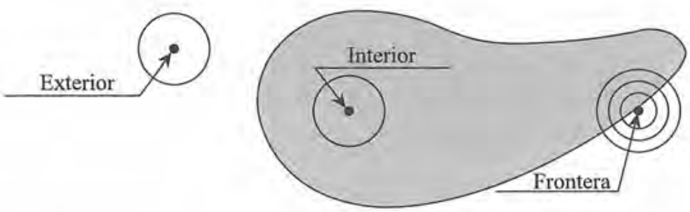
Nos vamos a limitar, ahora, a aquellas cuestiones, relativas a los conjuntos abiertos y a los conjuntos cerrados, que nos van a ser imprescindibles en lo que luego vendrá. Bien es verdad, no obstante, que después diremos algo más de estas cuestiones, pero lo será con carácter complementario o de ampliación.

[07]

Considérense un conjunto $C \subset \mathbb{R}^p$ y un punto $\boldsymbol{a} \in \mathbb{R}^p$. Se dice que:

- $\boldsymbol{a}$ es un *punto interior* de C si existe $\varepsilon > 0$ tal que $B(\boldsymbol{a}, \varepsilon) \subset C$.
- $\boldsymbol{a}$ es un *punto exterior* a C si existe $\varepsilon > 0$ tal que $B(\boldsymbol{a}, \varepsilon) \cap C = \varnothing$.
- $\boldsymbol{a}$ es un *punto frontera* de C si $B(\boldsymbol{a}, \varepsilon) \not\subset C$ y $B(\boldsymbol{a}, \varepsilon) \cap C \neq \varnothing$ para todo $\varepsilon > 0$.

- Un conjunto $C \subset \mathbb{R}^p$ se dice *abierto* si todos sus puntos son interiores; esto es, si ninguno de sus puntos frontera le pertenece. Un conjunto $C \subset \mathbb{R}^p$ se dice *cerrado* si todos sus puntos frontera le pertenecen.
- Dado un conjunto $C \subset \mathbb{R}^p$, se llama: 1.º) *interior* de C al conjunto $\overset{\circ}{C}$ que forman los puntos interiores de C; 2.º) *frontera* de C al conjunto $\mathrm{Fr}(C)$ que forman los puntos frontera de C; 3.º) *adherencia* de C al conjunto $\bar{C} = C \cup \mathrm{Fr}(C)$.



[07]₁ Observaciones

1. Dado un conjunto cualquiera $C \subset \mathbb{R}^p$, todo punto de $\mathbb{R}^p$ o es interior de C o es frontera de C o es exterior a C, sin que pueda ser dos de dichas cosas a la vez. Es decir, los conjuntos

$$\overset{\circ}{C} = \{\boldsymbol{x} \in \mathbb{R}^p \,/\, \boldsymbol{x} \text{ es interior de } C\}$$
$$\mathrm{Fr}(C) = \{\boldsymbol{x} \in \mathbb{R}^p \,/\, \boldsymbol{x} \text{ es frontera de } C\}$$
$$\mathrm{Ex}(C) = \{\boldsymbol{x} \in \mathbb{R}^p \,/\, \boldsymbol{x} \text{ es exterior de } C\}$$

forman una partición del espacio $\mathbb{R}^p$.

2. Dado un conjunto $C \subset \mathbb{R}^p$, para que un punto $\boldsymbol{a} \in \mathbb{R}^p$ sea interior de C es necesario (pero no suficiente) que $\boldsymbol{a}$ pertenezca a C. Para que $\boldsymbol{a}$ sea exterior a C es necesario (pero no suficiente) que $\boldsymbol{a}$ no pertenezca a C. Un punto frontera de C puede pertenecer a C o puede no pertenecerle. Si a C no le pertenece ninguno de sus puntos frontera, entonces se dice que C es abierto; si a C le pertenecen todos sus puntos frontera, se dice entonces que C es cerrado.

3. De las definiciones de interior, exterior y frontera de un conjunto se desprende que, dado un conjunto $C \subset \mathbb{R}^p$, para su complementario $\mathbb{R}^p - C$, se verifica que:

$$\begin{aligned} &\text{Interior de } (\mathbb{R}^p - C) = \text{Exterior de } C \\ &\text{Frontera de } (\mathbb{R}^p - C) = \text{Frontera de } C \\ &\text{Exterior de } (\mathbb{R}^p - C) = \text{Interior de } C \end{aligned}$$

4. De lo anterior se deduce fácilmente que los conjuntos cerrados son los complementarios de los abiertos y recíprocamente; es decir para cualquiera que sea $C \subset \mathbb{R}^p$, se verifica que (nótese que las dos relaciones siguientes son, en realidad, una sola):

$$\begin{aligned} &C \text{ es abierto} \iff \mathbb{R}^p - C \text{ es cerrado} \\ &C \text{ es cerrado} \iff \mathbb{R}^p - C \text{ es abierto} \end{aligned}$$

En efecto: la primera de estas relaciones (y, por tanto, la segunda) es consecuencia de la siguiente cadena de equivalencias:

$$C \text{ abierto} \iff \operatorname{Fr}(C) \cap C = \varnothing \iff \operatorname{Fr}(C) \subset \mathbb{R}^p - C \iff$$
$$\iff \operatorname{Fr}(\mathbb{R}^p - C) \subset \mathbb{R}^p - C \iff \mathbb{R}^p - C \text{ cerrado}$$

$[07]_2$ Ejemplo

Sea C el siguiente conjunto, de puntos de $\mathbb{R}^2$:

$$C = \left\{(x, y) \in \mathbb{R}^2 \ / \ x \neq \frac{1}{n}, \text{ para } n \in \mathbb{N}\right\}$$

El interior, la frontera y la adherencia de C son los conjuntos

$$\mathring{C} = C - \{(0, y) \in \mathbb{R}^2 \ / \ y \in \mathbb{R}\}$$
$$\operatorname{Fr}(C) = \left\{(x, y) \in \mathbb{R}^2 \ / \ x = 0 \text{ o } x = \frac{1}{n}, \text{ para } n \in \mathbb{N}\right\}$$
$$\overline{C} = \mathbb{R}^2$$

$[07]_3$ Propiedades de los abiertos y los cerrados

Para los abiertos y los cerrados, de $\mathbb{R}^p$, se verifican las siguientes propiedades:

1. *Propiedades de los conjuntos abiertos*

   a) $\mathbb{R}^p$ y $\varnothing$ son conjuntos abiertos.
   b) La unión de conjuntos abiertos es abierto.
   c) La intersección finita de conjuntos abiertos es abierto.

2. *Propiedades de los conjuntos cerrados*

   a) $\mathbb{R}^p$ y $\varnothing$ son conjuntos cerrados.
   b) La intersección de conjuntos cerrados es cerrado.
   c) La unión finita de conjuntos cerrados es cerrado.

Demostración

Las propiedades 1.a y 2.a son consecuencias obvias de la definición de abierto y cerrado. Para las demás propiedades, se tiene:

1.*b*) Si los conjuntos abiertos que se unen son los A_i, para $i \in I$ (familia de índices) y llamando A a su unión, entonces A es abierto ya que: si $\boldsymbol{a} \in A$, entonces $\boldsymbol{a} \in A_{i_0}$ para algún $i_0 \in I$ y, como A_{i_0} es abierto, existe cierto $\varepsilon > 0$ tal que $B(\boldsymbol{a}, \varepsilon) \subset A_{i_0}$, luego $B(\boldsymbol{a}, \varepsilon) \subset A$, lo que prueba que $\boldsymbol{a}$ es interior de A y de ahí que A sea abierto.

1.*c*) Si los conjuntos abiertos que se intersecan son los $A_1, A_2, ..., A_n$ y llamando A a su intersección, entonces A es abierto ya que: si $\boldsymbol{a} \in A$, entonces $\boldsymbol{a} \in A_i$ para $i = 1, 2, ..., n$ y, por ser A_i un conjunto abierto, existe $\varepsilon_i > 0$ tal que $B(\boldsymbol{a}, \varepsilon_i) \subset A_i$, luego llamando $\varepsilon = \text{mín}\,\{\varepsilon_1, ..., \varepsilon_n\}$ es $B(\boldsymbol{a}, \varepsilon) \subset A_i$ para $i = 1, 2, ..., n$ y, por tanto, $B(\boldsymbol{a}, \varepsilon) \subset A$, lo que prueba que $\boldsymbol{a}$ es interior de A y de ahí que A sea abierto.

2.*b*) Si los conjuntos cerrados que se intersecan son los C_i, para $i = 1, ..., n$, y llamando C a su intersección, entonces C es cerrado, pues su complementario es abierto, por ser la unión de los abiertos $\mathbb{R}^p - C_i$:

$$\mathbb{R}^p - C = \mathbb{R}^p - \bigcap_{i \in I} C_i = \bigcup_{i \in I} (\mathbb{R}^p - C_i)$$

2.*c*) Si los conjuntos cerrados que se unen son los $C_1, ..., C_n$ y llamando C a su unión, entonces C es cerrado, pues su complementario es abierto, por ser la intersección de los n abiertos $\mathbb{R}^p - C_i$:

$$\mathbb{R}^p - C = \mathbb{R}^p - \bigcup_{i=1}^{n} C_i = \bigcap_{i=1}^{n} (\mathbb{R}^p - C_i)$$

$[07]_4$ Observación (caracterización de los cerrados)

Un conjunto $C \subset \mathbb{R}^p$ es cerrado si, y sólo si, todo límite de oscilación (finito) de cualquier sucesión de puntos de C pertenece a C.

Demostración

1.º Supongamos primero que C es cerrado. Si $(\boldsymbol{x}_n)$ es una sucesión de puntos de C y si $\boldsymbol{a} \in \mathbb{R}^p$ es un límite de oscilación de $(\boldsymbol{x}_n)$, entonces en todo entorno de $\boldsymbol{a}$ hay infinitos puntos de $(\boldsymbol{x}_n)$, luego en todo entorno de $\boldsymbol{a}$ hay algún punto de C y, por ello, $\boldsymbol{a}$ es o interior o frontera de C, lo que nos permite asegurar que $\boldsymbol{a} \in C$, pues C es cerrado.

2.° Recíprocamente: suponiendo que todo límite de oscilación de cualquier sucesión de puntos de C es un punto de C, veamos que si $\boldsymbol{a} \in \mathbb{R}^p$ es un punto frontera de C, entonces, $\boldsymbol{a} \in C$. En efecto: como en todo entorno de $\boldsymbol{a}$ hay algún punto de C, en el entorno de $\boldsymbol{a}$ de radio $\varepsilon = 1/n$ para $n \in \mathbb{N}$ hay un punto $\boldsymbol{x}_n \in C$, al menos; la sucesión $(\boldsymbol{x}_n)$ tiene a $\boldsymbol{a}$ como límite y, de acuerdo con nuestra hipótesis, de ahí se deduce que $\boldsymbol{a} \in \mathbb{R}^p$.

CONJUNTOS COMPACTOS DE $\mathbb{R}^p$

Según tendremos ocasión de comprobar más adelante, los conjuntos de $\mathbb{R}^p$ que son cerrados y están acotados van a desempeñar en este curso un papel destacado. Estos conjuntos se caracterizan, en $\mathbb{R}^p$, por ser los que verifican a unas ciertas exigencias, de las que nos ocupamos enseguida, lo que da lugar a que se les llame conjuntos compactos de $\mathbb{R}^p$.

[08]

Las siguientes proposiciones, que están referidas a un conjunto $C \subset \mathbb{R}^p$, son equivalentes(*) entre sí:

a) El conjunto $C \subset \mathbb{R}^p$ es cerrado y está acotado.
b) Todo recubrimiento de $C \subset \mathbb{R}^p$ admite un sobrerrecubrimiento finito; esto es: si $\mathscr{A} = \{A_i \ / \ i \in I\}$ es una familia de conjuntos abiertos tal que $C \subset \bigcup_{i \in I} A_i$, entonces existen unos ciertos $A_{i_1}, ..., A_{i_k} \in \mathscr{A}$ tales que $C \subset A_{i_1} \cup \cdots \cup A_{i_k}$.
c) Toda sucesión de puntos de $C \subset \mathbb{R}^p$ tiene algún límite de oscilación que pertenece a C.

Se llaman *conjuntos compactos* de $\mathbb{R}^p$ a aquellos conjuntos $C \subset \mathbb{R}^p$ para los que se verifica una cualquiera(*) de las tres condiciones anteriores.

(*) En general (si, en lugar de $\mathbb{R}^p$, se considera un espacio métrico cualquiera), las condiciones *b*) y *c*) siguen siendo equivalentes y se llaman conjuntos compactos a los que las satisfacen. En el caso general se demuestra que (como aquí) todo conjunto compacto es cerrado y está acotado, sin que (en el caso general) se verifique necesariamente el recíproco, aunque éste sí que es cierto en $\mathbb{R}^p$ (teorema de Heine-Borel), como enseguida se prueba.

Demostración

1.° (*a* implica *b*) El conjunto C, por estar acotado, está incluido en algún intervalo compacto $I \subset \mathbb{R}^p$, esto es, en un cierto conjunto del tipo $I = [a_1, b_1] \times \cdots \times [a_p, b_p]$, siendo $-\infty < a_i < b_i < +\infty$ para $i = 1, ..., p$. Podemos suponer, y así lo haremos, que I tiene todos sus intervalos de igual longitud, a la que llamaremos l. Este intervalo, como cualquier otro, puede dividirse en 2^p subintervalos compactos, de lado mitad, sin más que dividir en dos partes iguales, mediante sus puntos medios, todos los intervalos $[a_i, b_i]$ para $i = 1, ..., p$.

Dado un recubrimiento abierto $\mathscr{A}$ de C, hemos de probar que $\mathscr{A}$ admite algún subrecubrimiento finito, lo que haremos por reducción al absurdo: suponiendo que éste no existe y obteniendo, de ello, una contradicción.

De acuerdo con dicha suposición, para cubrir C se precisan infinitos abiertos de $\mathscr{A}$ y, por ello, se precisan infinitos abiertos de $\mathscr{A}$ para cubrir los elementos de C que hay en uno, al menos, de los ya citados 2^p subintervalos de I, de lado $l/2$; sea I_1 un tal subintervalo. De ahí que, en consecuencia, se precisarían infinitos abiertos de $\mathscr{A}$ para cubrir los elementos de C que hay en uno, al menos, de los 2^p subintervalos de I_1 de lado $l/4$, obtenidos como ya se dijo; sea I_2 un tal subintervalo. Prosiguiendo con este proceso se obtiene una sucesión de intervalos, $I, I_1, I_2, ..., I_n, ...$, cada uno de ellos incluido en el anterior y de lado mitad que el de aquel (el lado de I_n es $l/2^{n+1}$ y su diámetro o diagonal es $\sqrt{p}l/2^{n+1}$); sea $\boldsymbol{x}_n$ el centro del intervalo I_n para $n \in \mathbb{N}$. La sucesión $(\boldsymbol{x}_n)$, de puntos de I, está acotada (por estarlo I) y por ello, según el teorema de Bolzano-Weierstrass (véase [05]), admite una subsucesión $\boldsymbol{x}_{n_1}, ..., \boldsymbol{x}_{n_i}, ...$ convergente hacia un cierto punto $\boldsymbol{x}_0 \in \mathbb{R}^p$. Para cualquiera que sea $\varepsilon > 0$, dentro del entorno $B(\boldsymbol{x}_0, \varepsilon)$ penetran a partir de uno los intervalos I_{n_i} (basta con que sea $\sqrt{p}l/2^{n_i+1} < \varepsilon$, o sea, $n_i > [L(\sqrt{p}l/\varepsilon)]/L2 - 1$), luego en dicho entorno hay infinitos elementos de C, por lo que $\boldsymbol{x}_0$ ha de pertenecer a C, ya que si no fuese así sería un punto frontera de C y, por ser C cerrado, entonces pertenecería a C. Por ser $\mathscr{A}$ un recubrimiento de C, existe un abierto $A \in \mathscr{A}$ tal que $\boldsymbol{x}_0 \in A$, de donde se infiere que es $\boldsymbol{x}_0 \in B(\boldsymbol{x}_0, r) \subset A$ para alguna bola abierta $B(\boldsymbol{x}_0, r)$ de centro $\boldsymbol{x}_0$ y un cierto radio $r > 0$. Cuando $i \to \infty$, el lado del intervalo I_{n_i} tiende a cero y el centro de I_{n_i} tiende a $\boldsymbol{x}_0$, por lo que se puede asegurar que, a partir de cierto i, el intervalo I_{n_i} llega a penetrar en la bola $B(\boldsymbol{x}_0, r)$ y, por tanto, se verifica que $I_{n_i} \subset A$. Este resultado es una contradicción; en efecto: se había dicho que para cubrir los elementos de C que había en I_{n_i} se precisaban infinitos de los abiertos de $\mathscr{A}$ y nos encontramos ahora que, como $I_{n_i} \subset A$, para ello es suficiente sólo el abierto $A \in \mathscr{A}$. Esta contradicción prueba que se verifica la propiedad b).

2.° (b implica c) Procedamos por reducción al absurdo. Supongamos que, siendo b) cierta, c) fuese falsa, esto es, que existiera una sucesión $(\boldsymbol{x}_n)$ de puntos de C que careciese de límites de oscilación en C (nótese que $(\boldsymbol{x}_n)$ tiene infinitos elementos distintos, pues, de no ser así, alguno de ellos se repetiría infinitas veces y él sería, entonces, límite de oscilación de la sucesión). Para cualquiera que sea el punto $\boldsymbol{x} \in C$, como él no puede ser límite de oscilación de $(\boldsymbol{x}_n)$, existe un $\delta_x > 0$ tal que en la bola abierta $A_x = B(\boldsymbol{x}, \delta_x)$ sólo hay, a lo sumo, un número finito de elementos de $(\boldsymbol{x}_n)$. Es evidente que la familia de dichas bolas abiertas $\mathscr{A} = \{A_x \mid \boldsymbol{x} \in C\}$ es un recubrimiento abierto de C; ahora bien, como se verifica la propiedad b), el conjunto C queda cubierto con sólo una colección finita $A_{x_1}, ..., A_{x_s}$ de tales bolas. Por tanto, los infinitos elementos de $(\boldsymbol{x}_n)$, como son de C, han de estar cubiertos por las bolas $A_{x_1}, ..., A_{x_s}$ lo cual no es posible, pues cada una de ellas cubre, como mucho, a un número finito de elementos de $(\boldsymbol{x}_n)$ luego, al haber sólo un número finito de bolas, entre todas no pueden cubrir a los infinitos elementos de $(\boldsymbol{x}_n)$. Hemos dado, pues, con una contradicción, lo que prueba que b) implica c).

3.° (c implica a) Suponiendo que se verifica c), probemos que C es cerrado y está acotado:

- Para ver que C es cerrado, comprobemos que, si $\boldsymbol{a} \in \mathbb{R}^p$ es un punto frontera de C, entonces $\boldsymbol{a} \in C$. En efecto: como en todo entorno de $\boldsymbol{a}$ hay algún punto de C, en el entorno de radio $1/n$ hay algún punto de C, al que llamaremos $\boldsymbol{x}_n$ para $n \in \mathbb{N}$; es claro que la sucesión $(\boldsymbol{x}_n)$ converge hacia $\boldsymbol{a}$, es decir, $\boldsymbol{a}$ es el único límite de oscilación de $(\boldsymbol{x}_n)$. Como se verifica la condición c), resulta que $\boldsymbol{a} \in C$.
- Si C no estuviera acotado, existiría una sucesión $(\boldsymbol{x}_n)$ de puntos de C que tendría límite ∞, la cual carecería, pues, de límites de oscilación en $\mathbb{R}^p$ y, por ello, no tendría ningún límite de oscilación en C, lo que es falso (pues se cumple c).

[08]$_1$ Ejercicio

En el espacio métrico $\mathbb{R}^p$ se consideran un conjunto abierto $A \neq \varnothing$, y un conjunto compacto $C \neq \varnothing$ incluido en A. Demuéstrese que la distancia ρ de C a la frontera de A es estrictamente positiva (se llama distancia entre dos conjuntos al ínfimo de las distancias entre puntos de uno y otro conjunto).

Resolución

La distancia ρ es el ínfimo de las distancias ρ_x de los puntos $\boldsymbol{x} \in C$ a la frontera de A. Para cada $\boldsymbol{x} \in C$, como $\boldsymbol{x} \in A$ y A es abierto, existe una bola centrada en $\boldsymbol{x}$ e incluida en A, por lo que la distancia ρ_x es positiva. Las bolas $B(\boldsymbol{x}, \rho_x)$ constituyen, para $\boldsymbol{x}$ recorriendo C, un recubrimiento abierto de C y, como C es compacto, basta con un número finito de estas bolas para cubrir C; sea ρ el menor de los radios de dichas bolas, que es positivo. Es evidente que ρ es la distancia de C a la frontera de A, la cual es, pues, positiva.

PUNTOS DE ACUMULACIÓN DE UN CONJUNTO

[09]

Dado un conjunto C, de infinitos puntos de $\mathbb{R}^p$, se dice que $\boldsymbol{a} \in \mathbb{R}^p$ es un *punto de acumulación* de C si se cumple una cualquiera de las siguientes condiciones, que son equivalentes entre sí:

1. En todo entorno del punto $\boldsymbol{a}$ hay infinitos puntos de C.
2. Existe una sucesión de puntos de $C - \{\boldsymbol{a}\}$ que converge hacia $\boldsymbol{a}$.

Demostración

- [1 ⇒ 2] Para cada $n \in \mathbb{N}$, en la bola $B(\boldsymbol{a}, 1/n)$ hay infinitos puntos de C, luego hay uno, al que llamaremos $\boldsymbol{x}_n$, que es distinto de $\boldsymbol{a}$; la sucesión $(\boldsymbol{x}_n)$ converge hacia $\boldsymbol{a}$ y está formada por puntos de $C - \{\boldsymbol{a}\}$.
- [2 ⇒ 1] Si la propiedad 1 fuese falsa, existiría un entorno de $\boldsymbol{a}$ en el que habría, a lo sumo, un número finito de puntos de C y, por tanto, existiría un entorno de $\boldsymbol{a}$ en el que, salvo quizá el propio punto $\boldsymbol{a}$, no habría ningún punto de C. De ser así, $\boldsymbol{a}$ no podría ser límite de ninguna sucesión de puntos de $C - \{\boldsymbol{a}\}$, lo que contradice a la hipótesis.

[09]$_1$ Ejercicio (caracterización de los puntos de acumulación)

Un punto $\boldsymbol{a} \in \mathbb{R}^p$ es de acumulación de un conjunto $C \subset \mathbb{R}^p$ si, y sólo si, en todo entorno reducido de $\boldsymbol{a}$ hay un punto, al menos, de C.

Resolución

1.º Si $\boldsymbol{a}$ es de acumulación de C y dado cualquier entorno reducido $B^*(\boldsymbol{a}, \varepsilon)$ de $\boldsymbol{a}$, como en el entorno $B(\boldsymbol{a}, \varepsilon)$ hay infinitos puntos de C, en $B^*(\boldsymbol{a}, \varepsilon)$ también hay infinitos puntos de C, luego hay al menos uno.

2.º Recíprocamente, supongamos ahora que en todo entorno reducido de $\boldsymbol{a}$ hay un punto, al menos, de C. Por ello, para cada $n \in \mathbb{N}$, en la bola reducida $B^*(\boldsymbol{a}, 1/n)$ hay, al menos, un punto $\boldsymbol{x}_n \in C$ y es $\boldsymbol{x}_n \neq \boldsymbol{a}$ (pues $\boldsymbol{a}$ no pertenece a sus entornos reducidos); es claro que la sucesión $(\boldsymbol{x}_n)$, que está formada por puntos de $C - \{\boldsymbol{a}\}$, converge hacia $\boldsymbol{a}$, luego $\boldsymbol{a}$ es punto de acumulación de C.

[09]$_2$ Teorema de Bolzano-Weierstrass para conjuntos

Todo conjunto infinito y acotado, de puntos de $\mathbb{R}^p$, tiene algún punto de acumulación, en $\mathbb{R}^p$.

Demostración

Como el conjunto, al que llamaremos C, es infinito, con sus puntos se puede formar una sucesión S con todos sus elementos distintos. Dado que C está acotado, entonces también lo está S. Según el teorema de Bolzano-Weierstrass (véase [05]), la sucesión S tiene algún límite de oscilación $\boldsymbol{a} \in \mathbb{R}^p$. Por tanto, hay alguna subsucesión S_1, de S, que converge hacia $\boldsymbol{a}$. Como los elementos de S_1 son distintos entre sí (pues lo eran los de S), de entre ellos no puede haber más que uno o ninguno que sea $\boldsymbol{a}$, lo que nos permite asegurar que $S_1 - \{\boldsymbol{a}\}$ es una sucesión de puntos de $C - \{\boldsymbol{a}\}$ que converge hacia $\boldsymbol{a}$, luego $\boldsymbol{a}$ es un punto de acumulación de C, con lo que concluye la demostración.

1.2. LÍMITE DE UNA FUNCIÓN EN UN PUNTO

Las funciones f que nos van a preocupar son aquellas que toman sus valores en $\mathbb{R}^q$, o sea, tales que $f(\boldsymbol{x}) \in \mathbb{R}^q$, donde «la variable» $\boldsymbol{x}$ es un punto de $\mathbb{R}^p$, y nos vamos a ocupar en estudiar si, cuando $\boldsymbol{x}$ «tiende» a un $\boldsymbol{a} \in \mathbb{R}^p$, dado, acontece que $f(\boldsymbol{x})$ «tiende» a un cierto $\boldsymbol{l} \in \mathbb{R}^q$. Cuando así sea, se dirá que $\boldsymbol{l}$ es el límite de f en el punto $\boldsymbol{a}$. Nótese que, para empezar, habrá que dar significado preciso a las expresiones «$\boldsymbol{x}$ tiende a $\boldsymbol{a}$» y «$f(\boldsymbol{x})$ tiende a $\boldsymbol{l}$». Esta definición, de límite de una función f en un punto $\boldsymbol{a}$, así como las propiedades de tales límites van a ser soporte básico del presente curso.

❑ ACERCA DE LAS FUNCIONES DE VARIAS VARIABLES

[10] Las funciones que van a ser objeto de estudio son las de $\mathscr{F}(C, \mathbb{R}^q)$, siendo $C \subset \mathbb{R}^p$, es decir, las $\boldsymbol{x} \mapsto f(\boldsymbol{x})$ donde $\boldsymbol{x} = (x_1, x_2, ..., x_p)$ recorre el conjunto $C \subset \mathbb{R}^p$ (se dice que hay p variables) y $f(\boldsymbol{x}) \in \mathbb{R}^q$. La función se llama real o vectorial según que sea $q = 1$ o $q \geqslant 2$. Cuando es $q \geqslant 2$ y si $f(\boldsymbol{x}) = (y_1, y_2, ..., y_q)$, se llaman coordenadas o componentes de f a las funciones reales $f_i : C \to \mathbb{R}$, $\boldsymbol{x} \mapsto f_i(\boldsymbol{x}) = y_i$ (para $i = 1, 2, ..., q$).

[10]$_1$ Observaciones

1. Cuando a una función, de p variables $x_1, x_2, ..., x_p$, se la conoce a través de una expresión analitica, $z = f(x_1, x_2, ..., x_p)$, que para cada sistema de valores que se asigna a $(x_1, x_2, ..., x_p)$ proporciona el valor z de la imagen, se suele entender que el campo de definición o dominio de dicha función es el conjunto más amplio de puntos $(x_1, x_2, ..., x_p)$ de $\mathbb{R}^p$ para el que existe un valor para z. Así, por ejemplo, la expresión

$$z = [1 - x^2 - y^2]^{-1/2} \text{ L } (xy)$$

«define a z como función de (x, y)» allí donde exista el segundo miembro; el primer factor existe en el círculo abierto de centro el origen y radio 1 y el segundo factor existe en el primer y tercer cuadrantes (sin incluir los ejes) del plano $\mathbb{R}^2$, luego el dominio de esta función es el conjunto

$$C = \{(x, y) \in \mathbb{R}^2 \ / \ xy > 0, x^2 + y^2 < 1\}$$

2. El conjunto $\mathscr{F}(C, \mathbb{R}^q)$, de las funciones definidas en el conjunto C y que toman sus valores en $\mathbb{R}^q$, es un espacio vectorial respecto de las operaciones usuales de suma de funciones y de producto de un número λ por una función:

$$(\boldsymbol{f} + \boldsymbol{g})(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{g}(\boldsymbol{x}) \qquad \text{y} \qquad (\lambda \boldsymbol{f})(\boldsymbol{x}) = \lambda \boldsymbol{f}(\boldsymbol{x}), \ \forall \ \boldsymbol{x} \in C$$

El conjunto $\mathscr{F}(C, \mathbb{R})$, caso particular del anterior ($q = 1$), es además un anillo (unitario, conmutativo y con divisores de cero) respecto de las operaciones usuales de suma y producto de funciones:

$$(f + g)(\boldsymbol{x}) = f(\boldsymbol{x}) + g(\boldsymbol{x}) \qquad \text{y} \qquad (fg)(\boldsymbol{x}) = f(\boldsymbol{x})g(\boldsymbol{x}), \qquad \forall \ \boldsymbol{x} \in C$$

3. Para este último conjunto $\mathscr{F}(C, \mathbb{R})$, la relación $\leqslant$ definida mediante

$$[f \leqslant g] \iff [f(\boldsymbol{x}) \leqslant g(\boldsymbol{x}), \ \forall \ \boldsymbol{x} \in C]$$

es un orden (parcial, salvo si C tuviera un solo elemento). Si o es la función nula (o sea $o(\boldsymbol{x}) = 0, \ \forall \ \boldsymbol{x} \in C$), se dice que una función f es positiva si es $o \leqslant f$ y se dice que es negativa si es $f \leqslant o$.

4. Se dice que una función $\boldsymbol{f}: C \to \mathbb{R}^q$ está acotada si su imagen $\boldsymbol{f}(C)$ es un conjunto acotado de $\mathbb{R}^q$, esto es, si existe $K > 0$ tal que $\|\boldsymbol{f}(\boldsymbol{x})\| \leqslant K$ para todo $\boldsymbol{x} \in C$. En particular, si $\boldsymbol{f}$ es real (es decir, para $q = 1$) y está acotada, entonces se llaman supremo de f (sup f) e ínfimo de f (inf f) a los valores supremo e ínfimo de $f(C)$.

CONCEPTO DE LÍMITE DE UNA FUNCIÓN

[11]

Sea $f: C \to \mathbb{R}^q$ una función, definida en un conjunto $C \subset \mathbb{R}^p$, y sea $\boldsymbol{a} \in \mathbb{R}^p$ un punto de acumulación de C. Se dice que $\boldsymbol{l} \in \mathbb{R}^q$ es *límite* de f en el punto $\boldsymbol{a}$ si se verifica una de las dos condiciones siguientes, que son equivalentes entre sí:

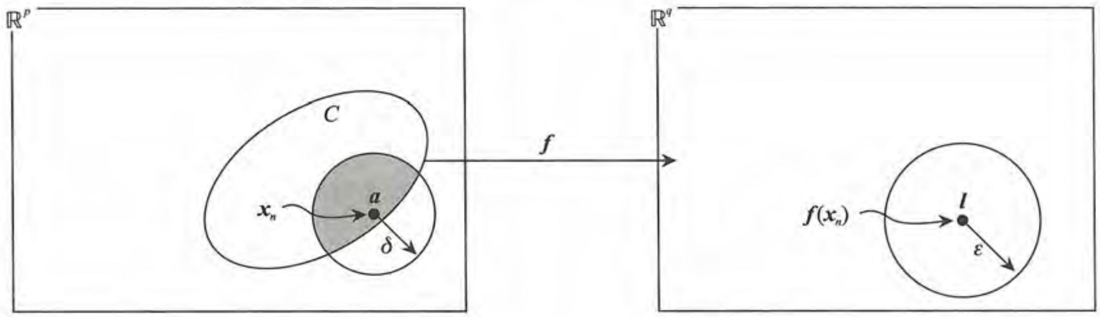
a) [Condición «$\varepsilon : \delta$»] Para cada $\varepsilon > 0$, existe un $\delta > 0$ tal que:

$$(\boldsymbol{x} \in C - \{\boldsymbol{a}\}, \|\boldsymbol{x} - \boldsymbol{a}\| < \delta) \quad \Rightarrow \quad \|f(\boldsymbol{x}) - \boldsymbol{l}\| < \varepsilon$$

b) Si $(\boldsymbol{x}_n)$ es una sucesión, de puntos de $C - \{\boldsymbol{a}\}$, que tiene límite $\boldsymbol{a}$, entonces la sucesión $(f(\boldsymbol{x}_n))$ tiene límite $\boldsymbol{l}$.

Para expresar que $\boldsymbol{l}$ es el límite de la función f en el punto $\boldsymbol{a}$, se pone:

$$\boldsymbol{l} = \lim_{\boldsymbol{x} \to \boldsymbol{a}} f(\boldsymbol{x})$$



Demostración

1.º [$a \Rightarrow b$] Suponiendo que se verifica la condición «$\varepsilon : \delta$» y si $(\boldsymbol{x}_n)$ es una sucesión de puntos de $C - \{\boldsymbol{a}\}$ que tiene límite $\boldsymbol{a}$, entonces $(f(\boldsymbol{x}_n))$ ha de tener límite $\boldsymbol{l}$. En efecto: dado $\varepsilon > 0$, por cumplirse la condición «$\varepsilon : \delta$» se sabe que existe $\delta > 0$ tal que, siempre que $\boldsymbol{x} \in C - \{\boldsymbol{a}\}$ sea tal que $\|\boldsymbol{x} - \boldsymbol{a}\| < \delta$, se verifica que $\|f(\boldsymbol{x}) - \boldsymbol{l}\| < \varepsilon$; ahora bien, como $\lim \boldsymbol{x}_n = \boldsymbol{a}$, existe $\nu \in \mathbb{N}$ tal que para $n \geqslant \nu$ es $\|\boldsymbol{x}_n - \boldsymbol{a}\| < \delta$, luego como además $\boldsymbol{x}_n \in C - \{\boldsymbol{a}\}$, es:

$$n \geqslant \nu \quad \Rightarrow \quad \|\boldsymbol{x}_n - \boldsymbol{a}\| < \delta \quad \Rightarrow \quad \|f(\boldsymbol{x}_n) - \boldsymbol{l}\| < \varepsilon$$

lo que prueba que $\lim f(\boldsymbol{x}_n) = \boldsymbol{l}$ (pues se cumple la condición «$\varepsilon : \nu$»), como había que comprobar.

2.º $[b \Rightarrow a]$ Comprobemos su contrarrecíproco [(no) $a \Rightarrow$ (no) b], para lo que, suponiendo que existe cierto $\varepsilon_0 > 0$ tal que para cada $\delta > 0$ hay algún $\boldsymbol{x} \in C - \{\boldsymbol{a}\}$ tal que $\|\boldsymbol{x} - \boldsymbol{a}\| < \delta$ y $\|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{l}\| \geqslant \varepsilon_0$, vamos a comprobar que existe cierta sucesión $(\boldsymbol{x}_n)$ de puntos de $C - \{\boldsymbol{a}\}$ tal que $\lim \boldsymbol{x}_n = \boldsymbol{a}$ y $(\boldsymbol{f}(\boldsymbol{x}_n))$ no tiene límite $\boldsymbol{l}$. En efecto: para cada $n \in \mathbb{N}$ y tomando $\delta = 1/n$, sabemos que existe $\boldsymbol{x}_n \in C - \{\boldsymbol{a}\}$ tal que $\|\boldsymbol{x}_n - \boldsymbol{a}\| < 1/n$ y $\|\boldsymbol{f}(\boldsymbol{x}_n) - \boldsymbol{l}\| \geqslant \varepsilon_0$; la sucesión $(\boldsymbol{x}_n)$ que así se obtiene resuelve, pues, la cuestión.

$[11]_1$ Observaciones

- La condición «$\varepsilon:\delta$» de límite puede expresarse, recurriendo a los entornos (o las bolas) de $\mathbb{R}^p$ y $\mathbb{R}^q$, diciendo que: para cada $\varepsilon > 0$ existe $\delta > 0$ tal que

$$\boldsymbol{x} \in B^*(\boldsymbol{a}, \delta) \cap C \quad \Rightarrow \quad \boldsymbol{f}(\boldsymbol{x}) \in B(\boldsymbol{l}, \varepsilon)$$

- Interesa resaltar que la definición de límite (de $\boldsymbol{f}$ en $\boldsymbol{a}$) es tal que, en ella, no tiene la menor influencia el posible valor de la función $\boldsymbol{f}$ en el punto $\boldsymbol{a}$.
- También hay que señalar que, en esta definición de límite, es fundamental que $\boldsymbol{a}$ sea un punto de acumulación de C. Si así no fuese, habría un $\delta_0 > 0$ tal que en la bola $B^*(\boldsymbol{a}, \delta_0)$ no habría puntos de C, lo que conduciría a que se verificase la condición «$\varepsilon:\delta$» (con $\delta = \delta_0$ para todo $\varepsilon > 0$) y ello tomando $\boldsymbol{l} \in \mathbb{R}^q$ cualquiera.

$[11]_2$ Ejemplo

La función real definida, en $\mathbb{R}^2 - \{(0, 0)\}$, mediante la expresión:

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

tiene límite 1 en (0, 0). En efecto: la condición «$\varepsilon:\delta$» de límite se cumple si se toma $\delta = \varepsilon$, ya que si $\sqrt{x^2 + y^2} < \delta$, entonces

$$|f(x, y) - 1| = \frac{|2x^2y + 2x^2 + y^2 - 2x^2 - y^2|}{|2x^2 + y^2|} = \frac{2x^2}{2x^2 + y}\,|y| \leqslant |y| < \delta = \varepsilon$$

$[11]_3$ Ejemplo

Considérese la función real $(x, y) \mapsto f(x, y)$ definida mediante

$$f(x, y) = \frac{x^2y}{x^2 - y}, \qquad \text{para } x^2 \neq y$$

(definida en el conjunto $C = \{(x, y) \in \mathbb{R}^2 \,/\, x^2 \neq y\}$). Esta función carece de límite en (0, 0), pues al hacer que $(x, y) \to (0, 0)$ de modos distintos se pueden obtener límites distintos. En efecto:

— Para la sucesión de los puntos $(x_n, y_n) = (1/n, 1/n)$, que tiene límite (0, 0), la función f tiene límite 0, pues:

$$f(x_n, y_n) = \frac{x_n^2 y_n}{x_n^2 - y_n} = \frac{1/n^3}{(1/n^2) - (1/n)} = \frac{1}{n - n^2} \to 0 \qquad \text{(cuando } n \to \infty)$$

— Para la sucesión de puntos $(x_n, y_n) = (1/\sqrt{n}, 1/n - 1/n^2)$, que tiene límite (0, 0), la función f tiene límite 1, pues:

$$f(x_n, y_n) = \frac{\dfrac{1}{n}\left(\dfrac{1}{n} - \dfrac{1}{n^2}\right)}{\dfrac{1}{n^2}} = 1 - \frac{1}{n} \to 1 \qquad \text{(cuando } n \to \infty)$$

[11]$_4$ Criterio de Cauchy

Sea $\boldsymbol{f}: C \to \mathbb{R}^q$ una función, definida en un conjunto $C \subset \mathbb{R}^p$, y sea $\boldsymbol{a} \in \mathbb{R}^p$ un punto de acumulación de C. Para que $\boldsymbol{f}$ tenga límite (finito) en $\boldsymbol{a}$ es necesario y suficiente que se cumpla la siguiente condición (de Cauchy): para cada $\varepsilon > 0$ existe un $\delta > 0$ tal que

$$[\boldsymbol{x}, \boldsymbol{x}' \in C - \{\boldsymbol{a}\}, \|\boldsymbol{x} - \boldsymbol{a}\| < \delta, \|\boldsymbol{x}' - \boldsymbol{a}\| < \delta] \;\Rightarrow\; \|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}')\| < \varepsilon$$

Demostración

1.º Suponiendo que $\boldsymbol{f}$ tiene límite $\boldsymbol{l} \in \mathbb{R}^p$ en $\boldsymbol{a}$, veamos que entonces se cumple la condición de Cauchy. En efecto: dado $\varepsilon > 0$, acudiendo a la condición «$\varepsilon : \delta$» de límite, se sabe que existe un $\delta > 0$ tal que, si $\boldsymbol{x} \in C - \{\boldsymbol{a}\}$ es tal que $\|\boldsymbol{x} - \boldsymbol{a}\| < \delta$, se verifica que $\|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{l}\| < \varepsilon/2$; por tanto, para $\boldsymbol{x}, \boldsymbol{x}' \in C - \{\boldsymbol{a}\}$, es:

$$[\|\boldsymbol{x} - \boldsymbol{a}\| < \delta, \|\boldsymbol{x}' - \boldsymbol{a}\| < \delta] \;\Rightarrow\; \|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}')\| \leqslant$$
$$\leqslant \|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{l}\| + \|\boldsymbol{f}(\boldsymbol{x}') - \boldsymbol{l}\| < \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon$$

2.º Supongamos ahora que se verifica la condición de Cauchy. Empecemos comprobando que, entonces, si una sucesión $(\boldsymbol{x}_n)$, de puntos de $C - \{\boldsymbol{a}\}$, tiene límite $\boldsymbol{a}$, entonces la sucesión $(\boldsymbol{f}(\boldsymbol{x}_n))$ tiene límite (finito); en efecto, dado $\varepsilon > 0$ y si es $\delta > 0$ el valor que da (para dicho ε) la condición de Cauchy, como $\lim \boldsymbol{x}_n = \boldsymbol{a}$, existe $\nu \in \mathbb{N}$ tal que $\|\boldsymbol{x}_n - \boldsymbol{a}\| < \delta$ para todo $n \geqslant \nu$ y, por cumplirse la condición de Cauchy, será:

$$p, q \geqslant \nu \;\Rightarrow\; [\|\boldsymbol{x}_p - \boldsymbol{a}\| < \delta, \|\boldsymbol{x}_q - \boldsymbol{a}\| < \delta] \;\Rightarrow\; \|\boldsymbol{f}(\boldsymbol{x}_p) - \boldsymbol{f}(\boldsymbol{x}_q)\| < \varepsilon$$

es decir, la sucesión $(f(\boldsymbol{x}_n))$ es de Cauchy, luego $(f(\boldsymbol{x}_n))$ es una sucesión convergente (véase [04]), como se había adelantado. Nos encontramos, pues, con que todas las tales sucesiones $(f(\boldsymbol{x}_n))$ son convergentes, lo que nos lleva a que todas tengan el mismo límite pues, de no ser así y si $(f(\boldsymbol{x}_n)) \to \boldsymbol{l}$ y $(f(\boldsymbol{x}'_n)) \to \boldsymbol{l}'$ con $\boldsymbol{l} \neq \boldsymbol{l}'$, resultaría que

$$\boldsymbol{x}_1, \boldsymbol{x}'_1, \ldots, \boldsymbol{x}_n, \boldsymbol{x}'_n, \cdots \to \boldsymbol{a} \qquad \text{y}$$

$$f(\boldsymbol{x}_1), f(\boldsymbol{x}'_1), \ldots, f(\boldsymbol{x}_n), f(\boldsymbol{x}'_n), \ldots \qquad \text{no converge,}$$

que es falso. Si todas estas sucesiones $(f(\boldsymbol{x}_n))$ tienen el mismo límite, al que llamaremos $\boldsymbol{l}$, entonces f tiene límite $\boldsymbol{l}$ en $\boldsymbol{a}$, de acuerdo con la definición de límite mediante sucesiones.

[11]$_5$ Límite infinito y límite en el infinito

Límite infinito. Sea $f: C \to \mathbb{R}^q$ una función, definida en un conjunto $C \subset \mathbb{R}^p$, y sea $\boldsymbol{a} \in \mathbb{R}^p$ un punto de acumulación de C. Se dice que f tiene límite infinito en $\boldsymbol{a}$ si para cada $K > 0$ existe un $\delta > 0$ tal que, para todo $\boldsymbol{x} \in B^*(\boldsymbol{a}, \delta) \cap C$, se verifica que $\|f(\boldsymbol{x})\| > K$. Cuando así ocurre, se pone:

$$\lim_{\boldsymbol{x} \to \boldsymbol{a}} f(\boldsymbol{x}) = \infty$$

Límite en el infinito. Sea $f: C \to \mathbb{R}^q$ una función, definida en un conjunto $C \subset \mathbb{R}^p$, y supóngase que C no está acotado. Se dice que f tiene límite $\boldsymbol{l} \in \mathbb{R}^q$ en el infinito si para cada $\varepsilon > 0$ existe un $H > 0$ tal que para los $\boldsymbol{x} \in C$ tales que $\|\boldsymbol{x}\| > H$ es $\|f(\boldsymbol{x}) - \boldsymbol{l}\| < \varepsilon$. Cuando así ocurre, se pone:

$$\lim_{\boldsymbol{x} \to \infty} f(\boldsymbol{x}) = \boldsymbol{l}$$

Límite infinito en el infinito. Sea $f: C \to \mathbb{R}^q$ una función, definida en un conjunto $C \subset \mathbb{R}^p$, y se supone que C no está acotado. Se dice que f tiene límite infinito en el infinito si para cada $K > 0$ existe un $H > 0$ tal que para los $\boldsymbol{x} \in C$ tales que $\|\boldsymbol{x}\| > H$ es $\|f(\boldsymbol{x})\| > K$. Cuando así ocurre, se pone

$$\lim_{\boldsymbol{x} \to \infty} f(\boldsymbol{x}) = \infty$$

En el caso de ser f una función real, $f: C \to \mathbb{R}$, cabe hablar de «límite $+\infty$» y de «límite $-\infty$». Así, por ejemplo, se dice que f tiene límite $+\infty$ en un punto $\boldsymbol{a}$, de acumulación de C, si para cada $K > 0$ existe un $\delta > 0$ tal que, para todo $\boldsymbol{x} \in B^*(\boldsymbol{a}, \delta) \cap C$, se verifica que $f(\boldsymbol{x}) > K$.

PRIMERAS PROPIEDADES DE LOS LÍMITES

[12]

Sea $f: C \to \mathbb{R}^q$ una función, definida en un conjunto $C \subset \mathbb{R}^p$, y sea $\boldsymbol{a} \in \mathbb{R}^p$ un punto de acumulación de C. Se verifica que:

1. Si f tiene límite en $\boldsymbol{a}$, entonces f tiene un solo límite en $\boldsymbol{a}$.
2. La función f tiene límite $\boldsymbol{l} \in \mathbb{R}^q$ en $\boldsymbol{a}$ si, y sólo si, la función $\boldsymbol{x} \mapsto f(\boldsymbol{x}) - \boldsymbol{l}$ tiene límite $\boldsymbol{o}$ en $\boldsymbol{a}$ (se dice entonces que $f(\boldsymbol{x}) - \boldsymbol{l}$ es un infinitésimo en $\boldsymbol{a}$).
3. Si f tiene límite en $\boldsymbol{a}$, entonces f está acotada «cerca de $\boldsymbol{a}$», esto es, existe un entorno reducido de $\boldsymbol{a}$ en el que f está acotada.
4. Dado $D \subset C$, si la rectricción de f a D tiene límite en $\boldsymbol{a}$, a este se le llama límite de f en $\boldsymbol{a}$ según D. Pues bien, si f tiene límite $\boldsymbol{l}$ en $\boldsymbol{a}$, entonces f tenga límite en $\boldsymbol{a}$ según todo $D \subset C$ (que tenga a $\boldsymbol{a}$ como punto de acumulación) y, en particular, según las rectas que pasan por $\boldsymbol{a}$.
5. Si las componentes de f son $f_1, \ldots, f_q$, para que f tenga límite $\boldsymbol{l} = (l_1, \ldots, l_q)$ en $\boldsymbol{a}$, es necesario y suficiente que f_i tenga límite l_i en $\boldsymbol{a}$ para $i = 1, 2, \ldots, q$.

Demostración

1. Acudiendo a la definición de límite mediante sucesiones (véase [11],b), nos encontramos con que esta propiedad es consecuencia obvia de la unicidad del límite de una sucesión convergente (véase [03],2).
2. Las condiciones «$\varepsilon:\delta$» que expresan que $f(\boldsymbol{x}) \to \boldsymbol{l}$ y que $f(\boldsymbol{x}) - \boldsymbol{l} \to \boldsymbol{o}$ (cuando $\boldsymbol{x} \to \boldsymbol{a}$) resultan ser la misma, lo que prueba que $f(\boldsymbol{x}) \to \boldsymbol{l}$ equivale a $f(\boldsymbol{x}) - \boldsymbol{l} \to \boldsymbol{o}$.
3. Llamando $\boldsymbol{l}$ al límite y tomando $\varepsilon = 1$ en la condición «$\varepsilon:\delta$» de límite, existe $\delta > 0$ tal que para $\boldsymbol{x} \in C$ en la bola reducida $B^*(\boldsymbol{a}, \delta)$ es $\|f(\boldsymbol{x}) - \boldsymbol{l}\| < \varepsilon$, luego f está acotada en el entorno reducido $B^*(\boldsymbol{a}, \delta)$, como había que comprobar.
4. Si la condición «$\varepsilon:\delta$» de límite se verifica para los puntos $\boldsymbol{x} \in C - \{\boldsymbol{a}\}$, entonces se verifican, en particular para los puntos de $D - \{\boldsymbol{a}\}$.
5. Si tenemos presente la definición de límite mediante sucesiones, resulta evidente que esta propiedad se verifica como consecuencia de su homólogo para sucesiones (véase [03],1).

[12]$_1$ Observación

Dada una función $f: C \to \mathbb{R}^q$, definida en un conjunto $C \subset \mathbb{R}^p$, y si $\boldsymbol{a} \in \mathbb{R}^p$ es un punto de acumulación de C, para comprobar que $\boldsymbol{l} \in \mathbb{R}^q$ es el límite de f en $\boldsymbol{a}$, es suficiente con encontrar una función real y positiva φ, definida «cerca de $\boldsymbol{a}$», que tenga límite 0 en $\boldsymbol{a}$ y sea tal que $\|f(\boldsymbol{x}) - \boldsymbol{l}\| < \varphi(\boldsymbol{x})$ para $\boldsymbol{x}$ cerca de $\boldsymbol{a}$. Es decir, se verifica la siguiente implicación (en la que $\boldsymbol{x}$ es un punto cualquiera de $U^* \cap C$, donde U^* es un cierto entorno reducido de $\boldsymbol{a}$):

$$\left.\begin{array}{l} \text{para «}\boldsymbol{x}\text{ cerca de }\boldsymbol{a}\text{» es } \|f(\boldsymbol{x}) - \boldsymbol{l}\| \leqslant \varphi(\boldsymbol{x}) \\ (\varphi(\boldsymbol{x}) \text{ es real y positivo}),\ \lim\limits_{\boldsymbol{x} \to \boldsymbol{a}} \varphi(\boldsymbol{x}) = 0 \end{array}\right\} \Rightarrow \lim_{\boldsymbol{x} \to \boldsymbol{a}} f(\boldsymbol{x}) = \boldsymbol{l}$$

Este resultado se comprueba fácilmente, viendo que la condición «$\varepsilon:\delta$» del límite de f en $\boldsymbol{a}$ se verifica, fijado un $\varepsilon > 0$ (cualquiera), si se toma para δ el valor que se obtiene al aplicar dicha condición «$\varepsilon:\delta$» al límite de φ en $\boldsymbol{a}$.

Ejemplo 1.º La función real f definida en $\mathbb{R}^2 - \{(0, 0)\}$ por:

$$f(x, y) = \frac{x^3 - y^7}{x^2 + y^4}$$

tiene límite 0 cuando $(x, y) \to (0, 0)$ ya que, para (x, y) «cerca» de $(0, 0)$ en concreto, para $|x| < 1$ y $|y| < 1$, se puede poner:

$$|f(x, y) - 0| = \left|\frac{x^3 - y^7}{x^2 + y^4}\right| = \left|x\,\frac{x^2}{x^2 + y^4} - y^3\,\frac{y^4}{x^2 + y^4}\right| \leqslant$$

$$\leqslant |x|\,\frac{x^2}{x^2 + y^4} + |y|^3\,\frac{y^4}{x^2 + y^4} < |x| + |y|^3 < |x| + |y|$$

y $\varphi(x, y) = |x| + |y|$ tiene límite 0 cuando $(x, y) \to (0, 0)$, como se comprueba trivialmente.

Ejemplo 2.º La función f definida en $\mathbb{R}^2 - \{(0, 0)\}$ mediante la expresión:

$$f(x, y) = \frac{x^{a+1}y^{b+1}}{x^2 + y^2 - xy} \qquad \text{(donde es } a > 0 \text{ y } b > 0)$$

tiene límite 0 cuando $(x, y) \to (0, 0)$, ya que (acudiendo a coordenadas polares: $x = \rho \cos \theta$, $y = \rho \operatorname{sen} \theta$):

$$|f(x, y) - 0| = |x^a y^b|\,\frac{|\cos\theta \operatorname{sen}\theta|}{|1 - \operatorname{sen}\theta\cos\theta|} = |x^a y^b|\,\frac{|\cos\theta \operatorname{sen}\theta|}{|1 - (1/2)\operatorname{sen} 2\theta|} < |x^a y^b|\,\frac{1}{1/2} = 2|x^a y^b|$$

y $\varphi(x, y) = 2|x^a y^b|$ tiene límite 0 cuando $(x, y) \to (0, 0)$, como se comprueba trivialmente.

$[12]_2$ Límites direccionales

Sea $f: C \to \mathbb{R}^q$ una función, definida en el conjunto $C \subset \mathbb{R}^p$, y sea $\boldsymbol{a} \in \mathbb{R}^p$ un punto de acumulación de C. Si R es una recta de $\mathbb{R}^p$ que pasa por el punto $\boldsymbol{a}$, consideremos la restricción de f a R, es decir la función $f_R : C \cap R \to \mathbb{R}^p$ definida por $f_R(\boldsymbol{x}) = f(\boldsymbol{x})$ para todo $\boldsymbol{x} \in C \cap R$, y supongamos que $\boldsymbol{a}$ es punto de acumulación de $C \cap R$. Pues bien, respondiendo a lo que ya se indicó al hablar de los límites según subconjuntos (véase [12],4), se dice que f tiene límite $\boldsymbol{l}$ en $\boldsymbol{a}$ según la recta R o según la dirección de R (límite direccional) si f_R tiene límite $\boldsymbol{l}$ en $\boldsymbol{a}$.

Es evidente que, si f tiene límite $\boldsymbol{l}$ en $\boldsymbol{a}$, entonces f tiene límite $\boldsymbol{l}$ en $\boldsymbol{a}$ según toda recta R que pase por $\boldsymbol{a}$ (se supone que $\boldsymbol{a}$ es punto de acumulación de $R \cap C$). Sin embargo, no es suficiente, con que f tenga límite $\boldsymbol{l}$ en $\boldsymbol{a}$ según todas las direcciones, para poder garantizar que f tiene límite $\boldsymbol{l}$ en $\boldsymbol{a}$, como veremos enseguida con algún ejemplo; en tal supuesto, lo que sólo se puede asegurar, con carácter general, es que, si f tuviera límite en $\boldsymbol{a}$, dicho límite sería $\boldsymbol{l}$. Señalemos también que, si no existiera el límite de f en $\boldsymbol{a}$ según una cierta recta R (se supone que $\boldsymbol{a}$ es punto de acumulación de $R \cap C$), entonces f no tendría límite en $\boldsymbol{a}$; a esta misma consecuencia se llega si f tuviera, en $\boldsymbol{a}$, límites direccionales distintos según dos rectas diferentes.

Ejemplo 1.º La función real f definida, en $C = \{(x, y) \in \mathbb{R}^2 \,/\, x \neq \pm y\}$, por

$$f(x, y) = \frac{xy + y^3}{x^2 - y^2}$$

no tiene límite en (0, 0), pues sus límites direccionales en (0, 0) son distintos. En efecto: las rectas que pasan por (0, 0) tienen por ecuaciones a las $y = mx$ (para $m \in \mathbb{R}$) y $x = 0$; para cada una de ellas es:

$$\lim_{x \to 0} f(x, mx) = \lim_{x \to 0} \frac{mx^2 + m^3x^3}{(1 - m^2)x^2} = \frac{m}{1 - m^2}, \qquad \text{para } m \neq \pm 1$$

$$\lim_{y \to 0} f(0, y) = \lim_{y \to 0} \frac{y^3}{-y^2} = 0$$

como estos límites son distintos, f carece de límite en (0, 0).

Ejemplo 2.º Consideremos la función real definida, en el conjunto $C = \{(x, y) \in \mathbb{R}^2 \,/\, y \neq x^2\}$, mediante la expresión

$$f(x, y) = \frac{x^3}{x^2 - y}$$

Esta función tiene, en (0, 0), límites según todas las direcciones y estos límites valen 0, ya que:

$$\lim_{x \to 0} f(x, mx) = \lim_{x \to 0} \frac{x^3}{x^2 - mx} = 0, \qquad \text{para } m \in \mathbb{R}$$

$$\lim_{y \to 0} f(0, y) = \lim_{y \to 0} \frac{0}{-y^2} = 0$$

Sin embargo, f no tiene límite en (0, 0), ya que según la curva $y = x^2 - x^3$ (que pasa por el origen) tiene límite distinto de 0 (si existiera el límite de f en (0, 0), éste tendría que ser 0, que es el valor que toman todos los direccionales), ya que:

$$\lim_{x \to 0} f(x, x^2 - x^3) = \lim_{x \to 0} \frac{x^3}{x^2 - x^2 + x^3} = 1 \neq 0$$

Ejemplo 3.º Sea $f: \mathbb{R}^2 \to \mathbb{R}$ una función que tiene límite l en (0, 0) y sea $\varphi: \mathbb{R}^2 \to \mathbb{R}$ la función definida mediante: $\varphi: (x, y) = f(x, y)$, si $y \neq x^2$, y $\varphi(x, x^2) = f(x, x^2) + a$, con $a \neq 0$. La función φ tiene límite l en (0, 0) según todas las direcciones, pero no tiene límite en (0, 0), pues según la parábola $y = x^2$ tiene límite $l + a$ en dicho punto (0, 0).

[12]$_3$ Desigualdades entre funciones y límites (en el caso real)

Sean f, g y h tres funciones reales, definidas en un mismo conjunto $C \subset \mathbb{R}^p$, y sea $\boldsymbol{a} \in \mathbb{R}^p$ un punto de acumulación de C.

1. En el supuesto de que exista el límite $l \in \mathbb{R}$ de f en $\boldsymbol{a}$, se verifica que:
   - Si es $k < l$ (es $k > l$), entonces es $k < f(\boldsymbol{x})$ (es $k > f(\boldsymbol{x})$) «para $\boldsymbol{x}$ cerca de $\boldsymbol{a}$»(*).
   - Si es $k < f(\boldsymbol{x})$ (es $k > f(\boldsymbol{x})$) «para $\boldsymbol{x}$ cerca de $\boldsymbol{a}$»(*), entonces es $k \leqslant l$ (es $k \geqslant l$).
2. En el supuesto de que existan los límites l, $m \in \mathbb{R}$ de f y g en $\boldsymbol{a}$, se verifica:
   - Si es $l < m$, entonces es $f(\boldsymbol{x}) < g(\boldsymbol{x})$ «para $\boldsymbol{x}$ cerca de $\boldsymbol{a}$»(*).
   - Si es $f(\boldsymbol{x}) < g(\boldsymbol{x})$ «para $\boldsymbol{x}$ cerca de $\boldsymbol{a}$»(*), entonces es $l \leqslant m$.
3. Si f y g tienen el mismo límite l, en el punto $\boldsymbol{a}$, y si es $f(\boldsymbol{x}) \leqslant h(\boldsymbol{x}) \leqslant g(\boldsymbol{x})$ «para $\boldsymbol{x}$ cerca de $\boldsymbol{a}$»(*), entonces h tiene límite l en $\boldsymbol{a}$ (regla del emparedado).

Demostración

1. La primera parte de esta propiedad se obtiene directamente de la condición «$\varepsilon:\delta$» de límite, tomando $\varepsilon = l - k$; para $\boldsymbol{x} \in C$ en el correspondiente entorno $B^*(\boldsymbol{a}, \delta)$ es $|f(\boldsymbol{x}) - l| < \varepsilon$ y, por tanto, $k < f(\boldsymbol{x})$. La segunda parte es la contrarrecíproca de la primera, por lo que también es cierta.
2. La primera parte de la propiedad se obtiene acudiendo a la condición «$\varepsilon:\delta$» de límite, en la que tomamos $\varepsilon = \frac{1}{2}(m - l)$; si son δ_l y δ_m los correspondientes valores de δ, para los límites l y m, la propiedad se cumple en el entorno de radio $\delta = \text{mín}\ \{\delta_l, \delta_m\}$. La segunda parte es la contrarrecíproca de la primera, por lo que también se verifica.
3. Dado $\varepsilon > 0$, sean $\delta_l > 0$ y $\delta_m > 0$ los valores de δ que se desprenden de la condición «$\varepsilon:\delta$» aplicada a l y m. Sea δ_0 el radio del entorno de $\boldsymbol{a}$ en el que es $f(\boldsymbol{x}) \leqslant h(\boldsymbol{x}) \leqslant g(\boldsymbol{x})$. Es evidente que la condición «$\varepsilon:\delta$» de límite se verifica para h si se toma $\delta = \text{mín}\ \{\delta_l, \delta_m, \delta_0\}$.

PROPIEDADES ARITMÉTICAS DE LOS LÍMITES (EN EL CASO REAL)

[13]

Sean f y g dos funciones reales, definidas en un mismo conjunto $C \subset \mathbb{R}^p$, y sea $\boldsymbol{a} \in \mathbb{R}^p$ un punto de acumulación de $\boldsymbol{a}$:

1. Si f es un infinitésimo (es decir, tiene límite 0) en $\boldsymbol{a}$ y si g está acotada en un entorno de $\boldsymbol{a}$, entonces su producto fg es un infinitésimo en $\boldsymbol{a}$.
2. Si f y g tienen límite finito en $\boldsymbol{a}$, entonces también lo tienen las siguientes funciones y dichos límites son los que se indican:

$$\left.\begin{array}{l} \lim\limits_{x \to a} f(x) = l \\ \lim\limits_{x \to a} g(x) = m \end{array}\right\} \Rightarrow \left\{\begin{array}{l} \lim\limits_{x \to a} [f(x) + g(x)] = l + m^{(*)} \\ \lim\limits_{x \to a} [f(x)g(x)] = lm \\ \lim\limits_{x \to a} [f(x)/g(x)] = l/m^{(**)} \end{array}\right.$$

Nota: Estas tres últimas igualdades siguen siendo ciertas cuando l y/o m son ∞, con tal de que $l + m$, lm, l/m no sean «indeterminaciones».

(*) Esta propiedad también es cierta cuando $\boldsymbol{f}$ y $\boldsymbol{g}$ son funciones vectoriales (o sea, si toman valor en $\mathbb{R}^q$ con $q > 1$), lo que se demuestra acudiendo a las componentes de $\boldsymbol{f}$ y $\boldsymbol{g}$.

(**) En este caso se necesita que sea $g(\boldsymbol{x}) \neq 0$ «cerca de $\boldsymbol{a}$» y $m \neq 0$.

(*) La expresión «para $\boldsymbol{x}$ cerca de $\boldsymbol{a}$» es un modo abreviado de decir «para todo $\boldsymbol{x} \in C \cap U^*$, donde U^* es un cierto entorno reducido de $\boldsymbol{a}$».

Demostración

Cuando se estudian los problemas de convergencia en $\mathbb{R}$, se ve que las anteriores propiedades se verifican si, en lugar de aludir a límites de funciones en un punto $\boldsymbol{a}$, se refieren a límites de sucesiones reales. Acudiendo a ello y a la definición de límite de una función mediante sucesiones (véase [11],b), se obtienen trivialmente las propiedades que aquí nos ocupan.

[13]$_1$ Observación

Si f y g son dos funciones reales, definidas en un mismo conjunto $C \subset \mathbb{R}^p$, y si $\boldsymbol{a}$ es un punto de acumulación de C, entonces se verifica que, siempre que las siguientes expresiones tengan sentido (no sean «indeterminaciones»), es:

$$\lim_{x \to a} (\log_u f(x)) = \log_u \left(\lim_{x \to a} f(x) \right)$$

$$\lim_{x \to a} (K^{f(x)}) = K^{\lim_{x \to a} f(x)}$$

$$\lim_{x \to a} (f(x)^{g(x)}) = \left(\lim_{x \to a} f(x) \right)^{\lim_{x \to a} g(x)}$$

Demostración

Como estas propiedades se verifican para los límites de las sucesiones de números reales, de ello y de la definición de límite de una función mediante sucesiones se deducen trivialmente las tres propiedades anteriores.

[13]$_2$ «Límites indeterminados»

1. *Casos de indeterminaciones.* La función que resulta de operar con funciones reales que tienen límite (en un cierto punto $\boldsymbol{a}$) no ha de tener, necesariamente, límite y, si lo tiene, puede ocurrir que éste no quede determinado en función de los límites de las funciones de partida. Cuando así ocurre, se suele decir que el límite en cuestión es un «límite indeterminado» o que es un «caso de indeterminación». Esos casos son (fundamentalmente) los que se representan poniendo $0/0$, ∞/∞, $0 \cdot \infty$, $\infty - \infty$, 0^0, ∞^0 y 1^∞, que corresponden a límites de la forma $f(x)/g(x)$, $f(x)/g(x)$, $f(x)g(x)$, $f(x) - g(x)$, $f(x)^{g(x)}$, $f(x)^{g(x)}$ y $f(x)^{g(x)}$ cuando $f(x)$ y $g(x)$ tienden, respectivamente a: 0 y 0, $\pm\infty$ y $\pm\infty$, 0 y $\pm\infty$, $+\infty$ y $+\infty$, 0 y 0, $\pm\infty$ y 0, 1 y $\pm\infty$.

   Para las indeterminaciones en forma de potencia, $f(x)^{g(x)}$, suele ser útil recurrir a poner

$$\lim_{x \to a} f(x)^{g(x)} = e^{\lambda}, \qquad \text{donde } \lambda = \lim_{x \to a} g(x) \, \mathrm{L}\, f(x)$$

2. *Infinitésimos e infinitésimos equivalentes.* Sean f y g dos funciones reales, definidas en un mismo conjunto $C \subset \mathbb{R}^p$, y sea $\boldsymbol{a} \in \mathbb{R}^p$ un punto de acumulación de C. Supongamos que f y g son ambas infinitésimos (o sea, que tienen límite 0) o que son ambas infinitos (o sea, que tienen límite ∞) en $\boldsymbol{a}$. Se define entonces la equivalencia, de f y g en $\boldsymbol{a}$, mediante:

$$\begin{pmatrix} f \text{ y } g \text{ son funciones} \\ \text{equivalentes en } \boldsymbol{a} \end{pmatrix} \Leftrightarrow \lim_{x \to a} \frac{f(\boldsymbol{x})}{g(\boldsymbol{x})} = 1$$

Si f y g son equivalentes en $\boldsymbol{a}$, entonces se verifica que (principio de sustitución)

$$\lim_{x \to a} F(\boldsymbol{x}) f(\boldsymbol{x}) = \lim_{x \to a} F(\boldsymbol{x}) g(\boldsymbol{x}) \qquad \text{y} \qquad \lim_{x \to a} \frac{F(\boldsymbol{x})}{f(\boldsymbol{x})} = \lim_{x \to a} \frac{F(\boldsymbol{x})}{g(\boldsymbol{x})}$$

(donde F está definida en C y se supone que existen los límites de los segundos miembros). Este resultado se prueba sin más que multiplicar y dividir a los primeros miembros por $g(\boldsymbol{x})$.

3. *Notación de Landau.* Sean $f: C \to \mathbb{R}^q$ y $g: C \to \mathbb{R}$ dos funciones (la primera vectorial, en general, y la segunda real), definidas en un mismo conjunto $C \subset \mathbb{R}^p$, y sea $\boldsymbol{a} \in \mathbb{R}^p$ un punto de acumulación de C; suponemos que es $g(\boldsymbol{x}) \neq 0$ «cerca de $\boldsymbol{a}$». Se conviene entonces en que:

$$f(\boldsymbol{x}) = \boldsymbol{o}[g(\boldsymbol{x})] \quad \text{significa que} \quad \lim_{x \to a} \frac{f(\boldsymbol{x})}{g(\boldsymbol{x})} = \boldsymbol{o}$$

$$f(\boldsymbol{x}) = \boldsymbol{O}[g(\boldsymbol{x})] \quad \text{significa que} \quad \frac{f(\boldsymbol{x})}{g(\boldsymbol{x})} \quad \text{está acotada «cerca de } \boldsymbol{a}\text{»}$$

(«cerca de $\boldsymbol{a}$» equivale a decir «para $\boldsymbol{x} \in C \cap U^*$, donde U^* es un cierto entorno reducido del punto $\boldsymbol{a}$). Nótese que, de acuerdo con la anterior definición de equivalencia (si f es también una función real, o sea, si $q = 1$), se tiene

$$\begin{pmatrix} f \text{ y } g \text{ son funciones} \\ \text{equivalentes en } \boldsymbol{a} \end{pmatrix} \Leftrightarrow f(\boldsymbol{x}) = g(\boldsymbol{x}) + o[g(\boldsymbol{x})]$$

$[13]_3$ Ejercicio

Hallar el límite en (0, 0), si existe, de la siguiente función real f definida, en $\mathbb{R}^2 - \{(0, 0)\}$, mediante la expresión:

$$f(x, y) = \frac{\sqrt{1 + x^3 - y^3} - 1}{\operatorname{arc\,sen}(x^2 + y^2 + x) - \operatorname{arc\,sen} x}$$

Resolución

Recuérdese que, si ε es un infinitésimo, entonces el infinitésimo $(1 + \varepsilon)^p - 1$ es equivalente al infinitésimo $p\varepsilon$. Acudiendo a ello, se obtiene que la función dada tiene el mismo límite que la función

$$g(x, y) = \frac{\frac{1}{2}(x^3 - y^3)}{\operatorname{arc\,sen}(x^2 + y^2 + x) - \operatorname{arc\,sen} x}$$

Si esta función tuviera límite en (0, 0), dicho límite tendrá que ser 0, ya que 0 es su límite según la recta $x = 0$. La función g tiene, efectivamente, límite 0 en (0, 0), ya que como

$$|g(x, y) - 0| < \frac{1}{2}\,\frac{|x^3 - y^3|}{|(x^2 + y^2 + x) - x|} = \frac{|x^3 - y^3|}{2(x^2 + y^2)}$$

recurriendo a coordenadas polares ($x = \rho\cos\theta$, $y = \rho\operatorname{sen}\theta$), se tiene que

$$|g(x, y) - 0| < \frac{1}{2}\,\rho|\cos^3\theta - \operatorname{sen}^3\theta|$$

y el último miembro tiene límite 0 en (0, 0) por ser el producto del infinitésimo $\frac{1}{2}\rho$ por la función acotada $|\cos^3\theta - \operatorname{sen}^3\theta|$.

LÍMITES REITERADOS

Esta cuestión de los límites reiterados (o sucesivos, como también se dice) tiene sentido cuando se consideran funciones de más de una variable. Aquí, en lugar de hacer que el punto variable $(x_1, x_2, ..., x_p)$ tienda al punto fijo $(a_1, a_2, ..., a_p)$, con lo que las variables x_i tienden simultáneamente a los valores a_i, se hace que las variables tiendan «sucesivamente» a los valores fijos: primero se hace que $x_1 \to a_1$, luego que $x_2 \to a_2$, etc. Aun cuando todo ello tiene pleno sentido para cualquiera que sea el número $p > 1$ de variables, sólo consideraremos el caso de dos variables, con lo que simplificamos la notación y abreviamos los razonamientos; la generalización al caso de más de dos variables es inmediata.

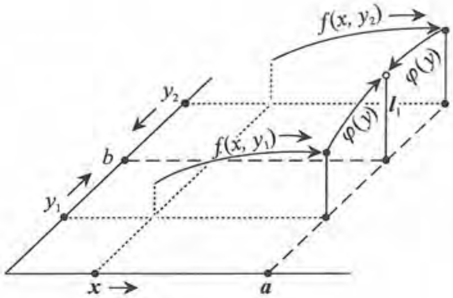
Sea $f: C \to \mathbb{R}^q$ una función, definida en un conjunto $C \subset \mathbb{R}^2$, y sea $(a, b) \in \mathbb{R}^2$ un punto de acumulación de C. Las expresiones

$$l_1 = \lim_{y \to b}\left(\lim_{x \to a} f(x, y)\right) \qquad \text{y} \qquad l_2 = \lim_{x \to a}\left(\lim_{y \to b} f(x, y)\right)$$

significa lo siguiente (hablemos de la primera, por ejemplo): 1.°, que para cada y, de un cierto entorno reducido de b, se considera la función $x \mapsto f(x, y)$; 2.°, que se supone que esta función tiene límite cuando x tiende a a, al que llamaremos (por depender de y)

$$\varphi(y) = \lim_{x \to a} f(x, y)$$

3.º, que también se supone que la función $y \mapsto \varphi(y)$ tiene límite l_1 cuando y tiende a b. En tal caso, a este l_1 se le llama límite reiterado de f cuando x tiende a a, primero, e y tiende a b, después.



Teorema del límite reiterado

[14]

Sea $f: C \to \mathbb{R}^q$ una función, definida en un conjunto $C \subset \mathbb{R}^2$, y sea $(a, b) \in \mathbb{R}^2$ un punto interior de C. Si existe y vale l el límite de f en (a, b) y si, para cada y de un entorno reducido de b, existe el límite de la función $x \mapsto f(x, y)$ cuando x tiene a a, entonces existe y vale l el límite reiterado

$$\lim_{y \to b} \left(\lim_{x \to a} f(x, y) \right)$$

(Para el otro límite reiterado se verifica un teorema análogo a éste.)

Demostración

Dado $\varepsilon > 0$, como l es el límite de f en (a, b), existe $\delta > 0$ tal que

$$\left. \begin{array}{l} 0 < |x - a| < \delta \\ 0 < |y - b| < \delta \end{array} \right\} \Rightarrow \quad \|f(x, y) - l\| < \frac{\varepsilon}{2}$$

Llamando U^* al entorno reducido de b del que habla el enunciado, sabemos que existe la función

$$y \mapsto \varphi(y) = \lim_{x \to a} f(x, y), \qquad \text{para } y \in U^*$$

Para cada $y \in U^*$ y de acuerdo con la condición «$\varepsilon : \delta$» de límite aplicada al límite $\varphi(y)$, existe un $\delta_y > 0$ tal que

$$0 < |x - a| < \delta_y \quad \Rightarrow \quad \|f(x, y) - \varphi(y)\| < \frac{\varepsilon}{2}, \qquad \text{para } y \in U^*$$

Así pues, para cualquiera que sea $y \in U^*$ tal que $0 < |y - b| < \delta$, recurriendo a cualquiera de los x tales que $0 < |x - a| < \text{mín}\ \{\delta, \delta_y\}$, se puede poner:

$$\|\varphi(y) - \boldsymbol{l}\| \leqslant \|\varphi(y) - \boldsymbol{f}(x, y)\| + \|\boldsymbol{f}(x, y) - \boldsymbol{l}\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

lo que prueba que $\lim\limits_{y \to b} \varphi(y) = \boldsymbol{l}$, como había que comprobar.

$[14]_1$ Observaciones

- Según tendremos ocasión de comprobar con los ejemplos que luego se analizan, respecto del límite de una función en un punto y de sus límites reiterados, pueden suceder cosas como las siguientes: 1) la función tiene límite en un punto, pero no existe, en dicho punto, ninguno de sus dos límites reiterados (o uno de ellos); 2) la función tiene en un punto sus dos límites reiterados y éstos son iguales, pero no existe su límite en el punto; 3) la función tiene, en un punto, sus dos límites reiterados y éstos son distintos.
- En ocasiones, el anterior teorema [14] permitirá asegurar que una función no tiene límite en un punto. Así, por ejemplo: 1.º, si una función tiene, en un punto, sus dos límites reiterados y éstos son distintos, entonces la función carece de límite en el punto; 2.º, dada la función $(x, y) \mapsto \boldsymbol{f}(x, y)$, si existe «para y cerca de b» la función $\varphi(y) = \lim\limits_{x \to a} \boldsymbol{f}(x, y)$, pero no existe $\lim\limits_{y \to b} \varphi(y)$, entonces $\boldsymbol{f}$ no tiene límite en (a, b).

$[14]_2$ Ejemplos

1.º La función real f definida, en $\mathbb{R}^2 - \{(0, 0)\}$, por la expresión

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

tiene iguales sus límites reiterados, pero carece de límites en (0, 0):

$$\lim_{y \to 0} \left(\lim_{x \to 0} f(xy) \right) = \lim_{y \to 0} \frac{0}{y^2} = \lim_{y \to 0} 0 = 0$$

$$\lim_{x \to 0} \left(\lim_{y \to 0} f(xy) \right) = \lim_{x \to 0} \frac{0}{x^2} = \lim_{x \to 0} 0 = 0$$

$$\lim_{\substack{x \to 0 \\ y \to 0}} f(x, y) \text{ no existe (sus direccionales son distintos)}$$

(el límite de f en (0, 0) según $y = mx$ es $m/(1 + m^2)$, que depende de m).

2.° La función real f definida, en $\mathbb{R}^2 - \{(0, 0)\}$, por la expresión

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

tiene distintos sus límites reiterados en el punto (0, 0), pues:

$$\lim_{y \to 0} \left(\lim_{x \to 0} f(xy) \right) = \lim_{y \to 0} \frac{y^2}{y^2} = \lim_{y \to 0} 1 = 1$$

$$\lim_{x \to 0} \left(\lim_{y \to 0} f(xy) \right) = \lim_{x \to 0} \frac{0}{x^2} = \lim_{x \to 0} 0 = 0$$

Nótese que, como consecuencia de ello, f carece de límite en el punto (0, 0).

3.° Considérese la función real $f: \mathbb{R}^2 \to \mathbb{R}$ definida por

$$f(x, y) = x \operatorname{sig}(y), \qquad \text{donde } \operatorname{sig}(y) = \begin{cases} 1, & \text{si } y > 0 \\ 0, & \text{si } y = 0 \\ -1, & \text{si } y < 0 \end{cases}$$

Esta función tiene límite 0 en (0, 0), pues es producto de un infinitésimo (x) por una función acotada (sig (y)); no obstante,

$$\nexists \lim_{x \to 0} \left(\lim_{y \to 0} f(x, y) \right), \qquad \text{pues } \nexists \lim_{y \to 0} f(x, y)$$

ya que los laterales de este último son distintos:

$$\lim_{y \to 0^+} f(x, y) = \lim_{y \to 0^+} x = x \qquad \text{y} \qquad \lim_{y \to 0^-} f(x, y) = \lim_{y \to 0^-} -x = -x$$

1.3. FUNCIONES CONTINUAS

La continuidad de una función $\boldsymbol{x} \mapsto \boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x})$, o la variación continua de $\boldsymbol{y}$ respecto de $\boldsymbol{x}$, viene a significar que incrementos pequeños de la variable $\boldsymbol{x}$ conducen a incrementos pequeños en el valor $\boldsymbol{y}$ de la función, de manera que, cuando aquel incremento tiende a anularse, entonces también es un infinitésimo el incremento de la función. Dicho de otra forma, $\boldsymbol{f}$ es continua en cierto punto $\boldsymbol{x}_0$ si $\boldsymbol{f}(\boldsymbol{x})$ tiende a $\boldsymbol{f}(\boldsymbol{x}_0)$ cuando $\boldsymbol{x}$ tiende a $\boldsymbol{x}_0$.

Las propiedades de los límites de funciones, que acabamos de estudiar, conducen de un modo natural a propiedades análogas para las funciones continuas; así, por ejemplo, veremos que la suma de funciones continuas es, a su vez, una función continua. Además de éstas, hay otras propiedades de especial relevancia, que se refieren a la continuidad global, es decir, a la continuidad en todos los puntos de un conjunto. Nos referimos a los «teoremas fundamentales de las funciones continuas», que se ocupan de estudiar qué propiedades se conservan por la continuidad. En concreto y a título de ejemplo, digamos que la compacidad es una de tales propiedades, es decir, se verifica que si f es una función continua en un conjunto C y este conjunto es compacto, entonces la imagen $f(C)$ es, también, un conjunto compacto.

CONTINUIDAD (EN UN PUNTO Y EN UN CONJUNTO)

[15]

Sea $f: C \to \mathbb{R}^q$ una función, definida en un conjunto $C \subset \mathbb{R}^p$, y considérese un punto $\boldsymbol{a} \in C$.

I. Se dice que f es *continua* en $\boldsymbol{a}$ si se verifica una de las dos condiciones siguientes, que son equivalentes entre sí:

a) [Condición «$\varepsilon : \delta$»] Para cada $\varepsilon > 0$, existe un $\delta > 0$ tal que:

$$[\boldsymbol{x} \in C,\ \|\boldsymbol{x} - \boldsymbol{a}\| < \delta] \quad \Rightarrow \quad \|f(\boldsymbol{x}) - f(\boldsymbol{a})\| < \varepsilon$$

b) Si $(\boldsymbol{x}_n)$ es una sucesión, de puntos de C, que tiene límite $\boldsymbol{a}$, entonces la sucesión $(f(\boldsymbol{x}_n))$ tiene límite $f(\boldsymbol{a})$.

En consecuencia, si el punto $\boldsymbol{a}$ es de acumulación de C(*), la función f es continua en $\boldsymbol{a}$ si, y sólo si, f tiene límite en $\boldsymbol{a}$ y dicho límite es $f(\boldsymbol{a})$, esto es, si, y sólo si, $\lim_{x \to a} f(x) = f(\boldsymbol{a})$.

II. Se dice que f es continua en un conjunto $D \subset C$ si f es continua en todo punto de D.

III. Si la función f no es continua en un punto $\boldsymbol{a}$ de acumulación de C, se dice entonces que es *discontinua* en $\boldsymbol{a}$. En tal caso, la discontinuidad se dice *evitable* o *esencial* según que exista o no el límite de f en $\boldsymbol{a}$.

(*) Si $\boldsymbol{a}$ es aislado de C, entonces las condiciones *a*) y *b*) se cumplen trivialmente, por lo que toda función es continua en los puntos aislados de su campo de definición.

[15]$_1$ Observaciones

1. La anterior definición de continuidad tiene interés cuando el punto $\boldsymbol{a}$ es de acumulación de C. En este caso, las dos condiciones *a*) y *b*) son equivalentes a sus correspondientes *a*) y *b*) de la definición de límite (véase [11]) tomando ahora $\boldsymbol{l} = f(\boldsymbol{a})$; por tanto, *a*) y *b*) son equivalentes, ambas, a la condición $\lim_{x \to a} f(\boldsymbol{x}) = f(\boldsymbol{a})$.

   Si $\boldsymbol{a}$ es un punto aislado de C (caso de poco interés), las dos condiciones *a*) y *b*) se verifican trivialmente, ya que: *a*) esta condición se verifica tomando $\delta > 0$ de manera que en la bola $B(\boldsymbol{a}, \delta)$ no hay más puntos de C que el $\boldsymbol{a}$, con independencia de quien sea ε; y *b*) la única sucesión de puntos de C que tiene límite $\boldsymbol{a}$ es la $\boldsymbol{a}, \boldsymbol{a}, ..., \boldsymbol{a}, ...$

2. La función $f: C \to \mathbb{R}^q$ es discontinua en el punto $\boldsymbol{a} \in C$, es decir, no es continua en $\boldsymbol{a} \in C$, si no se verifican las anteriores condiciones *a*) y *b*), lo que equivale a cualquiera de las siguientes:
   - Existe $\varepsilon_0 > 0$ tal que, para cada $\delta > 0$ existe algún $\boldsymbol{x} \in C$ para el que se cumple que $\|\boldsymbol{x} - \boldsymbol{a}\| < \delta$ y $\|\boldsymbol{f}(\boldsymbol{a}) - \boldsymbol{f}(\boldsymbol{a})\| \geqslant \varepsilon_0$.
   - Existe una sucesión $(\boldsymbol{x}_n)$ de puntos de C que tiene límite $\boldsymbol{a}$ y es tal que la sucesión $(\boldsymbol{f}(\boldsymbol{x}_n))$ no converge hacia $\boldsymbol{f}(\boldsymbol{a})$.
   - Existe una sucesión $(\boldsymbol{x}_n)$ de puntos de C que tiene límite $\boldsymbol{a}$ y es tal que la sucesión $(\boldsymbol{f}(\boldsymbol{x}_n))$ tiene límite (finito o infinito) distinto de $\boldsymbol{f}(\boldsymbol{a})$.
3. Si la función $f: C \to \mathbb{R}^q$ tiene una discontinuidad evitable en un punto $\boldsymbol{a}$ de acumulación de C, se dice que «se evita la discontinuidad» si se altera lo que con ella acontece en $\boldsymbol{a}$, tomando para $\boldsymbol{f}(\boldsymbol{a})$ el valor del límite de $\boldsymbol{f}$ en $\boldsymbol{a}$ (que existe); con esta manipulación se consigue que $\boldsymbol{f}$ pase a ser continua en $\boldsymbol{a}$.
4. La condición «$\varepsilon : \delta$» de continuidad, de $\boldsymbol{f}$ en $\boldsymbol{a}$, se puede expresar, acudiendo a los entornos o a las bolas, del siguiente modo: para cada $\varepsilon > 0$ existe un $\delta > 0$ tal que

$$\boldsymbol{f}(B(\boldsymbol{a}, \delta)) \subset B(\boldsymbol{f}(\boldsymbol{a}), \varepsilon)$$

5. La función $f: C \to \mathbb{R}^q$ se dice que es continua en $\boldsymbol{a} \in C$ según el conjunto $D \subset C$ si, siendo $\boldsymbol{a} \in D$, se verifica que la restricción de $\boldsymbol{f}$ a D es una función continua en $\boldsymbol{a}$. Si $\boldsymbol{f}$ es continua en $\boldsymbol{a}$, entonces es continua en $\boldsymbol{a}$ según todo subconjunto $D \subset C$ al que pertenezca el punto $\boldsymbol{a}$.

[15]$_2$ Ejercicio

Estudiar la continuidad de la siguiente función real f en todos los puntos $(a, b) \in \mathbb{R}^2$:

$$f: \mathbb{R}^2 \to \mathbb{R}, \qquad f(x, y) = \begin{cases} x^2 + 2y - 1, & \text{si } x \geqslant 0 \\ 3x + y^2, & \text{si } x < 0 \end{cases}$$

Resolución

- Si es $a > 0$, entonces f es continua en (a, b) ya que:

$$\lim_{\substack{x \to a \\ y \to b}} f(x, y) = \lim_{\substack{x \to a \\ y \to b}} (x^2 + 2y - 1) = a^2 + 2b - 1 = f(a, b)$$

- Si es $a < 0$, entonces f es continua en (a, b) ya que:

$$\lim_{\substack{x \to a \\ y \to b}} f(x, y) = \lim_{\substack{x \to a \\ y \to b}} (3x + y^2) = 3b + a^2 = f(a, b)$$

- Si es $a = 0$, llamando C y D a los semiplanos

$$C = \{(x, y) \in \mathbb{R}^2 \ / \ x \geqslant 0\} \qquad \text{y} \qquad D = \{(x, y) \in \mathbb{R}^2 \ / \ x < 0\}$$

f tendrá límite en $(0, b)$ si, y sólo si, $f|_C$ y $f|_D$ tienen el mismo límite en $(0, b)$; estos límites son:

$$\lim_{\substack{x \to 0 \\ y \to b}} f|_C(x, y) = \lim_{\substack{x \to 0 \\ y \to b}} (x^2 + 2y - 1) = 2b - 1 = f(0, b)$$

$$\lim_{\substack{x \to 0 \\ y \to b}} f|_D(x, y) = \lim_{\substack{x \to 0 \\ y \to b}} (3x + y^2) = b^2$$

Por tanto, f es continua en $(0, b)$ si, y sólo si, es $2b - 1 = b^2$, o sea, cuando es $b = 1$; luego f es continua en $(0, 1)$ y discontinua en los puntos $(0, b)$ para $b \neq 1$.

$[15]_3$ Ejercicio

Sea $f: \mathbb{R}^p \to \mathbb{R}^q$ una función lineal. Pruébese que f es continua en el punto $\boldsymbol{o} = (0, 0, ..., 0)$. Pruébese también que f es continua en cualquier punto $\boldsymbol{a} \in \mathbb{R}^p$.

Resolución

Sea $(\boldsymbol{e}_1, ..., \boldsymbol{e}_p)$ la base canónica de $\mathbb{R}^p$, de manera que llamando $x_1, ..., x_p$ a las coordenadas de $\boldsymbol{x} \in \mathbb{R}^p$, es

$$\boldsymbol{x} = \sum_{u=1}^{p} x_i \boldsymbol{e}_i \qquad \text{y} \qquad \boldsymbol{f}(\boldsymbol{x}) = \sum_{i=1}^{p} x_i \boldsymbol{f}(\boldsymbol{e}_i)$$

Llamemos k al mayor de los números $\|\boldsymbol{f}(\boldsymbol{e}_i)\|$, para $i = 1, 2, ..., p$.

- f es continua en el punto $\boldsymbol{o}$ ya que la condición «$\varepsilon : \delta$» de continuidad se verifica para $\delta_0 = \varepsilon/(pk)$, puesto que:

$$\|\boldsymbol{x} - \boldsymbol{o}\| < \delta_0 \quad \Rightarrow \quad \|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{o})\| = \left\| \sum_{i=1}^{p} x_i \boldsymbol{f}(\boldsymbol{e}_i) \right\| \leqslant \sum_{i=1}^{p} |x_i| \, \|\boldsymbol{f}(\boldsymbol{e}_i)\| <$$

$$< \|\boldsymbol{x}\| \sum_{i=1}^{p} \|\boldsymbol{f}(\boldsymbol{e}_i)\| < \delta_0 kp = \varepsilon$$

- f es continua en cualquier $\boldsymbol{a} \in \mathbb{R}^p$ ya que la condición «$\varepsilon : \delta$» de continuidad se verifica para el mismo δ_0 correspondiente al punto $\boldsymbol{o}$; en efecto

$$\|\boldsymbol{x} - \boldsymbol{a}\| < \delta_0 \quad \Rightarrow \quad \|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{a})\| = \|\boldsymbol{f}(\boldsymbol{x} - \boldsymbol{a})\| = \|\boldsymbol{f}(\boldsymbol{x} - \boldsymbol{a}) - \boldsymbol{f}(\boldsymbol{o})\| < \varepsilon$$

$[15]_4$ Ejercicio

Sea $f: A \to \mathbb{R}^q$ una función continua en el conjunto abierto $A \subset \mathbb{R}^p$. Si $V \subset \mathbb{R}^q$ es un conjunto abierto, pruébese que entonces $f^{-1}(V)$ también es abierto.

Resolución

Sea x un punto cualquiera de $U = f^{-1}(V)$; como $y = f(x)$ es de V y V es abierto, existe $\varepsilon_x > 0$ tal que $B(y, \varepsilon_x) \subset V$. La función f es continua en x (pues lo es en A y $x \in A$), por lo que se cumple la correspondiente condición «$\varepsilon : \delta$»; sea $\delta'_x > 0$ el valor correspondiente al anterior $\varepsilon_x > 0$. Por ser A abierto y como $x \in A$, existe $\delta''_x > 0$ tal que $B(x, \delta''_x) \subset A$; sea $\delta_x = \text{mín}\ \{\delta'_x, \delta''_x\}$. Resulta de ello que la bola abierta $B_x = B(x, \delta_x)$ está incluida en $U = f^{-1}(V)$, pues

$$f(B_x) = f(B_x \cap A) \subset f\big(B(x, \delta'_x) \cap A\big) \subset V$$

Así pues, se puede poner que $U = \bigcup_{x \in U} B_x$, luego U es unión de conjuntos abiertos, y, en consecuencia, él es también un abierto.

PRIMERAS PROPIEDADES DE LA CONTINUIDAD

[16]

Sea $f: C \to \mathbb{R}^q$ una función, definida en un conjunto $C \subset \mathbb{R}^p$, y considérese un punto $a \in C$. Se verifica que:

1.º La función f es continua en a si, y sólo si, son continuas en a las q componentes de f.
2.º Si f es continua en a, entonces f está acotada en un cierto entorno de a.
3.º Si f es una función real (esto es, si $q = 1$), si f es continua en a y si $f(a) \neq 0$, entonces f tiene signo constante en un cierto entorno de a.

Demostración

1.º La función f es continua en a, o sea, $f(x)$ tiende a $f(a)$ cuando x tiende a a si, y sólo si (véase [12],5), cada componente f_i de f ($i = 1, 2, ..., q$) tiene límite $f_i(a)$ cuando x tiende a a, es decir, si las q componentes de f son continuas en a.
2.º Si f es continua en a, entonces f tiene límite finito en a, luego (véase [12],3) f está acotada «cerca de a».
3.º Sea δ el valor que se obtiene al tomar $\varepsilon = \frac{1}{2}|f(a)|$ en la definición «$\varepsilon : \delta$» de continuidad (de f en a); esta propiedad 3.ª se verifica, pues, en el entorno de a que tiene radio δ.

[16]$_1$ Ejercicio

Sea $f: C \to \mathbb{R}^q$ una función, definida en un conjunto $C \subset \mathbb{R}^q$, y considérese la función $\|f\|$, definida mediante:

$$\|f\|: C \to \mathbb{R}, \qquad \|f\|(x) = \|f(x)\|$$

Si f es continua en a, pruébese que también $\|f\|$ es continua en a.

Resolución

Dado $\varepsilon > 0$, al aplicar la condición «$\varepsilon : \delta$» de continuidad a la función f en el punto $\boldsymbol{a}$, se sabe que existe un $\delta > 0$ tal que

$$[\boldsymbol{x} \in C, \ \|\boldsymbol{x} - \boldsymbol{a}\| < \delta] \Rightarrow \|f(\boldsymbol{x}) - f(\boldsymbol{a})\| < \varepsilon$$

Por tanto, si $\boldsymbol{x} \in C$ es tal que $\|\boldsymbol{x} - \boldsymbol{a}\| < \delta$, se verifica que:

$$|\|f\|(\boldsymbol{x}) - \|f\|(\boldsymbol{a})| = |\|f(\boldsymbol{x})\| - \|f(\boldsymbol{a})\|| \leqslant \|f(\boldsymbol{x}) - f(\boldsymbol{a})\| < \varepsilon$$

luego, para la función $\|f\|$ y en el punto $\boldsymbol{a}$, se verifica la condición «$\varepsilon : \delta$» de continuidad, es decir, $\|f\|$ es continua en $\boldsymbol{a}$.

[16]$_2$ Ejercicio

Sea $f: C \to \mathbb{R}^2$ la función definida, en el siguiente conjunto $C \subset \mathbb{R}^2$, mediante las expresiones que se indican:

$$C = \{(x, y) \in \mathbb{R}^2 \ / \ x^2 \neq y^2\}, \qquad f(x, y) = \left(\frac{\operatorname{sen} x - \operatorname{sen} y}{x - y}, \frac{e^x - e^{-y}}{x + y}\right)$$

Compruébese que f es continua en C y que se puede extender a todo $\mathbb{R}^2$ de manera que sea continua en todos los puntos de $\mathbb{R}^2$.

Resolución

Llamemos f_1 y f_2 a las componentes de f. Para $x \neq y$ es evidente que f_1 es continua y para $x \neq -y$ es continua f_2; así pues, f es continua en C. Veamos si f_1 tiene límite en los puntos (a, a), y si f_2 lo tiene en los puntos $(a, -a)$:

$$\lim_{\substack{x \to a \\ y \to a}} f_1(x, y) = \lim_{\substack{x \to a \\ y \to a}} \frac{\operatorname{sen} x - \operatorname{sen} y}{x - y} = \lim_{\substack{x \to a \\ y \to a}} \frac{2}{x - y} \operatorname{sen}\left(\frac{x - y}{2}\right) \cos\left(\frac{x + y}{2}\right) = \cos a$$

$$\lim_{\substack{x \to a \\ y \to -a}} f_2(x, y) = \lim_{\substack{x \to a \\ y \to -a}} \frac{e^x - e^{-y}}{x + y} = \lim_{\substack{x \to a \\ y \to -a}} e^{-y} \frac{e^{x+y} - 1}{x + y} = e^a$$

A la vista de estos resultados, se puede asegurar que f pasa a ser una función continua en todo $\mathbb{R}^2$ si se extiende tomando para cualquiera que sea $x \in \mathbb{R}$:

$$f_1(x, x) = \cos x \qquad \text{y} \qquad f_2(x, -x) = e^x$$

(nótese que $x \mapsto \cos x$ y $x \mapsto e^x$ son funciones continuas; nótese también que $\cos 0 = e^0$).

COMBINACIONES DE FUNCIONES CONTINUAS

[17]

I. *Operaciones aritméticas con funciones continuas.* Si f y g son dos funciones reales, definidas en un mismo conjunto $C \subset \mathbb{R}^q$, que son continuas en un punto $\boldsymbol{a} \in C$, entonces también son continuas en $\boldsymbol{a}$ su suma $f + g$(*), su producto fg y su cociente f/g (para esta última se exige que sea $g(\boldsymbol{a}) \neq 0$).

II. *Composición de funciones continuas.* Sean $\boldsymbol{f}: C \to \mathbb{R}^q$ y $\boldsymbol{g}: D \to \mathbb{R}^r$ dos funciones, definidas en $C \subset \mathbb{R}^p$ y $D \subset \mathbb{R}^q$ y tales que $\boldsymbol{f}(C) \subset D$. Si $\boldsymbol{f}$ es continua en un punto $\boldsymbol{a} \in C$ y $\boldsymbol{g}$ es continua en el punto $\boldsymbol{b} = \boldsymbol{f}(\boldsymbol{a}) \in D$, entonces la función compuesta $\boldsymbol{g} \circ \boldsymbol{f}$ es continua en el punto $\boldsymbol{a}$.

(*) Esta propiedad también es cierta cuando $\boldsymbol{f}$ y $\boldsymbol{g}$ son funciones vectoriales (o sea, si toman valores en $\mathbb{R}^q$ con $q > 1$), lo que se demuestra acudiendo a las componentes de $\boldsymbol{f}$ y $\boldsymbol{g}$.

Demostración

I. Como f y g tienen límites $f(\boldsymbol{a})$ y $g(\boldsymbol{a})$ cuando $\boldsymbol{x}$ tiende a $\boldsymbol{a}$, entonces (véase [13],2) $f + g$, fg y f/g tienen límites $f(\boldsymbol{a}) + g(\boldsymbol{a})$, $f(\boldsymbol{a})g(\boldsymbol{a})$ y $f(\boldsymbol{a})/g(\boldsymbol{a})$ (este último si $g(\boldsymbol{a}) \neq 0$) cuando $\boldsymbol{x}$ tiende a $\boldsymbol{a}$, es decir, $f + g$, fg y f/g son continuas en el punto $\boldsymbol{a}$.

II. Dado $\varepsilon > 0$, y como $\boldsymbol{g}$ es continua en el punto $\boldsymbol{b}$, existe un $\delta > 0$ tal que

$$[\boldsymbol{y} \in D,\ \|\boldsymbol{y} - \boldsymbol{b}\| < \delta] \ \Rightarrow\ \|\boldsymbol{g}(\boldsymbol{y}) - \boldsymbol{g}(\boldsymbol{b})\| < \varepsilon$$

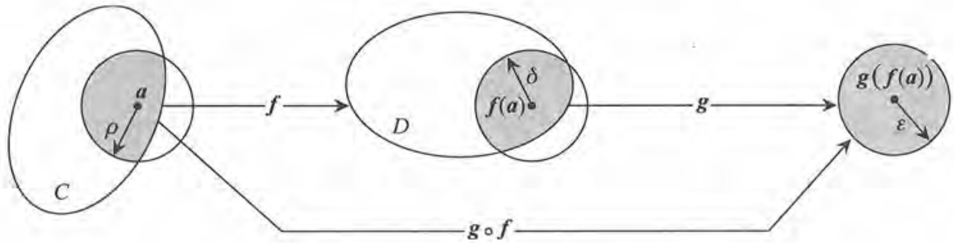
Como f es continua en el punto $\boldsymbol{a}$, existe un $\rho > 0$ (consecuencia del anterior δ) tal que

$$[\boldsymbol{x} \in C,\ \|\boldsymbol{x} - \boldsymbol{a}\| < \rho] \ \Rightarrow\ \|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{a})\| < \delta$$

Por tanto, de ello se concluye que (como $\boldsymbol{f}(C) \subset D$):

$$[\boldsymbol{x} \in C,\ \|\boldsymbol{x} - \boldsymbol{a}\| < \rho] \ \Rightarrow\ \|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{b}\| < \delta \ \Rightarrow$$
$$\Rightarrow\ \|\boldsymbol{g}(\boldsymbol{f}(\boldsymbol{x})) - \boldsymbol{g}(\boldsymbol{b})\| < \varepsilon \ \Rightarrow\ \|(\boldsymbol{g} \circ \boldsymbol{f})(\boldsymbol{x}) - (\boldsymbol{g} \circ \boldsymbol{f})(\boldsymbol{a})\| < \varepsilon$$

lo que prueba la continuidad de $\boldsymbol{g} \circ \boldsymbol{f}$ en el punto $\boldsymbol{a}$.



[17]$_1$ Ejercicio

Sean $\boldsymbol{f}$ y $\boldsymbol{g}$ dos funciones, de un conjunto $C \subset \mathbb{R}^p$ en $\mathbb{R}^q$, que son continuas en un punto $\boldsymbol{a} \in C$. Pruébese que entonces también es continua en $\boldsymbol{a}$ su producto escalar, es decir, la función

$$\boldsymbol{f} \cdot \boldsymbol{g} : C \to \mathbb{R}, \qquad \boldsymbol{x} \mapsto \boldsymbol{f}(\boldsymbol{x}) \cdot \boldsymbol{g}(\boldsymbol{x})$$

Resolución

Si f_i y g_i (para $i = 1, 2, ..., q$) son las componentes de $\boldsymbol{f}$ y $\boldsymbol{g}$, entonces

$$(\boldsymbol{f} \cdot \boldsymbol{g})(\boldsymbol{x}) = f_1(\boldsymbol{x})g_1(\boldsymbol{x}) + \cdots + f_q(\boldsymbol{x})g_q(\boldsymbol{x})$$

Como las funciones f_i y g_i son continuas en $\boldsymbol{a}$, entonces lo son las $f_i g_i$ y, como consecuencia, también lo es su suma, es decir $\boldsymbol{f} \cdot \boldsymbol{g}$ es continua en $\boldsymbol{a}$.

[17]$_2$ Ejercicio

Sea $f : \mathbb{R}^p \to \mathbb{R}$ la función definida, para todo $\boldsymbol{x} = (x_1, ..., x_p) \in \mathbb{R}^p$, mediante

$$f(\boldsymbol{x}) = \frac{\sum_{i=1}^{p} x_i \operatorname{sen} x_i}{\sum_{i=1}^{p} x_i^2}, \text{ si } \boldsymbol{x} \neq \boldsymbol{o}; \;\; f(\boldsymbol{o}) = 1$$

Compruébese que f es continua en $\mathbb{R}^p$.

Resolución

Si $\boldsymbol{x}_0 \neq \boldsymbol{o}$, entonces f es continua en $\boldsymbol{x}_0$, ya que son continuas en dicho punto las funciones $\boldsymbol{x} \mapsto x_i$, $x_i \mapsto \operatorname{sen} x_i$ y $\boldsymbol{x} \mapsto \Sigma\, x_i^2$ (y esta última es no nula «cerca» de $\boldsymbol{x}_0$). Veamos que f también es continua en $\boldsymbol{o}$, para lo que recurrimos a la función $\varphi : \mathbb{R} \to \mathbb{R}$ definida por

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

lo que nos permite poner (la suma se extiende para $i = 1, 2, ..., p$):

$$f(\boldsymbol{x}) = \frac{\Sigma\, x_i^2 \varphi(x_i)}{\Sigma\, x_i^2} = \varphi(x_1) + \sum_{i=1}^{p} \frac{x_i^2}{\Sigma\, x_i^2} [\varphi(x_i) - \varphi(x_1)]$$

(para todo $x \neq o$); como $x \mapsto \varphi(x_1)$ tiene límite 1 en o, como $x_i^2 / \sum x_i^2$ está acotada y como $x \mapsto \varphi(x_i) - \varphi(x_1)$ es un infinitésimo en o, concluimos que

$$\lim_{x \to o} f(x) = 1 + \sum_{i=1}^{p} 0 = 1 = f(o)$$

esto es, que f es continua en o.

IMAGEN CONTINUA DE UN COMPACTO. EXISTENCIA DE EXTREMOS

[18]

I. Si $C \subset \mathbb{R}^p$ es un conjunto compacto y si $f: C \to \mathbb{R}^q$ es una función continua en C, entonces el conjunto imagen $f(C) \subset \mathbb{R}^q$ también es compacto.
II. *Teorema de Weierstrass.* Si $C \subset \mathbb{R}^p$ es un conjunto compacto y si $f: C \to \mathbb{R}$ es una función (real) continua en C, entonces f tiene, en C, valores máximo y mínimo, es decir, existen $x_0, x_0' \in C$ tales que $f(x_0) \leqslant f(x) \leqslant f(x_0')$ para todo $x \in C$.

Demostración

I. Según la caracterización [08],c) de los conjuntos compactos hemos de probar que, si (y_n) es una sucesión de puntos de $f(C)$, entonces (y_n) tiene algún límite de oscilación que pertenece a $f(C)$. Así ocurre, en efecto: como $y_n \in f(C)$, existe $x_n \in C$ tal que $f(x_n) = y_n$ para $n \in \mathbb{N}$. Dado que C es compacto, se sabe que la sucesión (x_n) tiene algún límite de oscilación x que pertenece a C, luego existe una subsucesión (x_{n_i}) de (x_n) que tiene límite x. Ahora bien, como f es continua en x (pues $x \in C$ y f es continua en C) y por ser x el límite de (x_{n_i}), se puede asegurar que $f(x)$ es el límite de $(f(x_{n_i})) = (y_{n_i})$, luego $f(x)$ es un límite de oscilación de la sucesión dada (y_n), con lo que concluye la demostración, pues $f(x) \in f(C)$.

II. Por el teorema anterior, sabemos que $f(C)$ es un conjunto compacto (de $\mathbb{R}$), o sea, es cerrado y está acotado. Como $f(C)$ es un conjunto (no vacío) de la recta real que está acotado, sabemos que tiene supremo $h \in \mathbb{R}$ e ínfimo $k \in \mathbb{R}$; hay que demostrar que h y k son el máximo y el mínimo de $f(C)$, es decir, que h y k pertenecen a $f(C)$. Consideremos el caso de h; para k se puede razonar de modo análogo.

Para cada $n \in \mathbb{N}$, consideremos el intervalo $[h, h - 1/n[$; en él ha de haber algún punto de $f(C)$, al que llamaremos y_n, ya que de no ser así el número $h - 1/n$ sería cota superior de $f(C)$ y esto es falso, pues h es la menor de tales cotas. Es evidente que la sucesión (y_n) tiene límite h. Como $f(C)$ es compacto y los elementos de (y_n) son de $f(C)$, sabemos entonces que (y_n) tiene algún límite de oscilación que pertenece a $f(C)$; dado que h es el único límite de oscilación de (y_n), pues h es su límite, resulta que h pertenece a $f(C)$, como había que comprobar.

$[18]_1$ Ejercicios

Sea $f: \mathbb{R}^p \to \mathbb{R}^q$ una función y supongamos que existe una sucesión $(\boldsymbol{x}_n)$ acotada, de puntos de $\mathbb{R}^p$, tal que su sucesión imagen $(f(\boldsymbol{x}_n))$ no está acotada. Pruébese que f no puede ser continua en $\mathbb{R}^p$.

Resolución

Como $(\boldsymbol{x}_n)$ está acotada, existe $k > 0$ tal que $\|\boldsymbol{x}_n\| \leqslant k$ para todo $n \in \mathbb{N}$; considérese la bola cerrada $C = \bar{B}(\boldsymbol{o}, k)$ que, como está acotada, es un conjunto compacto. Como $\boldsymbol{x}_n \in C$ para $n \in \mathbb{N}$ y $(f(\boldsymbol{x}_n))$ no está acotada, $f(C)$ no está acotado, luego no es compacto. De ello se infiere que la función $f: C \to \mathbb{R}^q$ no puede ser continua en C, pues si lo fuese y por ser C compacto también sería compacto $f(C)$ y no lo es. Como f no es continua en C, no lo es en $\mathbb{R}^p$.

$[18]_2$ Continuidad de la función inversa

Sea $C \subset \mathbb{R}^p$ un conjunto compacto y sea $f: C \to \mathbb{R}^q$ una función continua en C que es inyectiva (con lo que existe la función inversa $f^{-1}: D \to \mathbb{R}^p$, donde $D = f(C)$). Se verifica, entonces, que f^{-1} es continua en D.

Demostración

Vamos a probar que f^{-1} es continua en cualquier punto $\boldsymbol{b} \in D$, esto es, que: si $(\boldsymbol{y}_n)$ es una sucesión de puntos de D que tiene límite $\boldsymbol{b}$, entonces la sucesión $(\boldsymbol{x}_n) = (f^{-1}(\boldsymbol{y}_n))$ tiene límite y éste es $\boldsymbol{a} = f^{-1}(\boldsymbol{b})$. Así ocurre, en efecto: como C es compacto y los puntos de $(\boldsymbol{x}_n)$ son de C, se puede asegurar entonces que $(\boldsymbol{x}_n)$ tiene algún límite de oscilación y que todos los límites de oscilación de $(\boldsymbol{x}_n)$ pertenecen a C. Hemos, pues, de comprobar que si $\boldsymbol{c} \in C$ es un límite de oscilación de $(\boldsymbol{x}_n)$, entonces debe ser $\boldsymbol{c} = \boldsymbol{a}$. Por ser $\boldsymbol{c}$ un límite de oscilación de $(\boldsymbol{x}_n)$, existe una subsucesión $(\boldsymbol{x}_{n_i})$ de $(\boldsymbol{x}_n)$ que tiene límite $\boldsymbol{c}$; como f es continua en $\boldsymbol{c}$ (pues es continua en C y $\boldsymbol{c} \in C$) y $(\boldsymbol{x}_{n_i})$ converge hacia $f(\boldsymbol{c})$, resulta que la sucesión $(\boldsymbol{y}_{n_i})$ converge hacia $f(\boldsymbol{c})$. Al ser $(\boldsymbol{y}_{n_i})$ una subsucesión de $(\boldsymbol{y}_n)$ y como esta sucesión tiene límite $\boldsymbol{b}$, resulta que $f(\boldsymbol{c}) = \boldsymbol{b}$, o sea, $f(\boldsymbol{c}) = f(\boldsymbol{a})$, de donde resulta que $\boldsymbol{c} = \boldsymbol{a}$ (pues f es inyectiva), como había que comprobar.

$[18]_3$ Homeomorfismos

- Se dice que dos conjuntos $C \subset \mathbb{R}^p$ y $D \subset \mathbb{R}^q$ son homeomorfos si existe una aplicación $f: C \to D$ biyectiva y tal que ella es continua en C y su inversa $f^{-1}: D \to C$ es continua en D; se dice que una tal aplicación f es un homeomorfismo (de C en D).
- Según se acaba de probar (en $[18]_2$), si $C \subset \mathbb{R}^p$ es compacto y $f: C \to \mathbb{R}^q$ es una aplicación inyectiva y continua en C, entonces f es un homeomorfismo de C en $D = f(C)$.

- Una propiedad, referente a un conjunto C (de un espacio euclídeo $\mathbb{R}^p$), se dice que es una *propiedad topológica* si es invariante por homeomorfismos, es decir si: para cualquiera que sea el homeomorfismo $f: C \to \mathbb{R}^q$ y siempre que C goce de la referida propiedad, se verifica que $f(C)$ también la posee. Nótese que de [18] se desprende que la compacidad es una propiedad topológica.

IMAGEN CONTINUA DE UN CONEXO. PROPIEDAD DE DARBOUX

Hay dos conceptos diferentes de conexión: la «conexión» (sin apellido) y la «conexión por arcos»; la primera es menos exigente que la segunda. Nosotros vamos a ocuparnos de la «conexión por arcos»; de la «conexión» nos limitaremos a definirla y a decir cuatro cosas sobre ella.

[19]

Curva continua. Dados un intervalo compacto $[\alpha, \beta] \subset \mathbb{R}$ y una función $\varphi[\alpha, \beta] \to \mathbb{R}^p$ continua en $[\alpha, \beta]$, se dice que el conjunto $\varphi([\alpha, \beta])$ es una curva continua, de $\mathbb{R}^p$, que une los puntos $\boldsymbol{a} = \varphi(\alpha)$ y $\boldsymbol{b} = \varphi(\beta)$.

Conjunto conexo por arcos (o por curvas). Se dice que un conjunto $C \subset \mathbb{R}^p$ es conexo por arcos si, para cualesquiera que sean los puntos $\boldsymbol{a}, \boldsymbol{b} \in C$, existe una curva continua que une $\boldsymbol{a}$ con $\boldsymbol{b}$ y está incluida en C.

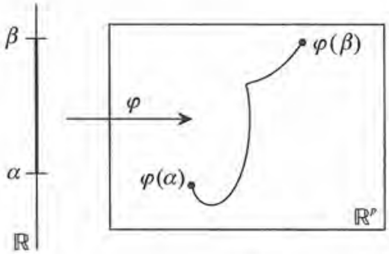
Imagen continua de un conexo(*). Si $C \subset \mathbb{R}^p$ es un conjunto conexo por arcos y si $f: C \to \mathbb{R}^q$ es una función continua en C, entonces el conjunto imagen $f(C) \subset \mathbb{R}^q$ es también conexo por arcos.

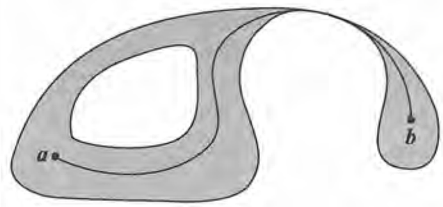
Propiedad de Darboux(*). Sea $C \subset \mathbb{R}^p$ un conjunto conexo por arcos y sea $f: C \to \mathbb{R}$ una función (real) continua en C. Dados $\boldsymbol{a}, \boldsymbol{b} \in C$ y si $f(\boldsymbol{a}) < k < f(\boldsymbol{b})$, entonces existe algún punto $\boldsymbol{c} \in C$ tal que $f(\boldsymbol{c}) = k$.

(*) Estas propiedades siguen siendo ciertas si C, en lugar de ser «conexo por arcos», fuese un conjunto «conexo», según la definición que luego iniciaremos (véase $[19]_2$).

Demostración

1. (Imagen continua de un conexo) Sean $\boldsymbol{a}'$ y $\boldsymbol{b}'$ dos puntos cualesquiera de $f(C)$. Se sabe, pues, que existen $\boldsymbol{a}, \boldsymbol{b} \in C$ tales que $\boldsymbol{a}' = f(\boldsymbol{a})$ y $\boldsymbol{b}' = f(\boldsymbol{b})$. Como C es conexo y $\boldsymbol{a}, \boldsymbol{b} \in C$, hay una curva continua incluida en C que une $\boldsymbol{a}$ con $\boldsymbol{b}$, es decir, existe una función $\varphi: [\alpha, \beta] \to \mathbb{R}^p$, que es continua en un intervalo compacto $[\alpha, \beta] \subset \mathbb{R}$, tal que $\varphi(\alpha) = \boldsymbol{a}$, $\varphi(\beta) = \boldsymbol{b}$ y $\varphi([\alpha, \beta]) \subset C$. Acudamos entonces a la función $f \circ \varphi : [\alpha, \beta] \to \mathbb{R}^q$, que es continua en $[\alpha, \beta]$, pues φ es continua en $[\alpha, \beta]$, f es continua en C y $\varphi([\alpha, \beta]) \subset C$; nótese que $(f \circ \varphi)(\alpha) = \boldsymbol{a}'$, que $(f \circ \varphi)(\beta) = \boldsymbol{b}'$ y que $(f \circ \varphi)([\alpha, \beta]) \subset f(C)$. Así pues, ha resultado que $(f \circ \varphi)([\alpha, \beta])$ es una curva continua incluida en $f(C)$ y que une a $\boldsymbol{a}'$ con $\boldsymbol{b}'$, lo que prueba que $f(C)$ es un conjunto conexo por arcos.





2. (Propiedad de Darboux) Como C es conexo por arcos, existe una función $\varphi: [\alpha, \beta] \to \mathbb{R}^p$, que es continua en un intervalo $[\alpha, \beta] \subset \mathbb{R}$ y tal que $\varphi(\alpha) = \boldsymbol{a}$, $\varphi(\beta) = \boldsymbol{b}$ y $\varphi([\alpha, \beta]) \subset C$. Recurramos a la función $f \circ \varphi: [\alpha, \beta] \to \mathbb{R}$, que es continua en $[\alpha, \beta]$, pues φ es continua en $[\alpha, \beta]$, f es continua en C y $\varphi([\alpha, \beta]) \subset C$. Como la función real $f \circ \varphi$ es continua en un intervalo real, el $[\alpha, \beta]$, su imagen es entonces un intervalo, al que pertenecen los puntos $f(\boldsymbol{a})$ y $f(\boldsymbol{b})$, ya que $f(\boldsymbol{a}) = (f \circ \varphi)(\alpha)$ y $f(\boldsymbol{b}) = (f \circ \varphi)(\beta)$; por tanto, son de este último intervalo todos los números comprendidos entre $f(\boldsymbol{a})$ y $f(\boldsymbol{b})$ y, en particular, a él pertenece el punto k, pues $f(\boldsymbol{a}) < k < f(\boldsymbol{b})$, es decir, existe $\gamma \in [\alpha, \beta]$ tal que $k = (f \circ \varphi)(\gamma)$, o sea, $k = f(\boldsymbol{c})$, siendo $\boldsymbol{c} = \varphi(\gamma)$ un punto de C, como había que comprobar.

[19]$_1$ Observaciones

- Supongamos que $S \subset \mathbb{R}^p$ es una curva, es decir, que existen un intervalo $[\alpha, \beta] \subset \mathbb{R}$ y una función $\varphi: [\alpha, \beta] \to \mathbb{R}^p$ continua en $[\alpha, \beta]$ tales que $\varphi([\alpha, \beta]) = S$. Se dice entonces que $\varphi: [\alpha, \beta] \to \mathbb{R}^p$ es una «representación» de la curva S. Si echamos mano de una aplicación $F: [\gamma, \delta] \to [\alpha, \beta]$, definida en un intervalo $[\gamma, \delta] \subset \mathbb{R}$ y que sea biyectiva y continua, resulta evidente que, entonces, la función $\varphi \circ F: [\gamma, \delta] \to \mathbb{R}^p$ es otra representación de la misma curva S. Toda curva tiene, pues, una infinidad de representaciones.
- Unas curvas continuas especialmente sencillas son las «poligonales». Una poligonal es la reunión de una sucesión finita de segmentos tales que el extremo de cada uno de ellos es el origen del siguiente; la poligonal «une» el origen del primero de sus segmentos con el extremo del último. Recuérdese que, en $\mathbb{R}^p$, el segmento que tiene por origen y extremo a los puntos $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^p$ es el conjunto

$$[\boldsymbol{a}, \boldsymbol{b}] = \{\boldsymbol{x} = \boldsymbol{a} + \lambda(\boldsymbol{b} - \boldsymbol{a}) \;/\; \lambda \in \mathbb{R},\, 0 \leqslant \lambda \leqslant 1\}$$

- Un conjunto $C \subset \mathbb{R}^p$ es «conexo por poligonales» si, para cualesquiera que sean los puntos $\boldsymbol{a}, \boldsymbol{b} \in C$, existe una poligonal que une $\boldsymbol{a}$ con $\boldsymbol{b}$ y está incluida en C. Si C es convexo por poligonales es, obviamente, conexo por arcos; el recíproco es falso (por ejemplo, una circunferencia es un conjunto conexo por arcos que no es conexo por poligonales).
- Las bolas abiertas, las bolas cerradas y las esferas (véase [02]) son conjuntos conexos por arcos. Si varios conjuntos $C_1, C_2, ..., C_n$, de $\mathbb{R}^p$, son conexos por arcos y $C_i \cap C_{i+1} \neq \varnothing$ para $i = 1, 2, ..., n - 1$, entonces la unión de todos ellos también es conexo por arcos. No cuesta mucho comprobar que si $C \subset \mathbb{R}^p$ es un conjunto conexo por arcos, entonces también lo es su adherencia $\overline{C}$ y, en general, todos los conjuntos $D \subset \mathbb{R}^p$ tales que $C \subset D \subset \overline{C}$.

- Según la definición, un conjunto es conexo por arcos si se le puede recorrer de punta a punta sin abandonarle y, en consecuencia no hay modo de dividirle en dos partes que estén «alejada» la una de la otra. Esta última propiedad, definida de manera precisa, da lugar al concepto de conjunto conexo (no por arcos), que generaliza al anterior (al de conexo por arcos). Nos ocupamos ahora, someramente, de esta nueva noción de conexión.

$[19]_2$ Algo acerca de los conjuntos conexos

- Un conjunto $C \subset \mathbb{R}^p$ se dice que es conexo si se verifica una de las dos condiciones siguientes, que son equivalentes (la demostración no es obvia):

1.ª Si C se divide en dos conjuntos cualesquiera $C_1 \neq \varnothing$ y $C_2 \neq \varnothing$ (es $C_1 \cup C_2 = C$ y $C_1 \cap C_2 = \varnothing$) se verifica siempre que hay algún punto de uno de los conjuntos C_1 o C_2 que es el límite de una sucesión de puntos del otro.

2.ª Si $A_1, A_2 \subset \mathbb{R}^p$ son dos conjuntos abiertos tales que

$$C \subset A_1 \cup A_2 \qquad \text{y} \qquad C \cap A_1 \cap A_2 = \varnothing$$

entonces uno de los $C \cap A_1$ o $C \cap A_2$ es el conjunto vacío.

- Se demuestra, cosa que nosotros no haremos, que se verifican las siguientes propiedades (análogas a las que se citaron en $[19]_1$ para los conjuntos conexos por arcos):

a) Si los conjuntos $C_1, C_2, ..., C_n$ son conexos de $\mathbb{R}^p$ y $C_i \cap C_{i+1} \neq \varnothing$, para $i = 1, 2, ..., n-1$, entonces la unión de todos ellos es un conjunto conexo.

b) Si $C \subset \mathbb{R}^p$ es un conjunto conexo, entonces también lo es su adherencia $\bar{C}$ y, en general, todos los conjuntos $D \subset \mathbb{R}^p$ tales que $C \subset D \subset \bar{C}$.

- Si $C \subset \mathbb{R}^p$ es un conjunto conexo por arcos, entonces C es un conjunto conexo; el recíproco es falso, en general. No obstante, si $C \subset \mathbb{R}^p$ es conexo y abierto (en cuyo caso se dice que es un dominio), entonces C es conexo por poligonales.
- Según ya se ha señalado (en la nota al pie de [19]), se verifica que:

1.º Si $C \subset \mathbb{R}^p$ es conexo y $f: C \to \mathbb{R}^q$ es una función continua en C, entonces $f(C)$ es conexo de $\mathbb{R}^q$.

2.º Sea $C \subset \mathbb{R}^p$ un conjunto conexo y sea $f: C \to \mathbb{R}$ una función (real) continua en C. Si $\alpha, \beta \in f(C)$ y es $\alpha < k < \beta$, entonces $k \in f(C)$ (propiedad de Darboux).

$[19]_3$ Ejercicio

Sea $f: C \to \mathbb{R}$ una función real, definida en el conjunto

$$C = \{(x, y) \mid x^2 + y^2 < 4\}$$

Sabiendo que $f(1, 1) = 1$, que $f(-1, -1) = -1$, y que

$$\{(x, y) \in C \mid f(x, y) = 0\} = \{(x, 0) \in C \mid 1 \leqslant |x| < 2\}$$

analícese si f puede ser continua en C.

Resolución

La función f no puede ser continua en C. Si lo fuese, entonces lo sería, en particular, en el segmento que une los puntos $\boldsymbol{a}(1, 1)$ y $\boldsymbol{b}(-1, -1)$ y, como $f(\boldsymbol{a}) > 0 > f(\boldsymbol{b})$, la función f debería anularse en algún punto del segmento $[\boldsymbol{a}, \boldsymbol{b}]$, lo cual es falso.

CONTINUIDAD UNIFORME

Una función $f: C \to \mathbb{R}^q$ es continua en el conjunto $C \subset \mathbb{R}^p$, es decir, en todo punto de C, si para cualesquiera que sean $\boldsymbol{x} \in C$ y $\varepsilon > 0$ existe un $\delta > 0$ (que dependerá de los $\boldsymbol{x}$ y ε que se consideren) tal que, siempre que se tome $\boldsymbol{x}' \in C$ verificando a $\|\boldsymbol{x} - \boldsymbol{x}'\| < \delta$, se verifica que $\|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}')\| < \varepsilon$. Para un mismo $\varepsilon > 0$, el valor de δ será distinto, en general, si se considera otro punto $\boldsymbol{y} \in C$, en lugar de $\boldsymbol{x}$. Dicho de otro modo, dado $\varepsilon > 0$, para poder segurar que se verifica que es $\|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}')\| < \varepsilon$ no bastará, en general, con saber que $\|\boldsymbol{x} - \boldsymbol{x}'\|$ es menor que cierto $\delta > 0$ único, sino que $\|\boldsymbol{x} - \boldsymbol{x}'\|$ deberá ser en unos sitios menor que en otros. Cuando exista un tal δ que no dependa más que de ε, se dirá que la continuidad de f en C es uniforme; más exactamente:

[20]

Sea $\boldsymbol{f}: C \to \mathbb{R}^q$ una función definida en un conjunto $C \subset \mathbb{R}^p$. Se dice que $\boldsymbol{f}$ es *uniformemente continua* en C si se cumple la siguiente condición: para cada $\varepsilon > 0$, existe un $\delta > 0$, tal que

$$(\boldsymbol{x}, \boldsymbol{x}' \in C, \ \ \|\boldsymbol{x} - \boldsymbol{x}'\| < \delta) \quad \Rightarrow \quad \|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}')\| < \varepsilon$$

I. Si la función $\boldsymbol{f}$ es uniformemente continua en C, entonces $\boldsymbol{f}$ es continua en C; el recíproco es falso, en general.

II. La función $\boldsymbol{f}$ *no* es uniformemente continua en C si se verifica una cualquiera de las dos condiciones siguientes:

   a) Existe $\varepsilon_0 > 0$ tal que, para cada $\delta > 0$, existen $\boldsymbol{x}, \boldsymbol{x}' \in C$ tales que $\|\boldsymbol{x} - \boldsymbol{x}'\| < \delta$ y $\|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}')\| \geqslant \varepsilon_0$.

   b) Existen $\varepsilon_0 > 0$ y dos sucesiones $(\boldsymbol{x}_n)$ y $(\boldsymbol{x}'_n)$, de puntos de C, tales que $\lim (\boldsymbol{x}_n - \boldsymbol{x}'_n) = \boldsymbol{o}$ y $\|f(\boldsymbol{x}_n) - f(\boldsymbol{x}'_n)\| \geqslant \varepsilon_0$ para $n \in \mathbb{N}$.

Demostración

I. Esta afirmación es evidentemente cierta. La falsedad del recíproco se comprueba, por ejemplo, con la función $f: \mathbb{R}^2 \to \mathbb{R}, f(x, y) = x^2 + y^2$, que es continua en $\mathbb{R}^2$, pero no es uniformemente continua en $\mathbb{R}^2$, como se comprueba fácilmente.

II. La condición *a*) es la negación (siguiendo estrictamente las reglas de la lógica) de la condición «$\varepsilon : \delta$» de la continuidad uniforme. Bastará con probar que las condiciones *a*) y *b*) son equivalentes:

- $[b \Rightarrow a]$ Para el ε_0 de b) y para cada $\delta > 0$, como $\boldsymbol{x}_n - \boldsymbol{x}'_n \to \boldsymbol{o}$, se sabe que existe $p \in \mathbb{N}$ tal que $\|\boldsymbol{x}_p - \boldsymbol{x}'_p\| < \delta$; luego la condición a) se cumple para $\boldsymbol{x} = \boldsymbol{x}_p$ y $\boldsymbol{x}' = \boldsymbol{x}'_p$.
- $[a \Rightarrow b]$ Tomando $\delta = 1/n$, se sabe que existen ciertos $\boldsymbol{x}_n$, $\boldsymbol{x}'_n \in C$ tales que $\|\boldsymbol{x}_n - \boldsymbol{x}'_n\| < 1/n$ y $\|\boldsymbol{f}(\boldsymbol{x}_n) - \boldsymbol{f}(\boldsymbol{x}'_n)\| \geqslant \varepsilon_0$. En consecuencia, la condición b) se verifica para las sucesiones $(\boldsymbol{x}_n)$ y $(\boldsymbol{x}'_n)$ así obtenidas.

[20]$_1$ Ejercicio

Sea $f: C \to \mathbb{R}$ la siguiente función real, definida en el conjunto $C \subset \mathbb{R}^2$ que se indica

$$C = \mathbb{R}^2 - \{(0, 0)\} \qquad f(x, y) = \frac{1}{x^2 + y^2}$$

Pruébese que la función f, que es continua en C, no es uniformemente continua en C.

Resolución

Para esta función f y en este conjunto C se verifica la anterior condición [20].II.b de «no continuidad uniforme» para

$$\varepsilon_0 = 1, \qquad \boldsymbol{x}_n = \left(\frac{1}{\sqrt{n}}, 0\right) \qquad \text{y} \qquad \boldsymbol{x}'_n = \left(0, \frac{1}{\sqrt{n+2}}\right)$$

ya que

$$\lim_{n \to \infty} (\boldsymbol{x}_n - \boldsymbol{x}'_n) = \lim_{n \to \infty} \left(\frac{1}{\sqrt{n}}, -\frac{1}{\sqrt{n+2}}\right) = (0, 0)$$

$$|f(\boldsymbol{x}_n) - f(\boldsymbol{x}'_n)| = |n - (n + 2)| = 2 < 1 = \varepsilon_0 ,$$

[20]$_2$ Ejercicio

Sea $\boldsymbol{f}: C \to \mathbb{R}^q$ una función uniformemente continua en el conjunto $C \subset \mathbb{R}^p$. Sea $\boldsymbol{a} \in \mathbb{R}^p$ un punto frontera de C. Pruébese que existe y es finito el límite de $\boldsymbol{f}$ en el punto $\boldsymbol{a}$ (se supone que $\boldsymbol{a} \notin C$; si $\boldsymbol{a} \in C$ la cuestión es obvia).

Resolución

Supongamos que ello no fuera cierto. Entonces se podría asegurar que (negando la condición de Cauchy [11]$_4$) existiría un $\varepsilon_0 > 0$ tal que, para cada $\delta > 0$, hay puntos $\boldsymbol{x}, \boldsymbol{x}' \in C$ tales que

$$\|\boldsymbol{x} - \boldsymbol{a}\| < \delta/2, \qquad \|\boldsymbol{x}' - \boldsymbol{a}\| < \delta/2 \qquad \text{y} \qquad \|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}')\| \geqslant \varepsilon_0$$

y, por tanto, para estos puntos x, $x' \in C$ sería:

$$\|x - x'\| \leqslant \|x - a\| + \|x' - a\| < \frac{\delta}{2} + \frac{\delta}{2} = \delta \qquad \text{y} \qquad \|f(x) - f(x')\| \geqslant \varepsilon_0$$

es decir, se verificaría la condición [20].II.a, o sea, f no sería uniformemente continua en C, que es falso.

$[20]_3$ Una condición suficiente para la continuidad uniforme

Una función $f: C \to \mathbb{R}^q$, definida en un conjunto $C \subset \mathbb{R}^p$, se dice que es *lipschitziana* en C si existe un $k > 0$ tal que, para cualesquiera x, $x' \in C$, se cumple la siguiente condición (de Lipschitz):

$$\|f(x) - f(x')\| \leqslant k\|x - x'\|$$

Se verifica que: si la función $f: C \to \mathbb{R}^q$ es lipschitziana en C, entonces f es uniformemente continua en C.

Demostración

Si es $k > 0$ la «constante de Lipschitz», para f en C, y dado $\varepsilon > 0$, la condición «$\varepsilon : \delta$» de continuidad uniforme, de f en C, se verifica para $\delta = \varepsilon/k$; en efecto: para cualesquiera x, $x' \in C$ es:

$$\|x - x'\| < \frac{\varepsilon}{k} \;\Rightarrow\; \|f(x) - f(x')\| \leqslant k\|x - x'\| < k\,\frac{\varepsilon}{k} = \varepsilon$$

❑ CONTINUIDAD UNIFORME EN UN COMPACTO

[21] TEOREMA DE HEINE. Si $f: C \to \mathbb{R}^q$ es una función continua en $C \subset \mathbb{R}^p$ y si el conjunto C es compacto, entonces f es uniformemente continua en C.

Demostración

Supongamos que f no es uniformemente continua en C, es decir (véase [20].2.b), que existen $\varepsilon_0 > 0$ y dos sucesiones (x_n) y (x'_n) de puntos de C tales que

$$\lim\,(x_n - x'_n) = 0 \qquad \text{y} \qquad \|f(x_n) - f(x'_n)\| \geqslant \varepsilon_0 \text{ para } n \in \mathbb{N}$$

Como C es compacto, existe una subsucesión $(\boldsymbol{x}_{n_i})$ de $(\boldsymbol{x}_n)$ que tiene límite y éste es un punto de C, al que llamaremos $\boldsymbol{a}$; es obvio que

$$\lim_{i\to\infty} \boldsymbol{x}_{n_i} = \lim_{i\to\infty} \boldsymbol{x}'_{n_i} = \boldsymbol{a} \qquad \text{y} \qquad \|\boldsymbol{f}(\boldsymbol{x}_{n_i}) - \boldsymbol{f}(\boldsymbol{x}'_{n_i})\| \geqslant \varepsilon_0, \qquad \forall\ i \in \mathbb{N}$$

Ahora bien, como $\boldsymbol{f}$ es continua en $\boldsymbol{a}$ (pues $\boldsymbol{a} \in C$ y $\boldsymbol{f}$ es continua en C), se verifica que

$$\lim_{i\to\infty} [\boldsymbol{f}(\boldsymbol{x}_{n_i}) - \boldsymbol{f}(\boldsymbol{x}'_{n_i})] = \lim_{i\to\infty} \boldsymbol{f}(\boldsymbol{x}_{n_i}) - \lim_{i\to\infty} \boldsymbol{f}(\boldsymbol{x}'_{n_i}) = \boldsymbol{f}(\boldsymbol{a}) - \boldsymbol{f}(\boldsymbol{a}) = \boldsymbol{o}$$

que entra en contradicción con la última de las relaciones anteriores, lo que prueba el teorema.

[21]$_1$ Ejercicio

Sea $\boldsymbol{f}: C \to \mathbb{R}^q$ una función continua en la bola abierta y reducida $C = B^*(\boldsymbol{o}, 2)$ de $\mathbb{R}^p$. Si $\boldsymbol{f}$ tiene límite finito en $\boldsymbol{o}$, pruébese que entonces $\boldsymbol{f}$ es uniformemente continua en la nueva bola abierta y reducida $B^*(\boldsymbol{o}, 1)$.

Resolución

Si es $\boldsymbol{l} \in \mathbb{R}^q$ el límite de $\boldsymbol{f}$ en $\boldsymbol{o}$, ampliando $\boldsymbol{f}$ de manera que sea $\boldsymbol{f}(\boldsymbol{o}) = \boldsymbol{l}$, se consigue que $\boldsymbol{f}$ sea continua en $\boldsymbol{a}$. Como $\boldsymbol{f}$ es continua en $B^*(\boldsymbol{o}, 2)$, con dicha ampliación se la ha convertido en una función continua en la bola $B(\boldsymbol{o}, 2)$ y, como $\bar{B}(\boldsymbol{o}, 1) \subset B(\boldsymbol{o}, 2)$, $\boldsymbol{f}$ es entonces continua en la bola cerrada $\bar{B}(\boldsymbol{o}, 1)$; ahora bien, por ser $\bar{B}(\boldsymbol{o}, 1)$ un conjunto compacto y según el teorema de Heine, $\boldsymbol{f}$ es entonces uniformemente continua en $\bar{B}(\boldsymbol{o}, 1)$ y, en consecuencia, lo es en cualquier subconjunto de $\bar{B}(\boldsymbol{o}, 1)$ y, en particular, en $B^*(\boldsymbol{o}, 1)$.

Ejercicios y problemas

ENUNCIADOS

1.1. Se llama distancia de un punto $\boldsymbol{a} \in \mathbb{R}^p$ a un conjunto $C \subset \mathbb{R}^p$ al número

$$d(\boldsymbol{a}, C) = \text{ínf } \{d(\boldsymbol{a}, \boldsymbol{c})/\boldsymbol{c} \in C\}$$

Caracterizar el conjunto C_1 de los puntos $\boldsymbol{x} \in \mathbb{R}^p$ tales que $d(\boldsymbol{x}, C) = 0$.

1.2. Se llama distancia entre los conjuntos A y B, de $\mathbb{R}^p$ al número

$$d(A, B) = \text{ínf } \{d(\boldsymbol{a}, \boldsymbol{b})/\boldsymbol{a} \in A, \boldsymbol{b} \in B\}$$

Se llama diámetro de un conjunto $A \subset \mathbb{R}^p$ a

$$\delta(A) = \sup \{d(\boldsymbol{a}, \boldsymbol{a}')/\boldsymbol{a}, \boldsymbol{a}' \in A\}$$

Pruébese que, para cualesquiera conjuntos no nulos A, $B \in \mathbb{R}^p$, es:

$$\delta(A \cup B) \leqslant d(A, B) + \delta(A) + \delta(B)$$

1.3. Sean dados un conjunto $C \subset \mathbb{R}^p$ y un número real $\rho > 0$. Se llama distancia de $\boldsymbol{x} \in \mathbb{R}^p$ a C al número

$$d(\boldsymbol{x}, C) = \text{ínf } \{d(\boldsymbol{x}, \boldsymbol{c})/\boldsymbol{c} \in C\}$$

Considérese el conjunto $C_\rho = \{\boldsymbol{x} \in \mathbb{R}^p / d(\boldsymbol{x}, C) < \rho\}$. Compruébese que C_ρ es abierto.

1.4. Considérese el siguiente conjunto $C \subset \mathbb{R}^2$:

$$C = (\mathbb{Q}_+ \times \mathbb{Q}_+) \cup (\mathbb{Q}_- \times \mathbb{Q}_-) \cup$$
$$\cup (\mathbb{R}_+ \times \mathbb{R}_+) \cup (\mathbb{R}_- \times \mathbb{R}_-)$$

Hallar $\mathring{C}$, $\text{Fr}(C)$ y $\overline{C}$ (recuérdese que, tanto si es $K = \mathbb{Q}$ como si es $K = \mathbb{R}$, se llaman $K_+ = \{x \in K/x \geqslant 0\}$ y $K_- = \{x \in K/x \leqslant 0\}$).

1.5. Si A es un conjunto abierto de $\mathbb{R}^2$, sea $C = A \cap (\mathbb{Q} \times \mathbb{Q})$. Hallar $\mathring{C}$, $\overline{C}$ y $\text{Fr}(C)$.

1.6. Sean (x_n) e (y_n) dos sucesiones acotadas de números reales que tienen, cada una, dos límites de oscilación. Si los límites de oscilación de (x_n) son a y a' y los de (y_n) son b y b', ¿qué se puede asegurar de los límites de oscilación de la sucesión $(\boldsymbol{u}_n)$ de $\mathbb{R}^2$, donde $\boldsymbol{u}_n = (x_n, y_n)$?

1.7. Dados A, $B \subset \mathbb{R}^p$, se llama distancia de A a B al número $d(A, B) = \text{ínf } \{d(\boldsymbol{x}, \boldsymbol{y})/\boldsymbol{x} \in A, \boldsymbol{y} \in B\}$. Si A y B son compactos, pruébese que existen $\boldsymbol{a} \in A$ y $\boldsymbol{b} \in B$ tales que $d(A, B) = d(\boldsymbol{a}, \boldsymbol{b})$.

1.8. Sean A y B dos conjuntos de puntos de $\mathbb{R}^p$. 1.° Hallar la relación de inclusión que liga a $\text{Fr}(A \cup B)$ con $\text{Fr}(A) \cup \text{Fr}(B)$. 2.° Estudiar la cuestión anterior en el caso de ser $\overline{A} \cap \overline{B} = \varnothing$.

1.9. Sea $(\boldsymbol{x}_n)$ una sucesión de puntos de $\mathbb{R}^p$. Pruébese que si $(\boldsymbol{x}_n)$ es tal que $\|\boldsymbol{x}_{n+1} - \boldsymbol{x}_n\| < 1/2^n$ para todo $n \in \mathbb{N}$, entonces $(\boldsymbol{x}_n)$ es una sucesión Cauchy. Si $(\boldsymbol{y}_n)$ es una sucesión convergente de puntos de $\mathbb{R}^p$, pruébese que hay alguna subsucesión $(\boldsymbol{y}_{n_i})$ de $(\boldsymbol{y}_n)$ que verifica a la anterior condición.

1.10. Sean A y B dos conjuntos no vacíos de puntos de $\mathbb{R}^p$. Se llama $A+B$ al siguiente conjunto $\{\boldsymbol{a}+\boldsymbol{b}/\boldsymbol{a} \in A, \boldsymbol{b} \in B\}$. Compruébese que si A es abierto y para cualquiera que sea B, entonces $A + B$ es abierto.

1.11. Hallar los campos de definición de las siguientes funciones $(x, y) \mapsto z$, reales de dos variables reales:

a) $z = \text{arc sen } \sqrt{2 - x - y}$

b) $z = \dfrac{\text{L }(1 + x) + \text{L }(1 + y)}{\text{L }(1 - x) + \text{L }(1 - y)}$

1.12. Hallar, si existen, los siguientes límites:

a) $l = \lim\limits_{\substack{x \to 0 \\ y \to 0}} (x + y^2) \text{ sen } \dfrac{1}{xy}$

b) $l = \lim\limits_{\substack{x \to 0 \\ y \to 0}} \dfrac{\text{sen }(xy)}{x}$

c) $l = \lim\limits_{\substack{x \to 0 \\ y \to 0}} \dfrac{y}{x + 2y}$

d) $l = \lim\limits_{\substack{x \to 0 \\ y \to 0}} \dfrac{x^3 + y^3}{x^2 + y^2}$

1.13. Hallar, si existe, el límite l para $(x, y) \to (0, 0)$ de la función:

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

1.14. Averiguar si existe el límite siguiente

$$\lim_{\substack{x \to 0 \\ y \to 0}} \frac{y^2}{x} \operatorname{sen}(x^2 + y^2)$$

1.15. Estudiar si existe el siguiente límite, para algún $n > 2$:

$$\lim_{\substack{x \to 0 \\ y \to 0}} \frac{x^n + y^n}{x^2 - y^2}$$

1.16. Hallar los límites según direcciones y los límites (si existen) cuando $(x, y) \to (0, 0)$ de las siguientes funciones:

1. $f(x, y) = \dfrac{xy^2}{x^2 + y^4}$
2. $g(x, y) = \dfrac{x^2 y}{x^2 + y^4}$

1.17. Analícese si existe el siguiente límite

$$\lim_{\substack{x \to 1 \\ y \to 1}} \frac{\mathrm{L}\, x + \mathrm{L}\, y}{x + y - 2}$$

1.18. Sea $\varepsilon : \mathbb{R} \to \mathbb{R}$ un infinitésimo en el origen tal que $\varepsilon(0) = 1$. Se considera la función $f: \mathbb{R}^2 \to \mathbb{R}$ definida por

$$f(x, y) = \frac{3x^2 \varepsilon(y) - y^2 \varepsilon(x)}{\mathrm{L}\,(x^2 + y^2 + 1)}$$

1. Analizar si f tiene límite en $(0, 0)$ según el conjunto $C = \{(x, y) \in \mathbb{R}^2 / xy \neq 0\}$.
2. Analizar si f tiene límite en $(0, 0)$.

1.19. Hallar el límite cuando $(x, y) \to (a, a)$, siendo $a \neq 0$, de la siguiente función:

$$f(x, y) = \frac{(x - y)a^n + (a - x)y^n - (a - y)x^n}{(x - y)(a - x)(a - y)}$$

1.20. Se considera la función, real de dos variables reales, definida por

$$f(x, y) = \frac{ax + y + bx^2}{\operatorname{sen} y + \mathrm{L}\,(1 + x)} \quad (a, b \in \mathbb{R} \text{ dados})$$

Se pide, para $(x, y) \to (0, 0)$:

1. Determinar $\boldsymbol{a}$ y b para que sus límites direccionales sean todos iguales.
2. Para los a y b del apartado anterior, comprobar que f carece de límite.

1.21. Considérese la función, real de dos variables reales, dada por

$$f(x, y) = \frac{x^2 + y^2}{\operatorname{sen}^2 y + \mathrm{L}\,(1 + x^2)}$$

Se pide, para $(x, y) \to (0, 0)$, hallar uno de sus límites direccionales, l, y comprobar que l es límite de f en $(0, 0)$ (acúdase a los desarrollos limitados de Mac-Laurin de $\operatorname{sen}^2 y$ y de $\mathrm{L}\,(1 + x^2)$).

1.22. Sean f y g las siguientes funciones, reales, de dos variables reales:

$$f(x, y) = \frac{x^3 - y^2}{1 - \cos x + y}$$

$$g(x, y) = \frac{x^3 - y^2}{1 - \cos x + |y|}$$

Para $(x, y) \to (0, 0)$, se pide:

1. Comprobar que f carece de límite.
2. Hallar el límite de g.

1.23. En $\mathbb{R}^2$ se considera una recta R y dos puntos $\boldsymbol{a}$ y $\boldsymbol{b}$ situados a distinto lado de la recta. Si γ es una curva continua de $\mathbb{R}^2$ que une $\boldsymbol{a}$ con $\boldsymbol{b}$, pruébese que γ tiene un punto, al menos, situado sobre la recta R.

1.24. Sea C el siguiente conjunto, de puntos de $\mathbb{R}^2$:

$$C = (\mathbb{R}_+ \times \mathbb{Q}_+) \cup (\mathbb{Q}_- \times \mathbb{R}_+) \cup$$
$$\cup\, (\mathbb{R}_- \times \mathbb{Q}_-) \cup (\mathbb{Q}_+ \times \mathbb{R}_-)$$

Hágase una representación esquemática de C y estúdiese si C es conexo. Si $\boldsymbol{a} \in C$ y llamando $C_a = C - \{a\}$, determinar en función de $\boldsymbol{a}$ el número de «componentes conexas» de C_a (se dice que $X \subset C_a$ es una componente conexa de C_a si X es conexo y no hay otro conexo Y tal que $X \subset Y \subset C_a$).

1.25. Sea $f: C \to \mathbb{R}^q$ una función continua en el conjunto cerrado $C \subset \mathbb{R}^p$. Pruébese que la gráfica de f, esto es, $G = \{(x, f(x))/x \in C\}$ es un cerrado de $\mathbb{R}^{p+q}$.

1.26. Sea $C \subset \mathbb{R}^p$ un conjunto dado. Demuéstrese que, si toda función $f: C \to \mathbb{R}$ que sea continua en C está acotada en C, entonces C es compacto.

1.27. Si $C \subset \mathbb{R}^p$ es un conjunto cerrado y no acotado y si $f: C \to \mathbb{R}$ es una función continua en C y que tiene límite $+\infty$ cuando $x \to \infty$, pruébese que f tiene mínimo en C.

1.28. Sea $\varphi: \mathbb{R} \to \mathbb{R}$ una función derivable y con derivada continua en $\mathbb{R}$. Se considera la nueva función $f: \mathbb{R}^2 \to \mathbb{R}$ definida por:

$$f(x, y) = \begin{cases} \dfrac{\varphi(x) - \varphi(y)}{x - y}, & \text{si} \quad x \neq y \\ \varphi'(x), & \text{si} \quad x = y \end{cases}$$

Estúdiese la continuidad de f en $\mathbb{R}^2$.

1.29. Sea f la función vectorial, de dos variables reales, cuyas componentes son f_1 y f_2:

$$f_1(x, y) = \frac{\cos\sqrt{xy} - 1}{y}$$

$$f_2(x, y) = \frac{\operatorname{sen} x - \operatorname{sen} y}{\operatorname{tg} x - \operatorname{tg} y}$$

Estudiar si es posible extender la función f al conjunto $\{(x, y) \in \mathbb{R}^2/|x| < \pi/2, |y| < \pi/2\}$ de manera que sea continua en él.

1.30. Sea $f: C \to \mathbb{R}^q$ una función definida en un conjunto $C \subset \mathbb{R}^p$. Pruébese que f es continua en C si, y sólo si, se cumple la siguiente condición: para cada abierto $A \subset \mathbb{R}^q$ existe otro abierto $A' \subset \mathbb{R}^p$ tal que $f^{-1}(A) = A' \cap C$.

1.31. Sean dados una función $f: \mathbb{R}^2 \to \mathbb{R}$ y un punto $a \in \mathbb{R}^2$. Considérese la función $F: \mathbb{R}^*_+ \to \mathbb{R}$, $r \mapsto F(r)$, donde $F(r)$ es la oscilación de f en la bola $B(a, r)$. Pruébese que:

1. Para todo $r_0 > 0$ existen los límites $F(r_0^+)$ y $F(r_0^-)$; en particular, existe $F(0^+)$.
2. $F(0^+)$ es 0 si, y sólo si, f es continua en a.

1.32. Sea f la función, real de dos variables reales, definida por

$$f(x, y) = \frac{x^2y^2}{x^2y^2 - (x^2 + y^2)}, \qquad f(0, 0) = a$$

1. Hallar el mayor valor del número real r tal que f está definida en la bola abierta $C = \{(x, y) \in \mathbb{R}^2 \ / \ x^2 + y^2 < r\}$.
2. Analizar si f es continua en C.
3. Analizar si es uniformemente continua en C.

1.33. Sean dadas dos funciones reales $\varphi: \mathbb{R} \to \mathbb{R}$ y $\psi: [0, \infty[\to \mathbb{R}$, con $\varphi \neq o$. Se considera la función $f: \mathbb{R}^2 \to \mathbb{R}$ definida por:

$$f(x, y) = \begin{cases} \varphi\left(\dfrac{y}{x}\right)\psi(x^2 + y^2), & \text{si} \quad x \neq 0 \\ \psi(y^2), & \text{si} \quad x = 0 \end{cases}$$

1. Obtener una condición necesaria y suficiente para que f tenga todos sus límites direccionales en (0, 0). Hallar éstos.
2. Si φ no es costante, obtener una condición para que los límites anteriores sean todos iguales. Hallarlos.
3. Si φ no es constante, y si todos los límites anteriores son iguales, hallar una condición necesaria y suficiente para que f tenga límites en (0, 0).
4. Si se cumple la condición del apar-tado anterior, ¿cuándo será f continua en (0, 0)?
5. Si es $\varphi(t) = |\operatorname{sen}^3 (u - 1)|$ y $|\psi(u)| \leqslant u^4$, analizar si f es continua en (0, 0).

1.34. Sea $f: C \to \mathbb{R}^q$ una función continua en el conjunto $C \subset \mathbb{R}^p$ y sea (x_n) una sucesión de Cauchy de puntos de C. Analícese si la sucesión $(f(x_n))$ ha de ser, necesariamente, de Cauchy. Estudiar la misma cuestión si, además, C es cerrado.

1.35. Se dice que una aplicación $f: \mathbb{R}^p \to \mathbb{R}^p$ es contractiva si existe $k \in \mathbb{R}$, $0 < k < 1$, tal que

$$\|f(x) - f(y)\| \leqslant k\|x - y\|, \quad \forall\, x, y \in \mathbb{R}^p$$

Si f es aplicación contractiva, pruébese que: 1) existe un punto $a \in \mathbb{R}^p$ tal que $f(a) = a$ (se dice que a es un punto fijo); 2) en $\mathbb{R}^p$ sólo hay un punto fijo de f; 3) si en la anterior relación fuese $k = 1$ (para $x \neq y$), puede suceder que no existe un punto fijo.

1.36. Sea $f: \mathbb{R}^p \to \mathbb{R}^p$ una aplicación tal que, para cualesquiera que sean $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^p$:

$$\|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{y})\| < \|\boldsymbol{x} - \boldsymbol{y}\|$$

a) Pruébese que la aplicación $\varphi: \mathbb{R}^p \to \mathbb{R}$, definida por $\varphi(\boldsymbol{x}) = \|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{x}\|$, es continua en $\mathbb{R}^p$.

b) Si $C \subset \mathbb{R}^p$ es un conjunto compacto y si $\boldsymbol{f}(C) \subset C$, pruébese que existe $\boldsymbol{a} \in C$ tal que $\boldsymbol{f}(\boldsymbol{a}) = \boldsymbol{a}$.

1.37. Sea $C \subset \mathbb{R}^p$ un conjunto dado y considérese la función $f: \mathbb{R}^p \to \mathbb{R}$ $f(\boldsymbol{x}) = d(\boldsymbol{x}, C)$, donde $d(\boldsymbol{x}, C)$ es la distancia de $\boldsymbol{x} \in C$, es decir,

$$d(\boldsymbol{x}, C) = \inf \{d(\boldsymbol{x}, \boldsymbol{c})/\boldsymbol{c} \in C\}$$

Pruébese que f es uniformemente continua en $\mathbb{R}^p$.

1.38. Sea $f: \mathbb{R}^2 \to \mathbb{R}$ la función definida mediante

$$f(x, y) = \begin{cases} (x+y) \cos \dfrac{\pi}{x} \cos \dfrac{\pi}{y}, & \text{si } x \neq 0 \text{ e } y \neq 0 \\ 0, & \text{si } x = 0 \text{ o } y = 0 \end{cases}$$

1. Estudiar la continuidad de f en $\mathbb{R}^2$.
2. Estudiar la continuidad uniforme de f en el conjunto $]0, 1] \times]0, 1]$.

1.39. Sea f la función real definida en $\mathbb{R}^2 - \{(0, 0)\}$ mediante

$$f(x, y) = \operatorname{sen}^2 \frac{1}{x^2 + y^2}$$

Estudiar la continuidad uniforme de f en los siguientes conjuntos:

1. $C = \{(x, y) \in \mathbb{R}^2 / 0 < xz^2 + y^2 < 1\}$.
2. $C = \{(x, y) \in \mathbb{R}^2 / x^2 + y^2 > 1\}$.

1.40. Sea $\boldsymbol{f}: C \to \mathbb{R}^q$ una función, definida en un conjunto $C \subset \mathbb{R}^p$. Pruébese que $\boldsymbol{f}$ es uniformemente continua en C si y sólo si lo son sus q componentes $f_1, f_2, \ldots, f_q$.

1.41. Si $\boldsymbol{f}: C \to \mathbb{R}^q$ es una función continua en el conjunto cerrado y no acotado $C \subset \mathbb{R}^p$ y si $\boldsymbol{f}$ tiene límite $\boldsymbol{l} \in \mathbb{R}^q$ en el infinito, pruébese que entonces $\boldsymbol{f}$ es uniformemente continua en C.

SOLUCIONES

1.1. $C_1 = C \cup \mathrm{Fr}(C) = \bar{C}$.

1.2. $d(\boldsymbol{a}, \boldsymbol{b}) \leqslant d(\boldsymbol{a}, \boldsymbol{a}_0) + d(\boldsymbol{a}_0, \boldsymbol{b}_0) + d(\boldsymbol{b}_0, \boldsymbol{b}) \leqslant \leqslant d(\boldsymbol{a}_0, \boldsymbol{b}_0) + \delta(A) + \delta(B)$.

1.3. $\boldsymbol{x} \in C_\rho \Rightarrow \exists \boldsymbol{c} \in C / d(\boldsymbol{x}, \boldsymbol{c}) < \rho \Rightarrow \boldsymbol{x} \in B(\boldsymbol{c}, \rho) \subset \subset C_\rho \Rightarrow \boldsymbol{x}$ interior de C_ρ.

1.4. $\mathring{C} = (\mathbb{R}^*_+ \times \mathbb{R}^*_-) \cup (\mathbb{R}^*_- \times \mathbb{R}^*_+)$, $\mathrm{Fr}(C) = (\mathbb{R}_+ \times \mathbb{R}_+) \cup (\mathbb{R}_- \times \mathbb{R}_-)$, $\bar{C} = \mathbb{R} \times \mathbb{R}$ (se denota por K^* a $\{x \in K / x \neq 0\}$).

1.5. $\mathring{C} = \varnothing$, $\bar{C} = \bar{A}$, $\mathrm{Fr}(C) = \bar{C}$.

1.6. Tiene 2, 3 ó 4 límites de oscilación, de entre los puntos (a, b), (a', b), (a, b') y (a', b'); de ellos, dos son los 1.° y 4.° o los 2.° y 3.°.

1.7. $d(\boldsymbol{x}_n, \boldsymbol{y}_n) \to d(A, B)$ con $\boldsymbol{x}_n \in A$, $\boldsymbol{y}_n \in B$; $\exists\, (\boldsymbol{x}'_n)$ subsucesión de $(\boldsymbol{x}_n)$ tal que $\boldsymbol{x}_n \to \boldsymbol{a} \in A$; $\exists\, (\boldsymbol{x}''_n)$ subsucesión de $(\boldsymbol{x}'_n)$ tal que $\boldsymbol{y}''_n \to \boldsymbol{b} \in B$.

1.8. 1.° $\mathrm{Fr}(A \cup B) \subset \mathrm{Fr}(A) \cup \mathrm{Fr}(B)$.

2.° $\mathrm{Fr}(A \cup B) = \mathrm{Fr}(A) \cup \mathrm{Fr}(B)$.

1.9. 1.° $\|\boldsymbol{x}_{n+p} - \boldsymbol{x}_n\| \leqslant \dfrac{(1 - 1/2^p)}{2^n - 1} < \dfrac{1}{2^n - 1}$.

2.° $\boldsymbol{a} = \lim \boldsymbol{y}_n$; $\forall\, i \in \mathbb{N}$, $\exists\, \boldsymbol{y}_{n_i} / \|\boldsymbol{y}_n - \boldsymbol{a}\| < 1/2^{i+1}$, $\forall\, n \geqslant n_i$.

1.10. Si $B(\boldsymbol{a}, \varepsilon) \subset A$, entonces $B(\boldsymbol{a} + \boldsymbol{b}, \varepsilon) \subset A + B$.

1.11. *a*) $1 < x + y < 2$.

b) $-1 < x < 1$, $-1 < y < 1$, $x \neq y$.

1.12. *a*) Infinitésimo por acotada, $l = 0$.

b) $\mathrm{sen}\,(xy) \sim xy$, $l = a$.

c) Las direcciones son distintas, no existe l.

d) En polares: infinitésimo por acotada, $l = 0$.

1.13. Polares; $\alpha + \beta \leqslant 0 \Rightarrow \nexists\ l$; $\alpha + \beta > 0$, $|f(x, y) - 0| < \rho^{\alpha+\beta-2}/(1 + \cos\theta\ \mathrm{sen}\,\theta) < < 2\rho^{\alpha+\beta-2}$, $l = 0$.

1.14. $y^2(x^2 + y^2)/x$ carece de límite: para $x \to 0$, según $x = y^4$ el límite es 1, según $x = y^5$ el límite es ∞.

1.15. No: según $y = 0$, el límite es 0; según $y = x + x^n$, el límite es ∞.

1.16. 1. Direccionales $= 0$; según $x = y^2$, límite $1/2$; no hay límite.

2. Direccionales $= 0$; $|g(x, y) - 0| < |y|$; límite $= 0$.

1.17. No: según direcciones hay límite 1; según la curva $y - 1 = -(x - 1) + (x - 1)^2$ hay límite -1.

1.18. 1. Si: polares $f(x, y) \sim 3\cos^2\theta\,\varepsilon(y) - - \mathrm{sen}^2\theta\,\varepsilon(x) \to 0 + 0$.

2. No: $f(0, y) \to -1 \neq 0$.

1.19. El numerador es divisible por el denominador; $f(x, y)$ coincide con una función polinómica para $x \neq y$, $x \neq a$ e $y \neq a$, luego tiene límite en todo punto, que es el límite según $y = 2a - x$; éste vale (L'Hôpital) $n(n - 1)a^{n-2}/2$.

1.20. 1. Direccionales $(y = mx)(a + m)/(1 + m)$, si $m \neq -1$; ∞ si $m = -1$ y $a \neq 1$; $2b$, si $m = -1$ y $a = 1$; son iguales para $a = 1$ y $b = 1/2$.

2. Según $y = y(x)$ con $y(0) = 0$, $y'(0) = -1$, $y''(0) = 1$ e $y'''(0) = \lambda$, el límite es $\lambda/(3 + \lambda)$, variable.

1.21. $l = 1$. Cerca de 0 es $\mathrm{sen}^2\, y > t^2/2$, $\mathrm{L}\,(1 + t^2) > > t^2/2$, $|t^2 - \mathrm{sen}^2\, t| < t^4$ y $|t^2 - \mathrm{L}\,(1 + t^2)| < t^4$, luego es también $|f(x, y) - 1| < 2(x^4 + y^4)/(x^2 + y^2)$; por ello, $f(x, y) - 1 \to 0$.

1.22. 1. Posible límite 0; según $y = \lambda x^2$ el límite es 0 si $\lambda \neq -1/2$ e ∞ si $\lambda = -1/2$.

2. Posible límite 0; cerca de $(0, 0)$ es $|f(x, y) - - 0| < |x^3 - y^2|/(x^2/4 + |y|) < 4|x^3|/x^2 + + |y^2|/|y| = 4|x| + |y| \to 0$.

1.23. $\gamma = f([\alpha, \beta])$, dividamos los puntos $t \in [\alpha, \beta]$ atendiendo al lado al que pertenece $f(t)$, formando una partición I_1, I_2 de I; si $t_0 = \sup I_1$, entonces $f(t)$ está en R.

1.24. C es conexo. C_a nunca es conexo. Sea $\boldsymbol{a} = (x, y)$; si $x \neq 0$ e $y \neq 0$, hay dos componentes conexas; si $x = 0$ e $y \notin \mathbb{Q}$, hay 2; si $x = 0$ e $y \in \mathbb{Q}^*$, hay 3; si $y = 0$ y $x \notin \mathbb{Q}$, hay 2; si $y = 0$ y $x \in \mathbb{Q}^*$, hay 3; si $x = 0$ e $y = 0$, hay 4.

1.25. $(\boldsymbol{a}, \boldsymbol{b}) \in \bar{G}$; $\boldsymbol{a} = \lim \boldsymbol{x}_n$ y $\boldsymbol{b} = \lim \boldsymbol{f}(\boldsymbol{x}_n)$ con $\boldsymbol{x}_n \in C$; $\boldsymbol{a} \in C$; $\boldsymbol{f}(\boldsymbol{x}_n) \to \boldsymbol{f}(\boldsymbol{a})$; $\boldsymbol{b} = \boldsymbol{f}(\boldsymbol{a})$; $(\boldsymbol{a}, \boldsymbol{b}) \in G$.

1.26. Contrarrecíproco. C no cerrado $\Rightarrow \exists\, \boldsymbol{a} \in C$, $\boldsymbol{a} \notin \bar{C}$, $f(\boldsymbol{x}) = d(\boldsymbol{a}, \boldsymbol{x})$ no acotada. C no acotado $\Rightarrow f(\boldsymbol{x}) = d(\boldsymbol{o}, \boldsymbol{x})$ no acotada.

1.27. Sea $k \in f(C)$, $\exists\, h > 0$ tal que $\|\boldsymbol{x}\| > h \Rightarrow f(\boldsymbol{x}) > k$; $C_1 = \{\boldsymbol{x} \in C / \|\boldsymbol{x}\| \leqslant k\}$ es compacto, $\exists \min f(C_1) = \min f(C)$.

1.28. Si $x \neq y$, f continua; si $(x, y) \to (a, a)$, $f(x, x) \to \varphi'(a) = f(a, a)$ y $f(x, y) - f(a, a) = f'(\xi) - f'(a) \to 0$; f es continua en $\mathbb{R}^2$.

1.29. Sí, tomando $f_1(x, 0) = x/2$ y $f_2(x, x) = \cos^3 x$.

1.30. $\boldsymbol{f}$ continua en $C \Rightarrow \forall\, \boldsymbol{x} \in \boldsymbol{f}^{-1}(A)$, $\exists\, \varepsilon_x > 0$ tal que $B(\boldsymbol{f}(\boldsymbol{x}), \varepsilon_x) \subset A$; de «$\varepsilon:\delta$» aplicado a $\boldsymbol{f}$ en $\boldsymbol{x}$, se tiene $\delta_x > 0$; A' es la unión de las $B(\boldsymbol{x}, \delta_a)$. Se cumple la condición $\Rightarrow \forall\, \boldsymbol{a} \in C$ y $\forall\, \varepsilon > 0$, sea $A = B(\boldsymbol{f}(\boldsymbol{a}), \varepsilon)$ y acudamos al A'; $\exists\, \delta > 0 \,/\, B(\boldsymbol{a}, \delta) \subset A'$; para este δ se cumple la «$\varepsilon:\delta$».

1.31. 1. F es monótona creciente.

2. La «$\varepsilon:\delta$» de límite 0 de F para $r \to 0^+$ coincide con la δ de «$\varepsilon:\delta$» del límite de f en $\boldsymbol{a}$.

1.32. 1. La distancia del origen a la curva $x^2y^2 - (x^2 + y^2) = 0$ (menos el origen) es mínima cuando la pendiente ($y = mx$) es $m = 1$; esta distancia vale $r = \sqrt{2}$.

2. En $(x, y) \neq (0, 0)$ continuidad (obvia); en $(0, 0)$ el límite es 0; pues llamando $x^2 + y^2 = \rho$, para $\rho < 1/2$ es $(x^2 + y^2) - x^2y^2 > \rho^2 - \rho^4 > \rho^2 - \rho^2/4 = 3\rho^2/4$, luego $|f(x, y) - 0| < \rho^4/(3\rho^2/4) \to 0$; tomar $a = 0$.

3. No, para $(x, y) \to (1, 1)$ es $f(x, y) \to \infty$.

1.33. 1. Existe el límite l de ψ en 0; según $y = mx$, $\varphi(m) = l$.

2. $l = 0$, los límites valen 0.

3. φ acotada.

4. $\psi(0) = 0$.

5. $\psi(0) = 0$, $\psi(0^+) = 0$, φ acotada, luego f tiene límite 0 en $(0, 0)$.

1.34. 1. No, $C = \{x \in \mathbb{R}/x > 0\}$, $f(x) = 1/x$, $x_n = 1/n$.

2. Sí; $(\boldsymbol{x}_n) \to \boldsymbol{l} \in C$, $(\boldsymbol{f}(\boldsymbol{x}_n)) \to \boldsymbol{f}(\boldsymbol{l})$.

1.35. 1. Sea $\boldsymbol{x}_0$ cualquiera y $\boldsymbol{x}_n = \boldsymbol{f}(\boldsymbol{x}_{n-1})$ para $n \in \mathbb{N}$; $\|\boldsymbol{x}_{n+1} - \boldsymbol{x}_n\| \leqslant k^n \|\boldsymbol{x}_1 - \boldsymbol{x}_0\|$; $(\boldsymbol{x}_n)$ de Cauchy; $\boldsymbol{a} = \lim \boldsymbol{x}_n$ es un punto fijo.

2. $\boldsymbol{a}$ y $\boldsymbol{b}$ fijos $\Rightarrow \|\boldsymbol{a} - \boldsymbol{b}\| < \|\boldsymbol{a} - \boldsymbol{b}\|$.

3. $f: \mathbb{R} \to \mathbb{R}$, $f(x) = \left(x + \sqrt{x^2 + 1}\right)/2$.

1.36. *a*) $\varepsilon > 0$, $\delta = \varepsilon/2$; $\|\varphi(\boldsymbol{x}) - \varphi(\boldsymbol{x}')\| \leqslant \|(\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{x}) - (\boldsymbol{f}(\boldsymbol{x}') - \boldsymbol{x}')\| < 2\|\boldsymbol{x} - \boldsymbol{x}'\| = \varepsilon$.

b) φ tiene mínimo $\varphi(\boldsymbol{a})$, $\boldsymbol{a} \in C$; $\boldsymbol{a} \neq \boldsymbol{f}(\boldsymbol{a}) \Rightarrow \varphi(\boldsymbol{f}(\boldsymbol{a})) < \boldsymbol{f}(\boldsymbol{a})$ falso.

1.37. Dados $\boldsymbol{x} \in \mathbb{R}^p$ y $\varepsilon > 0$, $\exists\, \boldsymbol{c} \in C \,/\, d(\boldsymbol{x}, \boldsymbol{c}) \leqslant d(\boldsymbol{x}, C) + \varepsilon$; $d(\boldsymbol{x}', A) \leqslant d(\boldsymbol{x}', \boldsymbol{c}) \leqslant d(\boldsymbol{x}', \boldsymbol{x}) + d(\boldsymbol{x}, \boldsymbol{c}) \leqslant d(\boldsymbol{x}', \boldsymbol{x}) + d(\boldsymbol{x}, C) + \varepsilon < d(\boldsymbol{x}', \boldsymbol{x}) + d(\boldsymbol{x}, C)$; $|d(\boldsymbol{x}', C) - d(\boldsymbol{x}, C)| < d(\boldsymbol{x}, \boldsymbol{x}')$; tómese $\delta = \varepsilon$.

1.38. 1. Continua en $\mathbb{R}^* \times \mathbb{R}^* \cup \{(0, 0)\}$.

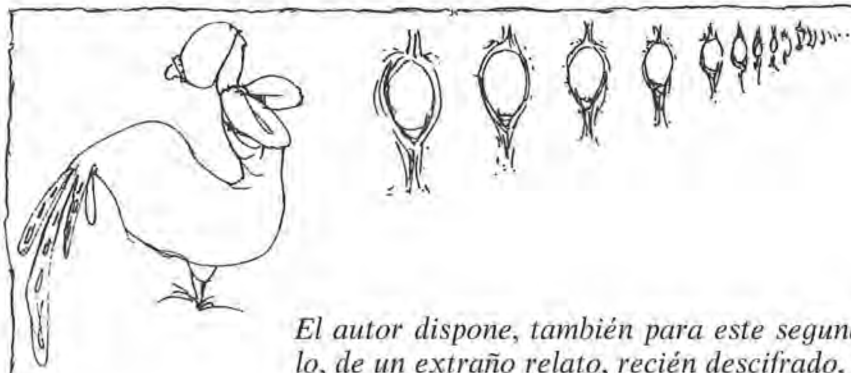
2. No es uniformemente continua en C: $(x_n, y_n) = (1/2, 1/(2n))$ y $(x'_n, y'_n) = (1/2, 1/(2n + 1))$.

1.39. 1. No: $(x_n, y_n) = \left(0, 1/\sqrt{2n\pi}\right)$ y $(x'_n, y'_n) = \left(0, 1/\sqrt{(2n + \pi/4)\pi}\right)$, $\varepsilon_0 = 1/2$.

2. Sí: $|f(x, y) - f(x', y')| < |1/\rho^2 - 1/\rho'^2| = \dfrac{|\rho - \rho'|(\rho + \rho')}{\rho^2\rho'^2} < 2|\rho - \rho'|$, $\delta = \varepsilon/2$.

1.40. Condición «$\varepsilon:\delta$» para $\boldsymbol{f} \Rightarrow$ «$\varepsilon:\delta$» para los f_i con igual δ. Condición «$\varepsilon:\delta$» para las $f_i \Rightarrow$ «$\varepsilon:\delta$» para $\boldsymbol{f}$ con el δ correspondiente al $\varepsilon/\sqrt{q}$ en las f_i.

1.41. $\exists\, k > 0 \,/\, \|\boldsymbol{x}\| \geqslant k \Rightarrow \|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{l}\| < \varepsilon/2$; $C_1 = \{\boldsymbol{x} \in C \,/\, \|\boldsymbol{x}\| \geqslant k\}$, $C_2 = \{\boldsymbol{x} \in C \,/\, \|\boldsymbol{x}\| \leqslant 2k\}$; C_2 compacto, «$\varepsilon:\delta$» en C_2; mín $\{\delta, k\}$.



El autor dispone, también para este segundo capítulo, de un extraño relato, recién descifrado, y se aventura a incluirle aquí, a modo de exordio. No quiere el autor ocultar al lector que el contenido del relato tiene muy poca enjundia matemática, pues sólo habla, y poco, de la carencia de plano tangencial en un punto un tanto singular; pero el autor entiende que este asunto, lo del plano tangente, es de tal fuste que bien merece uno y mil relatos.

EL CAMBIAZO DE UN HUEVO GALÁCTICO

Recordad el ya lejano viaje tripulado de una aeronave fuera de nuestro sistema planetario. Recordad que se nos dijo que algo raro pasó en aquel viaje, algo funesto, que cambió su ruta y llevó la nave a alejarse más y más de nosotros, sin posibilidad de regresar. Recordad que hubo mil rumores que decían que la nave retornó, que sobrevivió un solo tripulante, el astronauta Ívlov, el cual acabó perdiendo el juicio. Recordad que circularon mil versiones de las cosas disparatadas que se le atribuían a este desventurado.

Creo yo tener hoy pruebas fehacientes de que Ívlov regresó, de que Ívlov dijo que habían descubierto un mundo habitado por seres extraños, a los que llamó soef-soef, de que Ívlov, no sé si cuerdo o ya demente, contó muchas cosas sobre su estancia en el planeta de los soef-soef, de que las autoridades de la Tierra se han esforzado por ocultar estas noticias y han hecho desaparecer cuantas pruebas han podido de este asunto. No obstante, no se ha perdido el rastro de todo ello; yo dispongo, a buen recaudo, de uno de sus abundantes manuscritos que, a pesar de encontrarse bastante deteriorado, se ha podido interpretar satisfactoriamente.

Si bien es verdad que el manuscrito se ocupa de un asunto que da poca luz sobre las gentes de aquel planeta lejano y sus vidas, es una prueba de que Ívlov volvió y habló, de que los soef-soef existen y de que tienen hechuras muy distintas de las nuestras. Digamos ya, pues, lo que, en opinión de los expertos, se desprende de lo que aún puede leerse en el susodicho manuscrito.

Los soef-soef, al menos aquellos de los que tenemos noticias, se asemejan a los monstruos fabulosos de la mitología de la antigua Grecia: son unos seres con apariencia humana de cintura para arriba y el resto de su cuerpo se parece al de un ave. Se ignora si son alados, pero se sabe que son ovíparos; sus huevos son de forma y tamaño similares a los de los avestruces, y su período de incubación es larguísimo, de unos treinta meses. Para mantener los huevos a la temperatura pertinente durante todo este período, en lugar de ubicarse sobre ellos, como hacen aquí las aves, allá acuden a un método bastante original: colocan el huevo verticalmente con su parte más puntiaguda hacia abajo, apoyado en tres puntos situados algo por debajo de su ecuador, y sobre su parte más plana (la de arriba) se deja caer lentamente un delgado chorro de agua, a la temperatura adecuada, la cual desciende por el huevo lamiendo toda su cáscara. A nadie se le escapa que este método tiene el serio inconveniente de no ofrecer suficientes garantías sobre la maternidad del ser que va a nacer, pues se pueden producir, con gran facilidad, sustituciones de unos huevos por otros, ya que todos se encuentran, juntos, en una misma sala, que es una inmensa incubadora comunal. También suele ocurrir algo que es aún peor: que un huevo con vida sea sustituido por otro sin ella. Nada tiene, pues, de particular que los futuros padres y madres soef-soef acudan a las más variadas artimañas para evitar, a toda costa, que estas cosas ocurran.

2

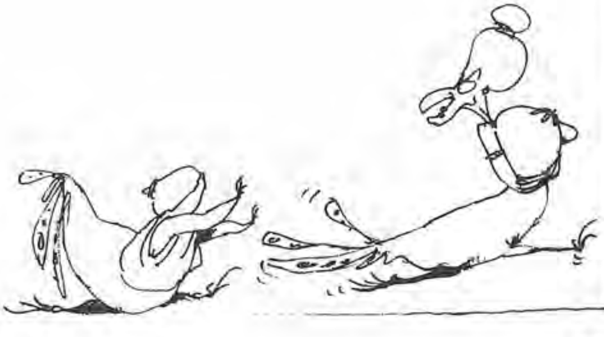
En nuestro manuscrito se relata el ardid al que acudió una madre soef-soef, llamada Osaca Is-Rop, para que nadie, subrepticiamente, se apropiase de su huevo, mejor dicho, para poderle recuperar en el caso de que alguien la diera el cambiazo. En su depósito de agua, aquel del que salía el agua caliente que bañaba a su huevo, Osaca Is-Rop añadió una sal, parecida a nuestro carbonato cálcico, de modo que el agua que de él salía contuviera gran cantidad de ella, en disolución, cosa esta que nacie apreciaría a simple vista. De esta disolución incolora, se iban depositando en el huevo pequeñas cantidades de sal, que quedaban adheridas fuertemente en su parte de abajo, a modo de incipiente estalactita. A los pocos días, sin más que pasar el dedo por allí, Osaca Is-Rop pudo apreciar un pequeño abultamiento de tipo esférico, algo así como un minúsculo grano que empieza a nacer, aún sin punta, redondeado en su cúspide, que no se distinguía a simple vista, pues además de su pequeño tamaño, era de color blanco y quedaba cubierto por el chorro de agua que abandonaba el huevo. Así pues, Osaca Is-Rop había provocado una deformación leve de la superficie del huevo, en las cercanías de su punto más bajo, la cual no se apreciaba aún con nitidez; ella, la superficie, seguía teniendo plano tangente en todos sus puntos, aunque ahora, al recorrer dicha zona, aquél variaba más rápidamente que antes.

Con el correr de los días, el abultamiento inferior del huevo, no sólo aumentaba lentamente de tamaño, sino que también variaba su configuración, pasando de ser redondeado a ser puntiagudo en su cima; parecía una púa o espina que, incluso, podía llegar a herir si se toqueteaba sin cuidado. Acontecía, pues, que la actual superficie del huevo había quedado deformada de modo que en su punto más bajo había dejado de tener plano tangente; este punto es de los que llaman angulosos. Y ya, cuando la longitud de la púa alcanzó entre medio y un milímetro, Osaca Is-Rop retiró la sal del depósito. Periódicamente se dirigía allí, reparaba en la púa y partía tranquila.

Cierto día, Osaca Is-Rop tuvo una corazonada: alguien había cambiado su huevo furtivamente. Sin perder un minuto se dirigió al sitio de su huevo, palpó el que allí había y apreció en él la púa que le distinguía de los demás; no obstante, algo en su interior le decía que aquel no era su huevo. Entonces pidió los permisos pertinentes y, acompañada de la autoridad, fue palpando, uno tras otro, todos los huevos de la inmensa incubadora. Al acercarse a uno de ellos, su espíritu maternal le dijo que aquel era el suyo; al palparlo, vio que no tenía la púa, pero luego, haciendo una inspección más detallada, en la parte baja de aquel huevo se detectaron unas pequeñas adherencias, lo que resultó ser los restos de aquella, que no pudo ser eliminada del todo por el ladrón. Se hicieron las pertinentes averiguaciones, de las que se dedujo que Osaca Is-Rop tenía razón; se le devolvió su huevo y se castigó severamente al culpable de todo aquello.

No parece desprenderse de lo dicho que dejarse guiar sólo por la intuición y por los pálpitos, olvidando el razonar sobre las cosas, deba ser nuestra norma de conducta. Pero tampoco es aconsejable despreciar ni los presentimientos ni los barruntos; a unos y a otros se les debe tener en cuenta para modificar, si hay precisión de ello, algunas de nuestras lógicas conclusiones.

Gudor Ben Jusá



CAPÍTULO 2

Diferenciación

Siguiendo con nuestro análisis de las funciones de varias variables, nos ocupamos ahora de evaluar la variación, de una tal función, en comparación con lo que varían aquéllas o, dicho de otro modo, con sus incrementos; se trata de relacionar aquella variación con estos incrementos y hacerlo cuando los cambios en los valores de las variables son pequeños o, aún mejor, cuando ellos tienden a cero.

Se van a estudiar funciones «lisas», que (además de ser continuas) no tengan cambios abruptos; es decir, en las que, con «pequeños» cambios de las variables, no sólo se obtienen «pequeños» cambios en el valor de la función, si no que también acontezca que la relación entre aquélla y estas variaciones sea «casi» lineal. Dicho de otro modo, vamos a considerar funciones que, cerca de cada uno de sus puntos, se comportan de modo muy parecido a como lo hacen las funciones lineales, esto es, que pueden aproximarse muy bien, en las cercanías de cada uno de sus puntos, por una función lineal (que dependerá del punto que se esté considerando). Si esta similitud con lo lineal se produce cuando se incrementa una sola de las variables, se dice que la función es derivable respecto de ella; si la «casi» linealidad lo es respecto de la variación conjunta de todas las variables, se dice que la función es diferenciable.

2.1. DERIVADAS PARCIALES

Según recordaremos, cuando se estudiaban las cuestiones relativas a la derivada de una función $x \mapsto \varphi(x)$, de una sola variable real x, se vio que la derivada $\varphi'(a)$, en un punto $a \in \mathbb{R}$, es el límite (si existe y es finito):

$$\varphi'(a) = \lim_{x \to a} \frac{\varphi(x) - \varphi(a)}{x - a} = \lim_{h \to 0} \frac{\varphi(a + h) - \varphi(a)}{h} \qquad [1]$$

Pero nosotros, ahora, vamos a considerar una función de varias variables $\boldsymbol{x} \mapsto f(\boldsymbol{x})$, donde $\boldsymbol{x} = (x_1, x_2, ..., x_p)$ es un vector de $\mathbb{R}^p$ (las variables son $x_1, x_2, ..., x_p$), y un punto $\boldsymbol{a} \in \mathbb{R}^p$.

En este caso, no se puede proceder como antes, pues ello nos llevaría a dividir por un vector, el $\boldsymbol{x} - \boldsymbol{a}$, lo que carece de sentido. No obstante, aquí se puede limitar la variación de $\boldsymbol{x}$ a una recta que pasa por $\boldsymbol{a}$, lo que conduce a las llamadas derivadas parciales, distintas, que dependen de la dirección según la que nos acerquemos al punto $\boldsymbol{a}$. Si $\boldsymbol{x}$ se acerca a $\boldsymbol{a}$ moviéndose a lo largo de la recta que tiene la dirección de un cierto vector $\boldsymbol{u} \neq \boldsymbol{o}$, esto es, si se toma $\boldsymbol{x} = \boldsymbol{a} + \lambda \boldsymbol{u}$, con $\lambda \in \mathbb{R}$, y se hace que $\lambda \to 0$, la definición [1] nos conduce de un modo natural a la siguiente definición de derivada parcial (de f en $\boldsymbol{a}$) respecto del vector $\boldsymbol{u}$:

$$f'_u(\boldsymbol{a}) = \lim_{\lambda \to 0} \frac{f(\boldsymbol{a} + \lambda \boldsymbol{u}) - f(\boldsymbol{a})}{\lambda}$$

De ahí que se den las siguientes definiciones de derivadas parciales de una función de varias variables:

DERIVADA SEGÚN UN VECTOR; DERIVADAS PARCIALES

[22]

Sea $f: C \to \mathbb{R}^q$ una función, definida en un abierto $C \subset \mathbb{R}^p$. Se llama *derivada* (o derivada parcial) de f en un punto $\boldsymbol{a} \in C$ y respecto de un vector no nulo $\boldsymbol{u} \in \mathbb{R}^p$ al siguiente límite, si existe y es finito, que se denota poniendo $f'_u(\boldsymbol{a})$, $D_u[f(\boldsymbol{a})]$ o $\partial f(\boldsymbol{a})/\partial \boldsymbol{u}$:

$$\frac{\partial f(\boldsymbol{a})}{\partial \boldsymbol{u}} = \lim_{\lambda \to 0} \frac{f(\boldsymbol{a} + \lambda \boldsymbol{u}) - f(\boldsymbol{a})}{\lambda}$$

Si $\|\boldsymbol{u}\| = 1$, se dice que $f'_u(\boldsymbol{a})$ es la derivada, de f en $\boldsymbol{a}$, respecto de la recta que pasa por $\boldsymbol{a}$ y está orientada por $\boldsymbol{u}$ (derivada direccional). Si $(\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_p)$ es la base canónica de $\mathbb{R}^p$, a la derivada $\partial f(\boldsymbol{a})/\partial \boldsymbol{e}_i$ (para $i = 1, 2, \ldots, p$) se la llama *derivada parcial*, de f en $\boldsymbol{a}$, de lugar i o respecto de su i-ésima variable y se la representa poniendo:

$$f'_i(\boldsymbol{a}), \qquad f'_{x_i}(\boldsymbol{a}), \qquad D_i[f(\boldsymbol{a})], \qquad D_{x_i}[f(\boldsymbol{a})] \qquad \text{o} \qquad \frac{\partial f(\boldsymbol{a})}{\partial x_i}$$

(aquí, x_i representa a la coordenada i-ésima de la variable $\boldsymbol{x} \in C$).

Si la función $f: C \to \mathbb{R}^q$ es derivable respecto de un vector $\boldsymbol{u} \neq \boldsymbol{o}$ en todos los puntos del abierto C, se llama *función derivada (parcial)* de f respecto de $\boldsymbol{u}$ a la aplicación f'_u, $D_u(f)$ o $\partial f/\partial \boldsymbol{u}$, de C en $\mathbb{R}^q$, definida por:

$$\frac{\partial f}{\partial \boldsymbol{u}} : \boldsymbol{x} \mapsto \frac{\partial f}{\partial \boldsymbol{u}}(\boldsymbol{x}) = \frac{\partial f(\boldsymbol{x})}{\partial \boldsymbol{u}}$$

[22]$_1$ Ejemplos

1.° Sea $f: \mathbb{R}^2 \to \mathbb{R}$ la función definida mediante las expresiones:

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

La derivada parcial de f en el punto (0, 0) respecto del vector $\boldsymbol{u} = (a, b)$, es:

$$\frac{\partial f(0, 0)}{\partial \boldsymbol{u}} = \lim_{\lambda \to 0} \frac{f(0 + \lambda a, 0 + \lambda b) - f(0, 0)}{\lambda} = \lim_{\lambda \to 0} \frac{a^2 \text{ sen } (\lambda b) + b^2 \text{ sen } (\lambda a)}{\lambda(a^2 + b^2)} =$$

$$\overset{(*)}{=} \lim_{\lambda \to 0} \frac{a^2 b \cos (\lambda b) + b^2 a \cos (\lambda a)}{a^2 + b^2} = \frac{a^2 b + b^2 a}{a^2 + b^2}$$

(*) Aplicando la regla de L'Hôpital.

2.° Sea $f: \mathbb{R}^2 \to \mathbb{R}$ la función definida por la expresión

$$f(x, y) = \varphi(x, y)(ax + by)$$

donde $\varphi: \mathbb{R}^2 \to \mathbb{R}$ es una función continua en (0, 0). Las dos derivadas parciales de f en el punto (0, 0) son:

$$\frac{\partial f(0, 0)}{\partial x} = \lim_{h \to 0} \frac{f(h, 0) - f(0, 0)}{h} = \lim_{h \to 0} \frac{\varphi(h, 0)ah}{h} = a\varphi(0, 0)$$

$$\frac{\partial f(0, 0)}{\partial y} = \lim_{k \to 0} \frac{f(0, k) - f(0, 0)}{k} = \lim_{k \to 0} \frac{\varphi(0, k)bk}{k} = b\varphi(0, 0)$$

[22]$_2$ Observaciones

1.° En la anterior definición [22], en lugar de suponer que $C \subset \mathbb{R}^p$ es un conjunto abierto, se podía haber exigido que $\boldsymbol{a}$ fuese un punto interior de C. De uno u otro modo, con ello se garantiza la existencia de

$$\frac{\boldsymbol{f}(\boldsymbol{a} + \lambda \boldsymbol{u}) - \boldsymbol{f}(\boldsymbol{a})}{\lambda}$$

para todo número real λ de un cierto entorno reducido de $\lambda = 0$, lo que posibilita la existencia de la derivada $\partial \boldsymbol{f}(\boldsymbol{a})/\partial \boldsymbol{u}$.

2.° Aquí en la definición [22], donde se considera una función $\boldsymbol{f}: C \to \mathbb{R}^q$ (con $C \subset \mathbb{R}^p$), no hemos podido utilizar la definición clásica de derivada de una función real de variable real a causa de que ahora hay más de una variable (o sea, de que es $p \geqslant 2$), pero en ello nada tiene que ver el hecho de que la función tome valores en el espacio $\mathbb{R}^q$ con $q > 1$.

3.º Una función $f: C \to \mathbb{R}^q$ (donde C es un abierto de $\mathbb{R}^p$) admite derivada respecto un vector no nulo $\boldsymbol{u} \in \mathbb{R}^p$ en un punto $\boldsymbol{a} \in C$ y dicha derivada es el vector $\boldsymbol{d} \in \mathbb{R}^q$ si para cada $\varepsilon > 0$ existe un $\delta > 0$ tal que

$$0 < |\lambda| < \delta \quad \Rightarrow \quad \left\| \frac{f(\boldsymbol{a} + \lambda \boldsymbol{u}) - f(\boldsymbol{a})}{\lambda} - \boldsymbol{d} \right\| < \varepsilon$$

4.º Las definiciones [22], de derivadas según vectores y, en particular, de derivadas parciales de la función $f: C \to \mathbb{R}^q$ (con $C \subset \mathbb{R}^p$) en un punto $\boldsymbol{a} \in C$, se pueden expresar de la siguiente manera, en la que se recurre a las coordenadas: poniendo $\boldsymbol{a} = (a_1, a_2, ..., a_p)$, $\boldsymbol{x} = (x_1, x_2, ..., x_p)$ y $\boldsymbol{u} = (u_1, u_2, ..., u_p)$, se tiene:

$$\frac{\partial f(\boldsymbol{a})}{\partial \boldsymbol{u}} = \lim_{\lambda \to 0} \frac{f(a_1 + \lambda u_1, a_2 + \lambda u_2, ..., a_p + \lambda u_p) - f(a_1, a_2, ..., a_p)}{\lambda}$$

$$\frac{\partial f(\boldsymbol{a})}{\partial x_i} = \lim_{\lambda \to 0} \frac{f(a_1, ..., a_{i-1}, a_i + \lambda, a_{i+1}, ..., a_p) - f(a_1, ..., a_{i-1}, a_i, a_{i+1}, ..., a_p)}{\lambda}$$

En particular, en el caso de dos variables, las derivadas parciales de $(x, y) \mapsto f(x, y)$ en el punto (a, b) son:

$$\frac{\partial f(a, b)}{\partial x} = \lim_{h \to 0} \frac{f(a + h, b) - f(a, b)}{h}$$

$$\frac{\partial f(a, b)}{\partial y} = \lim_{k \to 0} \frac{f(a, b + k) - f(a, b)}{k}$$

También en el caso de dos variables, la derivada según la dirección y sentido del vector unitario $\boldsymbol{u} = (\cos\theta, \operatorname{sen}\theta)$ es:

$$\frac{\partial f(a, b)}{\partial \boldsymbol{u}} = \lim_{\lambda \to 0} \frac{f(a + \lambda\cos\theta, b + \lambda\operatorname{sen}\theta) - f(a, b)}{\lambda}$$

5.º Cuando se considera una función $x \mapsto \varphi(x)$, de una variable real x, para denotar a su derivada se pone $d\varphi/dx$. Si la función $\boldsymbol{x} \mapsto f(\boldsymbol{x})$ que se considera es de varias variables, de manera que $\boldsymbol{x} = (x_1, x_2, ..., x_p)$, para denotar a su derivada parcial i-ésima se pone $\partial f/\partial x_i$. Nótese que se ha recurrido «∂» en lugar de acudir a la «d», como hacía en una variable. Ello no es algo caprichoso, sino que se hace para evitar equívocos, ya que: para la función de una variable φ, la derivada $d\varphi/dx$ resultó ser el cociente de $d\varphi$ (diferencial de φ) entre dx (diferencial de x), cosa que no ocurre cuando la función f es de más de una variable. Adviértase, pues, que $\partial f/\partial x_i$ es una notación que no debe interpretarse como cociente; se trata de un todo y no tiene sentido expresiones como ∂f o como ∂x_i, que sólo son partes del todo $\partial f/\partial x_i$.

[22]$_3$ Ejercicio

Sea $f: C \to \mathbb{R}^q$ una función, definida en un abierto $C \subset \mathbb{R}^p$, que admite derivada $\partial f(\boldsymbol{a})/\partial \boldsymbol{u}$, en un punto $\boldsymbol{a} \in C$ y respecto de un vector no nulo $\boldsymbol{u} \in \mathbb{R}^p$. Si $\boldsymbol{v} = \rho \boldsymbol{u}$, para cierto $\rho \in \mathbb{R}$ no nulo, hallar la derivada $\partial f(\boldsymbol{a})/\partial \boldsymbol{v}$.

Resolución

$$\frac{\partial f(\boldsymbol{a})}{\partial \boldsymbol{v}} = \lim_{\lambda \to 0} \frac{f(\boldsymbol{a} + \lambda \boldsymbol{v}) - f(\boldsymbol{a})}{\lambda} = \lim_{\lambda \to 0} \frac{f(\boldsymbol{a} + \lambda\rho \boldsymbol{u}) - f(\boldsymbol{a})}{\lambda} =$$

$$= \rho \lim_{\lambda\rho \to 0} \frac{f(\boldsymbol{a} + \lambda\rho \boldsymbol{u}) - f(\boldsymbol{a})}{\lambda\rho} = \rho \frac{\partial f(\boldsymbol{a})}{\partial \boldsymbol{u}}$$

PRIMERAS PROPIEDADES DE LA DERIVACIÓN

[23]

Sea $f: C \to \mathbb{R}^q$ una función, definida en un abierto $C \subset \mathbb{R}^p$, y considérense un punto $\boldsymbol{a} \in C$ y un vector no nulo $\boldsymbol{u} \in \mathbb{R}^p$.

1.° La función f es derivable en $\boldsymbol{a}$ según $\boldsymbol{u}$ si, y sólo si, lo son sus q componentes y, si dichas componentes son $f_1, f_2, ..., f_q$ (funciones de C en $\mathbb{R}$), el vector derivada $\partial f(\boldsymbol{a})/\partial \boldsymbol{u}$ tiene por coordenadas a los q números $\partial f_1(\boldsymbol{a})/\partial \boldsymbol{u}$, $\partial f_2(\boldsymbol{a})/\partial \boldsymbol{u}$, ..., $\partial f_q(\boldsymbol{a})/\partial \boldsymbol{u}$.

2.° Para la derivación parcial (respecto de la i-ésima variable, x_i) son de aplicación las reglas de la derivación de las funciones de una variable; el procedimiento operativo para el cálculo de derivadas parciales es el mismo que el que se sigue para hallar derivadas en el caso de una variable (aquí, la variable es x_i; las x_j con $j \neq i$ permanecen fijas en este proceso).

3.° La existencia de todas las derivadas de f en $\boldsymbol{a}$ (según los distintos vectores no nulos) no es ni necesario ni suficiente para que f sea continua en $\boldsymbol{a}$(*).

4.° Si llamamos φ a la función $\lambda \mapsto \varphi(t) = f(\boldsymbol{a} + \lambda \boldsymbol{u})$, cuya variable real λ recorre un cierto entorno de 0, se verifica que

$$\frac{\partial f(\boldsymbol{a})}{\partial \boldsymbol{u}} = \varphi'(0) \qquad \text{(si existen)}$$

(*) Ello no obstante, es evidente que la derivabilidad de f en $\boldsymbol{a}$ según $\boldsymbol{u}$ implica la continuidad de f en $\boldsymbol{a}$ según la recta que tiene la dirección de $\boldsymbol{u}$.

Comprobación

1.° El límite de una función vectorial es (véase [12],5) el vector cuyas coordenadas son los límites de las componentes de la función dada, los cuales existen si, y sólo si, existe aquel. Por ello y como

$$\frac{f_i(\boldsymbol{a} + \lambda \boldsymbol{u}) - f_i(\boldsymbol{a})}{\lambda} \qquad \text{es la componente } i\text{-ésima de} \qquad \frac{f(\boldsymbol{a} + \lambda \boldsymbol{u}) - f(\boldsymbol{a})}{\lambda}$$

al tomar límites para $\lambda \to 0$ se obtiene la propiedad que queríamos comprobar.

2.° Este resultado es consecuencia inmediata de la definición de derivada parcial (respecto de x_i), en la que las otras variables x_j (con $j \neq i$) no funcionan como tales, sino que permanecen invariantes. Por ello, calcular la derivada parcial $\partial f/\partial x_i$ equivale a, suponiendo que las demás variables se mantienen constantes, hallar la derivada de la función (de una variable) $x_i \mapsto f(x_1, ..., x_i, ..., x_p)$. Así pues, las reglas para la derivación parcial (respecto de x_i) son las de la derivación ordinaria (caso de variable única).

3.° Para comprobar este aserto, nos bastará con facilitar los correspondientes contraejemplos:

- La función $f: \mathbb{R}^2 \to \mathbb{R}$ definida mediante las expresiones:

$$f(x, y) = \frac{x^3}{y}, \text{ si } y \neq 0; \qquad f(x, 0) = 0$$

tiene derivadas en el punto (0, 0) según todos los vectores no nulos (dichas derivadas son nulas), pero no es continua en (0, 0) pues $f(0, 0) = 0$ y el límite de f en (0, 0) según la trayectoria $y = x^3$ es $1 \neq 0$.

- La función $f: \mathbb{R}^2 \to \mathbb{R}$ definida mediante $f(x, y) = \sqrt{x^2 + y^2}$ es continua en (0, 0), pero no admite derivada en (0, 0) según ningún vector $\boldsymbol{u} = (\alpha, \beta)$ no nulo, pues

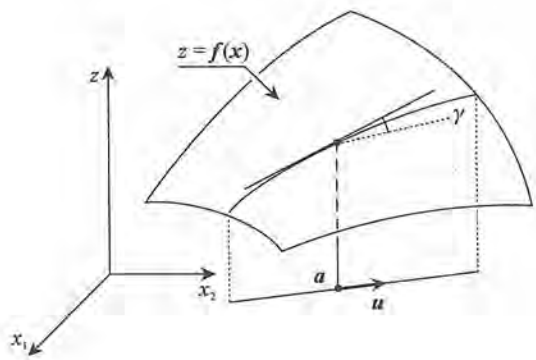
$$\lim_{\lambda \to 0} \frac{f(0 + \lambda\alpha,\ 0 + \lambda\beta) - f(0,\ 0)}{\lambda} = \lim_{\lambda \to 0} \sqrt{a^2 + b^2}\,\frac{|\lambda|}{\lambda} \qquad \text{no existe}$$

4.° Este resultado es evidente, ya que

$$\varphi'(0) = \lim_{\lambda \to 0} \frac{\varphi(\lambda) - \varphi(0)}{\lambda} = \lim_{\lambda \to 0} \frac{f(\boldsymbol{a} + \lambda\boldsymbol{u}) - f(\boldsymbol{a})}{\lambda} = \frac{\partial f(\boldsymbol{a})}{\partial \boldsymbol{u}}$$

[23]$_1$ Interpretación geométrica

En el caso de una función real de dos variables reales, es decir, al considerar una función $f: C \to \mathbb{R}$ donde C es un abierto de $\mathbb{R}^2$, el último de los resultados anteriores (véase [23],4.°) conduce a la siguiente interpretación geométrica, en la que se acude a la representación cartesiana usual de la función f, cuya gráfica es la superficie de ecuación $z = f(x_1, x_2)$



para $(x_1, x_2) \in C$, en ejes cartesianos rectangulares $x_1 x_2 z$. Dados el punto $\boldsymbol{a} \in C$ y el vector no nulo $\boldsymbol{u} \in \mathbb{R}^2$, consideremos el plano perpendicular al $z = 0$ que contiene a la recta que pasa por $\boldsymbol{a}$ y tiene la dirección de $\boldsymbol{u}$ (los puntos de esta recta son los $\boldsymbol{x} = \boldsymbol{a} + \lambda \boldsymbol{u}$ para $\lambda \in \mathbb{R}$); este plano corta a la gráfica de f según una cierta curva γ. Pues bien, la derivada $\partial f(\boldsymbol{a})/\partial \boldsymbol{u}$ es la pendiente, medida sobre el plano $z = 0$, de la curva γ en el punto $(x_1, x_2) = \boldsymbol{a}$, $z = f(\boldsymbol{a})$.

[23]$_2$ Ejercicios

1.° Hallar la derivada en el origen (0, 0) y respecto del vector unitario $\boldsymbol{u} = (\cos\theta, \operatorname{sen}\theta)$ de la siguiente función $\boldsymbol{f} = (f_1, f_2)$, cuyas componentes son las funciones $f_1 : \mathbb{R} \to \mathbb{R}$ y $f_2 : \mathbb{R}^2 \to \mathbb{R}$ definidas por las expresiones

$$\begin{cases} f_1(x,y) = \dfrac{\operatorname{sen} x^2 - \operatorname{sen} y^2}{x - y}, \ \text{si } x \neq y; & f_1(x, x) = 0 \\[2ex] f_2(x,y) = \dfrac{\mathrm{L}\ (x^2 + y^2 + 1)}{x - y}, \text{si } x \neq y; & f_2(x, x) = 0 \end{cases}$$

2.° Hallar las dos funciones derivadas parcial de la función $f : \mathbb{R}^2 \to \mathbb{R}$ definida por la expresión:

$$f(x, y) = \frac{(x^2 + y^2 + 1)^{\operatorname{sen} x}}{e^{2x + y^2}}$$

Resolución

1.° Para $\theta \neq \pi/4 + k\pi$ (o sea, para $x \neq y$) y recurriendo a la regla de L'Hôpital, se tiene:

$$\frac{\partial f_1(0,0)}{\partial \boldsymbol{u}} = \lim_{\lambda \to 0} \frac{\operatorname{sen}(\lambda^2 \cos^2\theta) - \operatorname{sen}(\lambda^2 \operatorname{sen}^2\theta)}{\lambda^2(\cos\theta - \operatorname{sen}\theta)} = 2(\cos\theta - \operatorname{sen}\theta)$$

$$\frac{\partial f_2(0,0)}{\partial \boldsymbol{u}} = \lim_{\lambda \to 0} \frac{\mathrm{L}\,(\lambda^2 + 1)}{\lambda^2(\cos\theta - \operatorname{sen}\theta)} = \frac{1}{(\cos\theta - \operatorname{sen}\theta)}$$

Para $\theta = \pi/4$, o sea si es $\boldsymbol{u} = \boldsymbol{u}_0 = (\sqrt{2}/2, \sqrt{2}/2)$, entonces:

$$\frac{\partial f_1(0,0)}{\partial \boldsymbol{u}_0} = \lim_{\lambda \to 0} \frac{0 - 0}{\lambda} \qquad \text{y} \qquad \frac{\partial f_2(0,0)}{\partial \boldsymbol{u}_0} = \lim_{\lambda \to 0} \frac{0 - 0}{\lambda} = 0$$

Así pues, la derivada pedida vale:

$$\frac{\partial \boldsymbol{f}(0,0)}{\partial \boldsymbol{u}} = \left(2(\cos\theta - \operatorname{sen}\theta), \frac{1}{\cos\theta - \operatorname{sen}\theta}\right), \qquad \text{si } \boldsymbol{u} \neq \boldsymbol{u}_0$$

$$\frac{\partial \boldsymbol{f}(0,0)}{\partial \boldsymbol{u}_0} = (0, 0), \qquad \text{donde } \boldsymbol{u}_0 = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$$

2.º Acudiendo a las reglas de derivación (que ya las conocemos: son las de funciones de una variable), se obtiene que:

$$\frac{\partial f(x,y)}{\partial x} = \frac{(x^2+y^2+1)^{\operatorname{sen} x}}{e^{2x+y^2}}\left(\cos x \operatorname{L}(x^2+y^2+1) + \frac{2x \operatorname{sen} x}{x^2+y^2+1} - 2\right)$$

$$\frac{\partial f(x,y)}{\partial y} = \frac{(x^2+y^2+1)^{\operatorname{sen} x}}{e^{2x+y^2}}\left(\frac{2y \operatorname{sen} x}{x^2+y^2+1} - 2y\right)$$

[23]$_3$ Ejercicio (función con una derivada parcial nula)

Sea $f: C \to \mathbb{R}^q$ una función, definida en un abierto $C \subset \mathbb{R}^p$. Si C es un conjunto convexo(*) y si existe y es nula en C la función f_i', derivada parcial i-ésima, entonces se verifica que $f(\boldsymbol{x}) = f(\boldsymbol{x}')$ para cualesquiera $\boldsymbol{x} = (x_1, ..., x_i, ..., x_p)$ y $\boldsymbol{x}' = (x_1, ..., x_i', ..., x_p)$ de C, es decir, la función f no depende de la variable i-ésima x_i.

Resolución

Si se demuestra esta propiedad para el caso de ser f una función real (esto es, para $q = 1$); acudiendo a la propiedad [23],1.º, se obtiene que la propiedad es entonces cierta para $q \in \mathbb{N}$ cualquiera. Suponemos, pues, que es $q = 1$.

Como C es convexo y $\boldsymbol{x}, \boldsymbol{x}' \in C$, también son de C todos los puntos del segmento $[\boldsymbol{x}, \boldsymbol{x}']$, luego la siguiente función φ está definida en el intervalo de extremos x_i y x_i':

$$t \mapsto \varphi(t) = f(x_1, ..., x_{i-1}, t, x_{i+1}, ..., x_p)$$

Esta función es derivable en dicho intervalo y su derivada en un punto t de él es $\varphi'(t) = f_i'(x_1, ..., t, ..., x_p)$. En consecuencia, a φ le es de aplicación el teorema de los incrementos finitos, en el intervalo en el que está definida, y dicho teorema dice que existe cierto ξ_i del susodicho intervalo tal que

$$\varphi(x_i') - \varphi(x_i) = (x_i' - x_i)\varphi'(\xi_i)$$

Como el punto $\xi = (x_1, ..., \xi_i, ..., x_p)$ es un punto de C (pues C es convexo) y por ello $f_i'(\xi) = 0$, la anterior igualdad nos conduce a:

$$f(\boldsymbol{x}') - f(\boldsymbol{x}) = (x_i' - x_i)f_i'(\xi) = 0$$

luego $f(\boldsymbol{x}') = f(\boldsymbol{x})$, como había que comprobar.

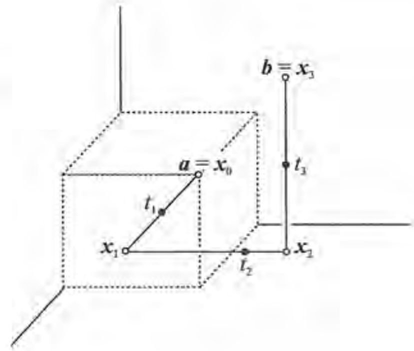
(*) Se dice que C es convexo si, siempre que $\boldsymbol{a} \in C$ y $\boldsymbol{b} \in C$, también son de C todos los puntos del segmento $[\boldsymbol{a}, \boldsymbol{b}]$. Para la cuestión que aquí se plantea, es suficiente con que la anterior condición se verifique sólo si el segmento $[\boldsymbol{a}, \boldsymbol{b}]$ es paralelo al eje i-ésimo de la referencia cartesiana.

$[23]_4$ Una primera fórmula de incrementos finitos

Dados dos puntos $\boldsymbol{a} = (a_1, a_2, ..., a_p)$ y $\boldsymbol{b} = (b_1, b_2, ..., b_p)$ de $\mathbb{R}^p$, sean $\boldsymbol{x}_0, \boldsymbol{x}_1, ..., \boldsymbol{x}_p$ los puntos $\boldsymbol{x}_i = (b_1, ..., b_i, a_{i+1}, ..., a_p)$, nótese que $\boldsymbol{x}_0 = \boldsymbol{a}$ y $\boldsymbol{x}_p = \boldsymbol{b}$, y considérense también los segmentos $S_i = [\boldsymbol{x}_{i-1}, \boldsymbol{x}_i]$, para $i = 1, 2, ..., p$.

Si $C \subset \mathbb{R}^p$ es un abierto que incluye a los segmentos $S_1, S_2, ..., S_p$ y si $f: C \to \mathbb{R}$ es una función real que admite sus p derivadas parciales $D_i(f)$, para $i = 1, 2, ..., p$, en todo punto de C, entonces existen p puntos $\boldsymbol{t}_i \in]\boldsymbol{x}_{i-1}, \boldsymbol{x}_i[$ (para $i = 1, ..., p$) tales que:

$$f(\boldsymbol{b}) - f(\boldsymbol{a}) = (b_1 - a_1)D_1[f(\boldsymbol{t}_1)] + \\ + (b_2 - a_2)D_2[f(\boldsymbol{t}_2)] + \cdots + (b_p - a_p)D_p[f(\boldsymbol{t}_p)]$$



O sea, existe (para $i = 1, 2, ..., p$) un número ξ_i del intervalo abierto de extremos a_i y b_i tal que:

$$f(b_1, ..., b_p) - f(a_1, ..., a_p) = \\ = \sum_{i=1}^{p} (b_i - a_i)D_i[f(b_1, ..., b_{i-1}, \xi_i, a_{i+1}, ..., a_p)]$$

Demostración

Empecemos expresando la diferencia $f(\boldsymbol{b}) - f(\boldsymbol{a})$ en la forma:

$$f(\boldsymbol{b}) - f(\boldsymbol{a}) = f(\boldsymbol{x}_p) - f(\boldsymbol{x}_0) = \sum_{i=1}^{p} [f(\boldsymbol{x}_i) - f(\boldsymbol{x}_{i-1})] \qquad [1]$$

Consideremos ahora, para cada $i = 1, 2, ..., p$, la función real φ_i, de variable real t, definida en el intervalo cerrado de extremos a_i y b_i por

$$t \mapsto \varphi_i(t) = f(b_1, ..., b_{i-1}, t, a_{i+1}, ..., a_p)$$

Como existe la derivada parcial f_i' en C y como $S_i \subset C$, resulta que φ_i es derivable en su intervalo de definición y su derivada vale

$$D[\varphi_i(t)] = D_i[f(b_1, ..., b_{i-1}, t, a_{i+1}, ..., a_p)]$$

En consecuencia, a la función φ_i le es de aplicación el teorema de los incrementos finitos, en su intervalo de definición, luego existe un número ξ_i del intervalo abierto de extremos a_i y b_i tal que $\varphi_i(b_i) - \varphi_i(a_i) = (b_i - a_i)D[\varphi_i(\xi_i)]$, esto es, verificándose que:

$$f(\boldsymbol{x}_i) - f(\boldsymbol{x}_{i-1}) = (b_i - a_i)D_i[f(b_1, ..., b_{i-1}, \xi_i, a_{i+1}, ..., a_p)]$$

Llevando esta expresión a la igualdad [1], se obtiene el resultado que buscábamos.

2.2. FUNCIONES DIFERENCIABLES

Para estudiar, al menos en primera aproximación, una función en las proximidades de uno de sus puntos, se acude a la función lineal que mejor aproxima a su incremento, si es que ella existe. A esta función lineal se la llama diferencial, de la función en el punto, y es tal que su diferencia con el incremento de la función es un infinitésimo de mayor orden que el incremento de la variable (mejor dicho: que la distancia del punto fijo al punto variable, en los que se considera la función).

DIFERENCIABILIDAD. DIFERENCIAL

[24]

Sea $f: C \to \mathbb{R}^q$ una función, definida en un abierto $C \subset \mathbb{R}^p$, y considérese también un punto $\boldsymbol{a} \in C$. Se dice que f es *diferenciable* en $\boldsymbol{a}$ si se verifica una cualquiera de las tres condiciones siguientes, equivalentes entre sí:

1.º Existe una aplicación lineal, de $\mathbb{R}^p$ en $\mathbb{R}^q$, que se denota por $df(\boldsymbol{a})$(*) y se llama *diferencial* de f en $\boldsymbol{a}$, tal que:

$$\lim_{\boldsymbol{x} \to \boldsymbol{a}} \frac{[f(\boldsymbol{x}) - f(\boldsymbol{a})] - df(\boldsymbol{a})(\boldsymbol{x} - \boldsymbol{a})}{\|\boldsymbol{x} - \boldsymbol{a}\|} = \boldsymbol{o}$$

Esto es, tal que $[f(\boldsymbol{x}) - f(\boldsymbol{a})] - df(\boldsymbol{a})(\boldsymbol{x} - \boldsymbol{a}) = \boldsymbol{o}(\|\boldsymbol{x} - \boldsymbol{a}\|)$.

2.º Existe una aplicación lineal $df(\boldsymbol{a}): \mathbb{R}^p \to \mathbb{R}^q$ y existe una función $\varepsilon: \boldsymbol{h} \mapsto \varepsilon(\boldsymbol{h})$ (definida para $\boldsymbol{x} = \boldsymbol{a} + \boldsymbol{h}$ recorriendo C) tales que:

$$f(\boldsymbol{a} + \boldsymbol{h}) - f(\boldsymbol{a}) = df(\boldsymbol{a})(\boldsymbol{h}) + \varepsilon(\boldsymbol{h})\|\boldsymbol{h}\| \qquad \text{y} \qquad \lim_{\boldsymbol{h} \to \boldsymbol{o}} \varepsilon(\boldsymbol{h}) = \boldsymbol{o}$$

3.º Existe una aplicación lineal $df(\boldsymbol{a}): \mathbb{R}^p \to \mathbb{R}^q$ y existen p funciones $\varepsilon_i: \boldsymbol{h} \mapsto \varepsilon_i(\boldsymbol{h})$ para $i = 1, 2, ..., p$ (definidas para $\boldsymbol{x} = \boldsymbol{a} + \boldsymbol{h}$ recorriendo C) tales que:

$$f(\boldsymbol{a} + \boldsymbol{h}) - f(\boldsymbol{a}) = df(\boldsymbol{a})(\boldsymbol{h}) + \sum_{i=1}^{p} \varepsilon_i(\boldsymbol{h})h_i, \qquad \text{siendo } \boldsymbol{h} = (h_1, h_2, ..., h_p)$$

$$\varepsilon_i(\boldsymbol{h}) \to \boldsymbol{o} \qquad \text{cuando } \boldsymbol{h} \to \boldsymbol{o} \qquad (\text{para } i = 1, 2, ..., p)$$

Si la función $f: C \to \mathbb{R}^q$ es diferenciable en todos los puntos de C, se dice entonces que f es diferenciable en C y a la aplicación df definida (en C) mediante $df: \boldsymbol{x} \mapsto df(\boldsymbol{x})$ se le llama diferencial de la función f.

(*) El valor que toma la función $df(\boldsymbol{a})$ cuando el incremento de la variable es $\boldsymbol{h} = \boldsymbol{x} - \boldsymbol{a}$, esto es, el vector $[df(\boldsymbol{a})](\boldsymbol{h})$, se denota poniendo simplemente $df(\boldsymbol{a})(\boldsymbol{h})$; hay quien prefiere poner $df(\boldsymbol{a}; \boldsymbol{h})$.

Comprobaciones

La equivalencia de las condiciones primera y segunda es inmediata, pues la relación $\varepsilon(\boldsymbol{h}) \to \boldsymbol{o}$ cuando $\boldsymbol{h} \to \boldsymbol{o}$ de la segunda equivale a $\varepsilon(\boldsymbol{h})\|\boldsymbol{h}\| = \boldsymbol{o}(\|\boldsymbol{h}\|)$, como se pide en la primera).

Para ver que las condiciones segunda y tercera son equivalentes, tómese

$$\varepsilon_i(\boldsymbol{h}) = \varepsilon(\boldsymbol{h}) \frac{h_i}{\|\boldsymbol{h}\|} \qquad \text{y} \qquad \varepsilon(\boldsymbol{h}) = \frac{1}{\|\boldsymbol{h}\|} \sum_{i=1}^{p} \varepsilon_i(\boldsymbol{h}) h_i$$

Como $h_i/\|\boldsymbol{h}\|$ está comprendido entre -1 y 1, resulta que: *a*) si $\varepsilon(\boldsymbol{h}) \to \boldsymbol{o}$, entonces $\varepsilon_i(\boldsymbol{h}) \to \boldsymbol{o}$ pues $h_i/\|\boldsymbol{h}\|$ está acotado; y *b*) si $\varepsilon_i(\boldsymbol{h}) \to \boldsymbol{o}$ (para $i = 1, 2, \ldots, p$) entonces

$$\lim_{h \to o} \|\varepsilon(\boldsymbol{h})\| \leqslant \lim_{h \to o} \frac{1}{\|\boldsymbol{h}\|} \sum_{i=1}^{p} \|\varepsilon_i(\boldsymbol{h})\| \leqslant \lim_{h \to 0} \sum_{i=1}^{p} \|\varepsilon_i(\boldsymbol{h})\| = 0$$

esto es, $\varepsilon(\boldsymbol{h}) \to \boldsymbol{o}$. Así pues, las condiciones segunda y tercera son equivalentes.

[24]$_1$ Ejemplos

1.° La función $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = e^{x+y} + 2 \operatorname{sen} (2x - y)$ es diferenciable en el punto (0, 0) y su diferencial es la aplicación lineal

$$(x, y) \mapsto df(0, 0)(x, y) = 5x - y$$

Así ocurre ya que, acudiendo a los desarrollos limitados (de funciones de una variable) $e^t = 1 + t + o(t)$ y $\operatorname{sen} t = t + o(t)$, se puede poner que

$$f(x, y) = 1 + (x + y) + o(x + y) + 2(2x - y) + o(2x - y) =$$
$$= 1 + 5x - y + o(\sqrt{x^2 + y^2}) = f(0, 0) + (5x - y) + o(\sqrt{x^2 + y^2})$$

puesto que los $o(x + y)$ y $o(2x - y)$ son, ambos, unos $o(\sqrt{x^2 + y^2})$. Esta última igualdad prueba que la diferencial es $5x - y$; nótese que el incremento $\boldsymbol{h}$ de la variable y su norma $\|\boldsymbol{h}\|$ son ahora

$$\boldsymbol{h} = (x - 0, y - 0) = (x, y) \qquad \text{y} \qquad \|\boldsymbol{h}\| = \sqrt{x^2 + y^2}$$

2.° La función $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = \sqrt{x^2 + y^2}$ no es diferenciable en el punto (0, 0). En efecto, para cualquiera que sea la aplicación lineal $(x, y) \mapsto \alpha x + \beta y$ (α y β son los coeficientes de tal aplicación; α y β son constantes), resulta que

$$\frac{[f(x, y) - f(0, 0)] - (\alpha x + \beta y)}{\|(x, y)\|} = \frac{\sqrt{x^2 + y^2} - (\alpha x + \beta y)}{\sqrt{x^2 + y^2}}$$

no puede tener límite 0 cuando $(x, y) \to (0, 0)$. En efecto: de tener límite 0, lo tendría en particular según $x = 0$ y según $y = 0$ y ello no es así, ya que:

$$\text{según } x = 0: \qquad \lim_{y \to 0}\left(1 - \alpha \frac{y}{|y|}\right) = \begin{cases} 1, & \text{si } \alpha = 0 \\ \nexists, & \text{si } \alpha \neq 0 \end{cases}$$

$$\text{según } y = 0: \qquad \lim_{x \to 0}\left(1 - \beta \frac{x}{|x|}\right) = \begin{cases} 1, & \text{si } \beta = 0 \\ \nexists, & \text{si } \beta \neq 0 \end{cases}$$

3.° Si $f: C \to \mathbb{R}^q$ es una función, definida en un abierto $C \subset \mathbb{R}^p$, que es diferenciable en un punto $\boldsymbol{a} \in C$ y dado un vector $\boldsymbol{v} \in \mathbb{R}^q$, la función $\varphi: C \to \mathbb{R}$, $\varphi(\boldsymbol{x}) = f(\boldsymbol{x}) \cdot \boldsymbol{v}$ (producto escalar usual en $\mathbb{R}^q$) es diferenciable en $\boldsymbol{a}$ y su diferencial es:

$$d\varphi(\boldsymbol{a}): \boldsymbol{h} \mapsto [df(\boldsymbol{a})(\boldsymbol{h})] \cdot \boldsymbol{v} \tag{1}$$

En efecto: acudiendo a la condición de diferenciabilidad, se tiene:

$$\varphi(\boldsymbol{a} + \boldsymbol{h}) - \varphi(\boldsymbol{a}) = [f(\boldsymbol{a} + \boldsymbol{h}) - f(\boldsymbol{a})] \cdot \boldsymbol{v} = [df(\boldsymbol{a})(\boldsymbol{h}) + \boldsymbol{o}(\boldsymbol{h})] \cdot \boldsymbol{v} = \\ = [df(\boldsymbol{a})(\boldsymbol{h})] \cdot \boldsymbol{v} + o(\boldsymbol{h})$$

como había que comprobar (nótese que la función definida en [1] es lineal).

[24]$_2$ Observación (notación)

Respecto de la notación, cuando se considera la diferencial de una función $\boldsymbol{x} \mapsto z = f(\boldsymbol{x})$ en un cierto punto fijo $\boldsymbol{x}(x_1, ..., x_p)$, es normal representar al punto variable por $\boldsymbol{x} + \Delta\boldsymbol{x}$ cuyas coordenadas son $(x_1 + \Delta x_1, ..., x_p + \Delta x_p)$, con lo que el «incremento» de la variable es $\Delta\boldsymbol{x}(\Delta x_1, ..., \Delta x_p)$. El correspondiente incremento de la función, al que denotaremos abreviadamente poniendo Δz, es $\Delta z = f(\boldsymbol{x} + \Delta\boldsymbol{x}) - f(\boldsymbol{x})$. El hecho de que f sea diferenciable en el punto $\boldsymbol{x}$ significa que Δz difiere «muy poco» de una expresión lineal en $\Delta\boldsymbol{x}$, a la que abreviadamente denotaremos poniendo dz, de manera que

$$\Delta z = dz + o(\|\Delta\boldsymbol{x}\|), \qquad \text{donde } \|\Delta\boldsymbol{x}\| = \sqrt{(\Delta x_1)^2 + \cdots + (\Delta x_p)^2}$$

Nótese que la anterior expresión adolece de imprecisión; debiera completarse poniendo algo como

$$\Delta z(\boldsymbol{x}; \Delta\boldsymbol{x}) = dz(\boldsymbol{x}; \Delta\boldsymbol{x}) + o(\|\Delta\boldsymbol{x}\|)$$

en la que se señalan el punto $\boldsymbol{x}$ y el incremento $\Delta\boldsymbol{x}$ a los que corresponden los Δz y dz. Esta última expresión puede ponerse, según ya sabemos, escribiendo:

$$\Delta z(\boldsymbol{x}; \Delta\boldsymbol{x}) = dz(\boldsymbol{x}; \Delta\boldsymbol{x}) + \varepsilon(\Delta\boldsymbol{x})\|\Delta\boldsymbol{x}\|$$

$$\Delta z(\boldsymbol{x}; \Delta\boldsymbol{x}) = dz(\boldsymbol{x}; \Delta\boldsymbol{x}) + \sum_{i=1}^{p} \varepsilon_i(\Delta\boldsymbol{x})\Delta x_i$$

para ciertos infinitésimos ε y ε_i $(i = 1, 2, ..., p)$ cuando $\Delta\boldsymbol{x} \to \boldsymbol{o}$.

[24]$_3$ Diferenciales de las variables independientes

Consideremos las siguientes p funciones (proyecciones):

$$\mathbb{R}^p \to \mathbb{R}, \qquad \boldsymbol{x} \mapsto p_i(\boldsymbol{x}) = x_i \qquad (\text{para } i = 1, 2, \ldots, p)$$

donde x_i es la coordenada i-ésima del punto $\boldsymbol{x} \in \mathbb{R}^p$. Estas funciones son diferenciables en todo punto $\boldsymbol{x}$, ya que sus incrementos, cuando la variable pasa de ser $\boldsymbol{x}$ a ser $\boldsymbol{x} + \Delta\boldsymbol{x}$ (llamaremos Δx_i a la coordenada i-ésima de $\Delta\boldsymbol{x}$), se pueden poner

$$p_i(\boldsymbol{x} + \Delta\boldsymbol{x}) - p_i(\boldsymbol{x}) = \Delta x_i = \Delta x_i + o(\|\Delta x_i\|)$$

con lo que su diferencial (para el $\Delta\boldsymbol{x}$ dado) vale Δx_i. A esta diferencial, es decir, a la diferencial de $\boldsymbol{x} \mapsto x_i$ en el punto $\boldsymbol{x}$, como ha resultado ser independiente de $\boldsymbol{x}$, se la denota poniendo dx_i; ello conduce a poner:

$$dx_i = \Delta x_i \qquad (\text{para } i = 1, 2, \ldots, p)$$

y se dice que la diferencial de la variable independiente x_i es igual a su incremento. Nótese que este resultado (el incremento es igual a la diferencial) no se verifica cuando se considera una función cualquiera (diferenciable) $\boldsymbol{x} \mapsto z$; aquí lo que ocurre es que $\Delta z - dz = o(\|\Delta\boldsymbol{x}\|)$ y este último infinitésimo no es nulo, en general.

PRIMERAS PROPIEDADES DE LA DIFERENCIABILIDAD

[25]

I. Sea $f: C \to \mathbb{R}^q$ una función, definida en un abierto $C \subset \mathbb{R}^p$, y considérese un punto $\boldsymbol{a} \in C$.

1.º Si f es diferenciable en $\boldsymbol{a}$, entonces f tiene una única diferencial en $\boldsymbol{a}$.

2.º La función f es diferenciable en $\boldsymbol{a}$ si, y sólo si, son diferenciables en $\boldsymbol{a}$ sus q componentes y, si dichas componentes son $f_1, f_2, \ldots, f_q$ (funciones de C en $\mathbb{R}$), entonces las componentes de la diferencial $df(\boldsymbol{a})$ son las diferenciales $df_i(\boldsymbol{a})$, para $i = 1, 2, \ldots, p$, es decir:

$$f = (f_1, f_2, \ldots, f_p) \quad \Rightarrow \quad df(\boldsymbol{a}) = (df_1(\boldsymbol{a}), df_2(\boldsymbol{a}), \ldots, df_p(\boldsymbol{a}))$$

3.º Si f es diferenciable en $\boldsymbol{a}$, entonces f es continua en $\boldsymbol{a}$. El recíproco es falso (en general).

II. Si f y g son funciones (de un abierto $C \subset \mathbb{R}^p$ en $\mathbb{R}^q$) diferenciables en un punto $\boldsymbol{a} \in C$, entonces también son diferenciables en $\boldsymbol{a}$ las funciones λf (para $\lambda \in \mathbb{R}$), $f + g$ y fg (esta última en el caso real, es decir, si es $q = 1$); se verifica entonces que:

$$d(\lambda f)(\boldsymbol{a}) = \lambda df(\boldsymbol{a})$$
$$d(f + g)(\boldsymbol{a}) = df(\boldsymbol{a}) + dg(\boldsymbol{a})$$
$$d(fg)(\boldsymbol{a}) = [df(\boldsymbol{a})]g(\boldsymbol{a}) + f(\boldsymbol{a})[dg(\boldsymbol{a})]$$

Demostración

1.º Supongamos que hay dos aplicaciones lineales, de $\mathbb{R}^p$ en $\mathbb{R}^q$, que son diferenciales de f en $\boldsymbol{a}$; llamando φ y ψ a dichas aplicaciones, se verifica que

$$f(a+h)-f(a)=\varphi(h)+o(\|h\|) \qquad \text{y} \qquad f(a+h)-f(a)=\psi(h)+o(\|h\|)$$

lo que permite poner que

$$\lim_{h\to o}\frac{(\varphi-\psi)(h)}{\|h\|}=\lim_{h\to o}\frac{[f(a+h)-f(a)]-\psi(h)}{\|h\|}-\lim_{h\to o}\frac{[f(a+h)-f(a)]-\varphi(h)}{\|h\|}=$$
$$=\lim_{h\to o}\frac{o(\|h\|)}{\|h\|}-\lim_{h\to o}\frac{o(\|h\|)}{\|h\|}=o-o=o$$

Para cualquiera que sea el vector no nulo $\boldsymbol{u}\in\mathbb{R}^p$, tomando en la relación anterior $\boldsymbol{h}=\lambda\boldsymbol{u}$ (para $\lambda\in\mathbb{R}$) y haciendo que λ tienda a 0, como φ y ψ son lineales, se obtiene que:

$$o=\lim_{\lambda\to 0}\frac{(\varphi-\psi)(\lambda u)}{\|\lambda u\|}=\lim_{\lambda\to 0}\frac{\lambda[(\varphi-\psi)(u)]}{|\lambda|\,\|u\|}=\frac{(\varphi-\psi)(u)}{\|u\|}\lim_{\lambda\to 0}\frac{\lambda}{|\lambda|}$$

de donde resulta que $(\varphi-\psi)(\boldsymbol{u})=\boldsymbol{o}$. Como esto ocurre para todo vector $\boldsymbol{u}$, acontece que $\varphi-\psi=\boldsymbol{o}$, como había que probar.

2.º La función f es diferenciable en $\boldsymbol{a}$ y su diferencial es $df(\boldsymbol{a})$ si, y sólo si, $df(\boldsymbol{a}):\mathbb{R}^p\to\mathbb{R}^q$ es una aplicación lineal tal que:

$$\lim_{h\to o}\frac{[f(a+h)-f(a)]-df(a)(h)}{\|h\|}=o$$

Ahora bien, llamando $[df(\boldsymbol{a})]_i$ a la componente i-ésima (para $i=1, 2, ..., p$) de $df(\boldsymbol{a})$, la anterior igualdad equivale (véase [12],5) a que sea:

$$\lim_{h\to o}\frac{[f_i(a+h)-f_i(a)]-[df(a)]_i(h)}{\|h\|}=0, \qquad \text{para } i=1, 2, ..., q$$

lo que significa que (para $i=1, 2, ..., q$) las funciones f_i son diferenciables en $\boldsymbol{a}$ y sus diferenciales son $df_i(\boldsymbol{a})=[df(\boldsymbol{a})]_i$, como había que comprobar.

3.º Si f es diferenciable en $\boldsymbol{a}$, entonces se puede poner

$$f(x)=f(a)+df(a)(x-a)+o(\|x-a\|)$$

Como $df(\boldsymbol{a})$ es continua, pues es lineal (véase el ejercicio $[15]_3$), tomando límites para $\boldsymbol{x}\to\boldsymbol{a}$ en la anterior igualdad, se tiene:

$$\lim_{x\to a}f(x)=f(a)+df(a)(o)+o=f(a)+o=f(a)$$

o sea, f es efectivamente continua en el punto $\boldsymbol{a}$.

Para comprobar que el recíproco es falso, basta con acudir a un contraejemplo: la función del ejemplo segundo de $[24]_1$ no es diferenciable en (0, 0), como ya se vio, lo que no impide que sea continua, como se comprueba trivialmente.

II. Comprobemos que, para cada una de las funciones λf, $f+g$ y fg, se cumple la condición de diferenciabilidad, siendo la correspondiente diferencial la que se señala en el enunciado:

[Para λf]:

$$\lim_{h\to o} \frac{[(\lambda f)(a+h)-(\lambda f)(a)]-\lambda df(a)(h)}{\|h\|} =$$
$$= \lambda \lim_{h\to o} \frac{[f(a+h)-f(a)]-df(a)(h)}{\|h\|} = \lambda o = o$$

[Para $f+g$]:

$$\lim_{h\to o} \frac{[(f+g)(a+h)-(f+g)(a)]-[df(a)](h)}{\|h\|} =$$
$$= \lim_{h\to o}\left(\frac{[f(a+h)-f(a)]-df(a)(h)}{\|h\|}+\frac{[g(a+h)-g(a)]-dg(a)(h)}{\|h\|}\right) = o+o=o$$

[Para fg]:

$$\lim_{h\to o} \frac{[(fg)(a+h)-(fg)(a)]-\{[df(a)]g(a)+f(a)[dg(a)]\}(h)}{\|h\|} =$$
$$= \lim_{h\to o}\left(\frac{[f(a+h)-f(a)]-df(a)(h)}{\|h\|}\,g(a)+f(a)\,\frac{[g(a+h)-g(a)]-dg(a)(h)}{\|h\|}\right)+$$
$$+\lim_{h\to o}\frac{[f(a+h)-f(a)][g(a+h)-g(a)]}{\|h\|} = o+o+\lim_{h\to o}\frac{[df(a)(h)][dg(a)(h)]}{\|h\|} =$$
$$= \lim_{h\to o}\left\{df(a)\left(\frac{h}{\|h\|}\right)\right\}\left\{dg(a)\left(\frac{h}{\|h\|}\right)\right\}\|h\| = o$$

(el último limite es nulo ya que: $h/\|h\|$ es un vector de la esfera de radio unidad (conjunto compacto) y $df(a)$ y $dg(a)$ son continuas (pues son lineales), luego las dos expresiones encerradas entre llaves están acotadas, véase [18]).

$[25]_1$ Ejercicio

Sea $f: C \to \mathbb{R}$ una función real, definida en un abierto $C \subset \mathbb{R}^p$, que no sea nula. Si f es diferenciable en un punto $a \in C$, pruébese que entonces también es diferenciable en a la función $1/f$ y se verifica que

$$d\left(\frac{1}{f}\right)(a) = -\frac{1}{f(a)^2}\,df(a)$$

Comprobación

$$\lim_{h \to o} \frac{[(1/f)(a+h) - (1/f)(a)] + [1/f(a)^2]\{[df(a)](h)\}}{\|h\|} =$$

$$\lim_{h \to o} \left\{ -\frac{[f(a+h) - f(a)] - [df(a)](h)}{f(a)f(a+h)\|h\|} + \frac{[f(a+h) - f(a)]\{[df(a)](h)\}}{f(a)^2 f(a+h)\|h\|} \right\} =$$

$$= -\frac{1}{f(a)^2} \lim_{h \to o} \frac{o(\|h\|)}{\|h\|} + \frac{1}{f(a)^3} \lim_{h \to o} \frac{\{[df(a)](h) + o(\|h\|)\}\{[df(a)](h)\}}{\|h\|} =$$

$$= o + \frac{1}{f(a)^3} \lim_{h \to o} \left\{[df(a)]\left(\frac{h}{\|h\|}\right) + \frac{o(\|h\|)}{\|h\|}\right\}\|h\|\left\{[df(a)]\left(\frac{h}{\|h\|}\right)\right\} = o$$

(obsérvese que $h/\|h\|$ es un vector de la esfera de radio unidad (conjunto compacto) y que $df(a)$ es una función continua en $\mathbb{R}^p$ (pues es lineal), luego las dos expresiones encerradas entre llaves en la última igualdad están acotadas, véase [18]).

[25]$_2$ Ejercicio

Sea $f: \mathbb{R}^2 \to \mathbb{R}$ la función definida por las siguientes expresiones:

$$f(x, y) = \frac{3x^2y - 2x^3}{x^2 + y^4}, \qquad \text{si } (x, y) \neq (0, 0); \qquad f(0, 0) = 0$$

1.° Compruébese que f admite derivada en (0, 0) según cualquier vector unitario $u =$ $= (\cos\theta, \operatorname{sen}\theta)$.

2.° Acudiendo a la definición, véase que f no es diferenciable en (0, 0).

Resolución

1.°

$$\frac{\partial f(0,0)}{\partial u} = \lim_{\lambda \to 0} \frac{f(\lambda\cos\theta,\ \lambda\operatorname{sen}\theta) - f(0,0)}{\lambda} = \lim_{\lambda \to 0} \frac{2\cos^3\theta\operatorname{sen}\theta - 2\cos^3\theta}{\cos^2\theta + \lambda^2\operatorname{sen}^4\theta} =$$

$$= \begin{cases} 3\operatorname{sen}\theta - 2\cos\theta, & \text{si } \cos\theta \neq 0 \\ 0, & \text{si } \cos\theta = 0 \end{cases}$$

2.° Para que f sea diferenciable en (0, 0) debe existir una expresión $\alpha x + \beta y$ (lineal en x e y) tal que

$$0 = \lim_{\substack{x \to 0 \\ y \to 0}} \frac{[f(x, y) - f(0, 0)] - [\alpha x + \beta y]}{\sqrt{x^2 + y^2}} = \lim_{\substack{x \to 0 \\ y \to 0}} \frac{3x^2y - 2x^3 - (\alpha x + \beta y)(x^2 + y^4)}{(x^2 + y^4)\sqrt{x^2 + y^2}} \qquad [1]$$

En particular, este límite debería ser 0 según las dos direcciones $x = 0$ e $y = 0$; obligando a que ello se cumpla, se obtiene que debiera ser $\alpha = -2$ y $\beta = 0$. Llevando estos valores a [1], resulta que f sería diferenciable en (0, 0) si

$$\lim_{\substack{x \to 0 \\ y \to 0}} \frac{xy(3x + 2y^3)}{(x^2 + y^4)\sqrt{x^2 + y^2}} \qquad \text{fuese } 0$$

Ello no es así, ya que según la dirección del vector $\boldsymbol{u} = (\cos\theta, \operatorname{sen}\theta)$ dicho límite es

$$\lim_{t \to 0} \frac{t}{|t|} \frac{\cos\theta \operatorname{sen}\theta(3\cos\theta + 2t^2 \operatorname{sen}^3\theta)}{(\cos^2\theta + t^2 \operatorname{sen}^4\theta)}$$

que no es 0 (salvo si $\operatorname{sen}\theta$ o $\cos\theta$ son nulos).

DERIVADAS DE LAS FUNCIONES DIFERENCIABLES

[26]

Sea $f: C \to \mathbb{R}^q$ una función, definida en un abierto $C \subset \mathbb{R}^p$, y considérese un punto $\boldsymbol{a} \in C$. *Si* f es diferenciable en $\boldsymbol{a}$, se verifica que:

1.º Entonces f es derivable en $\boldsymbol{a}$ según cualquier vector no nulo $\boldsymbol{u} \in \mathbb{R}^p$ y, en particular, f admite sus p derivadas parciales en $\boldsymbol{a}$; dichas derivadas son:

$$D_{\boldsymbol{u}}[f(\boldsymbol{a})] \equiv f'_{\boldsymbol{u}}(\boldsymbol{a}) = df(\boldsymbol{a})(\boldsymbol{u}) \qquad \text{y} \qquad D_i[f(\boldsymbol{a})] \equiv f'_i(\boldsymbol{a}) = df(\boldsymbol{a})(\boldsymbol{e}_i)$$

(donde $(\boldsymbol{e}_1, \boldsymbol{e}_2, ..., \boldsymbol{e}_p)$ es la base canónica de $\mathbb{R}^p$; para $i = 1, 2, ..., p$).

2.º Entonces, acudiendo a las derivadas parciales, se puede poner:

$$df(\boldsymbol{a})(\boldsymbol{h}) = D_1[f(\boldsymbol{a})]h_1 + \cdots + D_p[f(\boldsymbol{a})]h_p, \qquad \forall\, \boldsymbol{h} = (h_1, ..., h_p) \in \mathbb{R}^p$$
$$D_{\boldsymbol{u}}[f(\boldsymbol{a})] = D_1[f(\boldsymbol{a})]u_1 + \cdots + D_p[f(\boldsymbol{a})]u_p, \qquad \forall\, \boldsymbol{u} = (u_i) \in \mathbb{R}^p,\ \boldsymbol{u} \neq \boldsymbol{o}$$

3.º Si las componentes de f son $f_1, ..., f_q$, la matriz asociada a $df(\boldsymbol{a}): \mathbb{R}^p \to \mathbb{R}^q$, que se llama *matriz jacobiana* de f en $\boldsymbol{a}$ y se denota poniendo $Jf(\boldsymbol{a})$, tiene por elemento de lugar i, j a la derivada $D_j[f_i(\boldsymbol{a})]$ (para mayor detalle, véase $[26]_5$).

Demostración

1.º Acudiendo a la definición de derivada según un vector, como f es diferenciable en $\boldsymbol{a}$ y $df(\boldsymbol{a})$ es lineal, podemos poner:

$$D_{\boldsymbol{u}}[f(\boldsymbol{a})] = \lim_{\lambda \to 0} \frac{f(\boldsymbol{a} + \lambda\boldsymbol{u}) - f(\boldsymbol{a})}{\lambda} = \lim_{\lambda \to 0} \frac{df(\boldsymbol{a})(\lambda\boldsymbol{u}) + o(\|\lambda\boldsymbol{u}\|)}{\lambda} =$$
$$= \lim_{\lambda \to 0} \frac{\lambda df(\boldsymbol{a})(\boldsymbol{u})}{\lambda} + o = df(\boldsymbol{a})(\boldsymbol{u})$$

Repitiendo lo anterior para $\boldsymbol{u} = \boldsymbol{e}_i$, se obtiene la expresión correspondiente a la derivada parcial $D_i[f(\boldsymbol{a})]$.

2.° Acudiendo a que $df(\boldsymbol{a})$ es una función lineal y poniendo $\boldsymbol{h} = \Sigma h_i \boldsymbol{e}_i$ (la suma para $i = 1, ..., p$), se obtiene que

$$df(\boldsymbol{a})(\boldsymbol{h}) = df(\boldsymbol{a})(\Sigma h_i \boldsymbol{e}_i) = \Sigma h_i(df(\boldsymbol{a})(\boldsymbol{e}_i)) = \Sigma h_i D_i[f(\boldsymbol{a})]$$

Echando mano de este resultado y de la expresión de $D_{\boldsymbol{u}}[f(\boldsymbol{a})]$, ya obtenida, resulta finalmente que:

$$D_{\boldsymbol{u}}[f(\boldsymbol{a})] = df(\boldsymbol{a})(\boldsymbol{u}) = \Sigma u_i D_i[f(\boldsymbol{a})]$$

3.° Este apartado se analiza con detalle en $[26]_5$.

$[26]_1$ Observación

De acuerdo con lo que se acaba de obtener, si una función f es diferenciable en un punto $\boldsymbol{a}$, entonces f admite derivada $D_{\boldsymbol{u}}[f(\boldsymbol{a})]$, según cualquier vector $\boldsymbol{u} \neq \boldsymbol{o}$, admite en particular las derivadas parciales $D_i[f(\boldsymbol{a})]$ y entre todas ellas se verifica la relación:

$$D_{\boldsymbol{u}}[f(\boldsymbol{a})] = \sum_{i=1}^{p} u_i D_i[f(\boldsymbol{a})] \qquad [1]$$

Conviene notar que el recíproco es falso. No sólo pueden existir todas las derivadas, de f en $\boldsymbol{a}$, sin que f sea diferenciable en $\boldsymbol{a}$ (como puso de manifiesto el anterior ejercicio $[25]_2$), sino que la no diferenciabilidad puede darse también aún en el caso de que, existiendo todas las derivadas, de f en $\boldsymbol{a}$, éstas verifiquen a la anterior relación [1]. Considérese, en efecto, el siguiente ejemplo:

Sea $f: \mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto f(x, y)$, cualquier función diferenciable en $(0,0)$, con lo que admite en $(0,0)$ derivada según todo vector no nulo (u, v) y, en particular, admite las derivadas parciales, respecto de x y respecto de y, verificándose que:

$$\frac{\partial f(0,0)}{\partial(u,v)} = u\,\frac{\partial f(0,0)}{\partial x} + v\,\frac{\partial f(0,0)}{\partial y} \qquad [2]$$

Sea $\varphi: \mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto \varphi(x, y)$, la función definida, a partir de f, mediante

$$\begin{cases} \varphi(x,y) = f(x,y), & \text{si } y \neq x^2 \\ \varphi(x,x^2) = f(x,x^2) + 1, & \text{si } x \neq 0 \\ \varphi(0,0) = f(0,0) \end{cases}$$

Esta función φ no es continua en $(0, 0)$, ya que según la curva $y = x^2$ tiene, en $(0, 0)$, límite $f(0, 0) + 1$ y éste es distinto de $\varphi(0, 0) = f(0, 0)$. Por tanto, φ no es diferenciable en $(0, 0)$. Ahora bien, para cualquiera que sea el vector no nulo (u, v) y en particular para los vectores $(1, 0)$ y $(0, 1)$, se verifica que

$$\frac{\partial\varphi(0,0)}{\partial(u,v)} = \frac{\partial f(0,0)}{\partial(u,v)}, \qquad \frac{\partial\varphi(0,0)}{\partial x} = \frac{\partial f(0,0)}{\partial x}, \qquad \frac{\partial\varphi(0,0)}{\partial y} = \frac{\partial f(0,0)}{\partial y} \qquad [3]$$

Así ocurre debido a que según cualquier recta que pase por (0, 0) las funciones φ y f coinciden salvo, a lo sumo, en un solo punto distinto del (0, 0). Como consecuencia de [2] y [3] se verifica, pues, que

$$\frac{\partial\varphi(0,0)}{\partial(u,v)} = u\,\frac{\partial\varphi(0,0)}{\partial x} + v\,\frac{\partial\varphi(0,0)}{\partial y}$$

y ello sin necesidad de que φ sea diferenciable en (0, 0), que no lo es.

[26]$_2$ Observación

A la vista del anterior resultado [26], 2.°, la condición de diferenciabilidad [24], 1.ª se puede expresar de la forma:

Una función $f: C \to \mathbb{R}^q$, $\boldsymbol{x} \mapsto f(\boldsymbol{x})$, definida en un abierto $C \subset \mathbb{R}^p$, será diferenciable en un punto $\boldsymbol{x} \in C$ si, admitiendo todas sus p derivadas parciales $\partial f(\boldsymbol{x})/\partial x_i$ (respecto de sus p variables $x_1, x_2, \ldots, x_p$), se verifica que

$$f(\boldsymbol{x} + \Delta\boldsymbol{x}) - f(\boldsymbol{x}) = \frac{\partial f(\boldsymbol{x})}{\partial x_1}\,\Delta x_1 + \frac{\partial f(\boldsymbol{x})}{\partial x_2}\,\Delta x_2 + \cdots + \frac{\partial f(\boldsymbol{x})}{\partial x_p}\,\Delta x_p + \boldsymbol{o}(\|\Delta\boldsymbol{x}\|)$$

(donde $\Delta\boldsymbol{x} = (\Delta x_1, \Delta x_2, \ldots, \Delta x_p)$; nótese que $\|\Delta\boldsymbol{x}\| = \sqrt{(\Delta x_1)^2 + (\Delta x_2)^2 + \cdots + (\Delta x_p)^2}$).

[26]$_3$ Vector gradiente (de una función real)

Si $f: C \to \mathbb{R}$ es una función real, definida en un abierto $C \subset \mathbb{R}^p$, que es diferenciable en un punto $\boldsymbol{a} \in C$, se llama *gradiente* de f en $\boldsymbol{a}$ al vector de $\mathbb{R}^p$ que tiene por componentes a las p derivadas parciales de f en $\boldsymbol{a}$. Dicho gradiente se representa por grad $f(\boldsymbol{a})$ o $\nabla f(\boldsymbol{a})$, con lo que:

$$\nabla f(\boldsymbol{a}) = \left(f_1'(\boldsymbol{a}), f_2'(\boldsymbol{a}), \ldots, f_p'(\boldsymbol{a})\right) = \sum_{i=1}^{p} f_i'(\boldsymbol{a})\boldsymbol{e}_i$$

siendo $(\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_p)$ la base canónica de $\mathbb{R}^p$. Acudiendo al vector gradiente $\nabla f(\boldsymbol{a})$, las dos igualdades [26], 2.°, esto es, las relaciones $df(\boldsymbol{a})(\boldsymbol{h}) = \Sigma f_i'(\boldsymbol{a})h_i$ y $\partial f(\boldsymbol{a})/\partial\boldsymbol{u} = \Sigma f_i'(\boldsymbol{a})u_i$, se pueden expresar, utilizando el producto escalar usual(*) de $\mathbb{R}^p$, poniendo

$$df(\boldsymbol{a})(\boldsymbol{h}) = \nabla f(\boldsymbol{a})\cdot\boldsymbol{h} \qquad \text{y} \qquad \frac{\partial f(\boldsymbol{a})}{\partial\boldsymbol{u}} = \nabla f(\boldsymbol{a})\cdot\boldsymbol{u}$$

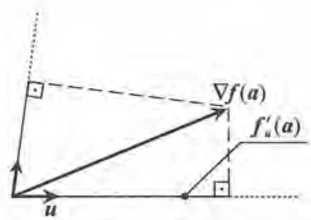
Nótese que si, en la última igualdad, se toma un vector $\boldsymbol{u} \in \mathbb{R}^p$ unitario, resulta que $\nabla f(\boldsymbol{a})\cdot\boldsymbol{u}$ es la proyección del vector gradiente sobre $\boldsymbol{u}$. Por tanto, la derivada de f en $\boldsymbol{a}$ según una dirección cualquiera (orientada) es igual a la proyección del gradiente sobre ella.

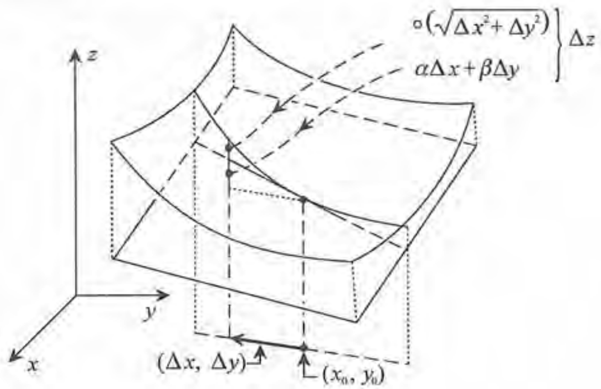
(*) El producto escalar usual de $\mathbb{R}^p$ queda definido por la relación

$$(u_1, u_2, \ldots, u_p)\cdot(v_1, v_2, \ldots, v_p) = u_1v_1 + u_2v_2 + \cdots + u_pv_p$$

En particular, la dirección (orientada) del vector gradiente es aquella según la que f tiene mayor derivada en $\boldsymbol{a}$; la función f tiene derivada nula, en $\boldsymbol{a}$, según las direcciones perpendiculares al gradiente $\nabla f(\boldsymbol{a})$.

Conviene insistir en que todo lo aquí dicho tiene sentido en el supuesto de que f sea diferenciable en $\boldsymbol{a}$. Si f no es diferenciable en $\boldsymbol{a}$ pero admite las p derivadas parciales $f_i'(\boldsymbol{a})$, entonces, aun cuando es posible construir el vector que tiene por componentes a las referidas derivadas, dicho vector no gozará de las anteriores propiedades del gradiente, por lo que no sería conveniente llamarle de tal modo.





[26]₄ Plano tangente

Consideremos aquí una función real de dos variables reales (definida en un abierto $C \subset \mathbb{R}^2$):

$$f: C \to \mathbb{R}, \qquad (x, y) \mapsto z = f(x, y)$$

Vamos a estudiar dicha función en las cercanías de un punto $(x_0, y_0) \in C$. A este respecto, comencemos señalando que:

- La función f es diferenciable en (x_0, y_0) si, según ya sabemos, existen dos constantes $\alpha, \beta \in \mathbb{R}$ (que son las derivadas parciales $\alpha = \partial f(x_0, y_0)/\partial x$ y $\beta = \partial f(x_0, y_0)/\partial y$) tales que:

$$f(x, y) - f(x_0, y_0) = \alpha(x - x_0) + \beta(y - y_0) + o(\rho), \qquad \text{donde } \rho = \sqrt{(x - x_0)^2 + (y - y_0)^2} \qquad [1]$$

- Se dice que la superficie de ecuación $z = f(x, y)$ tiene a un cierto plano π como plano tangente en el punto $P_0(x_0, y_0, z_0)$ de ella, donde $z_0 = f(x_0, y_0)$, si el ángulo que forma con dicho plano una secante P_0P, donde $P(x, y, z)$ con $z = f(x, y)$, tiende a cero cuando $(x, y) \to (x_0, y_0)$.

Pues bien, se verifica que la función f es diferenciable en (x_0, y_0) si, y sólo si, la superficie $z = f(x, y)$ tiene plano tangente en (x_0, y_0, z_0); además, acontece que

$$\begin{pmatrix} (\Delta x, \Delta y) \to \alpha\Delta x + \beta\Delta y \\ \text{es la diferencial de } f \text{ en } (x_0, y_0) \end{pmatrix} \Leftrightarrow \begin{pmatrix} \text{El plano } z - z_0 = \alpha(x - x_0) + \beta(y - y_0) \\ \text{es tangente a } z = f(x, y) \text{ en } (x_0, y_0, z_0) \end{pmatrix}$$

Vamos a comprobar que la diferenciabilidad implica la existencia de plano tangente; la implicación en el otro sentido puede construirse fácilmente a partir de la anterior. Así pues, sabiendo que se verifica la condición [1], veamos que $z = f(x, y)$ tiene por plano tangente al antes mencionado. Como el vector $(\alpha, \beta, -1)$ es perpendicular al susodicho plano, el ángulo $\theta(x, y)$ que él forma con la secante P_0P, que tiene la dirección del vector $(x - x_0, y - y_0, z - z_0)$, será tal que

$$\operatorname{sen}\,[\theta(x, y)] = \frac{\alpha(x - x_0) + \beta(y - y_0) - (z - z_0)}{\sqrt{\alpha^2 + \beta^2 + 1}\,\sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2}}$$

Acudiendo a la condición [1], de lo anterior se deduce que

$$|\operatorname{sen}\,[\theta(x, x)]| \leqslant \frac{|\alpha(x - x_0) + \beta(y - y_0) - (z - z_0)|}{\sqrt{1}\,\sqrt{(\Delta x)^2 + (\Delta y)^2}} = \frac{o(\rho)}{\rho}$$

Por tanto, de acuerdo con la regla del emparedado (véase $[12]_3$,3) se concluye que, como $o(\rho)/\rho$ tiene límite 0 cuando $(x, y) \to (x_0, y_0)$, también lo tiene $\theta(x, y)$, como había que comprobar. Nótese que el plano tangente del que estamos hablando no es otro que el

$$z - f(x_0, y_0) = \frac{\partial f(x_0, y_0)}{\partial x}(x - x_0) + \frac{\partial f(x_0, y_0)}{\partial y}(y - y_0)$$

$[26]_5$ Matriz jacobiana

Considérese una función $\boldsymbol{f}: C \to \mathbb{R}^q$, $\boldsymbol{x} \mapsto z = \boldsymbol{f}(\boldsymbol{x})$, definida en un abierto $C \subset \mathbb{R}^p$; sean $f_1, f_2, \ldots, f_q$ las componentes (funciones de C en $\mathbb{R}$) de $\boldsymbol{f}$. Si $\boldsymbol{f}$ es diferenciable en un punto $\boldsymbol{a} \in C$, entonces su diferencial $d\boldsymbol{f}(\boldsymbol{a}): \mathbb{R}^p \to \mathbb{R}^q$ tiene por ecuación a:

$$\Delta \boldsymbol{x} = (\Delta x_1, \Delta x_2, \ldots, \Delta x_n) \mapsto df(\boldsymbol{a})(\Delta \boldsymbol{x}) = (dz_1, dz_2, \ldots, dz_n)$$

siendo

$$\begin{bmatrix} dz_1 \\ dz_2 \\ \vdots \\ dz_q \end{bmatrix} = \begin{bmatrix} D_1[f_1(\boldsymbol{a})] & D_2[f_1(\boldsymbol{a})] & \cdots & D_p[f_1(\boldsymbol{a})] \\ D_1[f_2(\boldsymbol{a})] & D_2[f_2(\boldsymbol{a})] & \cdots & D_p[f_2(\boldsymbol{a})] \\ \cdots & \cdots & \cdots & \cdots \\ D_1[f_q(\boldsymbol{a})] & D_2[f_q(\boldsymbol{a})] & \cdots & D_p[f_q(\boldsymbol{a})] \end{bmatrix} \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \\ \vdots \\ \Delta x_p \end{bmatrix}$$

es decir, la matriz de la aplicación lineal $d\boldsymbol{f}(\boldsymbol{a}): \mathbb{R}^p \to \mathbb{R}^q$, respecto de las bases canónicas de $\mathbb{R}^p$ y $\mathbb{R}^q$, es la anterior matriz de tamaño $q \times p$, que recibe el nombre de *matriz jacobiana*(*) de $\boldsymbol{f}$ en $\boldsymbol{a}$ y a la que se denotará poniendo $J\boldsymbol{f}(\boldsymbol{a})$.

(*) Cuando es $p = q$, es decir, si la matriz jacobiana $J\boldsymbol{f}(\boldsymbol{a})$ es cuadrada, entonces a su determinante se le llama *jacobiano* de $\boldsymbol{f}$ en $\boldsymbol{a}$ y se le suele denotar poniendo

$$\det\,[J\boldsymbol{f}(\boldsymbol{a})] \qquad \text{o} \qquad \frac{\partial(f_1, f_2, \ldots, f_p)}{\partial(x_1, x_2, \ldots, x_p)}(\boldsymbol{a}),$$

donde $(x_1, x_2, \ldots, x_p)$ son las p variables o coordenadas de un punto genérico $\boldsymbol{x} \in C$.

Comprobación

La componente j-ésima de $df(\boldsymbol{a})$ es (véase [25], I,2.°) $df_i(\boldsymbol{a})$, con lo que la componente j-ésima de $df(\boldsymbol{a})(\Delta \boldsymbol{x})$, que hemos denotado por dz_j, es $dz_j = df_j(\boldsymbol{a})(\Delta \boldsymbol{x})$; la expresión [26], 2.° para la diferencial nos permite poner:

$$dz_j = df_i(\boldsymbol{a})(\Delta \boldsymbol{x}) = \sum_{i=1}^{p} D_i[f_j(\boldsymbol{a})]\Delta x_i$$

que al tomar $j = 1, 2, \ldots, q$ nos conduce a que, en efecto, $Jf(\boldsymbol{a})$ es la matriz de la aplicación lineal $df(\boldsymbol{a})$.

[26]$_6$ Ejercicio

Sabiendo que la siguiente función $(x, y, z) \mapsto (u, v)$ es diferenciable en el punto $(x, y, z) = (1, -1, 2)$, hallar su diferencial en dicho punto así como su derivada, en él, según el vector $\boldsymbol{a} = (2, 3, 1)$:

$$u = xy^2e^{3x+y-z} + 2z \qquad v = z^2 \text{ L } (\sqrt{x} + y^2 + 1)$$

Resolución

Las derivadas parciales de u y v son

$$u'_x = y^2(1 + 3x)e^{3x+y-z} \qquad v'_z = \frac{z^2}{2(\sqrt{x} + y^2 - 1)\sqrt{x}}$$

$$u'_y = xy(2 + y)e^{3x+y-z} \qquad v'_y = \frac{2z^2y}{\sqrt{x} + y^2 - 1}$$

$$u'_z = -xy^2e^{3x+y-z} + 2 \qquad v'_z = 2z \text{ L } (\sqrt{x} + y^2 - 1)$$

Por tanto

$$u'_x(1, -1, 2) = 4, \qquad u'_y(1, -1, 2) = -1, \qquad u'_z(1, -1, 2) = 1$$
$$v'_x(1, -1, 2) = 2, \qquad v'_y(1, -1, 2) = -8, \qquad v'_z(1, -1, 2) = 0$$

Así pues, la diferencial de la función dada es la aplicación:

$$\begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} \mapsto \begin{bmatrix} du \\ dv \end{bmatrix} = \begin{bmatrix} 4 & -1 & 1 \\ 2 & -8 & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} = \begin{bmatrix} 4\Delta x - \Delta y + \Delta z \\ 2\Delta x - 8\Delta y \end{bmatrix}$$

La derivada parcial según el vector $\boldsymbol{a}(2, 3, 1)$ será

$$\frac{\partial(u, v)}{\partial \boldsymbol{a}}(1, -1, 2) = (du(\boldsymbol{a}), dv(\boldsymbol{a})) = 4 \times 2 - 3 + 1, \ 2 \times 2 - 8 \times 3) = (4, -20)$$

2.3. PROPIEDADES DE LAS FUNCIONES DIFERENCIABLES

TEOREMA DEL VALOR MEDIO

[27]

Sea $f: C \to \mathbb{R}$ una función real, definida en un abierto $C \subset \mathbb{R}^p$, y sean $\boldsymbol{a}$ y $\boldsymbol{b} = \boldsymbol{a} + \boldsymbol{h}$ dos puntos de C tales que el segmento $[\boldsymbol{a}, \boldsymbol{b}]$, que tiene a $\boldsymbol{a}$ y $\boldsymbol{b}$ por extremos, está incluido en C. Si f es diferenciable en todos los puntos de $[\boldsymbol{a}, \boldsymbol{b}]$, entonces existe un punto $\xi \in]\boldsymbol{a}, \boldsymbol{b}[$ tal que

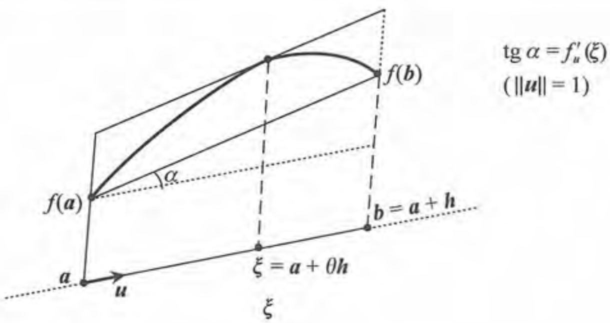
$$f(\boldsymbol{a} + \boldsymbol{h}) - f(\boldsymbol{a}) = df(\xi)(\boldsymbol{h}) \qquad (\xi = \boldsymbol{a} + \theta \boldsymbol{h}, \text{ con } 0 < \theta < 1)$$

Es decir, si $\boldsymbol{a} = (a_1, a_2, \ldots, a_p)$ y $\boldsymbol{b} = (b_1, b_2, \ldots, b_p)$, se verifica que:

$$f(\boldsymbol{b}) - f(\boldsymbol{a}) = f_1'(\xi)(b_1 - a_1) + f_2'(\xi)(b_2 - a_2) + \cdots + f_p'(\xi)(b_p - a_p)$$

LEMA. Sea $f: C \to \mathbb{R}^q$ una función, definida en un abierto $C \subset \mathbb{R}^p$, y sean $\boldsymbol{a}$ y $\boldsymbol{b} = \boldsymbol{a} + \boldsymbol{h}$ dos puntos tales que el segmento $[\boldsymbol{a}, \boldsymbol{b}]$ está incluido en C. Sean $\varphi: [0, 1] \to \mathbb{R}^q$ la función $t \mapsto \varphi(t) = f(\boldsymbol{a} + t\boldsymbol{h})$. Si f es diferenciable en todos los puntos de $[\boldsymbol{a}, \boldsymbol{b}]$, entonces φ es derivable en $[0, 1]$ y su derivada es:

$$\varphi'(t) = df(\boldsymbol{a} + t\boldsymbol{h})(\boldsymbol{h}) \equiv D_h[f(\boldsymbol{a} + t\boldsymbol{h})]$$



Demostración

1.º Empecemos comprobando el lema:

$$\varphi'(t) = \lim_{\lambda \to 0} \frac{\varphi(t + \lambda) - \varphi(t)}{\lambda} = \lim_{\lambda \to 0} \frac{f(\boldsymbol{a} + t\boldsymbol{h} + \lambda\boldsymbol{h}) - f(\boldsymbol{a} + t\boldsymbol{h})}{\lambda} =$$

$$= \lim_{\lambda \to 0} \frac{[df(\boldsymbol{a} + t\boldsymbol{h})](\lambda\boldsymbol{h}) + \boldsymbol{o}(\lambda\boldsymbol{h})}{\lambda} = \lim_{\lambda \to 0} \frac{\lambda[df(\boldsymbol{a} + t\boldsymbol{h})](\boldsymbol{h})}{\lambda} + \boldsymbol{o} =$$

$$= [df(\boldsymbol{a} + t\boldsymbol{h})](\boldsymbol{h}) = D_h[f(\boldsymbol{a} + t\boldsymbol{h})]$$

(la última igualdad es un resultado ya conocido; véase [26],1.º).

2.º Acudamos a la función $\varphi : [0, 1] \to \mathbb{R}$. A ella le es de aplicación en $[0, 1]$ el teorema del valor medio, para funciones reales de una variable real, pues es derivable en $[0, 1]$; por dicho teorema se sabe que existe $\theta \in]0, 1[$ tal que $\varphi(1) - \varphi(0) = \varphi'(\theta)$, esto es, que (según el lema):

$$f(\boldsymbol{a} + \boldsymbol{h}) - f(\boldsymbol{a}) = df(\boldsymbol{a} + \theta\boldsymbol{h})(\boldsymbol{h})$$

El teorema se cumple, pues, para $\xi = \boldsymbol{a} + \theta\boldsymbol{h}$.

[27]₁ Interpretación geométrica

Sea $f: C \to \mathbb{R}$ una función real, definida en un abierto $C \subset \mathbb{R}^2$, y sean $\boldsymbol{a}$, $\boldsymbol{b} = \boldsymbol{a} + \boldsymbol{h}$ dos puntos tales que $[\boldsymbol{a}, \boldsymbol{b}] \subset C$. Se supone que f es diferenciable en todos los puntos del segmento $[\boldsymbol{a}, \boldsymbol{b}]$. Llamando γ a la curva de intersección de la gráfica de f (superficie de $\mathbb{R}^3$) con el plano «vertical» que contiene al segmento $[\boldsymbol{a}, \boldsymbol{b}]$, el teorema del valor medio asegura que hay un punto de la curva γ en el que la tangente es paralela a la cuerda que une sus extremos. En efecto, si es P el punto de γ correspondiente a un punto ξ del segmento $[\boldsymbol{a}, \boldsymbol{b}]$, la tangente en P a γ tiene por pendiente a la derivada $f'_{\boldsymbol{u}}(\xi)$, donde $\boldsymbol{u}$ es el unitario de la dirección de $\boldsymbol{h} = \boldsymbol{b} - \boldsymbol{a}$; si tomamos para ξ el punto del teorema del valor medio, dicha pendiente valdrá

$$f'_{\boldsymbol{u}}(\xi) = df(\xi)(\boldsymbol{u}) = df(\xi)\left(\frac{\boldsymbol{h}}{\|\boldsymbol{h}\|}\right) = \frac{df(\xi)(\boldsymbol{h})}{\|\boldsymbol{h}\|} = \frac{f(\boldsymbol{b}) - f(\boldsymbol{a})}{\|\boldsymbol{h}\|}$$

que es la pendiente de la cuerda que une los extremos de la curva γ.

[27]₂ Observación

El anterior teorema del valor medio, que es de aplicación para funciones reales, no se verifica, en general, cuando se consideran funciones vectoriales; es decir, en las condiciones del teorema [27], para una función $\boldsymbol{f}: C \to \mathbb{R}^q$ puede no existir un punto $\xi \in]\boldsymbol{a}, \boldsymbol{b}[$ para el que sean iguales los vectores:

$$\boldsymbol{f}(\boldsymbol{b}) - \boldsymbol{f}(\boldsymbol{a}) \qquad \text{y} \qquad d\boldsymbol{f}(\xi)(\boldsymbol{b} - \boldsymbol{a})$$

Así ocurre, por ejemplo, con la siguiente función $\boldsymbol{f}: \mathbb{R} \to \mathbb{R}^2, \boldsymbol{f}(x) = (x^2 + x, x^3 + 1)$ en el intervalo $[a, b] = [0, 2]$. En este caso, es:

$$\boldsymbol{f}(b) - \boldsymbol{f}(a) = (6, 9) - (0, 1) = (6, 8)$$
$$d\boldsymbol{f}(\xi)(2) = [(2\xi + 1)2, (3\xi^2)2] = (4\xi + 2, 6\xi^2)$$

y estos dos vectores no pueden ser iguales para ningún valor de ξ, ya que el sistema de ecuaciones $6 = 4\xi + 2$, $8 = 6\xi^2$ es evidentemente incompatible.

No obstante lo anterior, existe un teorema del valor medio para funciones vectoriales, del que pasamos a ocuparnos.

$[27]_3$ Teorema del valor medio (caso vectorial)

Sea $f: C \to \mathbb{R}^q$ una función, definida en un abierto $C \subset \mathbb{R}^p$, y sean $\boldsymbol{a}$ y $\boldsymbol{b} = \boldsymbol{a} + \boldsymbol{h}$ dos puntos de C tales que el segmento $[\boldsymbol{a}, \boldsymbol{b}]$, que tiene a $\boldsymbol{a}$ y $\boldsymbol{b}$ por extremos, está incluido en C. Si f es diferenciable en todos los puntos de $[\boldsymbol{a}, \boldsymbol{b}]$, entonces existe un punto $\xi \in]\boldsymbol{a}, \boldsymbol{b}[$ tal que:

$$\|f(\boldsymbol{a} + \boldsymbol{h}) - f(\boldsymbol{a})\| \leqslant \|df(\xi)(\boldsymbol{h})\| \qquad (\xi = \boldsymbol{a} + \theta \boldsymbol{h}, \text{ con } 0 < \theta < 1)$$

Demostración

Llamemos $\boldsymbol{v} \in \mathbb{R}^q$ al vector $\boldsymbol{v} = f(\boldsymbol{b}) - f(\boldsymbol{a})$ y, acudiendo al producto escalar usual de $\mathbb{R}^q$, consideremos la siguiente función real

$$\varphi: C \to \mathbb{R}, \qquad \varphi(\boldsymbol{x}) = f(\boldsymbol{x}) \cdot \boldsymbol{v}$$

Esta función es diferenciable en todos los puntos del segmento $[\boldsymbol{a}, \boldsymbol{b}]$ (véase el ejercicio $[24]_1$, 3.º), ya que f es diferenciable en $[\boldsymbol{a}, \boldsymbol{b}]$, y se verifica (para $\boldsymbol{x} \in [\boldsymbol{a}, \boldsymbol{b}]$ y $\boldsymbol{u} \in \mathbb{R}^p$) que

$$d\varphi(\boldsymbol{x})(\boldsymbol{u}) = [df(\boldsymbol{x})(\boldsymbol{u})] \cdot \boldsymbol{v}$$

A la función real φ se la puede aplicar, pues, el teorema del valor medio [27], por el que se sabe que existe $\xi \in]\boldsymbol{a}, \boldsymbol{b}[$ tal que

$$\varphi(\boldsymbol{a} + \boldsymbol{h}) - \varphi(\boldsymbol{a}) = d\varphi(\xi)(\boldsymbol{h})$$

o sea

$$[f(\boldsymbol{a} + \boldsymbol{h}) - f(\boldsymbol{a})] \cdot \boldsymbol{v} = [df(\xi)(\boldsymbol{h})] \cdot \boldsymbol{v}$$

o

$$\|f(\boldsymbol{a} + \boldsymbol{h}) - f(\boldsymbol{a})\|^2 = [df(\xi)(\boldsymbol{h})] \cdot [f(\boldsymbol{a} + \boldsymbol{h}) - f(\boldsymbol{a})] \qquad [1]$$

Por otra parte, de acuerdo con la desigualdad de Schwarz del producto escalar, se verifica que

$$[df(\xi)(\boldsymbol{h})] \cdot [f(\boldsymbol{a} + \boldsymbol{h}) - f(\boldsymbol{a})] \leqslant \|df(\xi)(\boldsymbol{h})\| \, \|f(\boldsymbol{a} + \boldsymbol{h}) - f(\boldsymbol{a})\|$$

Llevando este resultado a [1] y dividiendo luego por $\|f(\boldsymbol{a} + \boldsymbol{h}) - f(\boldsymbol{a})\|$(*) se obtiene la conclusión del teorema.

$[27]_4$ Funciones con diferencial nula

Sea $f: C \to \mathbb{R}^q$ una función, definida en un abierto $C \subset \mathbb{R}^p$. Si C es un conjunto convexo(**) y si existe y es nula la diferencial de f en todo punto de C, entonces la función f es constante.

(*) Si fuese $f(\boldsymbol{a} + \boldsymbol{h}) = f(\boldsymbol{a})$, el teorema es una trivialidad: se verifica para todo $\xi \in]\boldsymbol{a}, \boldsymbol{b}[$.

(**) Se dice que C es convexo, si para cualesquiera que sean los puntos $\boldsymbol{a}$, $\boldsymbol{b}$, de C, acontece que el segmento $[\boldsymbol{a}, \boldsymbol{b}]$, que los tiene por extremos, está incluido en C.

Demostración

Vamos a demostrar esta propiedad para el caso $q = 1$, es decir, en el supuesto de que f es una función real; acudiendo a la propiedad [25], 2.º, se obtiene que la propiedad es entonces cierta para $q \in \mathbb{N}$ cualquiera. En lo que sigue, se toma, pues, $q = 1$.

Sean $\boldsymbol{x}$ y $\boldsymbol{x}'$ dos puntos cualesquiera de C. Como C es convexo, también pertenecen a C todos los puntos del segmento $[\boldsymbol{x}, \boldsymbol{x}']$ y nos encontramos, entonces con que a f le es de aplicación el teorema del valor medio [27] entre los puntos $\boldsymbol{x}$ y $\boldsymbol{x}'$. Según dicho teorema, se sabe que existe un punto $\xi \in]\boldsymbol{x}, \boldsymbol{x}'[$ tal que

$$f(\boldsymbol{x}') - f(\boldsymbol{x}) = df(\xi)(\boldsymbol{x}' - \boldsymbol{x})$$

Por ser C convexo, se sabe que $\xi \in C$, luego $df(\xi)$ es nula, por hipótesis, lo que nos permite asegurar que es $f(\boldsymbol{x}') - f(\boldsymbol{x}) = 0$, luego f es constante en C, como había que comprobar.

❑ DIFERENCIABILIDAD DE LAS FUNCIONES DE CLASE $\mathscr{C}^1$

Sabemos que la diferenciabilidad en un punto, de una cierta función, implica que existan sus derivadas parciales en él (véase [26], 1.º). También sabemos que la existencia de todas las derivadas parciales no es suficiente para poder asegurar la diferenciabilidad (véase $[26]_1$). Pues bien, vamos a ver ahora que, si existen las derivadas parciales y ellas son continuas en el punto, entonces la función es diferenciable en él.

[28]

Sea $f: C \to \mathbb{R}^q$ una función, definida en un abierto $C \subset \mathbb{R}^p$.

FUNCIÓN DE CLASE $\mathscr{C}^1$. Se dice que f es de clase $\mathscr{C}^1$ en un punto $\boldsymbol{a} \in C$ si f admite sus p derivadas parciales en un entorno de $\boldsymbol{a}$ y dichas derivadas son continuas en $\boldsymbol{a}$. Se dice que f es de clase $\mathscr{C}^1$ en C si es de clase $\mathscr{C}^1$ en todo punto de C. La función f es de clase $\mathscr{C}^1$ en $\boldsymbol{a}$ si, y sólo si, son de clase $\mathscr{C}^1$ en $\boldsymbol{a}$ todas sus componentes (funciones reales).

TEOREMA. Si f es de clase $\mathscr{C}^1$ en un punto $\boldsymbol{a} \in C$, entonces f es diferenciable en el punto $\boldsymbol{a}$.

GENERALIZACIÓN. Si la función $f: C \to \mathbb{R}^q$ ($C \subset \mathbb{R}^p$ abierto) admite $p - 1$ derivadas parciales en un entorno del punto $\boldsymbol{a} \in C$ y ellas son continuas en $\boldsymbol{a}$, y si admite en $\boldsymbol{a}$ la restante derivada parcial, entonces f es diferenciable en $\boldsymbol{a}$.

Demostración

Antes de ocuparnos del teorema, nótese que la equivalencia entre «f es de clase $\mathscr{C}^1$ en $\boldsymbol{a}$» y «las funciones componentes de f son de clase $\mathscr{C}^1$ en $\boldsymbol{a}$» es una consecuencia de [23], 1.º) y de [16], 1.º). El teorema del enunciado sólo es necesario demostrarle para el caso de ser f una función real ($q = 1$), pues de este caso se desprende el caso general ($q \geqslant 2$) sin más que acudir a [23], 1.º) y a la equivalencia de la que hablábamos en el párrafo anterior. Supondremos, pues, a lo largo de la demostración que f es una función real.

Vamos a probar la «generalización» del teorema, pues ello no nos supondrá un mayor esfuerzo, para lo que procederemos por inducción sobre p (número de variables). Para $p = 1$ la propiedad se verifica, ya que en este caso la hipótesis dice que existe la derivada (única) de f en un entorno del punto y, para las funciones de una variable, la diferenciabilidad en el punto equivale a la existencia de la derivada en él. Suponiendo, pues, que la propiedad es cierta para $p - 1$ ($p \geqslant 2$) variables, hemos de probar que entonces también lo es en el caso de p variables. Sabemos que una, al menos, de las derivadas parciales de f existe en un entorno de $\boldsymbol{a}$ y es continua en $\boldsymbol{a}$; supongamos que dicha derivada es la correspondiente a la primera variable. Llamando $\boldsymbol{a} = (a_1, a_2, ..., a_p)$ y si $\boldsymbol{x} = (x_1, x_2, ..., x_p)$ es un punto genérico de C, hay que comprobar que existen unos ciertos $c_1, c_2, ..., c_p \in \mathbb{R}^q$ tales que:

$$f(\boldsymbol{x}) - f(\boldsymbol{a}) = c_1(x_1 - a_1) + c_2(x_2 - a_2) + \cdots + c_p(x_p - a_p) + o(\|\boldsymbol{x} - \boldsymbol{a}\|) \qquad [1]$$

Para ello, empecemos poniendo

$$f(\boldsymbol{x}) - f(\boldsymbol{a}) = \Delta_1(\boldsymbol{x}) + \Delta_2(\boldsymbol{x}), \qquad \text{donde } \begin{cases} \Delta_1(\boldsymbol{x}) = f(\boldsymbol{x}) - f(\boldsymbol{x}_a) \\ \Delta_2(\boldsymbol{x}) = f(\boldsymbol{x}_a) - f(\boldsymbol{a}) \end{cases} \qquad [2]$$

$$\boldsymbol{x}_a = (a_1, x_2, ..., x_p)$$

Nótese que, como C es abierto, existe un cierto entorno U de $\boldsymbol{a}$ tal que $U \subset C$; para $\boldsymbol{x} \in C$, es $\boldsymbol{x}_a \in C$, por lo que existe $f(\boldsymbol{x}_a)$. Para la función de $p - 1$ variables

$$(x_2, ..., x_p) \mapsto f(\boldsymbol{x}_a) = f(a_1, x_2, ..., x_p)$$

se verifican, en torno del punto $(a_2, ..., a_p)$, las exigencias del teorema; de acuerdo, pues, con nuestra hipótesis de inducción, se puede asegurar que esta función es diferenciable en $(a_2, ..., a_p)$, esto es que:

$$\Delta_2(\boldsymbol{x}) = f(\boldsymbol{x}_a) - f(\boldsymbol{a}) = c_2(x_2 - a_2) + \cdots + c_p(x_p - a_p) + o(\|\boldsymbol{x}_a - \boldsymbol{a}\|) \qquad [3]$$

para ciertos $c_2, ..., c_p \in \mathbb{R}^q$. Nótese que, como $\|\boldsymbol{x}_a - \boldsymbol{a}\| \leqslant \|\boldsymbol{x} - \boldsymbol{a}\|$, el último sumando de la expresión anterior puede ponerse también de la forma $o(\|\boldsymbol{x} - \boldsymbol{a}\|)$.

Para cada $x_2, ..., x_p$, suficientemente próximas a $a_2, ..., a_p$, consideremos la función de la única variable x_1 (definida en un entorno de a_1):

$$x_1 \mapsto f(x_1, x_2, ..., x_p)$$

A esta función le es aplicable el teorema de los incrementos finitos en el intervalo de extremos x_1 y a_1 (para x_1 suficientemente pequeño), ya que ella es derivable en dicho intervalo (cerrado), pues su derivada es la derivada parcial f'_{x_1}. Según dicho teorema, existe ξ_1, comprendido entre x_1 y a_1 (y que dependerá también de quienes sean $x_2, ..., x_p$) tal que:

$$\Delta_1(\boldsymbol{x}) = f(\boldsymbol{x}) - f(\boldsymbol{x}_a) = (x_1 - a_1)f'_{x_1}(\xi_1, x_2, ..., x_p)$$

Ahora bien, como f'_{x_1} es una función continua en $\boldsymbol{a}$, se verifica que

$$f'_{x_1}(\xi_1, x_2, ..., x_p) = f'_{x_1}(\boldsymbol{a}) + \varepsilon(\boldsymbol{x}), \qquad \text{siendo } \lim_{\boldsymbol{x} \to \boldsymbol{a}} \varepsilon(\boldsymbol{x}) = o$$

Llamando $c_1 = f'_{x_1}(\boldsymbol{a})$, al llevar este resultado a la igualdad anterior, se obtiene que

$$\Delta_1(\boldsymbol{x}) = c_1(x_1 - a_1) + \varepsilon(\boldsymbol{x})(x_1 - a_1) = c_1(x_1 - a_1) + o(\|\boldsymbol{x} - \boldsymbol{a}\|) \qquad [4]$$

Llevando [3] y [4] a [2], se obtiene la relación [1], que había que probar.

[28]$_1$ Observaciones

1.° Si una función es de clase $\mathscr{C}^1$ en un punto, entonces es continua en dicho punto; en efecto: según acabamos de comprobar, la función es entonces diferenciable en el punto y, por ello (véase [25], 3.°), es continua en él.

2.° La diferenciabilidad, de una función en un punto, no es en general una condición suficiente para que la función sea de clase $\mathscr{C}^1$ en el punto. Así, por ejemplo, la función $f: \mathbb{R}^2 \to \mathbb{R}$ definida por las expresiones

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

Admite derivadas parciales en todo punto:

$$\frac{\partial f(x, y)}{\partial x} = 2x \text{ sen } \frac{1}{x^2 + y^2} - \frac{2x}{x^2 + y^2} \cos \frac{1}{x^2 + y^2}; \qquad \frac{\partial f(0, 0)}{\partial x} = 0$$

$$\frac{\partial f(x, y)}{\partial y} = 2y \text{ sen } \frac{1}{x^2 + y^2} - \frac{2y}{x^2 + y^2} \cos \frac{1}{x^2 + y^2}; \qquad \frac{\partial f(0, 0)}{\partial y} = 0$$

pero ninguna de ellas es continua en (0, 0), lo que no obsta para que la función sea diferenciable en (0, 0), pues

$$\lim_{\substack{x \to 0 \\ y \to 0}} \frac{[f(x, y) - f(0, 0)] - [x f'_x(0, 0) + y f'_y(0, 0)]}{\sqrt{x^2 + y^2}} = \lim_{\substack{x \to 0 \\ y \to 0}} \sqrt{x^2 + y^2} \text{ sen } \frac{1}{x^2 + y^2} = 0$$

3.° Insistimos aquí en que, en el supuesto de que una función $\boldsymbol{f}$ de p variables tenga todas sus derivadas en un punto $\boldsymbol{a}$, para garantizar la diferenciabilidad de $\boldsymbol{f}$ en $\boldsymbol{a}$ es suficiente con que $p - 1$ de sus derivadas parciales sean continuas en el punto (estando definidas en un entorno de él), pudiendo no serlo la otra derivada parcial. Así ocurre en el siguiente ejemplo. Considérese una función $\varphi: \mathbb{R}^2 \to \mathbb{R}$, $x \mapsto \varphi(x)$, que sea continua en $x = 0$ y que no sea derivable en ningún punto(*). A partir de φ, definimos la función $f: \mathbb{R}^2 \to \mathbb{R}$,

$$f(x, y) = \varphi(x) y$$

(*) Por ejemplo, sea $\varphi(x) = x$ si x es racional y $\varphi(x) = -x$ si x es irracional.

Las derivadas parciales de esta función son evidentemente:

$$f'_x(x, y) = \lim_{h \to 0} \frac{\varphi(x+h) - \varphi(x)}{h} y = \begin{cases} 0, & \text{si } y = 0 \\ \nexists, & \text{si } y \neq 0 \end{cases}$$

$$f'_y(x, y) = \varphi(x), \qquad \forall\ (x, y) \in \mathbb{R}^2$$

Como la primera existe en (0, 0) y la segunda existe en un entorno de (0, 0) y es continua en dicho punto, se puede asegurar que f es diferenciable en (0, 0). Comprobemos que, en efecto, así ocurre:

$$\lim_{\substack{x \to 0 \\ y \to 0}} \frac{[f(x, y) - f(0, 0)] - [x f'_x(0, 0) + y f'_y(0, 0)]}{\sqrt{x^2 + y^2}} = \lim_{\substack{x \to 0 \\ y \to 0}} (\varphi(x) - \varphi(0)) \frac{y}{\sqrt{x^2 + y^2}} = 0$$

DERIVABILIDAD Y DIFERENCIABILIDAD DE UNA FUNCIÓN COMPUESTA

En el supuesto de que las funciones que se consideran son suficientemente regulares (se supone que se verifican las condiciones de regularidad que en seguida se detallarán), la derivación de las funciones compuestas se rige por la «regla de la cadena» y su diferenciación responde a la «regla de invarianza». Adelantemos estos resultados, en un caso no muy complicado (pocas variables).

Ejemplo

Considérense dos funciones $(x, y) \mapsto (u, v)$ y $(u, v) \mapsto z$, de $\mathbb{R}^2$ en $\mathbb{R}^2$ y de $\mathbb{R}^2$ en $\mathbb{R}$, y también su función compuesta $(x, y) \mapsto z$, de $\mathbb{R}^2$ en $\mathbb{R}$:

$$\underbrace{(x, y) \xmapsto{(f,\, g)} (u, v) \overset{\varphi}{\mapsto} z}_{\psi} = \varphi(\overbrace{f(x, y)}^{u}, \overbrace{g(x, y)}^{v}) = \psi(x, y)$$

(aquí ψ es la composición de (f, g) con φ, o sea, $\psi = \varphi \circ (f, g)$).

- La «regla de la cadena», que en seguida demostraremos, asegura que las derivadas parciales de la función compuesta ψ vienen dadas, en función de las de f, g y φ, mediante las expresiones:

$$\frac{\partial \psi}{\partial x} = \frac{\partial \varphi}{\partial u}\frac{\partial f}{\partial x} + \frac{\partial \varphi}{\partial v}\frac{\partial g}{\partial x} \qquad \text{y} \qquad \frac{\partial \psi}{\partial y} = \frac{\partial \varphi}{\partial u}\frac{\partial f}{\partial y} + \frac{\partial \varphi}{\partial v}\frac{\partial g}{\partial y}$$

(las derivadas de ψ, f y g lo son en un cierto punto (x_0, y_0) y las derivadas de φ lo son en el correspondiente punto (u_0, v_0); esto se suele sobreentender siempre que, como aquí, se omite para simplificar las fórmulas y poder, así, resaltar los aspectos de ellas que son de interés en cada caso). Normalmente, se suele utilizar la misma letra para denotar a las funciones y a los valores que ellas toman; así, por ejemplo, en lugar de poner $u = f(x, y)$ o $z = \varphi(u, v)$, se acostumbra a escribir $u = u(x, y)$ y $z = z(u, v)$. Al proceder de este modo, las anteriores expresiones toman la forma:

$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial z}{\partial v}\frac{\partial v}{\partial x} \qquad \text{y} \qquad \frac{\partial z}{\partial y} = \frac{\partial z}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial z}{\partial v}\frac{\partial v}{\partial y}^{(*)}$$

Nótese que esta notación, aunque imprecisa, no se presta a equívocos. Así: la derivada $\partial z/\partial x$ lo es de la función $z = z(x, y)$ (no podría serlo de $z = z(u, v)$, pues aquí no figura la variable x respecto de la que se deriva); la derivada $\partial z/\partial v$ lo es de la función $z = z(u, v)$.

- La «regla de invarianza», que también probaremos de aquí a poco, permite poner la diferencial de la función compuesta ψ en la forma

$$d\psi(x, y) = \frac{\partial \varphi}{\partial u}\, df(x, y) + \frac{\partial \varphi}{\partial v}\, dg(x, y)$$

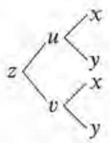
(las derivadas $\partial\varphi/\partial u$ y $\partial\varphi/\partial v$ lo son en el punto (u, v) correspondiente al (x, y) en el que se calculan las diferenciales). Si hacemos aquí, con la notación, lo mismo que en el caso de las derivadas, poniendo $z(x, y)$, $z(u, v)$, $u(x, y)$ y $v(x, y)$ en lugar de $\psi(x, y)$, $\varphi(x, y)$, $f(x, y)$ y $g(x, y)$, respectivamente, la «regla de invarianza» de la diferencial queda en la forma:

$$dz(x, y) = \frac{\partial z}{\partial u}\, du(x, y) + \frac{\partial z}{\partial v}\, dv(x, y)$$

A esta regla se la llama de «invarianza» debido a que la expresión que ella da, para la diferencial de la función compuesta $z(x, y)$, parece ser la misma (aunque no lo es) que la diferencial de la función $z(u, v)$. Nótese que la diferencial de esta última función es:

$$dz(u, v)(\Delta u, \Delta v) = \frac{\partial z}{\partial u}\,\Delta u + \frac{\partial z}{\partial v}\,\Delta v$$

(*) Hay quienes gustan de recordar estas fórmulas acudiendo a esquemas como el que aquí se indica (que llaman «árboles»):



En ellos, a la derecha de los trazos (ramas) se ponen las variables de las que depende la función que hay a su izquierda. Para hallar una derivada de la función (situada a la izquierda; la z) respecto de una de las variables últimas (situadas a la derecha; la x, por ejemplo), localizan todos los caminos que llevan de la función (z) a la variable (x), que figurará repetida varias veces a la derecha; para cada uno de los caminos aplican la regla de la cadena de las funciones de una variable y, después, suman todas las expresiones que así han ido obteniendo.

Cuando se escriben estas expresiones, la de $dz(x, y)$ y la de $dz(u, v)$, de forma abreviada, se pone (en ambos casos):

$$dz = \frac{\partial z}{\partial u} du + \frac{\partial z}{\partial v} dv$$

Nótese que en lugar de Δu e Δv, que son los incrementos arbitrarios de las variables independientes u y v, se puede poner du y dv (véase [24]$_3$). Con ello se crea un equívoco, pues du y dv denotan, al tiempo a Δu e Δv, por un lado, y a las diferenciales de las funciones $u(x, y)$ y $v(x, y)$, por otro. La «regla de la invarianza» hace que dicho equívoco no tenga consecuencias, ya que la expresión de dz resulta ser la misma (no varía), tanto si se considera a $z = z(u, v)$, en cuyo caso $du = \Delta u$ y $dv = \Delta v$, como si se supone que z es la función compuesta $z = z(x, y)$, en cuyo caso se debe tomar $du = du(x, y)$ y $dv = dv(x, y)$.

[29]

Sean $f: C \to \mathbb{R}^q$ y $g: D \to \mathbb{R}^r$ dos funciones, definidas en sendos abiertos $C \subset \mathbb{R}^p$ y $D \subset \mathbb{R}^q$, siendo $f(C) \subset D$; sean $f_1, ..., f_q$ las componentes de f; para cada $\boldsymbol{x} = (x_1, ..., x_p)$ recorriendo C, sea $\boldsymbol{y} = f(\boldsymbol{x}) = (y_1, ..., y_q)$.

1.° REGLA DE LA CADENA. Si f es derivable en un punto $\boldsymbol{x}$ respecto de $x_i^{(*)}$ y si g es diferenciable en el punto $\boldsymbol{y} = f(\boldsymbol{x})$, entonces $g \circ f$ es derivable en $\boldsymbol{x}$ respecto de $x_i^{(*)}$ y

$$\frac{\partial (g \circ f)(\boldsymbol{x})}{\partial x_i} = dg(\boldsymbol{y})(f'_{x_i}(\boldsymbol{x}))$$

esto es:

$$\frac{\partial (g \circ f)(\boldsymbol{x})}{\partial x_i} = \frac{\partial g(\boldsymbol{y})}{\partial y_1} \frac{\partial f_1(\boldsymbol{x})}{x_i} + \cdots + \frac{\partial g(\boldsymbol{y})}{\partial y_q} \frac{\partial f_q(\boldsymbol{x})}{\partial x_i}$$

2.° REGLA DE INVARIANZA. Si f es diferenciable en un punto $\boldsymbol{x} \in C$ y si g es diferenciable en el punto $\boldsymbol{y} = f(\boldsymbol{x})$, entonces $g \circ f$ es diferenciable en $\boldsymbol{x}$ y

$$d(g \circ f)(\boldsymbol{x}) = dg(\boldsymbol{y}) \circ df(\boldsymbol{x})$$

esto es:

$$d(g \circ f)(\boldsymbol{x}) = \frac{\partial g(\boldsymbol{y})}{\partial y_1} df_1(\boldsymbol{x}) + \cdots + \frac{\partial g(\boldsymbol{y})}{\partial y_q} df_q(\boldsymbol{x})$$

La matriz jacobiana de $g \circ f$ en $\boldsymbol{x}$ es, pues, $J(g \circ f)(\boldsymbol{x}) = Jg(\boldsymbol{y}) \cdot Jf(\boldsymbol{x})$.

3.° COMPOSICIÓN DE FUNCIONES DE CLASE $\mathscr{C}^1$. Si f es de clase $\mathscr{C}^1$ en el punto $\boldsymbol{x} \in C$ y si g es de clase $\mathscr{C}^1$ en el punto $\boldsymbol{y} = f(\boldsymbol{x})$, entonces $g \circ f$ es de clase $\mathscr{C}^1$ en $\boldsymbol{x}$.

(*) La propiedad sigue siendo cierta si la derivada lo es respecto de un vector no nulo $\boldsymbol{u} \in \mathbb{R}^p$, en cuyo caso queda:

$$\frac{\partial (g \circ g)(\boldsymbol{x})}{\partial \boldsymbol{u}} = dg(\boldsymbol{y})(f'_{\boldsymbol{u}}(\boldsymbol{x})) = \frac{\partial g(\boldsymbol{y})}{\partial y_1} \frac{\partial f_1(\boldsymbol{x})}{\partial \boldsymbol{u}} + \cdots + \frac{\partial g(\boldsymbol{y})}{\partial y_p} \frac{\partial f_q(\boldsymbol{x})}{\partial \boldsymbol{u}}$$

Demostración

1.º Comprobemos que la propiedad es cierta cuando se deriva $\boldsymbol{g} \circ \boldsymbol{f}$ respecto de un vector no nulo $\boldsymbol{u} \in \mathbb{R}^p$ cualquiera; tomando luego $\boldsymbol{u} = \boldsymbol{e}_i$ (vector i-ésimo de la base canónica), se obtiene como consecuencia el caso de la derivación respecto de x_i. Como $\boldsymbol{g}$ es diferenciable en el punto $\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x})$, llamando $\Delta \boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x} + \lambda \boldsymbol{u}) - \boldsymbol{f}(\boldsymbol{x})$ para $\lambda \in \mathbb{R}$, se tiene:

$$\frac{\partial(\boldsymbol{g} \circ \boldsymbol{f})(\boldsymbol{x})}{\partial \boldsymbol{u}} = \lim_{\lambda \to 0} \frac{\boldsymbol{g}[\boldsymbol{f}(\boldsymbol{x} + \lambda \boldsymbol{u})] - \boldsymbol{g}[\boldsymbol{f}(\boldsymbol{x})]}{\lambda} = \lim_{\lambda \to 0} \frac{\boldsymbol{g}(\boldsymbol{y} + \Delta \boldsymbol{y}) - \boldsymbol{g}(\boldsymbol{y})}{\lambda} =$$

$$= \lim_{\lambda \to 0} \frac{d\boldsymbol{g}(\boldsymbol{y})(\Delta \boldsymbol{y}) + \|\Delta \boldsymbol{y}\| \varepsilon(\Delta \boldsymbol{y})}{\lambda}$$

donde $\varepsilon(\Delta \boldsymbol{y}) \to \boldsymbol{o}$ cuando $\Delta \boldsymbol{y} \to \boldsymbol{o}$. Por tanto, se puede poner:

$$\frac{\partial(\boldsymbol{g} \circ \boldsymbol{f})(\boldsymbol{x})}{\partial \boldsymbol{u}} \lim_{\lambda \to 0} \left(d\boldsymbol{g}(\boldsymbol{y}) \left(\frac{\Delta \boldsymbol{y}}{\lambda} \right) + \left\| \frac{\Delta \boldsymbol{y}}{\lambda} \right\| \varepsilon(\Delta \boldsymbol{y}) \frac{|\lambda|}{\lambda} \right)$$

Como $\Delta \boldsymbol{y}/\lambda \to \boldsymbol{f}'_{\boldsymbol{u}}(\boldsymbol{x})$ (cuando $\lambda \to 0$), resulta que

$$\frac{\partial(\boldsymbol{g} \circ \boldsymbol{f})(\boldsymbol{x})}{\partial \boldsymbol{u}} = d\boldsymbol{g}(\boldsymbol{y})(\boldsymbol{f}'_{\boldsymbol{u}}(\boldsymbol{x})) + \boldsymbol{o}$$

(nótese que, cuando $\lambda \to 0$, $\Delta \boldsymbol{y} \to \boldsymbol{o}$ y por ello $\varepsilon(\Delta \boldsymbol{y}) \to \boldsymbol{o}$; además $|\lambda|/\lambda = \pm 1$, luego está acotado).

2.º Para simplificar las expresiones, llamemos $d\boldsymbol{f}(\boldsymbol{x}) = \varphi$ y $d\boldsymbol{g}(\boldsymbol{y}) = \psi$, con lo que se verificará (siempre que $\boldsymbol{x} + \Delta \boldsymbol{x} \in C$ e $\boldsymbol{y} + \Delta \boldsymbol{y} \in D$) que:

$$\begin{aligned} \boldsymbol{f}(\boldsymbol{x} + \Delta \boldsymbol{x}) &= \boldsymbol{f}(\boldsymbol{x}) + \varphi(\Delta \boldsymbol{x}) + \|\Delta \boldsymbol{x}\| \varepsilon(\Delta \boldsymbol{x}), \qquad \text{donde } \lim_{\Delta \boldsymbol{x} \to \boldsymbol{o}} \varepsilon(\Delta \boldsymbol{x}) = \boldsymbol{o} \\ \boldsymbol{g}(\boldsymbol{y} + \Delta \boldsymbol{y}) &= \boldsymbol{g}(\boldsymbol{y}) + \psi(\Delta \boldsymbol{y}) + \|\Delta \boldsymbol{y}\| \delta(\Delta \boldsymbol{y}), \qquad \text{donde } \lim_{\Delta \boldsymbol{y} \to \boldsymbol{o}} \delta(\Delta \boldsymbol{y}) = \boldsymbol{o} \end{aligned} \qquad [1]$$

Si en la segunda de estas relaciones se toma $\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x})$ e $\boldsymbol{y} + \Delta \boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x} + \Delta \boldsymbol{x})$ y como φ y ψ son funciones lineales, se obtiene:

$$\begin{aligned} \boldsymbol{g}[\boldsymbol{f}(\boldsymbol{x} + \Delta \boldsymbol{x})] &= \boldsymbol{g}(\boldsymbol{f}(\boldsymbol{x})) + \psi[\varphi(\Delta \boldsymbol{x}) + \|\Delta \boldsymbol{x}\| \varepsilon(\Delta \boldsymbol{x})] + \\ &+ \|\boldsymbol{f}(\boldsymbol{x} + \Delta \boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x})\| \delta[\boldsymbol{f}(\boldsymbol{x} + \Delta \boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x})] = \\ &= \boldsymbol{g}(\boldsymbol{f}(\boldsymbol{x})) + (\psi \circ \varphi)(\Delta \boldsymbol{x}) + \|\Delta \boldsymbol{x}\| \psi[\varepsilon(\Delta \boldsymbol{x})] + \\ &+ \|\boldsymbol{f}(\boldsymbol{x} + \Delta \boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x})\| \delta[\boldsymbol{f}(\boldsymbol{x} + \Delta \boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x})] \end{aligned}$$

Así pues, se verifica (siempre que $\boldsymbol{x} + \Delta \boldsymbol{x} \in C$) que:

$$(\boldsymbol{g} \circ \boldsymbol{f})(\boldsymbol{x} + \Delta \boldsymbol{x}) = (\boldsymbol{g} \circ \boldsymbol{f})(\boldsymbol{x}) + (\psi \circ \varphi)(\Delta \boldsymbol{x}) + \boldsymbol{r}(\Delta \boldsymbol{x}) \qquad [2]$$

donde

$$\boldsymbol{r}(\Delta \boldsymbol{x}) = \underbrace{\|\Delta \boldsymbol{x}\| \psi[\varepsilon(\Delta \boldsymbol{x})]}_{\boldsymbol{r}_1(\Delta \boldsymbol{x})} + \underbrace{\|\boldsymbol{f}(\boldsymbol{x} + \Delta \boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x})\| \delta[\boldsymbol{f}(\boldsymbol{x} + \Delta \boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x})]}_{\boldsymbol{r}_2(\Delta \boldsymbol{x})} \qquad [3]$$

El teorema quedará demostrado si se comprueba que es $\boldsymbol{r}(\Delta \boldsymbol{x}) = \boldsymbol{o}(\|\Delta \boldsymbol{x}\|)$, ya que entonces la anterior relación [2] asegurará que $\boldsymbol{g} \circ \boldsymbol{f}$ era diferenciable en el punto $\boldsymbol{x}$ y que su diferencial sería $\psi \circ \varphi$, como había que verificar. Vamos a comprobar, pues, que $\boldsymbol{r}_1(\Delta \boldsymbol{x}) = \boldsymbol{o}(\|\Delta \boldsymbol{x}\|)$ y $\boldsymbol{r}_2(\Delta \boldsymbol{x}) = \boldsymbol{o}(\|\Delta \boldsymbol{x}\|)$, con lo que la demostración habrá acabado.

- Como ψ es una aplicación lineal, se sabe(*) que existe una constante $K > 0$ tal que $\|\psi(\boldsymbol{v})\| \leqslant K\|\boldsymbol{v}\|$ para todo $\boldsymbol{v} \in \mathbb{R}^q$; por tanto, será:

$$\|\boldsymbol{r}_1(\Delta \boldsymbol{x})\| = \|\Delta \boldsymbol{x}\| \, \|\psi[\varepsilon(\Delta \boldsymbol{x})]\| \leqslant K\|(\Delta \boldsymbol{x})\| \, \|\varepsilon(\Delta \boldsymbol{x})\|$$

  de donde se desprende que $\boldsymbol{r}_1(\Delta \boldsymbol{x}) = \boldsymbol{o}(\|\Delta \boldsymbol{x}\|)$, pues $\varepsilon(\Delta \boldsymbol{x}) \to \boldsymbol{o}$ cuando $\Delta \boldsymbol{x} \to \boldsymbol{o}$.
- Para el sumando $\boldsymbol{r}_2(\Delta \boldsymbol{x})$, empecemos acotando el factor $\|\boldsymbol{f}(\boldsymbol{x} + \Delta \boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x})\|$. Para ello acudamos primero a la expresión de $\boldsymbol{f}(\boldsymbol{x} + \Delta \boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x})$ de [1] y, después, utilicemos que, por ser φ lineal, existe una constante $K' > 0$ tal que $\|\varphi(\boldsymbol{u})\| \leqslant K'\|\boldsymbol{u}\|$ para todo $\boldsymbol{u} \in \mathbb{R}^p$; de ello, se obtiene que

$$\begin{aligned}\|\boldsymbol{r}_2(\Delta \boldsymbol{x})\| &\leqslant \{\|\varphi(\Delta \boldsymbol{x})\| + \|\Delta \boldsymbol{x}\| \, \|\varepsilon(\Delta \boldsymbol{x})\|\}\delta[\boldsymbol{f}(\boldsymbol{x} + \Delta \boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x})] \leqslant \\ &\leqslant \|\Delta \boldsymbol{x}\|\{K' + \|\varepsilon(\Delta \boldsymbol{x})\|\}\delta[\boldsymbol{f}(\boldsymbol{x} + \Delta \boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x})]\end{aligned}$$

  de donde se desprende que $\boldsymbol{r}_2(\Delta \boldsymbol{x}) = \boldsymbol{o}\|(\Delta \boldsymbol{x})\|$, pues cuando $\Delta \boldsymbol{x} \to \boldsymbol{o}$ se verifica que $\varepsilon(\Delta \boldsymbol{x}) \to \boldsymbol{o}$ y $\delta[\boldsymbol{f}(\boldsymbol{x} + \Delta \boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x})] \to \boldsymbol{o}$ (pues $\boldsymbol{f}(\boldsymbol{x} + \Delta \boldsymbol{x}) \to \boldsymbol{f}(\boldsymbol{x})$). Queda, pues, comprobado que:

$$d(\boldsymbol{g} \circ \boldsymbol{f})(\boldsymbol{x}) = \psi \circ \varphi = d\boldsymbol{g}(\boldsymbol{y}) \circ d\boldsymbol{f}(\boldsymbol{x})$$

  Nótese que esta igualdad supone que la matriz de la aplicación lineal $d(\boldsymbol{g} \circ \boldsymbol{f})(\boldsymbol{x})$ es el producto de la matriz de $d\boldsymbol{g}(\boldsymbol{y})$ por la de $d\boldsymbol{f}(\boldsymbol{x})$; como estas matrices son las jacobianas respectivas, de $\boldsymbol{g} \circ \boldsymbol{f}$, $\boldsymbol{g}$ y $\boldsymbol{f}$, se puede poner:

$$J(\boldsymbol{g} \circ \boldsymbol{f})(\boldsymbol{x}) = J\boldsymbol{g}(\boldsymbol{y}) \cdot J\boldsymbol{f}(\boldsymbol{x})$$

3.° Ahora que las funciones son de clase $\mathscr{C}^1$, se verifican las hipótesis de los teoremas anteriores, luego existen las derivadas parciales de $\boldsymbol{g} \circ \boldsymbol{f}$, que están dadas por la regla de la cadena. Como las derivadas de $\boldsymbol{g} \circ \boldsymbol{f}$ son, entonces, sumas de productos de las derivadas de $\boldsymbol{f}$ por las de $\boldsymbol{g}$ y todas estas derivadas son continuas, pues $\boldsymbol{f}$ y $\boldsymbol{g}$ son de clase $\mathscr{C}^1$, resulta que también lo son todas las derivadas parciales de $\boldsymbol{g} \circ \boldsymbol{f}$, es decir, esta función es de clase $\mathscr{C}^1$.

[29]₁ Observación

Según se dijo en [29], 2.°, si $\boldsymbol{f}$ es una función diferenciable en un punto $\boldsymbol{x}$ y si $\boldsymbol{g}$ es una función diferenciable en el punto $\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x})$, entonces su composición $\boldsymbol{g} \circ \boldsymbol{f}$ es diferenciable en $\boldsymbol{x}$ y se verifica que

$$d(\boldsymbol{g} \circ \boldsymbol{f})(\boldsymbol{x}) = d\boldsymbol{g}(\boldsymbol{y}) \circ d\boldsymbol{f}(\boldsymbol{x}) \qquad \text{o} \qquad J(\boldsymbol{g} \circ \boldsymbol{f})(\boldsymbol{x}) = J\boldsymbol{g}(\boldsymbol{y}) \cdot J\boldsymbol{f}(\boldsymbol{x})$$

(*) Si la ecuación de la aplicación lineal $\psi : \mathbb{R}^q \to \mathbb{R}^r$, $\boldsymbol{v} \mapsto \boldsymbol{w} = \psi(\boldsymbol{v})$, es $w_i = \sum a_{ij}v_j$, llamando $a = \text{máx}\,\{|a_{ij}|\}$, es $|w_i| \leqslant \sum |a_{ij}||v_i| \leqslant ap\|\boldsymbol{v}\|$, luego $\|\boldsymbol{w}\| = (\sum w_i)^{1/2} \leqslant (ra^2p^2\|\boldsymbol{v}\|^2)^{1/2} = K\|\boldsymbol{v}\|$, donde $K = \sqrt{rap}$.

Si a las «variables» se las denota como en [29], 2.°, es decir, si se pone:

$$\boldsymbol{x} = (x_1, x_2, ..., x_p) \mapsto \boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x}) \qquad \text{e} \qquad \boldsymbol{y} = (y_1, y_2, ..., y_q) \mapsto \boldsymbol{g}(\boldsymbol{y})$$

y si, acudiendo a las coordenadas de las funciones, se pone $\boldsymbol{f} = (f_1, f_2, ..., f_q)$, $\boldsymbol{g} = (g_1, g_2, ..., g_q)$ y $\boldsymbol{g} \circ \boldsymbol{f} = ((\boldsymbol{g} \circ \boldsymbol{f})_1, (\boldsymbol{g} \circ \boldsymbol{f})_2, ..., (\boldsymbol{g} \circ \boldsymbol{f})_r)$, la anterior igualdad entre matrices jacobianas se puede escribir, de manera explícita, poniendo:

$$\begin{bmatrix} \dfrac{\partial(\boldsymbol{g}\circ\boldsymbol{f})_1}{\partial x_1}(\boldsymbol{x}) & \dfrac{\partial(\boldsymbol{g}\circ\boldsymbol{f})_1}{\partial x_2}(\boldsymbol{x}) & \cdots & \dfrac{\partial(\boldsymbol{g}\circ\boldsymbol{f})_1}{\partial x_p}(\boldsymbol{x}) \\ \dfrac{\partial(\boldsymbol{g}\circ\boldsymbol{f})_2}{\partial x_1}(\boldsymbol{x}) & \dfrac{\partial(\boldsymbol{g}\circ\boldsymbol{f})_2}{\partial x_2}(\boldsymbol{x}) & \cdots & \dfrac{\partial(\boldsymbol{g}\circ\boldsymbol{f})_2}{\partial x_p}(\boldsymbol{x}) \\ \cdots & \cdots & \cdots & \cdots \\ \dfrac{\partial(\boldsymbol{g}\circ\boldsymbol{f})_r}{\partial x_1}(\boldsymbol{x}) & \dfrac{\partial(\boldsymbol{g}\circ\boldsymbol{f})_r}{\partial x_2}(\boldsymbol{x}) & \cdots & \dfrac{\partial(\boldsymbol{g}\circ\boldsymbol{f})_r}{\partial x_p}(\boldsymbol{x}) \end{bmatrix} =$$

$$= \begin{bmatrix} \dfrac{\partial g_1}{\partial y_1}(\boldsymbol{y}) & \dfrac{\partial g_1}{\partial y_2}(\boldsymbol{y}) & \cdots & \dfrac{\partial g_1}{\partial y_q}(\boldsymbol{y}) \\ \dfrac{\partial g_2}{\partial y_1}(\boldsymbol{y}) & \dfrac{\partial g_2}{\partial y_2}(\boldsymbol{y}) & \cdots & \dfrac{\partial g_2}{\partial y_q}(\boldsymbol{y}) \\ \cdots & \cdots & \cdots & \cdots \\ \dfrac{\partial g_r}{\partial y_1}(\boldsymbol{y}) & \dfrac{\partial g_r}{\partial y_2}(\boldsymbol{y}) & \cdots & \dfrac{\partial g_r}{\partial y_q}(\boldsymbol{y}) \end{bmatrix} \cdot \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1}(\boldsymbol{x}) & \dfrac{\partial f_1}{\partial x_2}(\boldsymbol{x}) & \cdots & \dfrac{\partial f_1}{\partial x_p}(\boldsymbol{x}) \\ \dfrac{\partial f_2}{\partial x_1}(\boldsymbol{x}) & \dfrac{\partial f_2}{\partial x_2}(\boldsymbol{x}) & \cdots & \dfrac{\partial f_2}{\partial x_p}(\boldsymbol{x}) \\ \cdots & \cdots & \cdots & \cdots \\ \dfrac{\partial f_q}{\partial x_1}(\boldsymbol{x}) & \dfrac{\partial f_q}{\partial x_2}(\boldsymbol{x}) & \cdots & \dfrac{\partial f_q}{\partial x_p}(\boldsymbol{x}) \end{bmatrix} =$$

Nótese que esta igualdad de matrices equivale a las $p \times q$ igualdades:

$$\frac{\partial(\boldsymbol{g}\circ\boldsymbol{f})_i}{\partial x_j}(\boldsymbol{x}) = \frac{\partial g_i(\boldsymbol{y})}{\partial y_1}\,\frac{\partial f_1(\boldsymbol{x})}{\partial x_j} + \frac{\partial g_i(\boldsymbol{y})}{\partial y_2}\,\frac{\partial f_2(\boldsymbol{x})}{\partial x_j} + \cdots + \frac{\partial g_i(\boldsymbol{y})}{\partial y_q}\,\frac{\partial f_q(\boldsymbol{x})}{\partial x_j}$$

que es la regla de la cadena (para la derivada parcial j-ésima) correspondiente a la coordenada i-ésima de $\boldsymbol{g} \circ \boldsymbol{f}$.

$[29]_2$ Ejercicio

Sea $(x, y) \mapsto z(x, y)$ una función de clase $\mathscr{C}^1$ y considérese la expresión

$$E = x\,\frac{\partial z}{\partial y} - y\,\frac{\partial z}{\partial x} \qquad [1]$$

Si se hace el «cambio de variables» $(x, y) \mapsto (u, v)$ dado por:

$$x = u \cos v \qquad \text{e} \qquad y = u \operatorname{sen} v \qquad [2]$$

Obtener la forma en la que queda E cuando se expresa en función de u y v.

Resolución

Derivando la función $z(u, v)$, es decir, la función compuesta

$$(u, v) \mapsto z(u, v) = z\big(x(u, v), y(u, v)\big)$$

de acuerdo con la regla de la cadena, se obtiene:

$$\begin{cases} \dfrac{\partial z}{\partial u} = \dfrac{\partial z}{\partial x}\dfrac{\partial x}{\partial u} + \dfrac{\partial z}{\partial y}\dfrac{\partial y}{\partial u} = \dfrac{\partial z}{\partial x}\cos v + \dfrac{\partial z}{\partial y}\operatorname{sen} v \\ \dfrac{\partial z}{\partial v} = \dfrac{\partial z}{\partial x}\dfrac{\partial x}{\partial v} + \dfrac{\partial z}{\partial y}\dfrac{\partial y}{\partial v} = -\dfrac{\partial z}{\partial x}\, u \operatorname{sen} v + \dfrac{\partial z}{\partial y}\, u \cos v \end{cases}$$

Despejando $\partial z/\partial x$ y $\partial z/\partial y$, en función de $\partial z/\partial u$ y $\partial z/\partial v$, se obtiene (si es $u \neq 0$):

$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial u}\cos v - \frac{\partial z}{\partial v}\frac{\operatorname{sen} v}{u} \qquad \text{y} \qquad \frac{\partial z}{\partial y} = \frac{\partial z}{\partial u}\operatorname{sen} v + \frac{\partial z}{\partial v}\frac{\cos v}{u}$$

Llevando a [1] estas expresiones y las [2], se obtiene que

$$E = \frac{\partial z}{\partial v}$$

[29]$_3$ Funciones homogéneas. Teorema de Euler

FUNCIÓN HOMOGÉNEA. Sea $f: C \to \mathbb{R}$ una función real, definida en un abierto $C \subset \mathbb{R}^p$. Se dice que f es homogénea de grado $\alpha \neq 0$ en C si para cualesquiera $\boldsymbol{x} \in C$ y $t \in \mathbb{R}$ tal que $t\boldsymbol{x} \in C$, se verifica que

$$f(t\boldsymbol{x}) = t^\alpha f(\boldsymbol{x}), \qquad \text{o sea } f(tx_1, ..., tx_p) = t^\alpha f(x_1, ..., x_p) \tag{1}$$

TEOREMA DE EULER. Si la función $f: C \to \mathbb{R}$ es homogénea de grado α en C y, además, f es diferenciable en C, entonces

$$x_1 \frac{\partial f(\boldsymbol{x})}{\partial x_1} + x_2 \frac{\partial f(\boldsymbol{x})}{\partial x_2} + \cdots + x_p \frac{\partial f(\boldsymbol{x})}{\partial x_p} = \alpha f(\boldsymbol{x}) \tag{2}$$

para cualquiera que sea el punto $\boldsymbol{x} = (x_1, x_2, ..., x_p)$ de C.

Demostración

Expresando que las derivadas respecto de t del primero y segundo miembro de [1] son iguales (lo que ocurre para todo t tal que $t\boldsymbol{x} \in C$ y, en particular, para $t = 1$), acudiendo a la regla de la cadena se obtiene que:

$$\frac{\partial f(t\boldsymbol{x})}{\partial x_1} x_1 + \cdots + \frac{\partial f(t\boldsymbol{x})}{\partial x_p} x_p = \alpha t^{\alpha-1} f(\boldsymbol{x})$$

Tomando aquí $t = 1$ (para $t = 1$ se verifica la anterior igualdad), se obtiene la relación [2] que había que probar.

2.4. DERIVADAS SUCESIVAS

A las derivadas parciales de una función f se las suele llamar derivadas parciales de primer orden. Si f admite derivada parcial de primer orden $D_i f$, respecto de su i-ésima variable, y esta función fuese derivable, en un cierto punto $\boldsymbol{x}$, respecto de su j-ésima variable, de esta derivada de $D_i f$ se dice que es una derivada parcial segunda de f. Reiterando el proceso se definen de igual modo, cuando existen, las derivadas parciales de tercer, cuarto, ..., n-ésimo orden de la función f en el punto $\boldsymbol{x}$.

DERIVADAS PARCIALES DE ORDEN SUPERIOR AL PRIMERO

[30]

Sea $f: C \to \mathbb{R}^q$ una función, que está definida en un conjunto abierto $C \subset \mathbb{R}^p$.

DERIVADAS SUCESIVAS. Si f admite derivada parcial $f_i' : C \to \mathbb{R}$, respecto de la variable i-ésima, entonces a su derivada $(f_i')_j'$ (si existe) se la llama derivada parcial segunda de f, respecto de las variables i-ésima y j-ésima, y se la denota poniendo f_{ij}'', $D_{ij}^2 f$ o $\partial^2 f/\partial i \partial j$. Reiterando el proceso, a partir de la derivada parcial $D_{ij\ldots k}^r f$ (de orden r) y derivando de nuevo, se dice que $D_l[D_{ij\ldots k}^r f]$ (si existe) es derivada parcial de orden $r + 1$, de f, y se la representa poniendo $D_{ij\ldots kl}^{r+1} f$.

FUNCIONES DE CLASE $\mathscr{C}^r$. Si f es continua en C, se dice que f es de clase $\mathscr{C}^0$ en C. Si f admite todas sus p^r derivadas parciales de orden r (para $r \geqslant 1$) y ellas son continuas en C, se dice que f es de clase $\mathscr{C}^r$ en C; si f es de clase $\mathscr{C}^r$ en C, entonces también es de clase $\mathscr{C}^{r-1}$, $\mathscr{C}^{r-2}$, ..., $\mathscr{C}^1$ y $\mathscr{C}^0$ en C. Se dice que f es de clase $\mathscr{C}^r$ en un punto $\boldsymbol{a} \in C$ si f es de clase $\mathscr{C}^{r-1}$ en un cierto entorno de $\boldsymbol{a}$, si admite todas sus derivadas de orden r en un entorno de $\boldsymbol{a}$ y si éstas son continuas en el punto $\boldsymbol{a}$.

[30]$_1$ Ejemplo

Consideremos la función f, de dos variables reales y con valores en $\mathbb{R}^3$, definida por:

$$f(x, y) = \left(e^{-x^2} \operatorname{sen} y, \operatorname{arctg} \frac{x}{y}, \mathrm{L}\,(xy^3)\right)$$

(f está definida en el conjunto $C = \{(x, y) \in \mathbb{R}^2 / xy > 0\}$) y hallemos las primeras y segundas derivadas de f (en C):

$$\frac{\partial f(x, y)}{\partial x} = \left(-2xe^{-x^2} \operatorname{sen} y, \frac{y}{x^2 + y^2}, \frac{1}{x}\right)$$

$$\frac{\partial f(x, y)}{\partial y} = \left(e^{-x^2} \cos y, \frac{-x}{x^2 + y^2}, \frac{3}{y}\right)$$

$$\frac{\partial^2 f(x, y)}{\partial x^2} = \left((4x^2 - 2)e^{-x^2} \operatorname{sen} y, \frac{-2xy}{(x^2 + y^2)^2}, \frac{-1}{x^2}\right)$$

$$\frac{\partial^2 f(x, y)}{\partial x \partial y} = \left(-2xe^{-x^2} \cos y, \frac{x^2 - y^2}{(x^2 + y^2)^2}, 0\right)$$

$$\frac{\partial^2 f(x, y)}{\partial y \partial x} = \left(-2xe^{-x^2} \cos y, \frac{x^2 - y^2}{(x^2 + y^2)^2}, 0\right)$$

$$\frac{\partial^2 f(x, y)}{\partial y^2} = \left(-e^{-x^2} \operatorname{sen} y, \frac{2xy}{(x^2 + y^2)^2}, -\frac{3}{y^2}\right)$$

Obsérvese que las derivadas parciales $\partial^2 f/\partial x \partial y$ y $\partial^2 f/\partial y \partial x$, que se llaman derivadas «cruzadas» o «mixtas», han resultado ser iguales. Esto, que ocurre muy frecuentemente (como veremos pronto, en [31]), hay veces que no se verifica; para que las derivadas cruzadas sean diferentes en algún punto, la función debe ser «muy poco» regular en el punto.

[30]$_2$ Observaciones

1.º Para comprobar que si una función f es de clase $\mathscr{C}^r$ (es decir, admite todas sus derivadas parciales hasta las de orden r), entonces f es de clase $\mathscr{C}^s$ para todo $s < r$ (es decir, entonces son también continuas sus derivadas parciales de órdenes inferiores a r), basta acudir a [28]$_1$, 1.ª (según la que, si una función es de clase $\mathscr{C}^1$, también es de clase $\mathscr{C}^0$) y aplicársela a las derivadas parciales de orden r, de f, primero, a las de orden $r - 1$, después, y así hasta llegar a las derivadas parciales de primer orden.

2.º También se pueden definir, de análoga manera, las derivadas de orden superior respecto de vectores cualesquiera (no nulos): si la función $f: C \to \mathbb{R}^q$ (con $C \subset \mathbb{R}^p$ abierto) admite derivada $\partial f/\partial \boldsymbol{u}$, respecto de un vector $\boldsymbol{u}$, y esta función es derivable respecto de un vector $\boldsymbol{v}$, se dice que $\partial(\partial f/\partial \boldsymbol{u})/\partial \boldsymbol{v}$ es la derivada segunda de f respecto de $\boldsymbol{u}$, primero, y de $\boldsymbol{v}$, después, y se la representa poniendo $\partial^2 f/\partial \boldsymbol{u} \partial \boldsymbol{v}$; reiterando el proceso, se obtienen las derivadas de órdenes cualesquiera, respecto de una sucesión de vectores.

3.° La notación se suele simplificar lo máximo posible. Así, por ejemplo, para denotar a la derivada parcial tercera, de una función $(x, y) \mapsto z = f(x, y)$, respecto de y, primero, respecto de x, después, y respecto de x, finalmente, en lugar de poner

$$\frac{\partial^3 f}{\partial y \partial x \partial x} \qquad f'''_{yxx} \qquad D^3_{yxx}f \qquad \frac{\partial^3 z}{\partial y \partial x \partial x} \qquad z'''_{yxx} \qquad D^3_{yxx}z$$

se pondrá cualquiera de las siguientes expresiones:

$$\frac{\partial^3 f}{\partial y \partial x^2}, \qquad f'''_{yx^2}, \qquad D^3_{yx^2}f, \qquad \frac{\partial^3 z}{\partial y \partial x^2}, \qquad z'''_{yx^2}, \qquad D^3_{yx^2}z,$$

Cuando las «derivadas cruzadas» son iguales (cosa que ocurre «casi siempre»), a todas ellas se las denota de igual forma: así, se pone

$$\frac{\partial^3 f}{\partial x^2 \partial y} \qquad \text{para representar también a } \frac{\partial^3 f}{\partial x \partial y \partial x} \qquad \text{y} \qquad \frac{\partial^3 f}{\partial y \partial x^2}$$

4.° Según se señaló en $[30]_1$, las derivadas cruzadas no resultan ser siempre iguales. Veamos un ejemplo en el que son distintas (ejemplo de Schwarz). Sea $f: \mathbb{R}^2 \to \mathbb{R}$ la función definida por las expresiones

$$\begin{cases} f(x, y) = x^2 \operatorname{arctg} \dfrac{y}{x} - y^2 \operatorname{arctg} \dfrac{x}{y}, & \text{si } xy \neq 0 \\ f(x, 0) = 0 \qquad \text{y} \qquad f(0, y) = 0 & \end{cases}$$

Es fácil obtener (acudiendo a la definición de derivada parcial) que:

$$f'_x(x, 0) = 0, \qquad f'_x(0, y) = -y, \qquad f'_y(x, 0) = x, \qquad f'_y(0, y) = 0$$

Por tanto, las derivadas cruzadas en el punto (0, 0) son distintas, ya que:

$$f''_{xy}(0, 0) = \lim_{k \to 0} \frac{f'_x(0, k) - f'_x(0, 0)}{k} = \lim_{k \to 0} \frac{-k - 0}{k} = -1$$

$$f''_{yx}(0, 0) = \lim_{h \to 0} \frac{f'_y(h, 0) - f'_y(0, 0)}{h} = \lim_{h \to 0} \frac{h - 0}{h} = 1$$

Nótese que, para $xy \neq 0$, aplicando las reglas de derivación, se obtiene fácilmente que

$$f'_x(x, y) = 2x \operatorname{arctg} \frac{y}{x} - y \qquad \text{y} \qquad f'_y(x, y) = x - 2y \operatorname{arctg} \frac{x}{y}$$

$$f''_{xy}(x, y) = f''_{yx}(x, y) = \frac{x^2 - y^2}{x^2 + y^2} \qquad (\text{para } xy \neq 0)$$

Es fácil comprobar que las derivadas cruzadas f_{xy} y f_{yx} no son continuas en (0, 0). Según vamos a ver en seguida, si estas derivadas cruzadas hubieran sido continuas en (0, 0), entonces habrían resultado iguales las $f''_{xy}(0, 0)$ y $f''_{yx}(0, 0)$.

PERMUTABILIDAD DEL ORDEN DE DERIVACIÓN

[31]

TEOREMA DE SCHWARZ. Sea $(x, y) \mapsto f(x, y)$ una función definida, al menos, en un entorno $U \subset \mathbb{R}^2$ de un punto $(a, b) \in \mathbb{R}^2$ y que toma valores en $\mathbb{R}^q$. Si existen las derivadas parciales f'_x, f'_y y f''_{xy} en el entorno U y si f''_{xy} es continua en el punto (a, b), entonces existe $f''_{yx}(a, b)$ y se verifica que $f''_{yx}(a, b) = f''_{xy}(a, b)$.

COROLARIO. Si una función $f: U \to \mathbb{R}^q$, definida en un entorno $U \subset \mathbb{R}^p$ de un punto $\boldsymbol{a} \in \mathbb{R}^p$, es de clase $\mathscr{C}^r$ en $\boldsymbol{a}$ $(r \geqslant 2)$, entonces el valor de cualquiera de las derivadas parciales «mixtas» r-ésimas, de f en $\boldsymbol{a}$, no depende del orden en el que se tomen las r variables que intervienen en la derivación; esto es, entonces $f^{(r)}_{ij\ldots l}(\boldsymbol{a}) = f^{(r)}_{\alpha\beta\ldots\delta}(\boldsymbol{a})$, donde $\alpha\beta\ldots\delta$ es una permutación cualquiera de $ij\ldots l$.

Notas: Para esto de la permutabilidad del orden de derivación, en la mayoría de los casos suele ser suficiente con acudir al teorema de Bonnet, que se enuncia más abajo, en lugar de hechar mano del de Schwarz; aquél es un teorema más «débil» y más fácil de demostrar que éste. Las demostraciones de uno y otro las basaremos en el lema que se enuncia a continuación.

TEOREMA DE BONNET. Si la función $(x, y) \mapsto f(x, y)$ admite las derivadas $f'_x, f'_y, f''_{xy}, f''_{yx}$, en un entorno del punto (a, b) y si f''_{xy}, y f''_{yx}, son continuas en (a, b), entonces $f''_{xy}(a, b) = f''_{yx}(a, b)$.

LEMA. Si la función $(x, y) \mapsto f(x, y)$ está definida y admite las derivadas parciales f'_x y f''_{xy} en un entorno de un punto (a, b) y si f''_{xy} es continua en (a, b), entonces

$$f''_{xy}(a, b) = \lim_{\substack{h \to 0 \\ k \to 0}} \frac{\Delta(h, k)}{hk}, \qquad \text{donde } \Delta(h, k) = f(a + h, b + k) - f(a + h, b) - f(a, b + k) + f(a, b)$$

Demostración

En lo que sigue, para simplificar la notación, supondremos que se ha realizado el cambio de variables $x = a + x'$, $y = b + y'$ o, si se prefiere (y nos evitamos las «primas»), que es $a = 0$ y $b = 0$.

Nótese que, para demostrar el lema y los teoremas anteriores, basta con comprobar que son ciertos para el caso de ser f una función real $(q = 1)$, lo que así haremos. Para el caso vectorial $(q \geqslant 2)$ bastará, entonces, con acudir a las componentes de f y tener presente lo dicho en [12], 5.º), en [16], 1.º) y en [23], 1.º).

1.º Empecemos demostrando el «lema». Si el radio del entorno del enunciado es ρ, tomamos unos valores fijos para h y k de manera que sea $h < \rho/\sqrt{2}$ y $k < \rho/\sqrt{2}$, con lo que $\sqrt{h^2 + k^2} < \rho$ y todos los puntos que luego se consideran pertenecerán al referido entorno, en el que existen f, f'_x y f''_{xy}.

La expresión $\Delta(h, k)$ se puede poner en la forma

$$\Delta(h, k) = \varphi(h) - \varphi(0), \qquad \text{donde } \varphi(t) = f(t, k) - f(t, 0)$$

Como la función $t \mapsto \varphi(t)$ es derivable entre $t = 0$ y $t = h$ y su derivada es

$$\varphi'(t) = f'_x(t, k) - f'_x(t, 0)$$

resulta que a esta función φ le es de aplicación el teorema de los incrementos finitos (para funciones de una variable) entre los puntos $t = 0$ y $t = h$, del que se deduce que existe cierto ξ comprendido entre dichos puntos (con lo que $0 < |\xi| < |h|$) tal que

$$\Delta(h, k) = \varphi(h) - \varphi(0) = h\varphi'(\xi) = h[f'_x(\xi, k) - f'_x(\xi, 0)]$$

Recurriendo ahora a la función $y \mapsto \psi(t) = f'_x(\xi, t)$, podemos poner $\Delta(h, k) = h[\psi(k) - \psi(0)]$. Esta función ψ es derivable entre $t = 0$ y $t = k$ y su derivada es $\psi'(t) = f''_{xy}(\xi, t)$. A esta función se le puede aplicar, pues, el teorema de los incrementos finitos entre los puntos $t = 0$ y $t = k$, del que se obtiene que existe cierto ζ comprendido entre dichos puntos (con lo que $0 < |\zeta| < |k|$) tal que

$$\Delta(h, k) = h[\psi(k) - \psi(0)] = hk\psi'(\zeta) = hkf''_{xy}(\xi, \zeta)$$

Una vez llegados a este resultado, hagamos ahora que (h, k) tienda a $(0, 0)$, con lo que también (ξ, ζ) tiende a $(0, 0)$ y podemos poner que:

$$\lim_{\substack{h \to 0 \\ k \to 0}} \frac{\Delta(h, k)}{hk} = \lim_{\substack{h \to 0 \\ k \to 0}} f''_{xy}(\xi, \zeta) = \lim_{\substack{\xi \to 0 \\ \zeta \to 0}} f''_{xy}(\xi, \zeta) = f''_{xy}(0, 0)$$

(la última igualdad es cierta por ser f''_{xy} continua en el punto origen).

2.º Demostremos el teorema de Bonnet (que, aunque es consecuencia del de Schwarz, se demuestra más fácilmente que éste; quien no precise de mayores exigencias podría renunciar a la consideración del teorema de Schwarz). Si se verifican las hipótesis del teorema de Bonnet, entonces se verifican las del lema anterior y del que resulta de intercambiar los papeles de x e y; por tanto, de dichos lemas se desprende que $f''_{xy}(0, 0)$ y $f''_{yx}(0, 0)$ son, ambas, iguales al límite de $\Delta(h, k)/(h, k)$, con lo que son iguales entre sí, como había que comprobar.

3.º Demostremos ya el teorema de Schwarz. Como también ahora se cumplen las hipótesis del lema anterior, sabemos que

$$f''_{xy}(0, 0) = \lim_{\substack{h \to 0 \\ k \to 0}} \frac{\Delta(h, k)}{hk}$$

Dado, pues, $\varepsilon > 0$ (cualquiera), se sabe que existe $\delta > 0$ tal que

$$\left.\begin{array}{l} 0 < |h| < \delta \\ 0 < |k| < \delta \end{array}\right\} \Rightarrow \left|\frac{\Delta(h, k)}{hk} - f''_{xy}(0, 0)\right| < \varepsilon \qquad [1]$$

Nótese que, como existe f'_y en el entorno U, resulta que (según la definición de derivada respecto de y) es:

$$\lim_{k \to 0} \frac{\Delta(h,k)}{hk} = \frac{1}{h} \lim_{k \to 0} \left(\frac{f(h,k) - f(h,0)}{k} - \frac{f(0,k) - f(0,0)}{k} \right) = \frac{f'_y(h,0) - f'_y(0,0)}{h}$$

Por tanto, tomando límites en [1] cuando $k \to 0$, se obtiene que

$$0 < |h| < \delta \quad \Rightarrow \quad \left| \frac{f'_y(h,0) - f'_y(0,0)}{h} - f''_{xy}(0,0) \right| \leqslant \varepsilon$$

lo cual, como se verifica para todo $\varepsilon > 0$, significa que:

$$\exists\ l = \lim_{h \to 0} \frac{f'_y(h,0) - f'_y(0,0)}{h} \qquad \text{y es } l = f''_{xy}(0,0)$$

como $l = f''_{yx}(0,0)$, se concluye que existe $f''_{yx}(0,0)$ y es igual a $f''_{xy}(0,0)$, como había que ver.

4.° Finalmente, comprobemos el corolario. Para verificar que, en la derivación parcial, el resultado es independiente del orden en el que se consideren las variables respecto de las que se deriva, es suficiente con poner de manifiesto que la derivada no se altera si se cambia el orden en el que se realizan dos derivaciones parciales consecutivas. Aplicando este resultado sucesivas veces se llega trivialmente al resultado del corolario. Hemos, pues, de comprobar que:

$$\frac{\partial^r f}{\partial i_1 \cdots \partial i_h \partial i_{h+1} \cdots \partial i_r} = \frac{\partial^r f}{\partial i_1 \cdots \partial i_{h+1} \partial i_h \cdots \partial i_r}$$

llamemos φ, F y G a las siguientes derivadas parciales:

$$\varphi = \frac{\partial^{h-1} f}{\partial i_1 \cdots \partial i_{h-1}}, \qquad F = \frac{\partial^2 \varphi}{\partial i_h \partial i_{h+1}} \qquad \text{y} \qquad G = \frac{\partial^2 \varphi}{\partial i_{h+1} \partial i_h}$$

con lo que la relación [1], que debíamos demostrar, toma la forma

$$\frac{\partial^{r-h-1} F}{\partial i_{h+2} \cdots \partial i_r} = \frac{\partial^{r-h-1} G}{\partial i_{h+2} \cdots \partial i_r} \tag{2}$$

Ahora bien, la función φ es de clase $r - h + 1 \geqslant 1 + 1 = 2$ y, por ello, para φ se verifican (sobradamente) las hipótesis de los anteriores teoremas, luego (considerando a φ como función de las variables de lugares i_h y de i_{h+1}; las demás variables permanecen constantes en este proceso de derivación parcial) son iguales sus derivadas parciales cruzadas F y G. Por ser $F = G$, es obvio que se verifica [2], como debíamos comprobar.

[31]$_1$ Ejercicio

Acudir a la función $f\colon \mathbb{R}^2 \to \mathbb{R}$, que se define abajo, para comprobar que puede verificarse la igualdad $f''_{xy}(a, b) = f''_{yx}(a, b)$ aun en el caso de ser f''_{xy} y f''_{yx} discontinuas en (a, b):

$$f(x, y) = \begin{cases} xy^2 \operatorname{sen}(1/y), & \text{si } y \neq 0 \\ 0, & \text{si } y = 0 \end{cases}$$

Resolución

Es fácil comprobar que

$$f'_x(x, y) = y^2 \operatorname{sen}(1/y), \qquad \text{si } y \neq 0$$

$$f'_y(x, 0) = \lim_{k \to 0} \frac{f(x, k) - f(x, 0)}{k} = \lim_{k \to 0} xk \operatorname{sen}(1/k) = 0$$

$$f''_{xy}(0, 0) = \lim_{k \to 0} \frac{f'_x(0, k) - f'_x(0, 0)}{k} = \lim_{k \to 0} k \operatorname{sen}(1/k) = 0$$

$$f''_{yx}(0, 0) = \lim_{h \to 0} \frac{f'_y(h, 0) - f'_y(0, 0)}{h} = \lim_{h \to 0} 0 = 0$$

$$f''_{xy}(x, y) = f''_{yx}(x, y) = 2y \operatorname{sen}(1/y) - \cos(1/y), \qquad \text{si } y \neq 0$$

Resulta entonces que $f''_{xy}(0, 0) = f''_{yx}(0, 0)$ a pesar de que f''_{xy} y f''_{yx} no son continuas en (0, 0).

❑ DIFERENCIALES DE ORDEN SUPERIOR AL PRIMERO

La diferencial $df(\boldsymbol{a})$, que también llamaremos diferencial primera, de una función f en un punto $\boldsymbol{a}$, es la función lineal (su variable es el incremento $\Delta\boldsymbol{x} = \boldsymbol{x} - \boldsymbol{a}$) que mejor aproxima al incremento de la función $\Delta f = f(\boldsymbol{x}) - f(\boldsymbol{a})$, cuando se pasa del punto $\boldsymbol{a}$ al punto próximo $\boldsymbol{x}$. Para mejorar la aproximación de Δf (cerca de $\boldsymbol{a}$) se recurre a un segundo sumando, que se añade a df, que varía de forma cuadrática (respecto del incremento); para la obtención de este término cuadrático se introduce la diferencial segunda d^2f. Para una mejor aproximación local de f, se recurre a un tercer sumando, de tipo cúbico, que se obtiene a partir de la diferencial tercera d^3f y así sucesivamente.

Las diferenciales sucesivas (de orden mayor que 1) las vamos a necesitar nosotros para pocas cosas: además de su utilidad para obtener aproximaciones (locales) de las funciones, las usaremos como elementos de tipo formal que facilitan ciertos procesos de cálculo y proporcionan una notación que simplifica algunas expresiones. No ahondaremos demasiado en este asunto, obviaremos la definición de función sucesivamente diferenciable y nos limitaremos a hablar de las diferenciales sucesivas (de orden $r \geqslant 2$), suponiendo que la función cumple alguna condición suficiente (ser de clase $\mathscr{C}^r$) que dé sentido y utilidad a la diferencial (de orden r).

[32]

Sea $f: C \to \mathbb{R}^q$ una función definida en un abierto $C \subset \mathbb{R}^p$ y considérese un punto $\boldsymbol{a} \in C$.

I. • Si f es de clase $\mathscr{C}^2$ en el punto $\boldsymbol{a}$, se llama diferencial segunda de f en $\boldsymbol{a}$ a la siguiente aplicación cuadrática $d^2f(\boldsymbol{a})$, de $\mathbb{R}^p$ en $\mathbb{R}^q$:

$$\boldsymbol{h} \mapsto d^2f(\boldsymbol{a})(\boldsymbol{h}) = \sum_{i,j=1}^{p} f''_{ij}(\boldsymbol{a})h_ih_j \qquad (*)$$

(donde $\boldsymbol{h} = (h_1, h_2, ..., h_p)$ es el incremento arbitrario de la variable).

• Si f es de clase $\mathscr{C}^r$ en el punto $\boldsymbol{a}$, se llama diferencial de orden r de f en $\boldsymbol{a}$ a la siguiente aplicación $d^rf(\boldsymbol{a}): \mathbb{R}^p \to \mathbb{R}^q$:

$$\boldsymbol{h} \mapsto d^rf(\boldsymbol{a})(\boldsymbol{h}) = \sum_{i_1,\dots,i_r=1}^{p} f^{(r)}_{i_1\cdots i_r}(\boldsymbol{a})h_{i_1}\cdots h_{i_r} \qquad (*)$$

II. Si la función f es de clase $\mathscr{C}^r$ en el punto $\boldsymbol{a}$ y para cualquiera que sea el vector no nulo $\boldsymbol{u} \in \mathbb{R}^p$, entonces existe la derivada $\partial^r f(\boldsymbol{a})/\partial \boldsymbol{u}^r$, la cual vale

$$\frac{\partial^r f(\boldsymbol{a})}{\partial \boldsymbol{u}^r} = d^rf(\boldsymbol{a})(\boldsymbol{u})$$

(*) Si la variable es $\boldsymbol{x} = (x_1, ..., x_p)$, a su incremento se le representa, con frecuencia, poniendo $\boldsymbol{h} = d\boldsymbol{x}$ (diferencial de $\boldsymbol{x}$; véase la justificación de ello en $[24]_3$), con lo que se pondrá:

$$d^2f(\boldsymbol{a})(d\boldsymbol{x}) = \sum \frac{\partial^2 f(\boldsymbol{a})}{\partial x_i \partial x_j} dx_i dx_j, \qquad d^3f(\boldsymbol{a}) = \sum \frac{\partial^3 f(\boldsymbol{a})}{\partial x_i \partial x_j \partial x_k} dx_i dx_j dx_k, \qquad \text{etc.}$$

Estas expresiones se suelen representar, recurriendo a la siguiente notación «potencia simbólica», poniendo:

$$d^rf(\boldsymbol{a})(d\boldsymbol{x}) = \left(\frac{\partial f}{\partial x_1} dx_1 + \cdots + \frac{\partial f}{\partial x_p} dx_p\right)^{(r)}(\boldsymbol{a}), \qquad \text{para } r \in \mathbb{N}$$

Antes de hacer los oportunos comentarios a las anteriores definiciones, probemos la propiedad II, que se acaba de enunciar.

Comprobación

Procedamos por inducción sobre r. Para $r = 1$ la propiedad es cierta, según ya se comprobó en [26],1.º. Suponiendo, pues, que la propiedad es cierta para las derivadas de lugar $r - 1$, hemos de comprobar que la relación se verifica para el lugar r. Así es, ya que llamando $\boldsymbol{u} = (u_1, ..., u_p)$:

$$\frac{\partial^r f(\boldsymbol{a})}{\partial \boldsymbol{u}^r} = \frac{\partial^{r-1}[f'_{\boldsymbol{u}}(\boldsymbol{a})]}{\partial \boldsymbol{u}^{r-1}} = \sum_{i_1,\dots,i_{r-1}=1}^{p} \frac{\partial^{r-1}(f'_{\boldsymbol{u}}(\boldsymbol{a}))}{\partial i_1 \cdots \partial i_{r-1}} u_{i_1}\cdots u_{i_{r-1}} =$$

$$= \sum_{i_1,\dots,i_{r-1}=1}^{p} \frac{\partial^{r-1}\left(\sum_{i_r=1}^{p} f'_{i_r}(\boldsymbol{a})u_{i_r}\right)}{\partial i_1 \cdots \partial i_{r-1}} u_{i_1}\cdots u_{i_{r-1}} = \sum_{i_1,\dots,i_r=1}^{p} \frac{\partial^r f(\boldsymbol{a})}{\partial i_1 \cdots \partial i_r} u_1 \cdots u_r = d^rf(\boldsymbol{a})(\boldsymbol{u})$$

[32]$_1$ Observaciones

1.º Nosotros hemos definido a la diferencial segunda, de f en $\boldsymbol{a}$, como una aplicación cuadrática $\boldsymbol{h} \mapsto d^2 f(\boldsymbol{a})(\boldsymbol{h})$, de $\mathbb{R}^p$ en $\mathbb{R}^q$. Más preciso habría sido definir la diferencial segunda como la aplicación bilineal de la que se deduce la anterior aplicación cuadrática; esto es, en lugar de la expresión de $d^2 f(\boldsymbol{a})$ que se da en [32], se habría definido $d^2 f(\boldsymbol{a})$ como la aplicación de $\mathbb{R}^p \times \mathbb{R}^p$ en $\mathbb{R}^q$ dada por:

$$(\boldsymbol{h}, \boldsymbol{k}) \mapsto d^2 f(\boldsymbol{a})(\boldsymbol{h}, \boldsymbol{k}) = \sum_{i,j=1}^{p} f''_{ij}(\boldsymbol{a}) h_i k_j$$

donde $\boldsymbol{h} = (h_1, ..., h_p)$ y $\boldsymbol{k} = (k_1, ..., k_p)$ son dos incrementos arbitrarios de la variable. Otro tanto se puede decir de las diferenciales de mayores órdenes:

$$(\boldsymbol{h}^1, ..., \boldsymbol{h}^r) \mapsto d^r f(\boldsymbol{a})(\boldsymbol{h}^1, ..., \boldsymbol{h}^r) = \sum_{i_1, ..., i_r = 1}^{p} f^{(r)}_{i_1 ... i_r}(\boldsymbol{a}) h^1_{i_1} \cdots h^r_{i_r}$$

2.º De hacer las cosas como se acaba de indicar en el apartado precedente, la fórmula [32], II se puede generalizar y, razonando como en la anterior demostración, obtener que si f es de clase $\mathscr{C}^r$ en $\boldsymbol{a}$ y para cualesquiera vectores no nulos $\boldsymbol{u}_1, ..., \boldsymbol{u}_r$ de $\mathbb{R}^p$, entonces

$$\frac{\partial^r f(\boldsymbol{a})}{\partial \boldsymbol{u}_1 \cdots \partial \boldsymbol{u}_r} = d^r f(\boldsymbol{a})(\boldsymbol{u}_1, ..., \boldsymbol{u}_r)$$

3.º No cuesta ningún trabajo comprobar que, al igual que ocurre con la diferencial primera (véase [25], 2.º), se verifica que si $f: C \to \mathbb{R}^q$ es una función de clase $\mathscr{C}^r$ en el punto $\boldsymbol{a}$ y si las componentes de f son $f_1, ..., f_q$ (funciones de C en $\mathbb{R}$), entonces las componentes de la diferencial $d^r f(\boldsymbol{a})$, de orden r, son las diferenciales $d^r f_i(\boldsymbol{a})$, para $i = 1, 2, ..., q$; es decir:

$$f = (f_1, ..., f_q) \quad \Rightarrow \quad d^r f(\boldsymbol{a}) = \big(d^r f_1(\boldsymbol{a}), ..., d^r f_q(\boldsymbol{a})\big)$$

4.º Aunque sólo sea someramente, digamos algo acerca del concepto de función dos veces diferenciable. Sea $f: C \to \mathbb{R}^q$, donde $C \subset \mathbb{R}^p$ es abierto, una función diferenciable en C; se dice que f es dos veces diferenciable en $\boldsymbol{a} \in C$ si la diferencial df, que es una función de C en $\mathscr{L}(\mathbb{R}^p, \mathbb{R}^q)$ (espacio vectorial de las aplicaciones lineales de $\mathbb{R}^p$ en $\mathbb{R}^q$), es diferenciable en $\boldsymbol{a}$; si así ocurre, se llama diferencial segunda de f en $\boldsymbol{a}$ a la aplicación lineal, que se representa poniendo $d^2 f(\boldsymbol{a})$, definida por

$$d^2 f(a) = d(df)(\boldsymbol{a})$$

que es una aplicación lineal de $\mathbb{R}^p$ en $\mathscr{L}(\mathbb{R}^p, \mathbb{R}^q)$. Así pues, si f es dos veces diferenciable en $\boldsymbol{a}$, la diferencial segunda de f en $\boldsymbol{a}$ es la diferencial de df en $\boldsymbol{a}$.

5.º Las expresiones con las que se definen las diferenciales sucesivas, en [32], tienen desarrollos excesivamente largos. No obstante, vamos a explicitar dichos desarrollos para el caso de una función de dos variables. Suponemos que $(x, y) \mapsto z(x, y)$ es una función de clase $\mathscr{C}^r$ suficientemente avanzada ($r \geqslant 3$) en un punto (x, y), lo que dará validez a los siguientes desarrollos. Las diferenciales primera, segunda y tercera, de z en (x, y), correspondientes a un incremento (dx, dy) de las variables, son:

$$dz(x, y)(dx, dy) = \frac{\partial z(x, y)}{\partial x}\, dx + \frac{\partial z(x, y)}{\partial y}\, dy$$

$$d^2z(x, y)(dx, dy) = \left(\frac{\partial z}{\partial x}\, dx + \frac{\partial z}{\partial y}\, dy\right)^{(2)}(x, y) =$$

$$= \frac{\partial^2 z(x, y)}{\partial x^2}\, dx^2 + 2\,\frac{\partial^2 z(x, y)}{\partial x \partial y}\, dx\, dy + \frac{\partial^2 z(x, y)}{\partial y^2}\, dy^2$$

$$d^3z(x, y)(dx, dy) = \left(\frac{\partial z}{\partial x}\, dx + \frac{\partial z}{\partial y}\, dy\right)^{(3)}(x, y) =$$

$$= \frac{\partial^3 z(x, y)}{\partial x^3}\, dx^3 + 3\,\frac{\partial^3 z(x, y)}{\partial x^2 \partial y}\, dx^2 dy + 3\,\frac{\partial^3 z(x, y)}{\partial x \partial y^2}\, dx\, dy^2 + \frac{\partial^3 z(x, y)}{\partial y^3}\, dy^3$$

[32]$_2$ Ejercicio

Obtener las diferencias primera, segunda y tercera de la siguiente función $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ en el punto $(0, 0)$

$$f(x, y) = (e^{-x+2y}, (x + 2)^2 \operatorname{sen} y)$$

Resolución

Las derivadas parciales de f, hasta el tercer orden, son (téngase en cuenta que la función es de clase $\mathscr{C}^\infty$, por lo que las derivadas cruzadas son iguales):

$f'_x(x, y) = (-e^{-x+2y}, 2(x + 2) \operatorname{sen} y)$ $\qquad$ $f'_y(x, y) = (2e^{-x+2y}, (x + 2)^2 \cos y)$

$f''_{x^2}(x, y) = (e^{-x+2y}, 2 \operatorname{sen} y)$ $\qquad$ $f''_{xy}(x, y) = (-2e^{-x+2y}, 2(x + 2) \cos y)$

$f''_{y^2}(x, y) = (4e^{-x+2y}, -(x + 2)^2 \operatorname{sen} y)$ $\qquad$ $f'''_{x^3}(x, y) = (-e^{-x+2y}, 0)$

$f'''_{x^2y}(x, y) = (2e^{-x+2y}, 2 \cos y)$ $\qquad$ $f'''_{xy^2}(x, y) = (-4e^{-x+2y}, -2(x + 2) \operatorname{sen} y)$

$f'''_{y^3}(x, y) = (8e^{-x+2y}, -(x + 2)^2 \cos y)$

y, en particular, las derivadas parciales en $(0, 0)$ son:

$f'_x(0, 0) = (-1, 0)$, $\quad f'_y(0, 0) = (2, 4)$, $\quad f''_{x^2}(0, 0) = (1, 0)$, $\quad f''_{xy}(0, 0) = (-2, 4)$, $\quad f''_{y^2}(0, 0) = (4, 0)$

$f'''_{x^3}(0, 0) = (-1, 0)$, $\quad f'''_{x^2y}(0, 0) = (2, 2)$, $\quad f'''_{xy^2}(0, 0) = (-4, 0)$, $\quad f'''_{y^3}(0, 0) = (8, 4)$

Por tanto, las diferenciales primera, segunda y tercera de f en (0, 0) son las aplicaciones:

- $df(0, 0)$: $(dx, dy) \mapsto df(0, 0)(dx, dy) = (-dx + 2dy, 4dy)$
- $d^2f(0, 0)$: $(dx, dy) \mapsto d^2f(0, 0)(dx, dy) = (dx^2 - 4dx\,dy + 4dy^2, 8dx\,dy)$
- $d^3f(0, 0)$: $(dx, dy) \mapsto d^3f(0, 0)(dx, dy) = (-dx^3 + 6dx^2dy - 12dx\,dy^2 + 8dy^3, 6dx^2dy + 4dy^3)$

FÓRMULA Y DESARROLLO LIMITADO DE TAYLOR

No hace mucho, cuando empezábamos a hablar de las diferenciales de orden superior al primero, decíamos que de una función, si era suficientemente regular, se podrían obtener sucesivas aproximaciones de ella, en las cercanías de un punto, mediante polinomios. Las cosas funcionan de modo que, a medida que aumenta el grado del polinomio, se mejora la aproximación; de una aproximación (de grado n) se pasa a la siguiente (de grado $n + 1$), añadiendo, a aquella, un nuevo sumando (polinomio homogéneo de grado $n + 1$) que es, salvo una constante de proporcionalidad, el valor de la diferencial del correspondiente orden (orden $n + 1$).

[33]

TEOREMA DE TAYLOR. Sea $f: C \to \mathbb{R}$ una función real, definida en un abierto $C \subset \mathbb{R}^p$, y sean $\boldsymbol{a}$ y $\boldsymbol{b} = \boldsymbol{a} + \boldsymbol{h}$ dos puntos de C tales que el segmento $[\boldsymbol{a}, \boldsymbol{b}]$, que tiene a $\boldsymbol{a}$ y $\boldsymbol{b}$ por extremos, está incluido en C. Si f es de clase $\mathscr{C}^n$ en todos los puntos de $[\boldsymbol{a}, \boldsymbol{b}]$, entonces existe un punto $\xi \in]\boldsymbol{a}, \boldsymbol{b}[$ tal que:

$$f(\boldsymbol{a} + \boldsymbol{h}) = f(\boldsymbol{a}) + \frac{1}{1!}\,df(\boldsymbol{a})(\boldsymbol{h}) + \frac{1}{2!}\,d^2f(\boldsymbol{a})(\boldsymbol{h}) + \cdots + \frac{1}{(n-1)!}\,d^{n-1}f(\boldsymbol{a})(\boldsymbol{h}) + r_n(\boldsymbol{h})$$

donde

$$r_n(\boldsymbol{h}) = \frac{1}{n!}\,d^nf(\xi)(\boldsymbol{h})^{(*)} \qquad \begin{pmatrix}\text{Fórmula de Taylor;}\\ \text{resto de Lagrange}\end{pmatrix}$$

DESARROLLO LIMITADO DE TAYLOR. Sea $\boldsymbol{f}: C \to \mathbb{R}^q$ una función, definida en un abierto $C \subset \mathbb{R}^p$, y sea $\boldsymbol{a}$ un punto de C. Si $\boldsymbol{f}$ es de clase $\mathscr{C}^n$ en C, entonces se verifica que:

$$\boldsymbol{f}(\boldsymbol{a} + \boldsymbol{h}) = \boldsymbol{f}(\boldsymbol{a}) + \frac{1}{1!}\,d\boldsymbol{f}(\boldsymbol{a})(\boldsymbol{h}) + \frac{1}{2!}\,d^2\boldsymbol{f}(\boldsymbol{a})(\boldsymbol{h}) + \cdots + \frac{1}{n!}\,d^n\boldsymbol{f}(\boldsymbol{a})(\boldsymbol{h}) + \boldsymbol{o}[\|\boldsymbol{h}\|^n]$$

(*) El «punto intermedio» ξ se puede poner $\xi = \boldsymbol{a} + \theta\boldsymbol{h}$ para cierto $\theta \in]0, 1[$. Este «resto n-ésimo» puede expresarse (renunciando al punto intermedio ξ) en la forma:

$$r_n(\boldsymbol{h}) = \frac{1}{n!}\,d^nf(\boldsymbol{a})(\boldsymbol{h}) + o[\|\boldsymbol{h}\|^n]$$

Demostración

1.º Teorema de Taylor. Sea $\varphi : [0, 1] \to \mathbb{R}$ la función definida por la relación

$$t \mapsto \varphi(t) = f(\boldsymbol{a} + t\boldsymbol{h})$$

(esta función existe ya que $\boldsymbol{a} + t\boldsymbol{h} \in [\boldsymbol{a}, \boldsymbol{b}]$ para $0 < t < 1$). Si recurrimos a la propiedad [32], II (ahora, en lugar del punto $\boldsymbol{a}$ y del vector $\boldsymbol{u}$, se consideran el punto $\boldsymbol{a} + t\boldsymbol{h}$ y el vector $\boldsymbol{h}$), podemos poner que para $t \in [0, 1]$ es:

$$\varphi'(t) = \lim_{\lambda \to 0} \frac{f(\boldsymbol{a} + t\boldsymbol{h} + \lambda\boldsymbol{h}) - f(\boldsymbol{a} + t\boldsymbol{h})}{\lambda} = \frac{\partial f(\boldsymbol{a} + t\boldsymbol{h})}{\partial \boldsymbol{h}} = df(\boldsymbol{a} + t\boldsymbol{h})(\boldsymbol{h})$$

$$\varphi''(t) = \lim_{\lambda \to 0} \frac{D_{\boldsymbol{h}} f(\boldsymbol{a} + t\boldsymbol{h} + \lambda\boldsymbol{h}) - D_{\boldsymbol{h}} f(\boldsymbol{a} + t\boldsymbol{h})}{\lambda} = \frac{\partial^2 f(\boldsymbol{a} + t\boldsymbol{h})}{\partial \boldsymbol{h}^2} = d^2 f(\boldsymbol{a} + t\boldsymbol{h})(\boldsymbol{h})$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\varphi^{(n)}(t) = \lim_{\lambda \to 0} \frac{D^{n-1}_{\boldsymbol{h}^{n-1}} f(\boldsymbol{a} + t\boldsymbol{h} + \lambda\boldsymbol{h}) - D^{n-1}_{\boldsymbol{h}^{n-1}} f(\boldsymbol{a} + t\boldsymbol{h})}{\lambda} = \frac{\partial^n f(\boldsymbol{a} + t\boldsymbol{h})}{\partial \boldsymbol{h}^n} = d^n f(\boldsymbol{a} + t\boldsymbol{h})(\boldsymbol{h})$$

Así pues, φ es n veces derivable en el intervalo $[0, 1]$, por lo que le es de aplicación, en $[0, 1]$, la fórmula de Taylor (para funciones de una variable) con término de Lagrange en el lugar de la derivada n-ésima(*), es decir, se verifica que: existe $\theta \in]0, 1[$, tal que

$$\varphi(1) = \varphi(0) + \frac{1}{1!}\varphi'(0) + \frac{1}{2!}\varphi''(0) + \cdots + \frac{1}{(n-1)!}\varphi^{(n-1)}(0) + \frac{1}{n!}\varphi^{(n)}(\theta)$$

Llevando a esta última igualdad las expresiones antes obtenidas para $\varphi(t), \varphi'(t), \ldots, \varphi^{(n)}(t)$ (en las que se toma ahora $t = 1$, $t = 0$ y $t = \theta$) se llega a que:

$$f(\boldsymbol{a} + \boldsymbol{h}) = f(\boldsymbol{a}) + \frac{1}{1!} df(\boldsymbol{a})(\boldsymbol{h}) + \frac{1}{2!} d^2 f(\boldsymbol{a})(\boldsymbol{h}) + \cdots + \frac{1}{(n-1)!} d^{n-1} f(\boldsymbol{a})(\boldsymbol{h}) +$$
$$+ \frac{1}{n!} d^n f(\xi)(\boldsymbol{h}), \qquad \text{donde } \xi = \boldsymbol{a} + \theta\boldsymbol{h}$$

como debíamos comprobar (nótese que el punto $\xi = \boldsymbol{a} + \theta\boldsymbol{h}$, donde $\theta \in]0, 1[$, es un punto del segmento $[\boldsymbol{a}, \boldsymbol{b}]$, cuyos extremos son $\boldsymbol{a}$ y $\boldsymbol{b} = \boldsymbol{a} + \boldsymbol{h}$).

(*) Para funciones de una variable, la fórmula de Taylor (con término complementario de Lagrange) dice: Si $F : [\alpha, \beta] \to \mathbb{R}$ es una función de clase $\mathscr{C}^n$ en un intervalo compacto $[\alpha, \beta] \subset \mathbb{R}$, entonces existe un $\theta \in]\alpha, \beta[$ tal que

$$F(\beta) = F(\alpha) + (\beta - \alpha)\frac{F'(\alpha)}{1!} + (\beta - \alpha)^2 \frac{F''(\alpha)}{2!} + \cdots + (\beta - \alpha)^{n-1} \frac{F^{(n-1)}(\alpha)}{(n-1)!} + (\beta - \alpha)^n \frac{F^{(n)}(\theta)}{n!}$$

2.° Desarrollo limitado de Taylor en el «caso escalar» ($q = 1$). Como C es abierto y $\boldsymbol{a} \in C$, existe un cierto entorno de $\boldsymbol{a}$ que está incluido en C; sea $\delta > 0$ el radio de este entorno. Por tanto, a f le es ahora de aplicación el resultado anterior con tal de tomar $\boldsymbol{h}$ de modo que sea $\|\boldsymbol{h}\| < \delta$. Por otra parte, en relación con esta fórmula y respecto del resto n-ésimo, $r_n(\boldsymbol{h})$, nos interesa observar que, como las derivadas parciales n-ésimas de f son continuas en $\boldsymbol{a}$, se verifica que

$$\lim_{h \to o} \frac{d^n f(\xi)(\boldsymbol{h}) - d^n f(\boldsymbol{a})(\boldsymbol{h})}{\|\boldsymbol{h}\|^n} = \lim_{h \to o} \sum_{i_1, \ldots, i_n = 1}^{p} \left[\frac{\partial^n f(\xi)}{\partial i_1 \cdots \partial i_n} - \frac{\partial^n f(\boldsymbol{a})}{\partial i_1 \cdots \partial i_n} \right] \frac{h_{i_1}}{\|\boldsymbol{h}\|} \cdots \frac{h_{i_n}}{\|\boldsymbol{h}\|} = o$$

(nótese que $h_{i_1}/\|\boldsymbol{h}\|, \ldots, h_{i_n}/\|\boldsymbol{h}\|$ están acotados entre -1 y 1), por lo que

$$r_n(\boldsymbol{h}) = \frac{1}{n!} d^n f(\xi)(\boldsymbol{h}) = \frac{1}{n!} d^n f(\boldsymbol{a})(\boldsymbol{h}) + o[\|h\|^n]$$

llevando este resultado a la anterior fórmula de Taylor, se obtiene el «desarrollo limitado» que se quería probar.

3.° Desarrollo limitado de Taylor en el «caso vectorial» ($q \geqslant 2$). Si las componentes de la función $\boldsymbol{f}: C \to \mathbb{R}^q$ son $f_1, f_2, \ldots, f_q$ (funciones de C en $\mathbb{R}$), como estas funciones son reales y, para cada una de ellas, se cumplen las hipótesis del anterior apartado 2.°, se oconcluye que para $j = 1, 2, \ldots, q$ es

$$f_j(\boldsymbol{a} + \boldsymbol{h}) = f_j(\boldsymbol{a}) + \frac{1}{1!} df_j(\boldsymbol{a})(\boldsymbol{h}) + \frac{1}{2!} d^2 f_j(\boldsymbol{a})(\boldsymbol{h}) + \cdots + \frac{1}{n!} d^n f_j(\boldsymbol{a})(\boldsymbol{h}) + o[\|h\|^n] \qquad [1]$$

lo que prueba la validez del desarrollo limitado de Taylor que se da en el enunciado, para la función $\boldsymbol{f}$, pues el resultado obtenido en [1], que es válido para $j = 1, 2, \ldots, q$, es la componente j-ésima del referido desarrollo.

[33]₁ Observación

En la fórmula de Taylor con resto de Lagrange, de una función real $f: C \to \mathbb{R}$ (donde $C \subset \mathbb{R}^p$), según la cual (y si se cumplen las hipótesis del teorema de Taylor; véase [33]) existe $\xi \in]\boldsymbol{a}, \boldsymbol{a} + \boldsymbol{h}[$ tal que

$$f(\boldsymbol{a} + \boldsymbol{h}) - f(\boldsymbol{a}) = \sum_{k=1}^{n-1} \frac{1}{k!} d^k f(\boldsymbol{a})(\boldsymbol{h}) + \frac{1}{n!} d^n f(\xi)(\boldsymbol{h})$$

el sumando k-ésimo del segundo miembro es un polinomio homogéneo de grado k, en las variables $h_1, h_2, \ldots, h_p$ (coordenadas del incremento $\boldsymbol{h}$ de la variable independiente):

$$\frac{1}{k!} d^k f(\boldsymbol{a})(\boldsymbol{h}) = \frac{1}{k!} (f_1'(\boldsymbol{a})h_1 + f_2'(\boldsymbol{a})h_2 + \cdots + f_p'(\boldsymbol{a})h_p)^{(k)} = \frac{1}{k!} \sum_{i_1, \ldots, i_k = 1}^{p} f_{i_1 \cdots i_k}^{(k)}(\boldsymbol{a}) h_{i_1} \cdots h_{i_k}$$

Así, en el caso de ser $p = 2$ (dos variables independientes x e y), llamando (a, b) al punto fijo y (h, k) al incremento, la anterior fórmula de Taylor será:

$$\begin{aligned}
f(a+h, b+k) - f(a, b) &= \big(f'_x(a, b)h + f'_y(a, b)k\big) + \\
&+ \frac{1}{2!}\left(f''_{x^2}(a, b)h^2 + f''_{xy}(a, b)hk + f''_{y^2}(a, b)k^2\right) + \\
&+ \frac{1}{3!}\left(f'''_{x^3}(a, b)h^3 + 3f'''_{x^2y}(a, b)h^2k + 3f'''_{xy^2}(a, b)hk^2 + f'''_{y^3}(a, b)k^3\right) + \\
&+ \frac{1}{(n-1)!}\left(\binom{n-1}{0} f^{(n-1)}_{x^{n-1}}(a, b)h^{n-1} + \binom{n-1}{1} f^{(n-1)}_{x^{n-2}y}(a, b)h^{n-2}k + \cdots + \right. \\
&\left. + \binom{n-1}{n-1} f^{(n-1)}_{y^{n-1}}(a, b)k^{n-1}\right) + \\
&+ \frac{1}{n!}\left(\binom{n}{0} f^{(n)}_{x^n}(\xi, \zeta)h^n + \binom{n}{1} f^{(n)}_{x^{n-1}y}(\xi, \zeta)h^{n-1}k + \cdots + \binom{n}{n} f^{(n)}_{y^n}(\xi, \zeta)k^n\right)
\end{aligned}$$

donde $\xi = a + \theta h$ y $\zeta = b + \theta k$ para cierto $\theta \in]0, 1[$.

$[33]_2$ Funciones de clase $\mathscr{C}^\infty$ y funciones analíticas

Consideremos una función real $f\colon U \to \mathbb{R}$, definida en un entorno $U \subset \mathbb{R}^p$ de un punto $\boldsymbol{a} \in \mathbb{R}^p$. Se dice que f es de clase $\mathscr{C}^\infty$ en U si f es de clase $\mathscr{C}^n$ en U para todo $n \in \mathbb{N}$. Si f es de clase $\mathscr{C}^\infty$ en U, entonces para cualquiera que sea el punto $\boldsymbol{x} = \boldsymbol{a} + \boldsymbol{h}$ de U (esto es, si $\|\boldsymbol{h}\|$ es menor que el radio ρ de U), a f le es de aplicación en $[\boldsymbol{a}, \boldsymbol{x}]$ la fórmula de Taylor [33] para todo $n \in \mathbb{N}$; de manera que existe $\xi_n \in]\boldsymbol{a}, \boldsymbol{x}[$ (que depende de n) tal que:

$$f(\boldsymbol{x}) = f(\boldsymbol{a}) + \frac{1}{1!}\, df(\boldsymbol{a})(\boldsymbol{h}) + \frac{1}{2!}\, d^2 f(\boldsymbol{a})(\boldsymbol{h}) + \cdots + \frac{1}{(n-1)!}\, d^{n-1} f(\boldsymbol{a})(\boldsymbol{h}) + r_n(\boldsymbol{h})$$

$$\text{donde } r_n(\boldsymbol{h}) = \frac{1}{n!}\, d^n f(\xi_n)(\boldsymbol{h}) \qquad \text{(válida para todo } n \in \mathbb{N})$$

Si además se verifica que para cualquiera que sea $\boldsymbol{x} \in U$ el resto n-ésimo tiene límite 0 cuando n tiende a infinito, es decir, si

$$\lim_{n \to \infty} r_n(\boldsymbol{h}) = 0, \qquad \forall\, \boldsymbol{x} \in U \text{ (o sea, si } \|\boldsymbol{h}\| < \rho)$$

se dice que f es una función analítica en $\boldsymbol{a}$. En este supuesto, tomando límites, para $n \to \infty$, en ambos miembros de la fórmula de Taylor, resulta que

$$f(\boldsymbol{x}) = f(\boldsymbol{a}) + \lim_{n \to \infty} \sum_{k=1}^{n} \frac{1}{k!}\, d^k f(\boldsymbol{a})(\boldsymbol{h}) + \lim_{n \to \infty} r_n(\boldsymbol{h}) = f(\boldsymbol{a}) + \sum_{k=1}^{\infty} \frac{1}{k!}\, d^k f(\boldsymbol{a})(\boldsymbol{h})$$

$$\text{para todo } \boldsymbol{x} = \boldsymbol{a} + \boldsymbol{h} \text{ de } U \qquad \text{(o sea, si } \|\boldsymbol{h}\| < \rho)$$

El último miembro es otro modo de expresar el que le precede, y se dice que es una serie o suma de infinitos sumandos; en concreto, a esta serie se la llama serie de Taylor de f en torno del punto $\boldsymbol{a}$. Las funciones analíticas son, pues, aquellas que pueden «desarrollarse» en serie de Taylor.

$[33]_3$ Ejercicio

Sea $f: U \to \mathbb{R}$ una función real, definida en un entorno $U \subset \mathbb{R}^p$ de un punto $\boldsymbol{a} \in \mathbb{R}^p$. Sabiendo que f es de clase $\mathscr{C}^\infty$ en U y si existe una constante $K > 0$ tal que todas las derivadas parciales de orden n de f están acotadas en U por K^n, o sea, si

$$\left| \frac{\partial^n f(\boldsymbol{x})}{\partial x_i \cdots \partial i_n} \right| \leqslant K^n, \qquad \forall\, n \in \mathbb{N},\ \forall\, \boldsymbol{x} \in U \qquad [1]$$

pruébese que entonces f es analítica en $\boldsymbol{a}$.

Resolución

Sea ρ el radio del entorno U, con lo que llamando $h_1, \ldots, h_p$ a las componentes del incremento $\boldsymbol{h} = \boldsymbol{x} - \boldsymbol{a}$, si $\boldsymbol{x} \in U$ entonces $|h_1| \leqslant \rho, \ldots, |h_p| \leqslant \rho$. Nótese también que el «punto intermedio» ξ_n, del resto n-ésimo de la fórmula de Taylor de f en $\boldsymbol{a}$, pertenece al entorno U, siempre que pertenezca $\boldsymbol{x}$. Por tanto, si $\boldsymbol{x} \in U$ y si $n \in \mathbb{N}$, para dicho resto n-ésimo se tiene:

$$|r_n(\boldsymbol{h})| = \frac{1}{n!}\, |d^n f(\xi_n)| \leqslant \frac{1}{n!} \sum_{i_1, \ldots, i_n = 1}^{p} \left| \frac{\partial^n f(\xi_n)}{\partial i_1 \cdots \partial i_n} \right| |h_{i_1}| \cdots |h_{i_n}| <$$

$$< \frac{1}{n!}\, K^n \rho^n p^n = \frac{H^n}{n!} \qquad (\text{donde } H = K\rho p)$$

Tomando límites cuando $n \to \infty$, como $H^n/n!$ tiende a 0, resulta que

$$\lim_{n \to \infty} r_n(\boldsymbol{h}) = 0, \qquad \forall\, \boldsymbol{x} \in U$$

como había que comprobar.

$[33]_4$ Ejercicio

Sea $f(x, y, z)$ un polinomio de segundo grado; la superficie que tiene por ecuación a $f(x, y, z) = 0$ se llama cuádrica. Se pide: 1.°, obtener la ecuación de la cuádrica en los ejes $x'y'z'$ que resultan de trasladar los ejes xyz al punto (a, b, c) (nuevo origen de coordenadas); 2.° indicar cómo se podría hallar el centro de la cuádrica, esto es, un punto que sea centro de simetría de la misma.

Resolución

1.° A la función f le es de aplicación la fórmula de Taylor, centrada en el punto (a, b, c), con resto en el lugar de las derivadas terceras. Como f es un polinomio de segundo grado, tiene nulas todas sus derivadas parciales terceras, con lo que también es nulo el citado resto y, en consecuencia, la ecuación $f(x, y, z) = 0$, es decir, la ecuación $f(a + x', b + y', c + z') = 0$ toma la forma:

$$f(a, b, c) + x'f'_x(a, b, c) + y'f'_y(a, b, c) + z'f'_z(a, b, c) + x'^2 f''_{x^2}(a, b, c) + y'^2 f''_{y^2}(a, b, c) +$$
$$+ z'^2 f''_{z^2}(a, b, c) + 2x'y'f''_{xy}(a, b, c) + 2x'z'f''_{xz}(a, b, c) + 2y'z'f''_{yz}(a, b, c) = 0$$

2.º Para que un punto (a, b, c) sea el centro de la cuádrica, tiene que ocurrir que la ecuación de la cuádrica (en las coordenadas $x'y'z'$) no tenga términos de primer grado (al cambiar x', y', z' por $-x'$, $-y'$, $-z'$ la ecuación ha de seguir verificándose). Por tanto, el centro es el punto (a, b, c) solución (si existe) del sistema de ecuaciones:

$$f'_x(a, b, c) = 0, \qquad f'_y(a, b, c) = 0, \qquad f'_z(a, b, c) = 0$$

DERIVADAS SUCESIVAS DE UNA FUNCIÓN COMPUESTA

La aplicación reiterada de las reglas de derivación de una función compuesta (véase [29]) conducen fácilmente a los siguientes resultados.

[34]

Sean $\boldsymbol{f}: C \to \mathbb{R}^p$ y $\boldsymbol{g}: D \to \mathbb{R}^r$ dos funciones, definidas en sendos abiertos $C \subset \mathbb{R}^p$ y $D \subset \mathbb{R}^q$, siendo $\boldsymbol{f}(C) \subset D$; sean $f_1, ..., f_q$ las componentes de $\boldsymbol{f}$; si $\boldsymbol{x} = (x_1, ..., x_p)$ recorre C, sea $\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x}) = (y_1, ..., y_q)$.

COMPOSICIÓN DE FUNCIONES DE CLASE $\mathscr{C}^r$. Si $\boldsymbol{f}$ es de clase $\mathscr{C}^r$ (con $r \geqslant 2$) en un punto $\boldsymbol{x}$ y si $\boldsymbol{g}$ es de clase $\mathscr{C}^r$ en el punto $\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x})$, entonces $\boldsymbol{g} \circ \boldsymbol{f}$ es de clase $\mathscr{C}^r$ en $\boldsymbol{x}$. En tal caso, las derivadas parciales, hasta las de orden r, de $\boldsymbol{g} \circ \boldsymbol{f}$ en $\boldsymbol{x}$ se obtienen aplicando reiteradamente la regla de la cadena [29].

DERIVADAS Y DIFERENCIAL SEGUNDAS. En el caso particular de funciones de clase $\mathscr{C}^2$, al aplicar dos veces las reglas de derivación y diferenciación de las funciones compuestas, se obtiene:

$$\frac{\partial^2(\boldsymbol{g} \circ \boldsymbol{f})(\boldsymbol{x})}{\partial x_j \partial x_i} = \sum_{h,k=1}^{q} \frac{\partial^2 \boldsymbol{g}(\boldsymbol{y})}{\partial y_k \partial y_h} \frac{\partial f_k(\boldsymbol{x})}{\partial x_j} \frac{\partial f_h(\boldsymbol{x})}{\partial x_i} + \sum_{h=1}^{q} \frac{\partial \boldsymbol{g}(\boldsymbol{y})}{\partial y_h} \frac{\partial^2 f_h(\boldsymbol{x})}{\partial x_j \partial x_i}$$

$$d^2(\boldsymbol{g} \circ \boldsymbol{f})(\boldsymbol{x}) = \sum_{h,k=1}^{q} \frac{\partial^2 \boldsymbol{g}(\boldsymbol{y})}{\partial y_k \partial y_h} df_k(\boldsymbol{x}) df_h(\boldsymbol{x}) + \sum_{h=1}^{q} \frac{\partial \boldsymbol{g}(\boldsymbol{y})}{\partial y_h} d^2 f_h(\boldsymbol{x})$$

Nota: En estas fórmulas hay dos tipos de sumandos; en los primeros aparecen las derivadas segundas de $\boldsymbol{g}$ y en los segundos, aparecen las derivadas primeras de $\boldsymbol{g}$. A causa de estos últimos, la diferenciación no se rige ahora por una «regla de invarianza», como ocurría con la diferencial primera.

Comprobación

Como f y g son de clase $\mathscr{C}^1$ es de aplicación lo dicho en [29] acerca de la derivación de $g \circ f$; por tanto, para $i = 1, 2, ..., p$, es

$$\frac{\partial(g \circ f)(x)}{\partial x_i} = \sum_{h=1}^{q} \frac{\partial g(y)}{\partial y_h} \frac{\partial f_h(x)}{\partial x_i}$$

Como las derivadas del segundo miembro son, para $x \in C$ y para $y \in D$, funciones de clase $\mathscr{C}^1$, resulta que también lo es la derivada del primer miembro. Por ello esta derivada admite, a su vez, derivadas parciales, que serán (aplicando las reglas de derivación y, en particular, la regla de la cadena), para $j = 1, 2, ..., p$:

$$\begin{aligned}
\frac{\partial^2(g \circ f)(x)}{\partial x_j \partial x_i} &= \sum_{h=1}^{q} \left\{ \frac{\partial}{\partial x_j} \left(\frac{\partial g(f(x))}{\partial y_h} \right) \frac{\partial f_h(x)}{\partial x_i} + \frac{\partial g(y)}{\partial y_h} \frac{\partial}{\partial x_j} \left(\frac{\partial f_h(x)}{\partial x_i} \right) \right\} = \\
&= \sum_{h=1}^{q} \left\{ \left(\sum_{k=1}^{q} \frac{\partial^2 g(y)}{\partial y_k \partial y_h} \frac{\partial f_k(x)}{\partial x_j} \right) \frac{\partial f_h(x)}{\partial x_i} + \frac{\partial g(y)}{\partial y_h} \frac{\partial^2 f_h(x)}{\partial x_j \partial x_i} \right\} = \\
&= \sum_{h,k=1}^{q} \frac{\partial^2 g(y)}{\partial y_k \partial y_h} \frac{\partial f_k(x)}{\partial x_j} \frac{\partial f_h(x)}{\partial x_i} + \sum_{h=1}^{q} \frac{\partial g(h)}{\partial y_h} \frac{\partial^2 f_h(x)}{\partial x_j \partial x_i}
\end{aligned}$$

Como las derivadas del segundo miembro son todas continuas, para $x \in C$ e $y \in D$, también lo es la del primer miembro (para i y j variando de 1 a p), es decir $g \circ f$ es de clase $\mathscr{C}^2$. Reiterando el anterior razonamiento se obtiene, obviamente, que (si f y g son de clase $\mathscr{C}^r$ en x e y) la composición $g \circ f$ es de clase $\mathscr{C}^r$ en x. Nótese que la expresión que se acaba de obtener para la derivada parcial $\partial(g \circ f)/\partial x_j \partial x_i$ es, justamente, la que había que comprobar. La expresión de la diferencial segunda $d^2(g \circ f)$ se obtiene ya fácilmente recurriendo a las de las anteriores derivadas segundas. Para obtener $d^2(g \circ f)$, también se puede diferenciar en la expresión que da la diferencial primera $d(g \circ f)$ (véase [29]).

[34]₁ Observaciones

1.° Si, como es frecuente, a las funciones las representamos con las mismas letras que a sus valores, denotando así a las anteriores funciones f y g de [34]:

$$\begin{array}{l}
x = (x_1, ..., x_p) \overset{f}{\mapsto} y = y(x) = (y, ..., y_q) \overset{g}{\mapsto} \quad z = z(y) \\
\xrightarrow{\qquad\qquad\qquad g \circ f \qquad\qquad\qquad} z = z(y(x)) = z(x)
\end{array}$$

las derivadas parciales primeras y segundas y las diferenciales primera y segunda de la función compuesta $x \mapsto z(x)$ se escriben en la forma

$$\frac{\partial z(x)}{\partial x_i} = \sum_h \frac{\partial z(y)}{\partial y_h} \frac{\partial y_h(x)}{\partial x_i}, \qquad dz(x) = \sum_h \frac{\partial z(y)}{\partial y_h} dy_h(x)$$

$$\frac{\partial^2 z(x)}{\partial x_j \partial x_i} = \sum_{k,h} \frac{\partial^2 z(y)}{\partial y_k \partial y_h} \frac{\partial y_k(x)}{\partial x_j} \frac{\partial y_h(x)}{\partial x_i} + \sum_h \frac{\partial z(y)}{\partial y_h} \frac{\partial^2 y_h(x)}{\partial x_j \partial x_i}$$

$$d^2 z(x) = \sum_{k,h} \frac{\partial^2 z(y)}{\partial y_k \partial y_h} dy_k(x) dy_h(x) + \sum_h \frac{\partial z(y)}{\partial y_h} d^2 y_h(x) \qquad [1]$$

2.° Conviene señalar que, con la misma notación precedente, la diferencial segunda de la función $\mathbf{y} \mapsto z(\mathbf{y})$ es:

$$d^2z(\mathbf{y}) = \sum_{k,h} \frac{\partial^2 z(\mathbf{y})}{\partial y_h \partial y_h} \Delta y_h \Delta y_k \qquad [2]$$

Ahora, con las diferenciales segundas, entre las expresiones [1] y [2], de $d^2z(\mathbf{x})$ y de $d^2z(\mathbf{y})$, no se da la «invarianza» que se observa con las diferenciales primeras; ahora hay unas diferenciales segundas de las variables intermedias, $d^2y_h(\mathbf{x})$; éstas son nulas cuando las y_k son las variables independientes.

[34]₂ Ejercicio

Transformar la expresión (en la que se supone que $(x, y) \mapsto z(x, y)$ es una función de clase $\mathscr{C}^2$)

$$E = \frac{\partial^2 z}{\partial x^2} - 2\,\frac{\partial^2 z}{\partial x \partial y} + \frac{\partial^2 z}{\partial y^2} + \frac{2}{x}\left(\frac{\partial z}{\partial x} - \frac{\partial z}{\partial y}\right)$$

cuando se toman como nuevas variables a las

$$u = x + y \qquad \text{y} \qquad v = y/x$$

Resolución

La expresión E hay que expresarla referida a la función $(u, v) \mapsto z(u, v)$. Como

$$z(x, y) = z\big(u(x, y), v(x, y)\big), \qquad \text{con } \begin{cases} u(x, y) = x + y \\ v(x, y) = y/x \end{cases}$$

derivando en esta función compuesta, se obtiene que:

$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial u} + \frac{\partial z}{\partial v}\left(-\frac{y}{x^2}\right), \qquad \frac{\partial z}{\partial y} = \frac{\partial z}{\partial u} + \frac{\partial z}{\partial v}\left(\frac{1}{x}\right)$$

$$\frac{\partial^2 z}{\partial x^2} = \frac{\partial^2 z}{\partial u^2} + 2\,\frac{\partial^2 z}{\partial u \partial v}\left(-\frac{y}{x^2}\right) + \frac{\partial^2 z}{\partial v^2}\left(\frac{y^2}{x^4}\right) + \frac{\partial z}{\partial v}\left(\frac{2y}{x^3}\right)$$

$$\frac{\partial^2 z}{\partial y^2} = \frac{\partial^2 z}{\partial u^2} + 2\,\frac{\partial^2 z}{\partial u \partial v}\left(\frac{1}{x}\right) + \frac{\partial^2}{\partial v^2}\left(\frac{1}{x^2}\right)$$

$$\frac{\partial^2 z}{\partial x \partial y} = \frac{\partial^2 z}{\partial u^2} + \frac{\partial^2 z}{\partial u \partial v}\left(\frac{1}{x} - \frac{y}{x^2}\right) + \frac{\partial^2 z}{\partial v^2}\left(-\frac{y}{x^3}\right) + \frac{\partial z}{\partial v}\left(-\frac{1}{x^2}\right)$$

Llevando estos resultados a la expresión de E, se obtiene:

$$E = \frac{\partial^2 z}{\partial v^2}\left(\frac{y^2}{x^4} + \frac{2y}{x^3} + \frac{1}{x^2}\right) = \frac{(1 + v)^4}{u^3}\,\frac{\partial^2 z}{\partial v^2}$$

Ejercicios y problemas

ENUNCIADOS

2.1. Sea $f: \mathbb{R}^2 \to \mathbb{R}$ la función definida mediante

$$f(x, y) = \frac{(y - x^2)(y - 2x^2)}{|y - x^2||y - 2x^2|}$$

$$f(x, x^2) = f(x, 2x^2) = 1$$

1.º estudiar la continuidad de f; 2.º hallar las derivadas $f'_x(0, 0)$ y $f'_y(0, 0)$, si existen; 3.º hallar, si existe, la derivada $f'_u(0, 0)$, según cualquier vector $\boldsymbol{u} \in \mathbb{R}^2$.

2.2. Calcular $xz'_x + yz'_y$ en un punto interior del campo de definición de $z(x, y)$, en cada uno de los casos siguientes:

1.º $z(x, y) = \sqrt{\mathrm{L}\,(xy) + \operatorname{arcsen}(y/x)}$
2.º $z(x, y) = e^{y/x} \operatorname{sen}(y/x)$

2.3. Extender la siguiente función $f(x, y)$ a los puntos $(0, y)$ de manera que sea continua en ellos. Hallar $f'_x(0, y)$ y $f'_y(0, y)$ para $y \in \mathbb{R}$:

$$f(x, y) = y^2 \frac{\operatorname{sen} x}{x}$$

2.4. Hallar las derivadas parciales $\partial z/\partial x$ y $\partial z/\partial y$ de la función

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

2.5. Hallar las derivadas parciales z'_x y z'_y de las funciones $(x, y) \mapsto z$ definidas como sigue:

1.º $z = y \operatorname{sen}(xy)$
2.º $z = (e^{2x+y} - x^2)^{1/2}$
3.º $z = x/(x^2 + y^2)$
4.º $z^3 - y^2 z + x^2 y + y^3 = 2$

2.6. Calcular $dz(x, y)$ en un punto (x, y) interior del campo de definición de cada una de las funciones:

1.º $z(x, y) = \mathrm{L}\,\dfrac{x + \sqrt{x^2 - y^2}}{x - \sqrt{x^2 - y^2}}$
2.º $z(x, y) = \operatorname{arctg} \dfrac{x - y}{x + y}$

2.7. Hallar las diferenciales $dz(x, y)(dx, dy)$ de las siguientes funciones $(x, y) \mapsto z$, en los puntos en que sean diferenciables:

1.º $z = \dfrac{y}{x}$; 2.º $z = \mathrm{L}\,\sqrt{x^2 + y^2}$;
3.º $z = \operatorname{arctg} \dfrac{x}{y}$

2.8. Analizar si la siguiente función f tiene derivada según un vector $\boldsymbol{u} = (\cos \theta, \operatorname{sen} \theta)$ en $(0, 0)$ y si es diferenciable en dicho punto

$$f(x, y) = \frac{x^3 + y^3}{\operatorname{sen}(x^2 + y^2)}, \qquad f(0, 0) = 0$$

2.9. Sea $f: \mathbb{R}^2 \to \mathbb{R}$ la función definida por:

$$f(x, y) = \begin{cases} x^2 - y^2, & \text{si } x \text{ e } y \text{ son ambos racionales} \\ 0, & \text{si alguno de los } x \text{ e } y \text{ es irracional} \end{cases}$$

Estudiar:

1.º La existencia de f'_x y de f'_y.
2.º La diferenciabilidad de f; hállese df allí donde exista.

2.10. Sea $f: \mathbb{R}^2 \to \mathbb{R}$ la función definida por:

$$f(x, y) = (x^2 + y^2) \operatorname{sen} \frac{\pi}{x + y}, \qquad f(x, -x) = 0$$

1.º Estudiar la continuidad de f'_x y de f'_y en $(0, 0)$.
2.º Estudiar la diferenciabilidad de f en $(0, 0)$.

2.11. Sea $f: \mathbb{R}^2 \to \mathbb{R}$ la función definida por:

$$f(x, y) = x^2 \operatorname{sen} \frac{1}{x} + y^2 \operatorname{sen} \frac{1}{y},$$

$$f(x, 0) = f(0, y) = 0$$

1.° Estudiar la diferenciabilidad de f en $(0, 0)$.
2.° Analizar la continuidad de f'_x y de f'_y en $(0, 0)$.

2.12. Hallar la diferencial de la función $(x, y) \mapsto z$ en los casos:

1.° $z = \mathrm{L} \operatorname{sen} (x^2y^2 - 1)$
2.° $z = x^2y^2\sqrt{1 - x^2 - y^2}$

2.13. Hallar las derivadas de las siguientes funciones f en los puntos y según las direcciones que se indican:

1.° $f(x, y, z) = x^2 - y^2 + 2z^2$; $(3, 2, 1)$; la de la recta que pasa por $(0, 5, 3)$.
2.° $f(x, y, z) = 2x - 5y + 3z$; $(1, 2, 2)$; normal exterior a $x^2 + y^2 + z^2 = 9$.

2.14. Sea $f: \mathbb{R}^3 \to \mathbb{R}$ una función diferenciable y supóngase que γ es una curva de ecuaciones $f(x, y, z) = h$, $g(x, y, z) = k$, que tiene tangente $\boldsymbol{t}$ (vector unitario) en un cierto punto (x_0, y_0, z_0). Hallar la derivada $\partial f/\partial \boldsymbol{t}$ en dicho punto.

2.15. Sea $f: \mathbb{R}^2 \to \mathbb{R}$ la función definida por:

$$f(x, y) = ax^2y + by^2z + cz^2x$$

Hallar a, b y c para que, en el punto $P(1, 1, 1)$, la derivada máxima de f se produzca en la dirección del vector $(1, 5, 0)$ y valga 13.

2.16. Sea $f(x, y)$ una función diferenciable en $(0, 0)$ y tal que $f'_x(0, 0) = f'_y(0, 0) = 1$. Calcular las derivadas de f en $(0, 0)$ según los siguientes vectores: 1.° $\boldsymbol{u} = (\cos\,\theta, \operatorname{sen}\,\theta)$ con $\theta = 30°$; 2.° el unitario que dé máxima derivada; 3.° el unitario que dé derivada nula, hallando dicho vector.

2.17. Sean f y g dos funciones reales diferenciables en un punto $\boldsymbol{x}$ interior de su campo de definición. Calcular $\nabla(fg)(\boldsymbol{x})$ en función de $\nabla f(\boldsymbol{x})$ y $\nabla g(\boldsymbol{x})$.

2.18. Sea $f: \mathbb{R}^2 \to \mathbb{R}$ la función definida mediante:

$$f(x, y) = \frac{x^2 - y^2}{x^2 + y^2}\, xy, \qquad f(0, 0) = 0$$

Se pide: 1.° estudiar la continuidad de f en $(0, 0)$; 2.° f'_x y f'_y y estudiar su continuidad en $(0, 0)$; 3.° estudiar la diferenciabilidad de f en $(0, 0)$; 4.° $f''_{xy}(0, 0)$ y $f''_{yx}(0, 0)$.

2.19. Estudiar la derivabilidad y la diferenciabilidad de la siguiente función $f(x, y)$ en el punto $(0, 0)$. Hallar f''_{xy} y $f''_{yx}(0, 0)$

$$f(x, y) = \frac{x^4 - y^4}{x^2 + xy + y^2}, \qquad f(0, 0) = 0$$

2.20. Sea $f: \mathbb{R}^2 \to \mathbb{R}$ la función definida por:

$$f(x, y) = xy \operatorname{sen}\left(\frac{\pi(x - y)}{2(x + y)}\right), \qquad f(x, -x) = 0$$

1.° Hallar $f'_x(0, 0)$, $f'_y(0, 0)$, $f''_{xy}(0, 0)$ y $f''_{yx}(0, 0)$.
2.° Estudiar la diferenciabilidad de f en $(0, 0)$.

2.21. Hallar $f''_{xy}(0, 0)$ y $f''_{yx}(0, 0)$, siendo f la función:

$$f(x, y) = \frac{\operatorname{sen} x - y}{x - \operatorname{sen} y}, \qquad f(0, 0) = 1$$

2.22. Hallar, si existen, las derivadas parciales $f'_x(0, 0)$, $f'_y(0, 0)$, $f''_{xy}(0, 0)$ y $f''_{yx}(0, 0)$, siendo:

$$f(x, y) = \frac{|x|\sqrt{2x^2 + y^2 - 1}}{\sqrt{x^2 + 2y^2}}, \qquad f(0, 0) = 0$$

2.23. Hallar las derivadas parciales z'_x, z''_{xy}, z'''_{xyy} de la función $z = \operatorname{sen} xe^{a^y}$.

2.24. Hállese $\Delta z \equiv \partial^2 z/\partial x^2 + \partial^2 z/\partial y^2$ para las siguientes funciones $(x, y) \mapsto (x, y)$:

1.° $z(x, y) = \cos x\, e^y$.
2.° $z(x, y) = \mathrm{L}\, \sqrt{x^2 + y^2}$.

2.25. Sea $z = f(xy) + \sqrt{xy}g(y/x)$, donde f y g son funciones de clase $\mathscr{C}^2$. Hallar:

$$E = x^2 z''_{xx} - y^2 z''_{yy}$$

2.26. Hallar $\partial^2 z/\partial x^2 + \partial^2 z/\partial y^2$ en los dos casos siguientes:

1.° $z = \operatorname{arctg} \dfrac{y}{x}$.
2.° $z = \mathrm{L}\, \sqrt{(x - a)^2 + (y - b)^2}$.

2.27. Sean $u(x, y)$ y $v(x, y)$ funciones de clase $\mathscr{C}^2$. Obténgase $\Delta(uv)$ en función de Δu, Δv, ∇u y ∇v (se conviene en que $\Delta z = \partial^2 z/\partial x^2 + \partial^2 z/\partial y^2$, laplaciana de z, para $z = z(x, y)$ de clase $\mathscr{C}^2$).

2.28. Hallar la función $t \mapsto f(t)$ más general de clase $\mathscr{C}^2$ tal que la función $(x, y) \mapsto f(x^2 - y^2)$ tenga iguales sus derivadas segundas respecto de x dos veces y respecto de y dos veces.

2.29. Hallar la diferencial segunda de

$$z = xy + \operatorname{arctg}(xy)$$

2.30. Calcular la expresión (potencia simbólica):

$$\left(\frac{\partial f(x, y)}{\partial x} + \frac{\partial f(x, y)}{\partial y}\right)^{(n)}$$

en los casos siguientes:

1.° $f(x, y) = \operatorname{sen}(x - y)e^{x+y}$.
2.° $f(x, y) = \cos(x - y)e^{x+y}$.

2.31. Calcular el jacobiano de las siguientes funciones de $(x_1, x_2, ..., x_n)$:

$$y_1 = x_1 e^{x_2},\ y_2 = x_2 e^{x_3},\ ...,\ y_n = x_n e^{x_1}$$

2.32. Si $x = \rho \cos \theta$ e $y = \rho \operatorname{sen} \theta$, hallar los determinantes jacobianos de (x, y) respecto de (ρ, θ) y de (ρ, θ) respecto de (x, y).

2.33. Se llama «laplaciana» de una función $(x, y) \mapsto z(x, y)$ de clase $\mathscr{C}^2$ a:

$$\Delta z = \frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2}$$

Expresar Δz en coordenadas polares, esto es, en función de las derivadas (primeras y segundas) de z respecto de ρ y θ, siendo $x = \rho \cos \theta$ e $y = \rho \operatorname{sen} \theta$.

2.34. Si $(x, y, z) \mapsto w(x, y, z)$ es una función de clase $\mathscr{C}^2$, expresar la «laplaciana» de w, esto es,

$$\Delta w = \frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} + \frac{\partial^2 w}{\partial z^2}$$

en coordenadas esféricas ρ, φ y θ (con $x = \rho \operatorname{sen} \varphi \cos \theta$, $y = \rho \operatorname{sen} \varphi \operatorname{sen} \theta$, $z = \rho \cos \varphi$), es decir, en función de las derivadas (primeras y segundas) de w respecto de ρ, φ y θ.

2.35. Si u, v y w son las funciones de (x, y, z) definidas por:

$$u = e^x \cos y, \quad v = e^x \operatorname{sen} y, \quad w = x + y + z$$

hallar el jacobiano de (u, v, w) respecto de las coordenadas esféricas ρ, φ y θ:

$$x = \rho \cos \varphi \cos \theta, \qquad y = \rho \cos \varphi \operatorname{sen} \theta,$$
$$z = \rho \operatorname{sen} \varphi$$

2.36. Sea $f: \mathbb{R}^2 \to \mathbb{R}^2$ la función $(x, y) \mapsto f(x, y) = (y + e^{xy}, x - e^{xy})$:

1.° Comprobar que f es diferenciable en $\mathbb{R}^2$ y hallar su matriz jacobiana.
2.° Comprobar que $f \circ f$ es diferenciable y calcular su matriz jacobiana en (0, 0).

2.37. Sea $z(x, y)$ una función, de clase $\mathscr{C}^2$, considérese la transformación:

$$(u, v) \mapsto (x, y), \qquad \text{con} \begin{cases} x = u + \operatorname{L} v \\ y = v + \operatorname{L} u \end{cases}$$

y sea $z(u, v)$ la composición de las funciones anteriores. Expresar z''_{xy} en función de las derivadas parciales de esta última función.

2.38. Sea $f: \mathbb{R}^2 \to \mathbb{R}$ una función de clase $\mathscr{C}^2$ y considérese la composición de las siguientes funciones:

$$(x, y) \mapsto \left(u = \operatorname{arctg}\frac{x}{y},\ v = \sqrt{x^2 + y^2}\right)$$
$$(u, v) \mapsto f(\operatorname{sen} u, \cos u)$$

Hallar z''_{xy}.

2.39. Obtener el desarrollo de Taylor, en torno del punto (1, −1) con término complementario en el lugar de las derivadas terceras, de y^2/x^3.

2.40. Obtener el desarrollo limitado de Mac Laurin de $e^{x+y} \cos(x + y)$.

2.41. Obtener el desarrollo limitado de Mac Laurin de:

$$f(x, y) = x \operatorname{sen} y + y \operatorname{sen} x$$

SOLUCIONES

2.1. 1.º f continua en $\mathbb{R}^2$ menos en los puntos de las parábolas $y = x^2$ e $y = 2x^2$.
2.º $f'_x(0, 0) = f'_y(0, 0) = 0$.
3.º $f'_u(0, 0) = 0$.

2.2. 1.º $1/z$.
2.º 0.

2.3. $f(0, y) = y^2$; $f'_x(0, y) = 0$, $f'_y(0, y) = 2y$.

2.4. $z'_x = \dfrac{y^2 - |xy|}{(x^2 + y^2)^{3/2}}$ (signo de x)

$z'_y = \dfrac{x^2 - |xy|}{(x^2 + y^2)^{3/2}}$ (signo de y)

2.5. 1.º $z'_z = y^2 \operatorname{sen}(xy)$, $z'_y = \operatorname{sen}(xy) + xy \cos(xy)$
2.º $z'_x = (e^{2x+y} - x)(e^{2x+y} - x^2)^{-1/2}$
$z'_y = \frac{1}{2}e^{2x+y}(e^{2x+y} - x^2)^{-1/2}$
3.º $z'_x = (y^2 - x^2)/(x^2 + y^2)^2$
$z'_y = -2xy/(x^2 + y^2)^2$
4.º $z'_x = (2xy)/(y^2 - 3z^2)$
$z'_y = (x^2 + 3y^2 - 2yz)/(y^2 - 3z^2)$

2.6. 1.º $dz(x, y)(dx, dy) = \dfrac{1}{\sqrt{x^2 - y^2}}\left(2dx - \dfrac{2x}{y}\,dy\right)$

2.º $dz(x, y)(dx, dy) = \dfrac{1}{x^2 + y^2}(ydx - xdy)$

2.7. 1.º $(-ydx + xdy)/y^2$
2.º $(xdx + ydy)/(x^2 + y^2)$
3.º $(ydx - xdy)/(x^2 + y^2)$

2.8. $f'_u = \cos^3\theta + \operatorname{sen}^3\theta$; no existe $df(0, 0)$.

2.9. 1.º $f'_x(x_r, y_r)$ no existe salvo $f'_x(0, 0) = 0$; $f'_x(x_r, y_i) = 0$; $f'_x(x_i, y_r)$ no existe; $f'_x(x_i, y_i) = 0$; $f'_y(x_r, y_r)$ no existe salvo $f'_y(0, 0) = 0$; $f'_y(x_i, y_r) = 0$; $f'_y(x_r, y_i)$ no existe; $f'_y(x_i, y_i) = 0$.
2.º Sólo es diferenciable en $(0, 0)$; $df(0, 0)(dx, dy) = 0$.

2.10. 1.º f'_x y f'_y no son continuas en $(0, 0)$.
2.º f es diferenciable en $(0, 0)$.

2.11. 1.º f es diferenciable en $(0, 0)$.
2.º f'_x y f'_y no son continuas en $(0, 0)$.

2.12. 1.º $dx = 2(xy^2\,dx + x^2y\,dy)/\operatorname{tg}(x^2y^2 - 1)$.
2.º $dz = \dfrac{1}{\sqrt{1 - x^2 - y^2}}\,[(2xy^2 - 3x^3y^2 - 2xy^4)\,dx + (2x^2y - 3x^2y^3 - 2x^4y)dy]$.

2.13. 1.º $-\sqrt{22}$.
2.º $-2/3$.

2.14. ∇f y $\boldsymbol{t}$ son perpendiculares; $\dfrac{\partial f(x_0, y_0, z_0)}{\partial \boldsymbol{t}} = 0$.

2.15. $\nabla f(1, 1, 1) = (2a + c,\ a + 2b,\ b + 2c)$; $a = \rho$, $b = 2\rho$, $c = -\rho$; $\rho = \sqrt{13/2}$.

2.16. 1.º $(1 + \sqrt{3})/2$.
2.º $\sqrt{2}$.
3.º $\boldsymbol{u} = (\cos\theta, \operatorname{sen}\theta)$ con $\theta = 3\pi/4$ y $\theta = 7\pi/4$.

2.17. $\nabla fg(\boldsymbol{x}) = f(\boldsymbol{x})\nabla g(\boldsymbol{x}) + g(\boldsymbol{x})\nabla f(\boldsymbol{x})$.

2.18. 1.º f es continua en $(0, 0)$.
2.º $f'_x(0, 0) = 0$, $f'_y(0, 0) = 0$; f'_x y f'_y continuas.
3.º f es diferenciable en $(0, 0)$.
4.º $z''_{xy}(0, 0) = -1$, $z''_{yx}(0, 0) = 1$.

2.19. $f'_x(0, 0) = f'_y(0, 0) = 0$; $df(0, 0) = o$; $f''_{xy}(0, 0) = 1$; $f''_{yx}(0, 0) = -1$.

2.20. 1.º $f'_x(0, 0) = f'_y(0, 0) = 0$, $f''_{xy}(0, 0) = -1$, $f''_{yx}(0, 0) = 1$.
2.º f es diferenciable en $(0, 0)$.

2.21. $f''_{xy}(0, 0) = 1/6$, $f''_{yx} = -1/6$.

2.22. $f'_x(0, 0)$ no existe, $f'_y(0, 0) = 0$, $f''_{xy}(0, 0)$ no existe, $f''_{yx}(0, 0) = 0$.

2.23. $z'_x = e^{a^y}\cos x$, $z''_{xy} = e^{a^y}a^y \operatorname{L} a \cos x$, $z'''_{xyy} = e^{a^y}(a^y + 1)a^y(\operatorname{L} a)^2 \cos x$.

2.24. 1.º $\Delta z = 0$.
2.º $\Delta z = 0$.

2.25. $E = 0$.

2.26. En ambos casos es nulo.

2.27. $\Delta(uv) = u\Delta v + v\Delta u + 2\nabla u \cdot \nabla v$.

2.28. $tf''(t) = -f'(t)$; $f'(t) = k/t$; $f(t) = k \operatorname{L} t + k'$.

2.29. $-\dfrac{2xy^3}{(1+x^2y^2)^2}\,dx^2+2\left(1+\dfrac{1-x^2y^2}{(1+x^2y^2)^2}\right)dx\,dy-$
$-\dfrac{2yx^3}{(1+x^2y^2)^2}\,dy^2.$

2.30. 1.° $2^n \operatorname{sen}(x-y)e^{x+y}$.
2.° $2^n \cos(x-y)e^{x+y}$.

2.31. $e^{x_1+x_2+\cdots+x_n}[1-(-1)^n x_1x_2\cdots x_n]$.

2.32. ρ y $1/\rho$.

2.33. $\Delta z = \dfrac{\partial^2 z}{\partial\rho^2}+\dfrac{1}{\rho^2}\dfrac{\partial^2 z}{\partial\theta^2}+\dfrac{1}{\rho}\dfrac{\partial z}{\partial\rho}.$

2.34. $\Delta w = \dfrac{\partial^2 w}{\partial\rho^2}+\dfrac{1}{\rho^2}\dfrac{\partial^2 w}{\partial\varphi^2}+\dfrac{1}{\rho^2\operatorname{sen}^2\varphi}\dfrac{\partial^2 w}{\partial\theta^2}+$
$+\dfrac{2}{\rho}\dfrac{\partial w}{\partial\rho}+\dfrac{1}{\rho^2\operatorname{tg}\varphi}\dfrac{\partial w}{\partial\varphi}$

2.35. $\left|\dfrac{\partial(u,v,w)}{\partial(\rho,\varphi,\theta)}\right| = \left|\dfrac{\partial(u,v,w)}{\partial(x,y,z)}\right|\left|\dfrac{\partial(x,y,z)}{\partial(\rho,\varphi,\theta)}\right| =$
$= -\rho^2\cos\varphi$

2.36. 1.° $\begin{bmatrix} ye^{xy} & 1+xe^{xy} \\ 1-ye^{xy} & -xe^{xy}\end{bmatrix}.$
2.° $\begin{bmatrix} 1+1/e & -1/e \\ -1/e & 1+1/e\end{bmatrix}.$

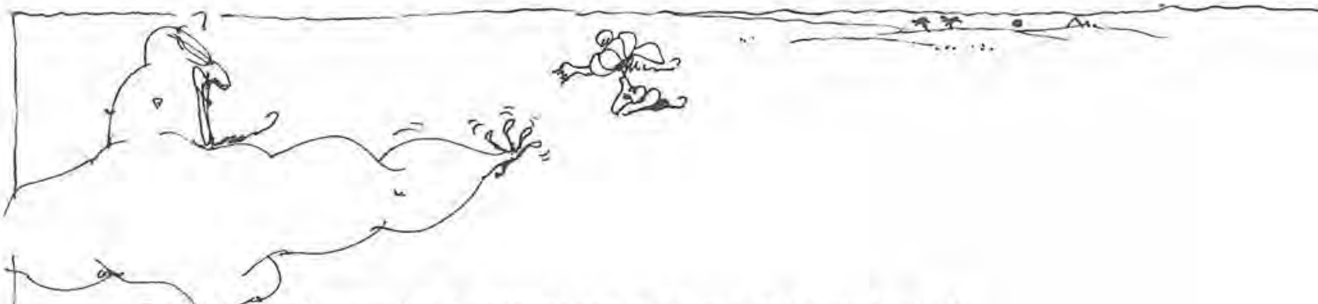
2.37. $z''_{xy} = \dfrac{-uv}{(uv-1)^2}\,[uz''_{uu}+vz''_{vv}-(1+v)z''_{uv}]+$
$+\dfrac{uv}{(uv-1)^3}\,[(1-u)z'_u+(1-v)z'_v].$

2.38. $z''_{xy} = \dfrac{f''_{uu}\cos^2 u - f'_u \operatorname{sen} u}{(x^2+y^2)^2}\,xy+$
$+\dfrac{f''_{uv}\cos u\operatorname{sen} v}{(x^2+y^2)^{3/2}}\,(y^2-x^2)+$
$+\dfrac{f''_{vv}\operatorname{sen}^2 u - f'_v\cos v}{(x^2+y^2)^2}\,xy-$
$-\dfrac{f'_u\cos u}{(x^2+y^2)^2}\,(y^2-x^2)+\dfrac{f'_v\operatorname{sen} v}{(x^2+y^2)^{3/2}}\,xy.$

2.39. $1-3(x-1)-2(y+1)+6(x-1)^2+$
$+6(x-1)(y+1)+(y+1)^2+$
$+\dfrac{1}{3!}\left((x-1)\dfrac{\partial}{\partial x}+(y+1)\dfrac{\partial}{\partial y}\right)^{(3)}$
$f(1+\theta(x-1),\,-1+\theta(y+1)).$

2.40. $\displaystyle\sum_{i=0}^{n}\frac{1}{i!}(\sqrt{2})^i(x+y)^i\cos\frac{n\pi}{4}.$

2.41. $\dfrac{4}{2!}\,xy-\dfrac{4}{4!}\,(x^3y+xy^3)+\cdots$
$+(-1)^{n-1}\dfrac{2n}{(2n)!}\,(x^{2n-1}y+xy^{2n-1}).$



El autor ha de reconocer que tiene más suerte de lo normal, pues de nuevo aquí, para comenzar el capítulo tercero, ha llegado a su mano un relato ad hoc. En él se ve cómo, desde antiguo, lo relacionado con el crecer y el decrecer de las cosas han tenido grande interés para abordar los más insospechables asuntos.

LOS SOTITINANE, ESPÍAS OBLIGADOS DE LO OCULTO(*)

No podía yo sospechar que mis habilidades en lo tocante al control mental pudieran llegarme a ser de tanta utilidad. Gracias a ellas, he podido defenderme de la agresión, vil agresión, de unos seres pequeñísimos y extraños que, aprovechando el menor descuido, espían en los pensamientos de las gentes para enterarse de cuanto hay en sus cerebros. En más de una ocasión, he visto sobre mi cabeza, levitando, a uno de esos minúsculos personajes que me miraba con un mirar tan penetrante que producía escalofríos de terror; al sentirse descubierto, torcía el gesto y desaparecía. Mis conocimientos sobre el dominio de la mente, me han permitido desvelar algunos de los enigmas que rodean a estos pequeños grandes ladrones de la intimidad. Por el interés que, obviamente, tiene para todos, me apresuro a relatar cuanto sobre este asunto he llegado a conocer.

Los sotitinane, que así se llaman estos inquietantes seres, no olvidan que, en tiempos remotos, fueron los pequeños y resignados sirvientes de los nibelungos; hoy son los esclavos del temible brujo Olam Ogam, que les obliga a llevar una existencia abyecta y llena de indignidad. Desde que guardan recuerdo, se han ocupado de quehaceres que desarrollan con la mente, se ocupan de actividades relacionadas con el pensamiento, con el intelecto. Esforzándose, pueden intuir, primero, y averiguar, finalmente, cosas ocultas y desconocidas; concentrándose, llegan a desvelar los secretos más guardados. Otra de sus peculiaridades radica en que sus energías son muy escasas, casi carecen de fuerzas, tienen muy poca capacidad para los trabajos físicos; esto les impide procurarse el sustento por su cuenta y les obliga a depender de quienes se lo proporcionan. Por este motivo, los sotitinane han sido dominados, en múltiples ocasiones, por gentes que se convertían en sus amos; a cambio de mantenerles, obtenían de ellos información sobre asuntos que les eran de interés. Casi siempre, tales informaciones han sido sobre cuestiones de dineros y riquezas.

Los orígenes de los sotitinane son desconocidos; las primeras noticias que se tienen de ellos se remontan a los tiempos en los que los nibelungos, aquellos enanos descendientes del rey de la oscuridad, se afanaban en acumular riquezas, lo que hacían con el auxilio de los sotitinane. A éstos los arrastraban por lóbregos bosques para que, haciendo uso de sus poderes mentales, averiguasen dónde había riquezas ocultas bajo la tierra. Cuando un sotitinane detectaba una zona rica en metales preciosos, era obligado por los nibelungos a elevarse sobre el terreno, quedando suspendido en el aire tanto más arriba cuanto mayores eran los bienes ocultos; si debajo de un punto P del bosque había una cantidad de riqueza igual a $f(P)$ (una función del punto P), entonces el sotitinane se situaba a la altura $z = f(P)$, es decir su desplazamiento por el aire lo era a lo largo de la superficie de ecuación $z = f(P)$. Los nibelungos exigían aún más a sus diminutos siervos: éstos tenian que llevar una varilla de la mano, manteniéndola con aquella inclinación que, en cada momento, mejor señalase la variación que, al avanzar, se producía en los tesoros que, soterrados, habia por donde ellos pasaban; así pues, la varilla ocupaba la posición de la tangente a la trayectoria que, en levitación, iban describiendo.

(*) Esta narración se ha tomado del libro *Los relatos de Gudor Ben Jusá*, de Juan de Burgos, publicado por la Fundación de la Universidad Politécnica de Madrid.

3

En cuanto los nibelungos veían que el sotitinane recorría una trayectoria ascendente, con la varilla hacia arriba, permanecían atentos y, cuando el ascenso cesaba, la varilla quedaba horizontal y él comenzaba a descender, entonces señalaban el punto del terreno sobre el que esto ocurría, esperando encontrar, bajo él, grandes riquezas. Pero con el tiempo descubrieron que en los puntos que así determinaban no solían acumularse los mayores tesoros; bien pensado, aquello era razonable: sólo se podía asegurar que los tesoros soterrados bajo un tal punto eran mayores que bajo los otros puntos de la trayectoria seguida por el sotitinane, pero nada se sabía respecto del resto de los puntos. A la vista de ello, decidieron cambiar de táctica.

Después de darle vueltas a cómo funciona el asunto de los valores máximos en las superficies, los nibelungos obligaron a los sotitinane a que, situados en una cierta posición, avanzasen según aquella dirección en la que mayor era el aumento de las riquezas del subsuelo; los sotitinane debían, pues, recorrer las líneas de máxima pendiente de la superficie $z = f(P)$ sobre la que se desplazaban. Cuando la varilla, pasando de estar inclinada hacia arriba y empezando a estarlo hacia abajo, quedaba horizontal, los nibelungos señalaban el punto P_0 del terreno sobre el que esto ocurría y obligaban al sotitinane a que se acercara a P_0 según otras trayectorias, a fin de averiguar si según ellas también acontecía lo mismo. Si así era, buscaban bajo P_0 esperando, razonablemente, que allí su búsqueda fuese muy fructuosa, pues $z_0 = f(P_0)$ debiera ser mayor que todos los $z = f(P)$ de las cercanías.

Cuando Sigfredo, el héroe, se apoderó del fabuloso tesoro que los nibelungos habían reunido, entonces y aprovechando el desconcierto que ello produjo, los sotitinane lograron escapar, librándose del pesado yugo que les habían impuesto.

Poco sé yo de los sotitinane desde los anteriores acontecimientos hasta los tiempos presentes. Sólo he llegado a enterarme de que tuvieron otras épocas de esclavitud; también sé que, de entre ellas, resultaron especialmente penosas las que permanecieron en poder de Alí Babá, un famoso ladrón, y de Craso, un gobernante asaz avariento.

Hoy los sotitinane están dominados por Olam Ogam, que haciendo uso de sus malignos poderes ocultos, los utiliza para enterarse de los pensamientos, deseos, recuerdos y todo tipo de secretos que las gentes guardamos, tanto de forma consciente como, incluso, en nuestros inconscientes. Elegida la víctima, Olam Ogam envía a un sotitinane, el cual, sin ser visto, permanece sobre ella, mentalmente atento al menor de sus descuidos. Esos momentos en que quedamos ensimismados, con la mente en blanco, sin pensar en nada, como enajenados, estos momentos, digo, son los que aprovechan los sotitinane para, con sus poderes mentales y ayudados por la perversa habilidad de Olam Ogam, instalarse en nuestro cerebro, adueñándose de cuanto en él encuentran.

Después de mucho esforzarme, creo haber encontrado la manera de evitar la agresión de Olam Ogam. Para anular los poderes de los sotitinane, basta con llevar con nosotros dos cosas. En primer lugar, debemos ir provistos de una moneda antigua de oro o, en su defecto, una joya u objeto valioso; con ello, los sotitinane, recordando la época en la que les dominaron los nibelungos, se concentran en la moneda y olvidan su dependencia de Olam Ogam. En segundo lugar, para contrarrestar la magia de este perverso brujo, hay que llevar escrito con letra muy clara, el siguiente conjuro: *Ned et Adreim Ahcum Ogam Otidlam.*

A todos os exhorto, vivamente, a que difundáis cuanto aquí he revelado.

Gudor Ben Jusá

CAPÍTULO 3

Aplicaciones de la derivación

Nos ocupamos aquí de aplicaciones de la derivación al propio Cálculo Infinitesimal. Dichas aplicaciones son, fundamentalmente, los teoremas de existencia y regularidad de las funciones implícita e inversa, el estudio de los extremos relativos y de los extremos «condicionados» de las funciones reales, la dependencia funcional y el cambio de variables. Cada una de estas cuestiones la iremos presentando en su momento oportuno.

3.1. EXISTENCIA Y REGULARIDAD DE LA FUNCIÓN IMPLÍCITA

Antes de entrar en materia, hagamos unas observaciones y comentarios acerca del tema que nos ocupa, que nos servirán de presentación del mismo.

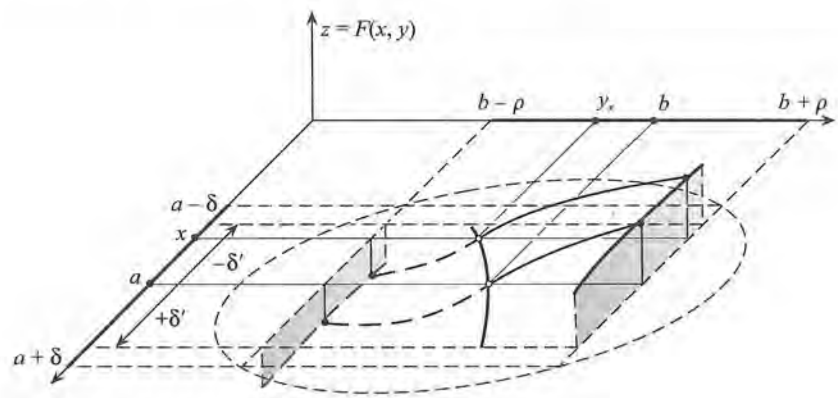
Dada una función $(x, y, z) \mapsto f(x, y, z)$, real de varias variables, suficientemente regular, nos preguntamos si la ecuación $f(x, y, z) = 0$ tendrá solución para la variable z, en función de las x e y; es decir, si existirá una función $(x, y) \mapsto z = \varphi(x, y)$ tal que $f(x, y, \varphi(x, y)) \equiv 0$, para todos los x e y. Cuando así sea, se dirá que la ecuación $f(x, y, z) = 0$ define implícitamente a $z = \varphi(x, y)$, aunque esta definición precise de algunas matizaciones.

Nótese, en primer lugar, que el problema planteado puede no tener solución, como ocurre, por ejemplo, en la ecuación $x^2 + y^2 + z^2 + 3 = 0$, donde el primer miembro es mayor o igual que 3 (para cualesquiera x, y, z). Para evitar que ocurra lo anterior, tomaremos como punto de partida que se verifica la igualdad $f(x_0, y_0, z_0) = 0$, para un cierto punto (x_0, y_0, z_0), y se buscará la $z = \varphi(x, y)$ para (x, y) en las proximidades de (x_0, y_0) y de manera que sea $z_0 = \varphi(x_0, y_0)$. Es más, interesarán aquellas $z = \varphi(x, y)$ que sean regulares en el supuesto de que lo es f. Si, por ejemplo, la ecuación $f(x, y, z) = 0$ es ahora la $x^2 + y^2 + z^2 - 1 = 0$ (ecuación de una esfera), ella tiene por «solución» a la $z = \pm\sqrt{1 - x^2 - y^2}$, siempre que (x, y) sea un punto del círculo $x^2 + y^2 \leqslant 1$. Si el punto de partida es el $(x_0, y_0, z_0) = (1/\sqrt{3}, 1/\sqrt{3}, -1/\sqrt{3})$, la función implícita buscada sería la $z = -\sqrt{1 - x^2 - y^2}$, que cerca de (x_0, y_0) es de clase $\mathscr{C}^\infty$; nótese, no obstante, que cerca de (x_0, y_0) hay otras muchas soluciones, como la

$$z = s(x, y)\sqrt{1 - x^2 - y^2}, \qquad \text{donde } s(x, y) = \begin{cases} 1, & \text{si } x + y \in \mathbb{Q} \\ -1, & \text{si } x + y \notin \mathbb{Q} \end{cases}$$

si bien es verdad que ellas no son ni siquiera continuas en (x_0, y_0). La situación se complica si se toma (x_0, y_0) en la circunferencia $x^2 + y^2 = 1$ (con lo que $z_0 = 0$), ya que entonces cerca de (x_0, y_0, z_0) hay dos soluciones (y no una sola) de la ecuación $x^2 + y^2 + z^2 = 1$, pues ahora las dos $z = \pm\sqrt{1 - x^2 - y^2}$ conducen, ambas, a puntos próximos al $(x_0, y_0, 0)$; pero es más, las soluciones $z = \pm\sqrt{1 - x^2 - y^2}$ no existen en todo un entorno de (x_0, y_0), sino sólo en la parte de él que está dentro de la circunferencia $x^2 + y^2 = 1$. Todas estas dificultades se presentan debido a que, en este punto $(x_0, y_0, 0)$, el plano tangente a la superficie $f(x, y, z) = 0$, esto es, el plano tangente a la esfera $x^2 + y^2 + z^2 = 1$, es paralelo al «eje de la z», es decir, a causa de que $f'_z(x_0, y_0, z_0) = 0$; por ello, cuando se estudian las funciones implícitas, se parte de la hipótesis $f'_z(x_0, y_0, z_0) \neq 0$. Si además de lo ya dicho, la función f es diferenciable, se puede entonces asegurar que la función φ, definida implícitamente por la ecuación $f(x, y, z) = 0$, es diferenciable; si f fuese de clase $\mathscr{C}^r$, entonces φ sería también de clase $\mathscr{C}^r$, como pasamos a comprobar.

TEOREMA DE LA FUNCIÓN IMPLÍCITA (CASO DE UNA SOLA ECUACIÓN)



[35]

FUNCIÓN IMPLÍCITA. Sea $U \subset \mathbb{R}^{p+1}$ un entorno de un punto $(a_1, ..., a_p; b) \in \mathbb{R}^{p+1}$; los puntos de U se denotarán poniendo $(\boldsymbol{x}; y)$ con $\boldsymbol{x} = (x_1, ..., x_p)$ y, en particular, $\boldsymbol{a} = (a_1, ..., a_p)$. Sea dada una función real $F : U \to \mathbb{R}$ tal que $F(\boldsymbol{a}; b) = 0$. Diremos que la ecuación $F(\boldsymbol{x}; y)$ define implícitamente a y como función continua de $\boldsymbol{x}$, en torno de $\boldsymbol{x} = \boldsymbol{a}$ e $y = b$, si existe un entorno $U_a \subset \mathbb{R}^p$ de $\boldsymbol{a}$ y una única función continua $f: U_a \to \mathbb{R}$, $\boldsymbol{x} \mapsto y = f(\boldsymbol{x})$, tales que

$$f(\boldsymbol{a}) = b \qquad \text{y} \qquad F(\boldsymbol{x}; f(\boldsymbol{x})) = 0 \qquad \text{para todo } \boldsymbol{x} \in U_a^{(*)}$$

I. EXISTENCIA DE LA FUNCIÓN IMPLÍCITA. Si $F : U \to \mathbb{R}$, $(\boldsymbol{x}; y) \mapsto F(\boldsymbol{x}; y)$, es una función continua en el entorno U del punto $(\boldsymbol{a}; b)$, si es $F(\boldsymbol{a}; b) = 0$ y si F es estrictamente monótona respecto de $y^{(**)}$, entonces la ecuación $F(\boldsymbol{x}; y) = 0$ define implícitamente a una única función continua $\boldsymbol{x} \mapsto y = f(\boldsymbol{x})$ en torno de $\boldsymbol{x} = \boldsymbol{a}$ e $y = b$.

II. DIFERENCIABILIDAD DE LA FUNCIÓN IMPLÍCITA. Si $F : U \to \mathbb{R}$, $(\boldsymbol{x}; y) \mapsto F(\boldsymbol{x}; y)$, es una función continua en el entorno U del punto $(\boldsymbol{a}; b)$, si es $F(\boldsymbol{a}; b) = 0$, si la derivada parcial F'_y existe y es no nula en U y si F es diferenciable en $(\boldsymbol{a}; b)$, entonces la ecuación $F(\boldsymbol{x}; y) = 0$ define implícitamente a una única función $\boldsymbol{x} \mapsto y = f(\boldsymbol{x})$ en torno de $\boldsymbol{x} = \boldsymbol{a}$ e $y = b$, la cual es diferenciable; las derivadas parciales de f son

$$\frac{\partial f(\boldsymbol{x})}{\partial x_i} = -\frac{F'_{x_i}(\boldsymbol{x}; f(\boldsymbol{x}))}{F'_y(\boldsymbol{x}; f(\boldsymbol{x}))} \qquad (i = 1, ..., p)$$

(*) Aunque puede sobreentenderse con lo ya dicho, conviene insistir en que, para cualquiera que sea $\boldsymbol{x} \in U_a$, el punto $(\boldsymbol{x}; f(\boldsymbol{x}))$ pertenece a U.

(**) Es decir, que para cualquier $\boldsymbol{x}$ tal que $(\boldsymbol{x}; b) \in U$, la función $y \mapsto F(\boldsymbol{x}; y)$ es estrictamente monótona (para y variando de modo que $(\boldsymbol{x}; y)$ recorra U). Esto se cumple, en particular, si la derivada parcial F'_y existe y es no nula en U.

Demostración

I. Existencia y continuidad de la función implícita. Sea $\varepsilon_0 > 0$ el radio de U y tómese cualquier $\rho \in]0, \varepsilon_0[$; llamando $\delta = \sqrt{\varepsilon_0^2 - \rho^2}$, se verifica que

$$[\|\boldsymbol{x} - \boldsymbol{a}\| < \delta \text{ y } |y - b| < \rho] \quad \Rightarrow \quad (\boldsymbol{x}; y) \in U$$

ya que es $\|(\boldsymbol{x}; y) - (\boldsymbol{a}; b)\|^2 < \delta^2 + \rho^2 = \varepsilon_0^2$. Nótese que, en particular, si es $|y - b| < \rho$, se verifica que $(\boldsymbol{a}; y) \in U$.

Como la función $y \mapsto F(\boldsymbol{a}; y)$ es estrictamente monótona, para $y \in]b - \varepsilon_0, b + \varepsilon_0[$, y debido a que $F(\boldsymbol{a}; b) = 0$, resulta que $F(\boldsymbol{a}; b - \rho)$ y $F(\boldsymbol{a}; b + \rho)$ tienen distinto signo; supondremos, pues, que es:

$$F(\boldsymbol{a}; b - \rho) < 0 \qquad \text{y} \qquad F(\boldsymbol{a}; b + \rho) > 0$$

(igual se razonaría en el caso contrario). Como F es continua en $(\boldsymbol{a};\ b-\rho)$ y en $(\boldsymbol{a};\ b+\rho)$, existe cierto $\delta' \in]0, \delta[$ (que depende de ρ) tal que

$$\|\boldsymbol{x}-\boldsymbol{a}\| < \delta' \quad \Rightarrow \quad [F(\boldsymbol{x};\ b-\rho) < 0 \quad \text{y} \quad F(\boldsymbol{x};\ b+\rho) > 0] \qquad [1]$$

Para cada $\boldsymbol{x} \in B(\boldsymbol{a}, \delta')$, la función $y \mapsto F(\boldsymbol{x};\ y)$ es continua para $y \in [b-\rho, b+\rho]$, por lo que, de [1] y según el teorema de Bolzano(*), se desprende que existe algún $y_x \in]b-\rho, b+\rho[$ tal que $F(\boldsymbol{x};\ y_x) = 0$ y este y_x es único (para cada $\boldsymbol{x}$) debido a la monotonía estricta de la función $y \mapsto F(\boldsymbol{x};\ y)$. Así pues, llamando $f(\boldsymbol{x}) = y_x$, se acaba de comprobar la existencia de una única función

$$f : B(\boldsymbol{a}, \delta') \to \mathbb{R}, \qquad \boldsymbol{x} \mapsto y = f(\boldsymbol{x})$$

definida (en el entorno $U_a = B(\boldsymbol{a}, \delta')$ de $\boldsymbol{a}$) implícitamente por la ecuación $F(\boldsymbol{x};\ y) = 0$ en torno de $\boldsymbol{x} = \boldsymbol{a}$ e $y = b$.

Para acabar, hemos de verificar que f es continua en todo punto $\boldsymbol{x} \in B(\boldsymbol{a}, \delta')$. Nótese que, si probamos la continuidad de f en el punto $\boldsymbol{a}$, la continuidad de f en un punto $\boldsymbol{x} \in B(\boldsymbol{a}, \delta')$ se puede comprobar de igual modo que en $\boldsymbol{a}$, pues f en $\boldsymbol{x}$ verifica a las mismas propiedades que en $\boldsymbol{a}$ (en $\boldsymbol{x}$, habría que sustituir el entorno $U = B(\boldsymbol{a}, \delta')$ por un entorno $U_x \subset U_a$). La continuidad de f en $\boldsymbol{a}$ es algo obvio pues, como sabemos, para cada $\rho > 0$ existe un $\delta' > 0$ tal que

$$\|\boldsymbol{x}-\boldsymbol{a}\| < \delta' \quad \Rightarrow \quad |b - y_x| < \rho \quad \Leftrightarrow \quad |b - f(\boldsymbol{x})| < \rho$$

II. Diferenciabilidad de la función implícita. En el supuesto de que se cumplen las exisgencias del enunciado II, se cumplen entonces, sobradamente, lo que se pide en el enunciado I (nótese que, como F'_y existe y es no nula en U, entonces F es estrictamente monótona respecto de y) y, por ello, la ecuación $F(\boldsymbol{x};\ y) = 0$ define implícitamente a una única función continua $\boldsymbol{x} \mapsto y = f(\boldsymbol{x})$ en torno de $\boldsymbol{x} = \boldsymbol{a}$ e $y = b$. Así pues, sólo nos resta comprobar que esta función $f : U_a \to \mathbb{R}$ (donde U_a es el entorno de $\boldsymbol{a}$ en el que f está definida) es diferenciable en cualquier punto $\boldsymbol{x} \in U_a$ y que sus derivadas parciales se obtienen aplicando la regla de la cadena, esto es, que vienen dadas por la fórmula del enunciado. Así ocurre, en efecto: tomando $\boldsymbol{x}_0 = (x_1, ..., x_p)$ en U_a y para todo $\Delta\boldsymbol{x} = (\Delta x_1, ..., \Delta x_p)$ suficientemente pequeño (esto es, tal que $\boldsymbol{x}_0 + \Delta\boldsymbol{x} \in U_a$), es evidente que el siguiente incremento Δ de F es nulo:

$$\Delta = F(\boldsymbol{x}_0 + \Delta\boldsymbol{x}_0;\ f(\boldsymbol{x}_0 + \Delta\boldsymbol{x})) - F(\boldsymbol{x}_0;\ f(\boldsymbol{x}_0)) = 0 - 0 = 0$$

Ahora bien, como $(\boldsymbol{x}_0;\ f(\boldsymbol{x}_0))$ y $(\boldsymbol{x}_0 + \Delta\boldsymbol{x};\ f(\boldsymbol{x}_0 + \Delta\boldsymbol{x}))$ son dos puntos de U y F es diferenciable en U, se verifica que, llamando $y_0 = f(\boldsymbol{x}_0)$ e $y_0 + \Delta y = f(\boldsymbol{x}_0 + \Delta\boldsymbol{x})$, es (véase [24], 3.° y [26], 2.°):

$$0 = \Delta = \sum_{i=1}^{p} (F'_{x_i}(\boldsymbol{x}_0;\ y_0) + \varepsilon_i(\Delta\boldsymbol{x};\ \Delta y))\Delta x_i + (F'_y(\boldsymbol{x}_0;\ y_0) + \varepsilon_0(\Delta\boldsymbol{x};\ \Delta y))\Delta y$$

donde $\varepsilon_i(\Delta\boldsymbol{x};\ \Delta y) \to 0$ cuando $(\Delta\boldsymbol{x};\ \Delta y) \to (\boldsymbol{o};\ 0)$, para $i = 0, 1, ..., p$

(*) Si $\varphi : [\alpha, \beta] \to \mathbb{R}$ es una función continua, en el intervalo compacto $[\alpha, \beta] \subset \mathbb{R}$, y si $\varphi(\alpha)$ y $\varphi(\beta)$ tienen distinto signo, entonces existe algún punto $\xi \in]\alpha, \beta[$ tal que $\varphi(\xi) = 0$.

Obsérvese que, por la continuidad de f en $\boldsymbol{x}_0$, acontece que $\Delta y \to 0$ cuando $\Delta\boldsymbol{x} \to \boldsymbol{o}$, por lo que $\varepsilon_i(\Delta\boldsymbol{x}; \Delta y) \to 0$ cuando $\Delta\boldsymbol{x} \to \boldsymbol{o}$ (para $i = 0, 1, ..., p$). Despejando Δy en la anterior igualdad, se obtiene que

$$\Delta y = \frac{-1}{F'_y(\boldsymbol{x}_0; y_0) + \varepsilon_0(\Delta\boldsymbol{x}; \Delta y)} \sum_{i=1}^{p} (F'_{x_i}(\boldsymbol{x}_0; y_0) + \varepsilon_i(\Delta\boldsymbol{x}; \Delta y))\Delta x_i \qquad [1]$$

Ahora bien, como $F'_y(\boldsymbol{x}_0; y_0) \neq 0$ y como $\varepsilon_0(\Delta\boldsymbol{x}; \Delta y) \to 0$ cuando $\Delta\boldsymbol{x} \to \boldsymbol{o}$, es evidente que se puede poner:

$$\frac{-1}{F'_y(\boldsymbol{x}_0; y_0) + \varepsilon_0(\Delta\boldsymbol{x}; \Delta y)} = \frac{-1}{F'_y(\boldsymbol{x}_0; y_0)} + \varepsilon(\Delta\boldsymbol{x}; \Delta y), \qquad \text{donde } \lim_{\Delta\boldsymbol{x} \to \boldsymbol{o}} \varepsilon(\Delta\boldsymbol{x}; \Delta y) = 0$$

Llevando este resultado a la igualdad [1], se llega a:

$$\Delta y = \sum_{i=1}^{p} -\frac{F'_{x_i}(\boldsymbol{x}_0; y_0)}{F'_y(\boldsymbol{x}_0; y_0)} \Delta x_i + \sum_{i=1}^{p} \delta_i(\Delta\boldsymbol{x})\Delta x_i, \qquad \text{donde } \lim_{\Delta\boldsymbol{x} \to \boldsymbol{o}} \delta_i(\Delta\boldsymbol{x}) = 0 \qquad [2]$$

$$\delta_i(\Delta\boldsymbol{x}) = F'_{x_i}(\boldsymbol{x}_0; y_0)\varepsilon(\Delta\boldsymbol{x}; \Delta y) + \varepsilon_i(\Delta\boldsymbol{x}; \Delta y) \frac{-1}{F'_y(\boldsymbol{x}_0; y_0)} + \varepsilon_i(\Delta\boldsymbol{x}; \Delta y)\varepsilon(\Delta\boldsymbol{x}; \Delta y)$$

(nótese que ponemos $\delta_i(\Delta\boldsymbol{x})$ en lugar de $\delta_i(\Delta\boldsymbol{x}; \Delta y)$, ya que Δy depende de $\Delta\boldsymbol{x}$). La expresión [2], de $\Delta y = f(\boldsymbol{x}_0 + \Delta\boldsymbol{x}) - f(\boldsymbol{x}_0)$, nos permite asegurar que f es diferenciable en el punto $\boldsymbol{x}_0$ (punto cualquiera de U_a) y que sus derivadas parciales son (para $i = 1, 2, ..., p$):

$$f'_{x_i}(\boldsymbol{x}_0) = \frac{F'_{x_i}(\boldsymbol{x}_0; y_0)}{F'_y(\boldsymbol{x}_0; y_0)} = -\frac{F'_{x_i}(\boldsymbol{x}_0; f(\boldsymbol{x}_0))}{F'_y(\boldsymbol{x}_0; f(\boldsymbol{x}_0))}$$

como había que comprobar.

[35]₁ Ejemplo

La ecuación $F(x, y, z) = 0$, donde F es la función real definida por

$$F(x, y, z) = xe^{y+z} + (z - x)\operatorname{sen}(y + 1) + \sqrt{2xyz} \qquad (\text{para } (x, y, z) \in \mathbb{R}_- \times \mathbb{R}_- \times \mathbb{R}_+)$$

define implícitamente a una única función continua $(x, y) \mapsto z = z(x, y)$ en un cierto entorno del punto $(x, y) = (-2, -1)$ y tal que $z(-2, -1) = 1$. Así ocurre, en efecto: la función F está definida y es diferenciable en un entorno de $(x, y, z) = (-2, -1, 1)$, se verifica que $F(-2, -1, 1) = 0$ y, en un entorno de $(-2, -1, 1)$, es no nula la derivada parcial F'_z, puesto que

$$F'_z(x, y, z) = xe^{y+z} + \operatorname{sen}(y + 1) + \sqrt{\frac{xy}{2z}}, \qquad F'_z(-2, -1, 1) = -1 \neq 0$$

(nótese que F'_z es continua en $(-2, -1, 1)$ y que en este punto es no nula, de lo que se infiere que F'_z es no nula en un entorno del susodicho punto). Esto es, para la ecuación $F(x, y, z) = 0$ se verifican todas las hipótesis del anterior teorema [35], II, de lo que se infiere la existencia de la función diferenciable $z = z(x, y)$ que es solución de la ecuación, o sea, tal que $F(x, y, z(x, y)) \equiv 0$ en torno de $(x, y) = (-2, -1)$ y $z = 1$.

Como las otras dos derivadas parciales de $F(x, y, z)$ son:

$$F'_x(x, y, z) = e^{y+z} - \operatorname{sen}(y + 1) + \sqrt{\frac{yz}{2x}}, \qquad F'_x(-2, -1, 1) = 3/2$$

$$F'_y(x, y, z) = xe^{y+z} + (z - x)\cos(y + 1) + \sqrt{\frac{xz}{2y}}, \qquad F'_y(-2, -1, 1) = 2$$

resulta que las derivadas parciales de $z(x, y)$ en el punto $(x, y) = (-2, -1)$ son

$$z'_x(-2, -1) = -\frac{F'_x(-2, -1, 1)}{F'_z(-2, -1, 1)} = \frac{3}{2} \qquad \text{y} \qquad z'_y(-2, -1) = -\frac{F'_y(-2, -1, 1)}{F'_z(-2, -1, 1)} = 2$$

$[35]_2$ Observación

Los anteriores teoremas [35], I y II, tienen carácter local. Se sabe que la función implícita $\boldsymbol{x} \mapsto y = f(\boldsymbol{x})$, a la que ellos hacen referencia, está definida en un cierto entorno $U_{\boldsymbol{a}}$ del punto $\boldsymbol{x} = \boldsymbol{a}$; nada se sabe, con carácter general, sobre el radio de este entorno, que puede ser «grande» o «pequeño», dependiendo de los casos. Usualmente, esta función $f\colon U_{\boldsymbol{a}} \to \mathbb{R}$ se puede prolongar hasta dar con un punto $\boldsymbol{x}$ en el que deje de cumplirse alguna de las hipótesis, como por ejemplo, en el que la derivada parcial $F'_y(\boldsymbol{x}; f(\boldsymbol{x}))$ sea nula; esto puede ocurrir muy cerca de $\boldsymbol{a}$, pero también puede acontecer que f llegue a prolongarse hasta el infinito.

Consideremos, a este respecto, un caso sencillo: el de la ecuación

$$y^2 - \operatorname{sen}(\pi/x) = 0, \qquad x > 0 \tag{1}$$

que es satisfecha por todos los puntos de las dos gráficas $y = \sqrt{\operatorname{sen}(\pi/x)}$ e $y = -\sqrt{\operatorname{sen}(\pi/x)}$, para $x > 0$. Dado un número real positivo a comprendido entre $1/(2n - 1)$ y $1/(2n - 2)$, donde $n \in \mathbb{N}$, la ecuación [1] define implícitamente, en un entorno de $x = a$, a una única función continua $y = y(x)$ tal que $y(a) = \operatorname{sen}(\pi/a)$ (también define otra función continua, también única en un entorno de $x = a$, tal que $y(a) = -\text{sen}\,(\pi/a))^{(*)}$. El entorno en cuestión es el $]1/(2n-1),\ 1/(2n-2)[$, cuya amplitud varía en función de a; si a está próximo a 0, entonces el entorno es pequeño; si a es mayor que 1, el entorno es $]1, +\infty[$.

$[35]_3$ Función implícita de clase $\mathscr{C}^r$

Sea $F: U \to \mathbb{R}$, $(\boldsymbol{x}; y) \mapsto F(\boldsymbol{x}; y)$, una función de clase $\mathscr{C}^r$, con $r \geqslant 1$, en un entorno $U \subset \mathbb{R}^{p+1}$ del punto $(\boldsymbol{a}; b) \in \mathbb{R}^p \times \mathbb{R}$. Si $F'_y(\boldsymbol{a}; b) \neq 0$ y $F(\boldsymbol{a}; b) = 0$, entonces la ecuación $F(\boldsymbol{x}; y) = 0$ define implícitamente, en torno de $\boldsymbol{x} = \boldsymbol{a}$ e $y = b$, a la variable y como función única de $\boldsymbol{x}$, $\boldsymbol{x} \mapsto y = f(\boldsymbol{x})$, siendo ésta una función de clase $\mathscr{C}^r$ en un entorno de $\boldsymbol{x} = \boldsymbol{a}$. Esto es, al verificarse las anteriores hipótesis, existen sendos entornos $U_a \subset \mathbb{R}^p$ de $\boldsymbol{a}$ y $U_b \subset \mathbb{R}$ de b, tales que $U_a \times U_b \subset U$, y existe una única función $f: U_a \to U_b$, $\boldsymbol{x} \mapsto y = f(\boldsymbol{x})$, de clase $\mathscr{C}^r$ en U_a tal que $b = f(\boldsymbol{a})$ y $F(\boldsymbol{x}; f(\boldsymbol{x})) = 0$ para todo $\boldsymbol{x} \in U_a$. Las derivadas parciales de esta función f se obtienen derivando o diferenciando, con la ayuda de la regla de la cadena, en $F(\boldsymbol{x}; f(\boldsymbol{x})) \equiv 0$.

Comprobación

Las anteriores hipótesis son más exigentes$^{(**)}$ que las del teorema [35], II, de existencia y diferenciabilidad de la función implícita; por lo que podemos asegurar que la tal función implícita $\boldsymbol{x} \mapsto y = f(\boldsymbol{x})$ existe y es diferenciable. Las derivadas parciales de f vienen, pues, dadas por la expresión de [35], II, es decir, son (para $\boldsymbol{x} \in U_a$):

$$\frac{\partial f(\boldsymbol{x})}{\partial x_i} = -\frac{F'_{x_i}(\boldsymbol{x}; f(\boldsymbol{x}))}{F'_y(\boldsymbol{x}; f(\boldsymbol{x}))} \qquad (i = 1, 2, \ldots, p)$$

Como el numerador y el denominador, del anterior segundo miembro, son funciones continuas de $\boldsymbol{x}$, también lo es $\partial f(\boldsymbol{x})/\partial x_i$ (para $i = 1, 2, \ldots, p$), luego f es de clase $\mathscr{C}^1$. Nótese que dichos numerador y denominador son, pues, funciones de clase $\mathscr{C}^1$ (si es $r \geqslant 2$), ya que son composición de funciones de clase $\mathscr{C}^1$, luego las derivadas parciales $\partial f(\boldsymbol{x})/\partial x_i$ son de clase $\mathscr{C}^1$, luego f es de clase $\mathscr{C}^2$ (si es $r \geqslant 2$). Reiterando este último razonamiento se llega fácilmente a que f es, en efecto, de clase $\mathscr{C}^r$.

$[35]_4$ Observación

Dada la ecuación $F(\boldsymbol{x}; y) = 0$, de la que se habla en el apartado anterior, para obtener las derivadas de la función $\boldsymbol{x} \mapsto y = y(\boldsymbol{x})$ que dicha ecuación define implícitamente, bastará con igualar a cero la derivada del primer miembro de la identidad

$$F(\boldsymbol{x}; y(\boldsymbol{x})) \equiv 0 \qquad \text{o} \qquad F(x_1, \ldots, x_p; y(x_1, \ldots, x_p)) \equiv 0 \tag{1}$$

(*) Para $a \in]1/(2n), 1/(2n-1)[$ no existe función implícita $y = y(x)$ definida por [1] y ello para cualquiera que sea el valor que se tome para $y(a)$.

(**) Nótese que como F'_y es continua en $(\boldsymbol{a}; b)$ y como $F'_y(\boldsymbol{a}; b) \neq 0$, se puede entonces asegurar que $F'_y(\boldsymbol{x}; y) \neq 0$ para todo $(\boldsymbol{x}; y)$ de un cierto entorno de $(\boldsymbol{a}; b)$.

que se calculan acudiendo a la regla de la cadena. Derivando, según se ha dicho, en [1] una vez respecto de x_i (para $i = 1, 2, ..., p$) se obtiene:

$$F'_{x_i}(\boldsymbol{x}; y) + F'_y(\boldsymbol{x}; y)y'_{x_i}(\boldsymbol{x}) \equiv 0, \qquad \text{luego } y'_{x_i}(\boldsymbol{x}) = -\frac{F'_{x_i}(\boldsymbol{x}; y)}{F'_y(\boldsymbol{x}; y)} \qquad [2]$$

Derivando nuevamente, ahora en la primera relación de [2] respecto de x_j (para $j = 1, 2, ..., p$), se obtiene que

$$F''_{x_ix_j} + F''_{x_iy}y'_{x_j} + F''_{yx_j}y'_{x_i} + F''_{yy}y'_{x_i}y'_{x_j} + F'_y y''_{x_ix_j} \equiv 0$$

de donde

$$y''_{x_ix_j} = -\frac{1}{F'_y}\left[F''_{x_ix_j} - F''_{x_iy}\frac{F'_{x_j}}{F'_y} - F''_{yx_j}\frac{F'_{x_i}}{F'_y} + F''_{yy}\frac{F'_{x_i}F'_{x_j}}{F'^2_y}\right]$$

[35]$_5$ Ejercicio

Sea $z = z(x, y)$ la función definida implícitamente mediante

$$y^2z + x(\mathrm{L}\, z - 1) = 0, \qquad z(1, -1) = 1 \qquad [1]$$

Obtener la aproximación local de $z(x, y)$, cerca de $(x, y) = (1, -1)$, que proporciona el desarrollo limitado de Taylor con polinomio de segundo grado.

Resolución

Nótese primeramente que dicha función $z = z(x, y)$ existe y es de clase $\mathscr{C}^\infty$, puesto que $F(x, y, z) = y^2z + x(\mathrm{L}\, z - 1)$ es de clase $\mathscr{C}^\infty$ en un entorno de $(x, y, z) = (1, -1, 1)$, es $F(1, -1, 1) = 0$ y $F'_z(1, -1, 1) \neq 0$, pues

$$F'_z(x, y, z) = y^2 + \frac{x}{z}, \qquad \text{luego } F'_z(1, -1, 1) = 1 + 1 = 2 \neq 0$$

Las derivadas parciales de $z(x, y)$ en $(x, y) = (1, -1)$, que se precisan para hallar el desarrollo pedido, se obtienen derivando en la ecuación [1], de lo que resulta:

$$\partial x: \quad y^2z'_x + \mathrm{L}\, z - 1 + \frac{x}{z}z'_x = 0 \qquad \rightarrow z'_x(1, -1) = 1/2$$

$$\partial y: \quad 2yz + y^2z'_y + \frac{x}{z}z'_y = 0 \qquad \rightarrow z'_y(1, -1) = 1$$

$$\partial x\partial x: \quad y^2z''_{xx} + 2\frac{z'_x}{z} + \frac{xz'^2_x}{z^2} + \frac{xz''_{xx}}{z} = 0 \qquad \rightarrow z''_{xx}(1, -1) = -3/8$$

$$\partial x\partial y: \quad 2yz'_x + y^2z''_{xy} + \frac{z'_y}{z} - \frac{xz'_xz'_y}{z^2} + \frac{xz''_{xy}}{z} = 0 \rightarrow z''_{xy}(1, -1) = -1/4$$

$$\partial y\partial y: \quad 2z + 4yz'_y + y^2z''_{yy} - \frac{xz'^2_y}{z^2} + \frac{xz''_{yy}}{z} = 0 \quad \rightarrow z''_{yy}(1, -1) = 3/2$$

Por tanto, el desarrollo pedido es:

$$z(x, y) = z(1, -1) + \frac{1}{1!}\,[z'_z(1, -1)(x-1) + z'_y(1, -1)(y+1)] +$$

$$+ \frac{1}{2!}\,[z''_{xx}(1, -1)(x-1)^2 + 2z''_{xy}(1, -1)(x-1)(y+1) + z''_{yy}(1, -1)(y+1)^2] +$$

$$+ o\left(\sqrt{(x-1)^2 + (y+1)^2}\right)^3 =$$

$$= 1 + \frac{1}{2}(x-1) + (y-1) - \frac{3}{16}(x-1)^2 - \frac{1}{4}(x-1)(y+1) + \frac{3}{4}(y+1)^2 +$$

$$+ o\left((x-1)^2 + (y+1)^2\right)^{3/2}$$

TEOREMA DE LA FUNCIÓN IMPLÍCITA (CASO DE UN SISTEMA DE ECUACIONES)

Para fijar las ideas, comenzaremos con unos comentarios sobre el problema que nos ocupa, tomando como referencia un caso particular. Consideremos el sistema de dos ecuaciones

$$\left.\begin{array}{l} f(x, y, z, u, v) = 0 \\ g(x, y, z, u, v) = 0 \end{array}\right\} \qquad [1]$$

que liga a las cinco variables x, y, z, u, v, donde f y g son funciones «suficientemente regulares». Nos preguntamos si, por satisfacerse estas dos ligaduras, las u y v son entonces variables dependientes de las x, y, z (variables independientes). Si así ocurriese, diríamos que el sistema de ecuaciones [1] define implícitamente a u y v como funciones de x, y y z. Salvo en casos sencillos, el problema planteado no suele tener solución de tipo global; hay que conformarse con soluciones locales, esto es, en las que (x, y, z) se mueve en un entorno de un cierto punto (x_0, y_0, z_0) y a partir de una determinada solución $(x_0, y_0, z_0, u_0, v_0)$ del sistema [1]. En concreto, podemos resumir la situación diciendo que si f y g son funciones reales de clase $\mathscr{C}^1$ en un entorno de un punto $(x_0, y_0, z_0, u_0, v_0)$, si f y g se anulan en dicho punto y si, en él, el determinante jacobiano de f y g respecto de u y v es no nulo, entonces existen dos funciones $u = u(x, y, z)$ y $v = v(x, y, z)$, que son de clase $\mathscr{C}^1$ en un cierto entorno de (x_0, y_0, z_0), tales que

$$\left.\begin{array}{l} u_0 = u(x_0, y_0, z_0) \\ v_0 = u(x_0, y_0, z_0) \end{array}\right\} \quad \text{y} \quad \left.\begin{array}{l} f(x, y, z, u(x, y, z), v(x, y, z)) \equiv 0 \\ g(x, y, z, u(x, y, z), v(x, y, z)) \equiv 0 \end{array}\right\} \qquad [2]$$

estas últimas identidades lo son para todo (x, y, z) del referido entorno de (x_0, y_0, z_0). Las expresiones de las derivadas y de las diferenciales de $u(x, y, z)$ y $v(x, y, z)$ se obtienen derivando y diferenciando en las identidades [2], siguiendo la regla de la cadena.

[36]

Sea $U \subset \mathbb{R}^{p+q}$ un entorno de un punto $(a_1, ..., a_p; b_1, ..., b_q) \in \mathbb{R}^{p+q}$; los puntos de U se denotarán poniendo $(\boldsymbol{x}; \boldsymbol{y})$ con $\boldsymbol{x} = (x_1, ..., x_p) \in \mathbb{R}^p$ e $\boldsymbol{y} = (y_1, ..., y_q) \in \mathbb{R}^q$ y, en particular, $\boldsymbol{a} = (a_1, ..., a_p)$ y $\boldsymbol{b} = (b_1, ..., b_q)$. Sea dada una función $\boldsymbol{F} : U \to \mathbb{R}^q$, $(\boldsymbol{x}; \boldsymbol{y}) \mapsto \boldsymbol{F}(\boldsymbol{x}; \boldsymbol{y})$, a cuyas funciones componentes las denotaremos por $F_j : U \to \mathbb{R}$, para $j = 1, ..., q$. Si $\boldsymbol{F}$ es de clase $\mathscr{C}^1$ en U, si $\boldsymbol{F}(\boldsymbol{a}; \boldsymbol{b}) = \boldsymbol{o}$ y si el jacobiano (esto es, el determinante de la matriz jacobiana; véase $[26]_5$) de $\boldsymbol{F}$ respecto de $\boldsymbol{y}$ es no nulo en $(\boldsymbol{a}; \boldsymbol{b})$, o sea, si $\det [J_y \boldsymbol{F}(\boldsymbol{a}; \boldsymbol{b})] \neq 0$ o

$$\frac{\partial(F_1, ..., F_q)}{\partial(y_1, ..., y_q)}(\boldsymbol{a}; \boldsymbol{b}) = \begin{vmatrix} D_{y_1}[F_1(\boldsymbol{a}; \boldsymbol{b})] & D_{y_2}[F_1(\boldsymbol{a}; \boldsymbol{b})] & \cdots & D_{y_q}[F_1(\boldsymbol{a}; \boldsymbol{b})] \\ D_{y_1}[F_2(\boldsymbol{a}; \boldsymbol{b})] & D_{y_2}[F_2(\boldsymbol{a}; \boldsymbol{b})] & \cdots & D_{y_q}[F_2(\boldsymbol{a}; \boldsymbol{b})] \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ D_{y_1}[F_q(\boldsymbol{a}; \boldsymbol{b})] & D_{y_2}[F_q(\boldsymbol{a}; \boldsymbol{b})] & \cdots & D_{y_q}[F_q(\boldsymbol{a}; \boldsymbol{b})] \end{vmatrix} \neq 0$$

entonces la ecuación $\boldsymbol{F}(\boldsymbol{x}; \boldsymbol{y}) = \boldsymbol{o}$ define implícitamente a $\boldsymbol{y}$ como función de clase $\mathscr{C}^1$ de $\boldsymbol{x}$ en torno de $\boldsymbol{x} = \boldsymbol{a}$ e $\boldsymbol{y} = \boldsymbol{b}$, esto es: existe una única función $\boldsymbol{f} : U_a \to \mathbb{R}^q$, $\boldsymbol{x} \mapsto \boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x})$, donde $U_a \subset \mathbb{R}^p$ es un cierto entorno de $\boldsymbol{a}$, la cual es de clase $\mathscr{C}^1$ en U_a, es tal que $\boldsymbol{b} = \boldsymbol{f}(\boldsymbol{a})$ y satisface a la identidad $\boldsymbol{F}(\boldsymbol{x}; \boldsymbol{f}(\boldsymbol{x})) \equiv \boldsymbol{o}$ para $\boldsymbol{x} \in U_a$.

La derivada parcial $\partial \boldsymbol{f}(\boldsymbol{x})/\partial x_j$ y la diferencial $d\boldsymbol{f}(\boldsymbol{x})$, de la función implícita $\boldsymbol{x} \mapsto \boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x})$ en un punto $\boldsymbol{x} \in U_a$, quedan definidas por las siguientes relaciones (las matrices con una «t» de supraíndice, que significa «traspuesta de», son matrices columna):

$$\left[\frac{\partial \boldsymbol{F}(\boldsymbol{x}; \boldsymbol{y})}{\partial x_j}\right]^t + J_y \boldsymbol{F}(\boldsymbol{x}; \boldsymbol{y}) \left[\frac{\partial \boldsymbol{f}(\boldsymbol{x})}{\partial x_j}\right]^t$$

$$J_x \boldsymbol{F}(\boldsymbol{x}; \boldsymbol{y})[d\boldsymbol{x}]^t + J_y \boldsymbol{F}(\boldsymbol{x}; \boldsymbol{y})[d\boldsymbol{f}(\boldsymbol{x})(d\boldsymbol{x})]^t$$

Estas fórmulas se analizan, desarrollándolas en términos de sus componentes, en $[36]_1$

Demostración

Procedamos por inducción sobre q (número de variables dependientes o coordenadas de $\boldsymbol{y}$). Para $q = 1$ el teorema se verifica, según ha sido ya probado anteriormente (véase $[35]_3$). Así pues, debemos demostrar que, si el teorema es cierto en el caso de que $\boldsymbol{y}$ tenga $q - 1$ coordenadas, entonces también lo es cuando $\boldsymbol{y}$ tiene q coordenadas. Vamos a ello: Como sabemos que $\det [J_y \boldsymbol{F}(\boldsymbol{a}; \boldsymbol{b})] \neq 0$, se puede asegurar que uno, al menos, de los q adjuntos de los elementos de la primera columna, de este determinante, es no nulo. Supongamos, pues, que no es nulo el primero de dichos adjuntos (de ser otro, cambiaríamos la ordenación de las componentes de $\boldsymbol{F}$); nótese que tal adjunto, al que llamaremos J, es el determinante jacobiano de la función $\boldsymbol{F}^* = (F_2, ..., F_q)$ respecto de $\boldsymbol{y}^* = (y_2, ..., y_q)$ en el punto $(\boldsymbol{a}; b_1, \boldsymbol{b}^*)$, donde $\boldsymbol{b}^* = (b_2, ..., b_q)$:

$$J = \det [J_{y^*} \boldsymbol{F}^*(\boldsymbol{a}; b_1, \boldsymbol{b}^*)] = \frac{\partial(F_2, ..., F_q)}{\partial(y_2, ..., y_q)}(\boldsymbol{a}; b_1, \boldsymbol{b}^*) \neq 0$$

Nos encontramos entonces con que para la función $F^* = (F_2, ..., F_q)$, de U en $\mathbb{R}^{q-1}$, se verifican todas las exigencias de nuestro teorema, al considerar como variables independientes a $(x; y_1) = (x_1, ..., x_p; y_1)$ y como variables dependientes a $y^* = (y_2, ..., y_q)$; como aquí tenemos $q - 1$ variables dependientes y según la hipótesis de inducción, resulta que se verifica, para esta F^*, la tesis del teorema, es decir, que existe una sola función $f^*: U^* \to \mathbb{R}^{q-1}$, $(x; y_1) \mapsto y^* = f^*(x; y_1)$, donde $U^* \subset \mathbb{R}^{p+1}$ es en cierto entorno de $(a; b_1) \in \mathbb{R}^{p+1}$, la cual es de clase $\mathscr{C}^1$ en U^*, es tal que $b^* = f^*(a; b_1)$ y satisface a la identidad

$$F^*(x; y_1, f^*(x, y_1)) \equiv 0, \qquad \forall\ (x; y_1) \in U^* \qquad [1]$$

Llamemos $f_2, ..., f_q$ a las funciones (de U^* en $\mathbb{R}$) componente de f^*, esto es, pongamos $f^* = (f_2, ..., f_q)$. Con ello, la anterior identidad [1] se puede poner, recurriendo a las coordenadas de F^* y de f^*, en la forma: para todo $(x, y_1) \in U^*$ es

$$\left.\begin{array}{c} F_2(x;\ y_1,\ f_2(x,\ y_1),\ ...,\ f_q(x,\ y_1)) \equiv 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ F_q(x;\ y_1,\ f_2(x,\ y_1),\ ...,\ f_q(x,\ y_1)) \equiv 0 \end{array}\right\} \qquad [1']$$

El teorema quedará entonces demostrado si se prueba la existencia de una única función $f_1: U_a \to \mathbb{R}, x \mapsto y_1 = f_1(x)$, donde $U_a \subset \mathbb{R}^p$ es un entorno de $a \in \mathbb{R}^p$, que sea de clase $\mathscr{C}^1$ en U_a, que verifique a $b_1 = f_1(a)$ y sea tal que se satisfaga la identidad

$$F_1(x; f_1(x), f^*(x; f_1(x)) \equiv 0, \qquad \forall\ x \in U_a \qquad [2]$$

que, si llamamos $G: U^* \to \mathbb{R}$ a la función definida por:

$$(x; y_1) \mapsto G(x; y_1) = F_1(x; y_1, f^*(x, y_1)) = F_1(x; y_1, f_2(x, y_1), .., f_q(x, y_1)) \qquad [3]$$

(que es de clase $\mathscr{C}^1$ en U^*, pues es composición de funciones de clase $\mathscr{C}^1$), puede expresarse (la identidad [2]) en la forma:

$$G(x; f_1(x)) \equiv 0, \qquad \forall\ x \in U_a$$

Es obvio que, de ser así las cosas, nuestro teorema se verificaría entonces para la función $f: U_a \to \mathbb{R}^q$ definida mediante:

$$x \mapsto f(x) = (y_1, y^*) = (f_1(x), f^*(x; f_1(x)))$$

pues para ella se verifican, como se comprueba fácilmente, todas las exigencias de la tesis. Para probar que, en efecto, existe la tal función f_1 acudiremos al teorema, ya probado, de la función implícita para el caso de una sola ecuación (véase [35], II y $[35]_3$), que aplicamos a la ecuación $G(\boldsymbol{x}; y_1) = 0$, donde $G: U^* \to \mathbb{R}$ es la función definida en [3]. Las hipótesis de este teorema se verifican todas, ya que:

- G es de clase $\mathscr{C}^1$ en U^*, ya que (véase [3]) F_1 es de clase $\mathscr{C}^1$ en U y $\boldsymbol{f}^*$ es de clase $\mathscr{C}^1$ en U^*.
- Se verifica que $G(\boldsymbol{a}; b_1) = 0$, ya que:

$$G(\boldsymbol{a}; b_1) = F_1(\boldsymbol{a}; b_1, \boldsymbol{f}^*(\boldsymbol{a}, b_1)) = F_1(\boldsymbol{a}; b_1, \boldsymbol{b}^*) = F_1(\boldsymbol{a}; \boldsymbol{b}) = 0$$

- Se verifica que $G'_{y_1}(\boldsymbol{a}; b_1) \neq 0$; en efecto: derivando parcialmente respecto de y_1, en el punto $(\boldsymbol{a}; b_1)$, en la expresión [3] de G y en las relaciones [1′], se obtienen las relaciones

$$\left.\begin{array}{r}
\dfrac{\partial G}{\partial y_1}(\boldsymbol{a};\, b_1) = \dfrac{\partial F_1}{\partial y_1}(\boldsymbol{a};\, \boldsymbol{b}) + \dfrac{\partial F_1}{\partial y_2}(\boldsymbol{a};\, \boldsymbol{b})\,\dfrac{\partial f_2}{\partial y_1}(\boldsymbol{a};\, b_1) + \cdots + \dfrac{\partial F_1}{\partial y_q}(\boldsymbol{a};\, \boldsymbol{b})\,\dfrac{\partial f_q}{\partial y_1}(\boldsymbol{a};\, b_1) \\
0 = \dfrac{\partial F_2}{\partial y_1}(\boldsymbol{a};\, \boldsymbol{b}) + \dfrac{\partial F_2}{\partial y_2}(\boldsymbol{a};\, \boldsymbol{b})\,\dfrac{\partial f_2}{\partial y_1}(\boldsymbol{a};\, b_1) + \cdots + \dfrac{\partial F_2}{\partial y_q}(\boldsymbol{a};\, \boldsymbol{b})\,\dfrac{\partial f_q}{\partial y_1}(\boldsymbol{a};\, b_1) \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
0 = \dfrac{\partial F_q}{\partial y_1}(\boldsymbol{a};\, \boldsymbol{b}) + \dfrac{\partial F_q}{\partial y_2}(\boldsymbol{a};\, \boldsymbol{b})\,\dfrac{\partial f_2}{\partial y_1}(\boldsymbol{a};\, b_1) + \cdots + \dfrac{\partial F_q}{\partial y_q}(\boldsymbol{a};\, \boldsymbol{b})\,\dfrac{\partial f_q}{\partial y_1}(\boldsymbol{a};\, b_1)
\end{array}\right\}$$

que pueden considerarse como un sistema de q ecuaciones lineales, en el que las incógnitas son los q valores

$$1,\ \frac{\partial f_2}{\partial y_1}(\boldsymbol{a};\, b_1),\ \ldots,\ \frac{\partial f_q}{\partial y_1}(\boldsymbol{a};\, b_1) \qquad [4]$$

(valores que existen; nos son ya conocidos), cuya matriz de coeficientes es la matriz jacobiana $J_y\boldsymbol{F}(\boldsymbol{a};\, \boldsymbol{b})$. Si la derivada parcial $G'_{y_1}(\boldsymbol{a};\, b_1)$ fuese nula, entonces el anterior sistema de ecuaciones lineales sería homogéneo y como tiene solución no nula (al menos el primero de los q valores [4] es no nulo) resultaría que el determinante de dicho sistema habría de ser cero, lo que no es cierto, pues tal determinante es el jacobiano de $\boldsymbol{F}$ respecto de $\boldsymbol{y}$ en $(\boldsymbol{a}; \boldsymbol{b})$ que no es nulo, según la hipótesis. Así pues, debe ser $G'_{y_1}(\boldsymbol{a}; b_1) \neq 0$.

Como ha resultado que se cumplen todas las hipótesis, podemos asegurar que realmente existe la función $\boldsymbol{f}$ que andábamos buscando.

El cálculo de las derivadas parciales $\partial \boldsymbol{f}/\partial x_i$ y de la diferencial $d\boldsymbol{f}$ se aborda en el siguiente apartado $[36]_1$.

[36]$_1$ Derivación y diferenciación de la función implícita

En lo que sigue, suponemos que, para el sistema de ecuaciones (se usa la misma notación que en [36])

$$\boldsymbol{F}(\boldsymbol{x};\boldsymbol{y})=\boldsymbol{o} \qquad \text{o} \qquad \begin{cases} F_1(x_1, \ldots, x_p; y_1, \ldots, y_q)=0 \\ F_2(x_1, \ldots, x_p; y_1, \ldots, y_q)=0 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ F_q(x_1, \ldots, x_p; y_1, \ldots, y_q)=0 \end{cases} \qquad [1]$$

se verifican las hipótesis del teorema [36] de la función implícita en torno de un punto $(\boldsymbol{x};\boldsymbol{y})=(\boldsymbol{a};\boldsymbol{b})$, por lo que dicho sistema define implícitamente (en un entorno de $\boldsymbol{x}=\boldsymbol{a}$) a una función $\boldsymbol{x}\mapsto\boldsymbol{y}=\boldsymbol{f}(\boldsymbol{x})$, de clase $\mathscr{C}^1$, tal que $\boldsymbol{b}=\boldsymbol{f}(\boldsymbol{a})$ y que satisface a la ecuación [1], es decir, verificando a:

$$\boldsymbol{F}(\boldsymbol{x};\boldsymbol{f}(\boldsymbol{x}))\equiv\boldsymbol{o} \qquad \text{o} \qquad \begin{cases} F_1(\boldsymbol{x}; f_1(\boldsymbol{x}), \ldots, f_q(\boldsymbol{x}))\equiv 0 \\ F_2(\boldsymbol{x}; f_1(\boldsymbol{x}), \ldots, f_q(\boldsymbol{x}))\equiv 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ F_q(\boldsymbol{x}; f_1(\boldsymbol{x}), \ldots, f_q(\boldsymbol{x}))\equiv 0 \end{cases} \qquad [2]$$

(para todo $\boldsymbol{x}$ de un cierto entorno del punto $\boldsymbol{a}$).

DERIVADAS PARCIALES DE LA FUNCIÓN IMPLÍCITA. De acuerdo con la regla de la cadena (véase [29], 1.º), derivando parcialmente en [2], respecto de la variable x_j (componente j-ésima de $\boldsymbol{x}$, para $j=1, 2, \ldots, p$), se obtiene:

$$\left.\begin{matrix} \dfrac{\partial F_1}{\partial x_j}+\dfrac{\partial F_1}{\partial y_1}\dfrac{\partial f_1}{\partial x_j}+\cdots+\dfrac{\partial F_1}{\partial y_q}\dfrac{\partial f_q}{\partial x_j}\equiv 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \dfrac{\partial F_q}{\partial x_j}+\dfrac{\partial F_q}{\partial y_1}\dfrac{\partial f_1}{\partial x_j}+\cdots+\dfrac{\partial F_q}{\partial y_q}\dfrac{\partial f_q}{\partial x_j}\equiv 0 \end{matrix}\right\} \qquad \text{o} \qquad \frac{\partial \boldsymbol{F}}{\partial x_j}+\sum_{k=1}^{q}\frac{\partial \boldsymbol{F}}{\partial y_k}\frac{\partial f_k}{\partial x_j}\equiv 0 \qquad [3]$$

luego expresándole en forma matricial, el anterior sistema [3] es:

$$\begin{bmatrix} \dfrac{\partial F_1}{\partial x_j} \\ \vdots \\ \dfrac{\partial F_q}{\partial x_j} \end{bmatrix} = -\begin{bmatrix} \dfrac{\partial F_1}{\partial y_1} & \cdots & \dfrac{\partial F_1}{\partial y_q} \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ \dfrac{\partial F_q}{\partial y_1} & \cdots & \dfrac{\partial F_q}{\partial y_q} \end{bmatrix} \begin{bmatrix} \dfrac{\partial f_1}{\partial x_j} \\ \vdots \\ \dfrac{\partial f_q}{\partial x_j} \end{bmatrix} \qquad \text{o} \qquad \left[\frac{\partial \boldsymbol{F}}{\partial x_j}\right]^t = -\left[\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{y}}\right]\left[\frac{\partial \boldsymbol{f}}{\partial x_j}\right]^t \qquad [3']$$

donde $[\partial\boldsymbol{F}/\partial\boldsymbol{y}]$ representa a la matriz jacobiana de $\boldsymbol{F}$ respecto de $\boldsymbol{y}$, que también se denota poniendo $J_y\boldsymbol{F}$, y donde la «t» como supraíndice significa «matriz traspuesta».

Como la referida matriz jacobiana es regular por hipótesis (no se olvide que, según lo supuesto, el correspondiente jacobiano es no nulo en el entorno en el que nos movemos), de [3′] se deduce que

$$\left[\frac{\partial f}{\partial x_j}\right]^t = -\left[\frac{\partial F}{\partial y}\right]^{-1}\left[\frac{\partial F}{\partial x_j}\right]^t \qquad \text{o} \qquad \begin{bmatrix} \dfrac{\partial f_1}{\partial x_j} \\ \vdots \\ \dfrac{\partial f_q}{\partial x_j} \end{bmatrix} = -\left[\frac{\partial(F_1, \ldots, F_q)}{\partial(y_1, \ldots, y_q)}\right]^{-1} \begin{bmatrix} \dfrac{\partial F_1}{\partial x_j} \\ \vdots \\ \dfrac{\partial F_q}{\partial x_j} \end{bmatrix} \qquad [4]$$

Esta expresión [4] facilita las derivadas parciales de la función implícita $f = (f_1, \ldots, f_q)$ respecto de x_j (una cualquiera de sus p variables $x_1, x_2, \ldots, x_p$).

DIFERENCIAL DE LA FUNCIÓN IMPLÍCITA. De acuerdo con la regla de derivación de las funciones compuestas (véase [29], 2.º), al diferenciar en las anteriores identidades [2], se obtiene:

$$\sum_{j=1}^{p} \frac{\partial F}{\partial x_j}\,dx_j + \sum_{i=1}^{q} \frac{\partial F}{\partial y_i}\,df_i$$

esto es:

$$\left\{\begin{array}{c} \dfrac{\partial F_1}{\partial x_1}\,dx_1 + \cdots + \dfrac{\partial F_1}{\partial x_p}\,dx_p + \dfrac{\partial F_1}{\partial y_1}\,df_1 + \cdots + \dfrac{\partial F_1}{\partial y_q}\,df_q \equiv 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \dfrac{\partial F_q}{\partial x_1}\,dx_1 + \cdots + \dfrac{\partial F_q}{\partial x_p}\,dx_p + \dfrac{\partial F_q}{\partial y_1}\,df_1 + \cdots + \dfrac{\partial F_q}{\partial y_q}\,df_q \equiv 0 \end{array}\right. \qquad [5]$$

Acudiendo a la notación matricial y haciendo uso de las matrices jacobianas de $\boldsymbol{F}$ respecto de $\boldsymbol{x}$ y respecto de $\boldsymbol{y}$, que se denotan por:

$$\left[\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{x}}\right] = J_x\boldsymbol{F} \qquad \text{y} \qquad \left[\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{y}}\right] = J_y\boldsymbol{F}$$

(que son de tamaños $q \times p$ y $q \times q$, respectivamente), las anteriores identidades [5] pueden ponerse en la forma (la «t» significa «matriz traspuesta»):

$$J_x\boldsymbol{F}\begin{bmatrix} dx_1 \\ \vdots \\ dx_p \end{bmatrix} + J_y\boldsymbol{F}\begin{bmatrix} df_1 \\ \vdots \\ df_q \end{bmatrix} \equiv O \qquad \text{o} \qquad -\left[\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{x}}\right][d\boldsymbol{x}]^t \equiv \left[\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{y}}\right][d\boldsymbol{f}]^t \qquad [5']$$

Como la matriz jacobiana $[\partial \boldsymbol{F}/\partial \boldsymbol{y}]$ es regular (pues su determinante es el jacobiano de $\boldsymbol{F}$ respecto de $\boldsymbol{y}$, que es no nulo, por hipótesis, en el entorno en el que estamos operando), de la anterior ecuación [5'] se puede despejar $[d\boldsymbol{f}]^t$; en efecto: premultiplicando por la matriz $[\partial \boldsymbol{F}/\partial \boldsymbol{y}]^{-1}$, se obtiene:

$$[d\boldsymbol{f}]^t = -\left[\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{y}}\right]^{-1}\left[\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{x}}\right][d\boldsymbol{x}]^t,$$

esto es:

$$\begin{bmatrix} df_1 \\ \vdots \\ df_q \end{bmatrix} = -\left[\frac{\partial(F_1, \ldots, F_q)}{\partial(y_1, \ldots, y_q)}\right]^{-1}\left[\frac{\partial(F_1, \ldots, F_q)}{\partial(x_1, \ldots, x_p)}\right]\begin{bmatrix} dx_1 \\ \vdots \\ dx_p \end{bmatrix} \qquad [6]$$

Esta expresión [6] facilita la diferencial $d\boldsymbol{f} = (df_1, \ldots, df_q)$ de la función implícita $\boldsymbol{f} = (f_1, \ldots, f_q)$.

$[36]_2$ Ejercicio

Considérese el sistema de ecuaciones:

$$\begin{cases} xu^2 + yv^2 + z = 5 \\ xyz + uv = 0 \end{cases} \qquad [1]$$

Analícese si dicho sistema define implícitamente a u y v como funciones de (x, y, z) en torno de $(x, y, z) = (x_0, y_0, z_0)$ con $(u, v) = (u_0, v_0)$, en los dos casos siguientes:

a) $(x_0, y_0, z_0, u_0, v_0) = (0, 1, 1, 0, 2)$.

b) $(x_0, y_0, z_0, u_0, v_0) = (1, 0, 5, 0, 2)$.

Resolución

1.º Debemos verificar si existen y son únicas tales funciones

$$u = u(x, y, z) \qquad \text{y} \qquad v = v(x, y, z)$$

definidas en todo un entorno $U \subset \mathbb{R}^3$ del punto (x_0, y_0, z_0) y si son, al menos, continuas en U. Acudiendo al teorema de la función implícita (véase [36]), como

$$F(x, y, z, u, v) = xu^2 + yv^2 + z - 5 \qquad \text{y} \qquad G(x, y, z) = xyz + uv$$

son, al menos, de clase $\mathscr{C}^1$ en $\mathbb{R}^5$, como se anulan en $(x_0, y_0, z_0, u_0, v_0)$, y como el jacobiano de F y G respecto de u y v es:

$$\det \frac{\partial(F, G)}{\partial(u, v)} = \begin{vmatrix} F'_u & F'_v \\ G'_u & G'_v \end{vmatrix} = \begin{vmatrix} 2xu & 2yv \\ v & u \end{vmatrix} = 2xu^2 - 2yv^2$$

resulta que, para los dos casos del enunciado, es

$$\det \frac{\partial(F, G)}{\partial(u, v)}(0, 1, 1, 0, 2) = -8 \neq 0 \qquad \text{y} \qquad \det \frac{\partial(F, G)}{\partial(u, v)}(1, 0, 5, 0, 2) = 0$$

Nos encontramos, pues, con que: *a*) en el primer caso, el teorema segura que existen las funciones que se buscan, las cuales son, al menos, de clase $\mathscr{C}^1$ en un entorno de $(x, y, z) = (0, 1, 1)$; y *b*) en el segundo caso, el teorema no asegura nada.

2.° Estudiando directamente este segundo caso, es decir, el correspondiente al punto $(x_0, y_0, z_0, u_0, v_0) = (1, 0, 5, 0, 2)$, comprobaremos que las funciones $u(x, y, z)$ y $v(x, y, z)$ del enunciado no existen. Para ello, es suficiente con que veamos que no existen sus restricciones a la recta $x = x_0 = 1$, $y = y_0 = 0$; así ocurre, ya que para tal restricción debiera ser, de acuerdo con [1]:

$$\begin{cases} u^2 = 5 - z \\ uv = 0, \end{cases} \qquad \text{o sea} \begin{cases} u = \pm\sqrt{5 - z} \\ uv = 0 \end{cases}$$

Por consiguiente: *a*) no existe u para $z > 5 = z_0$, es decir, no hay u en un semientorno de z_0; *b*) para $z < 5 = z_0$ hay dos valores de u, en lugar de haber sólo uno; *c*) por si lo ya dicho no fuera bastante, como u sólo vale 0 para $z = 5$, de la segunda ecuación se desprende que debiera ser $v = 0$ para $z \neq 5 = z_0$, siendo así que sabemos que $v = 2$ para $z = z_0$, con lo que v no puede ser continua en (x_0, y_0, z_0).

3.° En el primer caso, esto es, para $x_0 = 0$, $y_0 = 1$, $z_0 = 1$, $u_0 = 0$ y $v_0 = 2$, las derivadas respecto de x de las $u(x, y, z)$ y $v(x, y, z)$ definidas implícitamente (que existen, según ya se indicó) se obtienen derivando, respecto de x, en [1], de lo que se obtiene:

$$u^2 + 2xuu'_x + 2yvv'_x = 0 \qquad \text{e} \qquad yz + u'_xv + uv'_x = 0$$

de donde, al particularizar para el punto dado, $(x, y, z, u, v) = (0, 1, 1, 0, 2)$, se llega a que

$$4v'_x(0, 1, 1) = 0 \qquad \text{y} \qquad 1 + 2u'_x(0, 1, 1) = 0$$

o sea, $v'_x(0, 1, 1) = 0$ y $u'_x(0, 1, 1) = -1/2$.

[36]$_3$ Función implícita de clase $\mathscr{C}^r$

Sea $\boldsymbol{F}: U \to \mathbb{R}^q$, $(\boldsymbol{x}; \boldsymbol{y}) \mapsto \boldsymbol{F}(\boldsymbol{x}; \boldsymbol{y})$ una función de clase $\mathscr{C}^r$, con $r \geqslant 1$, en un entorno $U \subset \mathbb{R}^{p+q}$ de un punto $(\boldsymbol{a}; \boldsymbol{b}) \in \mathbb{R}^{p+q}$. Si $\det [J_y\boldsymbol{F}(\boldsymbol{a}; \boldsymbol{b})] \neq 0$ y $\boldsymbol{F}(\boldsymbol{a}; \boldsymbol{b}) = \boldsymbol{o}$, entonces la ecuación $\boldsymbol{F}(\boldsymbol{x}; \boldsymbol{y}) = \boldsymbol{o}$ define implícitamente, en torno de $\boldsymbol{x} = \boldsymbol{a}$ e $\boldsymbol{y} = \boldsymbol{b}$, a la variable $\boldsymbol{y}$ como función única de $\boldsymbol{x}$, $\boldsymbol{x} \mapsto \boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x})$, siendo esta una función de clase $\mathscr{C}^r$ en un entorno de $\boldsymbol{x} = \boldsymbol{a}$. Es decir, al verificarse las anteriores hipótesis, existen sendos entornos $U_a \subset \mathbb{R}^p$ de $\boldsymbol{a}$ y $U_b \subset \mathbb{R}^q$ de $\boldsymbol{b}$, tales que $U_a \times U_b \subset U$, y existe una única función $\boldsymbol{f}: U_a \to U_b$, $\boldsymbol{x} \mapsto \boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x})$, de clase $\mathscr{C}^r$ en U_a tal que $\boldsymbol{b} = \boldsymbol{f}(\boldsymbol{a})$ y $\boldsymbol{F}(\boldsymbol{x}; \boldsymbol{f}(\boldsymbol{x})) = \boldsymbol{o}$ para todo $\boldsymbol{x} \in U_a$. Las derivadas parciales de esta función $\boldsymbol{f}$ se obtienen derivando o diferenciando, con la ayuda de la regla de la cadena, en $\boldsymbol{F}(\boldsymbol{x}; \boldsymbol{f}(\boldsymbol{x})) \equiv \boldsymbol{o}$ para $\boldsymbol{x} \in U_a$.

Comprobación

Las anteriores hipótesis son más exigentes que las del teorema de la función implícita (véase [36]), ya que allí se exigía que $\boldsymbol{F}$ fuese de clase $\mathscr{C}^1$ y aquí se pide que sea de clase $\mathscr{C}^r$, con $r \geqslant 1$. Ahora se verifica, pues, la tesis de dicho teorema de la función implícita, así como todas sus consecuencias. Por ello, podemos asegurar que existe una única función $\boldsymbol{f}$ definida implícitamente por $\boldsymbol{F}(\boldsymbol{x}; \boldsymbol{y}) = \boldsymbol{o}$, que $\boldsymbol{f}$ es de clase $\mathscr{C}^1$, al menos, en torno de $\boldsymbol{x} = \boldsymbol{a}$ y que sus derivadas parciales están dadas por la expresión [4] de $[36]_1$, es decir, por

$$\left[\frac{\partial f}{\partial x_j}\right]^t = \left[\frac{\partial(F_1, \ldots, F_q)}{\partial(y_1, \ldots, y_q)}\right]^{-1} \left[\frac{\partial \boldsymbol{F}}{\partial x_j}\right]^t$$

(x_j es la componente j-ésima de $\boldsymbol{x}$; la inversa de la matriz jacobiana existe, pues ella es regular). Nótese que, como $\boldsymbol{F} = (F_1, F_2, \ldots, F_q)$ es de clase $\mathscr{C}^r$, es evidente que todos los elementos de las matrices del segundo miembro son de clase $\mathscr{C}^{r-1}$, luego también lo son los elementos de la matriz del primer miembro, es decir, las derivadas parciales de $\boldsymbol{f}$ son de clase $\mathscr{C}^{r-1}$, lo que significa que $\boldsymbol{f}$ es de clase $\mathscr{C}^r$, todo ello en los correspondientes entornos del enunciado.

$[36]_4$ Observación

Para obtener las derivadas parciales de todo orden de la función $\boldsymbol{x} \mapsto \boldsymbol{y} = \boldsymbol{y}(\boldsymbol{x})$ definida implícitamente por la ecuación $\boldsymbol{F}(\boldsymbol{x}; \boldsymbol{y}) = \boldsymbol{o}$, de la que se habla en el apartado anterior, basta con igualar a cero las derivadas sucesivas del primer miembro de la identidad

$$\boldsymbol{F}(\boldsymbol{x}; \boldsymbol{y}(\boldsymbol{x})) \equiv \boldsymbol{o}, \qquad \text{o bien} \begin{cases} F_i(x_1, \ldots, x_p;\ y_1(x_1, \ldots, x_p), \ldots, y_q(x_1, \ldots, x_q)) \equiv 0 \\ (q \text{ identidades: para } i = 1, 2, \ldots, q) \end{cases}$$

Al derivar una vez (respecto de la variable x_h, para $h = 1, 2, \ldots, p$) se obtiene (según ya vimos en $[36]_1$, [3]) que:

$$\begin{cases} \dfrac{\partial F_i}{\partial x_h} + \dfrac{\partial F_i}{\partial y_1}\dfrac{\partial y_1}{\partial x_h} + \cdots + \dfrac{\partial F_i}{\partial y_q}\dfrac{\partial y_q}{\partial x_h} \equiv 0 \\ (q \text{ identidades: para } i = 1, 2, \ldots, q) \end{cases} \tag{1}$$

de donde, como ya se hizo anteriormente (véase $[36]_1$, [4]), se obtienen las q derivadas parciales $\partial y_1/\partial x_h, \ldots, \partial y_q/\partial x_h$. Si se deriva de nuevo, ahora en [1] y respecto de la variable x_k (para $k = 1, 2, \ldots, p$), se obtiene:

$$\begin{aligned} &\frac{\partial^2 F_i}{\partial x_h \partial x_k} + \frac{\partial^2 F_i}{\partial x_h \partial y_1}\frac{\partial y_1}{\partial x_k} + \cdots + \frac{\partial^2 F_i}{\partial x_h \partial y_q}\frac{\partial y_q}{\partial x_k} + \\ &+ \left(\frac{\partial^2 F_i}{\partial y_1 \partial x_k} + \frac{\partial^2 F_i}{\partial y_1 \partial y_1}\frac{\partial y_1}{\partial x_k} + \cdots + \frac{\partial^2 F_i}{\partial y_1 \partial y_q}\frac{\partial y_q}{\partial x_k}\right)\frac{\partial y_1}{\partial x_h} + \frac{\partial F_i}{\partial y_1}\frac{\partial^2 y_1}{\partial x_h \partial x_k} + \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &+ \left(\frac{\partial^2 F_i}{\partial y_q \partial x_k} + \frac{\partial^2 F_i}{\partial y_q \partial y_1}\frac{\partial y_1}{\partial x_k} + \cdots + \frac{\partial^2 F_i}{\partial y_q \partial y_q}\frac{\partial y_q}{\partial x_k}\right)\frac{\partial y_q}{\partial x_h} + \frac{\partial F_i}{\partial y_q}\frac{\partial^2 y_q}{\partial x_h \partial x_k} \equiv 0 \end{aligned} \tag{2}$$

Suponiendo que ya nos son conocidas las derivadas parciales primeras (de las $y_1, ..., y_q$ respecto de x_h y respecto de x_k), las q identidades [2] permiten calcular las derivadas parciales segundas $\partial^2 f_j/\partial x_h \partial x_k$ (para $j = 1, 2, ..., q$). Nótese que dicho sistema [2] es lineal en las incógnitas $\partial^2 f_j/\partial x_h \partial x_k$ y que el determinante del sistema (de q ecuaciones y q incógnitas) no es otro que el jacobiano de $\boldsymbol{F}$ respecto de $\boldsymbol{y}$, que es no nulo, lo que permite, en efecto, despejar las referidas incógnitas.

[36]$_5$ Ejercicio

Hallar las derivadas segundas, en el punto $(x, y) = (1, 2)$, de las funciones $u(x, y)$ y $v(x, y)$ definidas implícitamente por:

$$\begin{cases} e^{x-u} + yv + 3 = 0 \\ e^{y+v} - xu = 0 \end{cases} \qquad \begin{matrix} u(1, 2) = 1 \\ v(1, 2) = -2 \end{matrix}$$

Resolución

Derivando respecto de x, primero, y respecto de y, después, en las dos ecuaciones dadas, se obtiene:

$$\begin{cases} e^{x-u}(1 - u'_x) + yv'_x = 0 \\ e^{y+v}v'_x - u - xu'_x = 0 \end{cases} \qquad \begin{cases} e^{x-u}(-u'_y) + v + yv'_x = 0 \\ e^{y+v}(1 + v'_y) - xu'_y = 0 \end{cases} \qquad [1]$$

Particularizando para $(x, y, u, v) = (1, 2, 1, -2)$, se llega a que:

$$\left.\begin{matrix} 1 - u'_x + 2v'_x = 0 \\ v_x - 1 - u'_x = 0 \end{matrix}\right\} \rightarrow \begin{matrix} u'_x = -3 \\ v'_x = -2 \end{matrix} \qquad \left.\begin{matrix} -u'_y - 2 - 2v'_y = 0 \\ 1 + v'_y - u'_y = 0 \end{matrix}\right\} \rightarrow \begin{matrix} u'_y = 4 \\ v'_y = 3 \end{matrix}$$

Derivando nuevamente en [1] respecto de x y respecto de y, se obtiene:

$$\left.\begin{matrix} e^{x-u}(1 - u'_x)^2 - e^{x-u}u''_{xx} + yv''_{xx} = 0 \\ e^{y+v}(v'_x)^2 + e^{y+v}v''_{xx} - 2u'_x - xu''_{xx} = 0 \end{matrix}\right\} \rightarrow \left.\begin{matrix} 16 - u''_{xx} + 2v''_{xx} = 0 \\ 4 + v''_{xx} + 6 - u''_{xx} = 0 \end{matrix}\right\} \rightarrow \begin{matrix} u''_{xx} = 4 \\ v''_{xx} = -6 \end{matrix}$$

$$\left.\begin{matrix} -e^{x-u}(1 - u'_x)u'_y - e^{x-u}u''_{xy} + v'_x + yv''_{xy} = 0 \\ e^{y+v}v'_x(1 + v'_y) + e^{y+v}v''_{xy} - u'_y - xu''_{xy} = 0 \end{matrix}\right\} \rightarrow \left.\begin{matrix} -18 - u''_{xy} + 2v''_{xy} = 0 \\ 4 + v''_{xy} - u''_{xy} = 0 \end{matrix}\right\} \rightarrow \begin{matrix} u''_{xy} = 26 \\ v''_{xy} = 22 \end{matrix}$$

$$\left.\begin{matrix} e^{x-u}(u'_y)^2 - e^{x-u}u''_{yy} + 2v'_y + yv''_{yy} = 0 \\ e^{y+v}(1 + v'_y)^2 + e^{y+v}v''_{yy} - xu''_{yy} = 0 \end{matrix}\right\} \rightarrow \left.\begin{matrix} 22 - u''_{yy} + 2v''_{yy} = 0 \\ 16 + v''_{yy} - u''_{yy} = 0 \end{matrix}\right\} \rightarrow \begin{matrix} u''_{yy} = 10 \\ v''_{yy} = -6 \end{matrix}$$

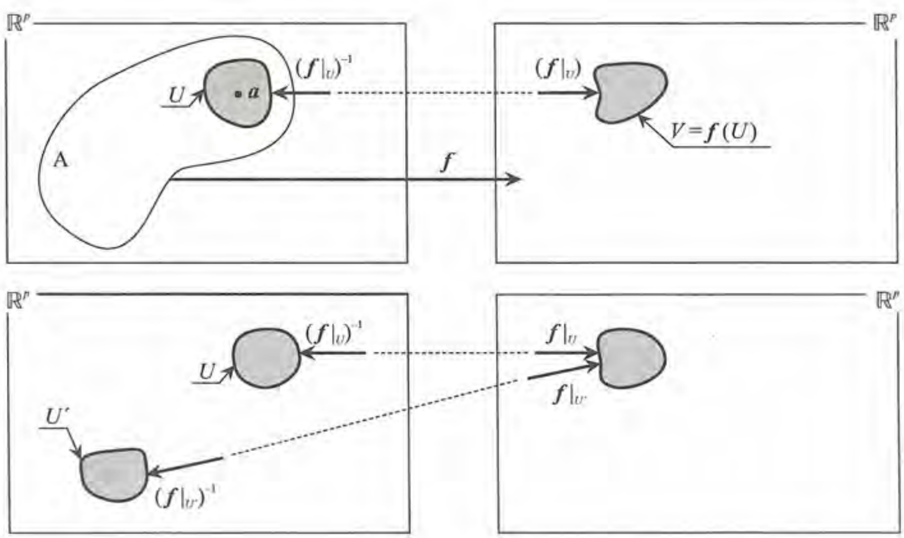
3.2. EXISTENCIA Y REGULARIDAD DE LA FUNCIÓN INVERSA

Según es conocido, se dice que una función $f: C \to D$, definida en un conjunto C y que toma sus valores en un conjunto D, es biyectiva si cada elemento de D es la imagen, por f, de uno y sólo uno elemento de C; en tal caso, se llama función inversa o recíproca de f a la $f^{-1}: D \to C$ que asigna a cada $d \in C$ el único $c \in C$ tal que $f(c) = d$, con lo que $f^{-1}(d) = c$ equivale a $f(c) = d$.

Son pocas las funciones que admiten inversa global, en el sentido que acabamos de considerar, lo que no es necesariamente un impedimento para que exista «inversa local», de una función $f: C \to D$, en las «proximidades» de un punto $c_0 \in C$, en el que la función alcanza un valor $d_0 \in D$. Este concepto de «función inversa local», del que en seguida nos ocuparemos, es suficiente para el estudio de los asuntos que, al respecto, se consideran en Cálculo Infinitesimal.

INVERSA LOCAL DE UNA FUNCIÓN

[37] Sea dada una función $f: A \to \mathbb{R}^p$, donde $A \subset \mathbb{R}^p$; esto es, f es una transformación entre puntos de $\mathbb{R}^p$. Se dice que f es «localmente biyectiva» en un punto $\boldsymbol{a} \in \mathbb{R}^p$, interior de A, si existe un entorno $U \subset A$ de $\boldsymbol{a}$ tal que, llamando $V = f(U)$, la función $f: U \to V$ es biyectiva; en tal caso, se llama «inversa local» de f en $\boldsymbol{a}$ a la función $f^{-1}: V \to U$, para la que se verifica que $f^{-1}(f(\boldsymbol{x})) = \boldsymbol{x}$ para todo $\boldsymbol{x} \in U$, y también $f(f^{-1}(\boldsymbol{y})) = \boldsymbol{y}$ para todo $\boldsymbol{y} \in V$. Si A es un abierto de $\mathbb{R}^p$, se dice que f es localmente biyectiva en A si es localmente biyectiva en todo punto $\boldsymbol{a} \in A$.



[37]₁ Ejemplo

Considérese la transformación $f: C \to \mathbb{R}^2$, donde $C = \{(x, y) \in \mathbb{R}^2 / y \geqslant 0\}$, definida mediante

$$(x, y) \mapsto f(x, y) = (u, v), \qquad \text{siendo} \begin{cases} u = \dfrac{1}{2}(|x| - y) \\ v = \sqrt{|x|y} \end{cases} \tag{1}$$

Nótese que la correspondencia recíproca de f viene dada por

$$|x| = u + \sqrt{u^2 + v^2} \qquad \text{e} \qquad y = -u + \sqrt{u^2 + v^2} \qquad (\text{para } v \geqslant 0)$$

La imagen de f es $f(C) = \{(u, v) \in \mathbb{R}^2 / v \geqslant 0\} = C$. La aplicación f no es inyectiva ya que los puntos (x, y) y $(-x, y)$ tienen la misma imagen. Sin embargo, f es localmente biyectiva en muchos de los puntos de C, pero no lo es en todos; en concreto:

- En un punto $(x_0, y_0) \in C$ con $x_0 > 0$ e $y_0 > 0$ la función f es localmente biyectiva en el entorno de radio $r = \text{mín}(x_0, y_0)$ de (x_0, y_0) y su inversa local está dada por:

$$(u, v) \mapsto (x, y) = (u + \sqrt{u^2 + v^2}, -u + \sqrt{u^2 + v^2})$$

- En un punto $(x_0, y_0) \in C$ con $x_0 < 0$ e $y_0 > 0$, la función f es localmente biyectiva en el entorno de radio $r = \text{mín}(-x_0, y_0)$ de (x_0, y_0) y su inversa local está dada por:

$$(u, v) \mapsto (x, y) = (-u - \sqrt{u^2 + v^2}, -u + \sqrt{u^2 + v^2})$$

- En un punto $(x_0, 0) \in C$, la función f no es localmente biyectiva pues $(x_0, 0)$ no es un punto interior de C.
- En un punto $(0, y_0) \in C$ con $y_0 > 0$, la función f no es localmente biyectiva pues, por muy pequeño que sea el radio $\varepsilon > 0$ de un entorno que se tome de $(0, y_0)$, los puntos $(\varepsilon/2, y_0)$ y $(-\varepsilon/2, y_0)$ son de dicho entorno y tienen la misma imagen por f, luego f no es intectiva en ningún entorno de $(0, y_0)$.

[37]₂ Observación

El hecho de que una función $f: A \to \mathbb{R}^p$, donde $A \subset \mathbb{R}^p$, tenga inversa local en todos los puntos de A no implica, en general, que f tenga función inversa en A. Para comprobar lo dicho, acudamos al siguiente ejemplo.

Considérese la función $f: A \to \mathbb{R}^2$, donde $A = \{(x, y) \in \mathbb{R}^2 / x > 0\}$, definida mediante

$$(x, y) \mapsto f(x, y) = (u, v), \qquad \text{donde} \begin{cases} u = x \cos y \\ v = x \operatorname{sen} y \end{cases}$$

Nótese que la recta de ecuación $x = x_0$ (para cualquiera que sea $x_0 > 0$, fijo) se transforma en la circunferencia $u^2 + v^2 = x_0^2$ (de centro el origen y radio x_0); para cualquiera que sea $y_0 \in \mathbb{R}$, el segmento $x = x_0$, $y_0 \leqslant y < y_0 + 2\pi$ se transforma biyectivamente en la citada circunferencia $u^2 + v^2 = x_0^2$. De lo dicho se desprende que, dado cualquier $(x_0, y_0) \in A$, si es $\varepsilon = \text{mín}\ \{x_0, \pi\}$, la restricción de f al entorno de radio ε del punto (x_0, y_0) es inyectiva, por lo que f es localmente biyectiva en todo punto de A. No obstante, es claro que f no es globalmente biyectiva, de A en $f(A)$, ya que $f(x, y) = f(x, y + 2\pi)$ para cualquiera que sea $(x, y) \in A$.

TEOREMA DE LA FUNCIÓN INVERSA

Vamos a ocuparnos aquí de una generalización natural del teorema que, para una función real f de una variable definida en torno de un punto a, asegura la existencia de su inversa o recíproca, en las proximidades del punto $b = f(a)$, en el caso de que la función dada sea continua y estrictamente monótona, cosa que seguro acontece si f es de clase $\mathscr{C}^1$ y $f'(a) \neq 0$.

El siguiente teorema da condiciones suficientes para que una transformación f suficientemente regular (de clase $\mathscr{C}^1$, al menos), entre puntos de $\mathbb{R}^p$, admita inversa local, en torno de un cierto punto $\boldsymbol{a}$. Conviene fijarse en que lo que este teorema dice, de f en las proximidades de $\boldsymbol{a}$, es la versión local de lo que acontece, globalmente, a la diferencial de f en $\boldsymbol{a}$, $df(\boldsymbol{a})$. Al suponer que esta aplicación lineal, $df(\boldsymbol{a})$, es biyectiva de $\mathbb{R}^p$ en $\mathbb{R}^p$, esto es, en la hipótesis de que su matriz $Jf(\boldsymbol{a})$ es regular (o sea, si el jacobiano de f en $\boldsymbol{a}$ es no nulo), nuestro teorema asegura que también hay, entonces, biyectividad para f en $\boldsymbol{a}$, pero ésta tiene carácter local.

[38]

Sea $f: U \to \mathbb{R}^p$, $\boldsymbol{x} \mapsto \boldsymbol{y} = f(\boldsymbol{x})$, una transformación entre puntos de $\mathbb{R}^p$, donde $U \subset \mathbb{R}^p$ es un entorno de un punto $\boldsymbol{a} \in \mathbb{R}^p$. Si f es de clase $\mathscr{C}^1$ en U y si el determinante jacobiano de f en $\boldsymbol{a}$ es no nulo, det $[Jf(\boldsymbol{a})] \neq 0$, entonces f admite inversa local en $\boldsymbol{a}$ o, más exactamente, existen un entorno U_a del punto $\boldsymbol{a}$ y un entorno V_b del punto $\boldsymbol{b} = f(\boldsymbol{a})$ tales que $f: U_a \to V_b$ es biyectiva.

Además se verifica, entonces, que la función inversa local $f^{-1}: V_b \to U_a$ es de clase $\mathscr{C}^1$; la matriz jacobiana y la diferencial de f^{-1} en un punto $\boldsymbol{y} = f(\boldsymbol{x})$ de V_b son las inversas de las de f en el punto $\boldsymbol{x} \in U_a$, es decir:

$$Jf^{-1}(\boldsymbol{y}) = [Jf(\boldsymbol{x})]^{-1} \qquad \text{y} \qquad df^{-1}(\boldsymbol{y}) = [df(\boldsymbol{x})]^{-1}$$

Así pues, la derivada parcial $\partial f^{-1}(\boldsymbol{y})/\partial y_i$ (donde y_i es la coordenada i-ésima de $\boldsymbol{y}$) y la diferencial $df^{-1}(\boldsymbol{y})$, en un punto $\boldsymbol{y} = f(\boldsymbol{x})$ de V_b, quedan definidas por las siguientes relaciones(*):

$$Jf(\boldsymbol{x})\left[\frac{\partial f^{-1}(\boldsymbol{y})}{\partial y_i}\right]^t = I_i$$

$$Jf(\boldsymbol{x})[df^{-1}(\boldsymbol{y})(d\boldsymbol{y})]^t = [d\boldsymbol{y}]^t$$

Estas fórmulas se analizan, expresándolas en términos de sus componentes, en [38]₁

(*) Las matrices con una «t» de supraíndice, que significa «traspuesta de», son matrices columna. La atriz I_i es la columna que tiene todos sus elementos nulos excepto el de lugar i, que es un 1.

Nota: Bajo las hipótesis de este teorema y si el jacobiano de f es no nulo en todo punto $\boldsymbol{a}$ de un abierto $A \subset U$, entonces $f(A)$ es un abierto.

Demostración

Este teorema puede reducirse al de la función implícita (véase [36]), ya que son equivalentes las dos proposiciones que se enuncian a continuación:

1.ª La función $\boldsymbol{x} \mapsto \boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x})$ admite inversa local cerca de $\boldsymbol{a}$ (a la que llamamos $\boldsymbol{g}$, en lugar de $\boldsymbol{f}^{-1}$), $\boldsymbol{y} \mapsto \boldsymbol{x} = \boldsymbol{g}(\boldsymbol{y})$ de clase $\mathscr{C}^1$, para $\boldsymbol{y}$ en un cierto entorno del punto $\boldsymbol{b} = \boldsymbol{f}(\boldsymbol{a})$, de manera que $\boldsymbol{f}(\boldsymbol{g}(\boldsymbol{y})) = \boldsymbol{y}$ para todo $\boldsymbol{y}$ de dicho entorno; es obvio que $\boldsymbol{a} = \boldsymbol{g}(\boldsymbol{b})$.

2.ª La ecuación $\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{y} = \boldsymbol{o}$ define implícitamente a $\boldsymbol{x}$ como función de $\boldsymbol{y}$ en torno de $\boldsymbol{y} = \boldsymbol{b}$ y $\boldsymbol{x} = \boldsymbol{a}$, es decir, existe la función $\boldsymbol{y} \mapsto \boldsymbol{x} = \boldsymbol{g}(\boldsymbol{y})$ de clase $\mathscr{C}^1$, para $\boldsymbol{y}$ en un cierto entorno del punto $\boldsymbol{b} = \boldsymbol{f}(\boldsymbol{a})$, tal que satisface a la identidad $\boldsymbol{f}(\boldsymbol{g}(\boldsymbol{y})) - \boldsymbol{y} \equiv \boldsymbol{o}$, para todo $\boldsymbol{y}$ de dicho entorno, siendo $\boldsymbol{a} = \boldsymbol{g}(\boldsymbol{b})$.

Así pues, para cerciorarnos de que, en efecto, se verifica la afirmación primera, vamos a comprobar que es cierta la segunda, que equivale a aquélla. Dicha afirmación segunda es cierta ya que, para ella, se verifican todas las hipótesis del teorema [36], de la función implícita. En efecto: nótese en primer lugar que los papeles de $\boldsymbol{x}$ e $\boldsymbol{y}$ se intercambian (aquí es $\boldsymbol{y}$ lo que allí es $\boldsymbol{x}$ y viceversa); la que allí es función $(\boldsymbol{x}; \boldsymbol{y}) \mapsto \boldsymbol{F}(\boldsymbol{x}; \boldsymbol{y})$, aquí es la función $(\boldsymbol{y}; \boldsymbol{x}) \mapsto \boldsymbol{F}(\boldsymbol{y}; \boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{y}$, que es de clase $\mathscr{C}^1$ en $\mathbb{R}^p \times U$ (entorno del punto $(\boldsymbol{b}; \boldsymbol{a})$); el jacobiano det $[J_y\boldsymbol{F}(\boldsymbol{a}; \boldsymbol{b})]$ que debe ser no nulo, aquí es el

$$\det\,[J_x\boldsymbol{F}(\boldsymbol{b}; \boldsymbol{a})] = \det\,[J_x\boldsymbol{f}(\boldsymbol{a})] = \det\,[J\boldsymbol{f}(\boldsymbol{a})] \neq 0$$

y finalmente la función $\boldsymbol{F}$ se anula en $(\boldsymbol{b}; \boldsymbol{a})$ ya que $\boldsymbol{F}(\boldsymbol{b}; \boldsymbol{a}) = \boldsymbol{f}(\boldsymbol{a}) - \boldsymbol{b} = \boldsymbol{b} - \boldsymbol{b} = 0$. Luego la proposición segunda es verdadera, lo que prueba la existencia local, cerca de $\boldsymbol{x} = \boldsymbol{a}$, de la función inversa y que esta es de clase $\mathscr{C}^1$ en un cierto entorno V_b del punto $\boldsymbol{b} = \boldsymbol{f}(\boldsymbol{a})$.

Si llamamos $U_a = \boldsymbol{f}^{-1}(V_b)$ hemos de comprobar ahora que U_a es un entorno de $\boldsymbol{a}$, es decir, que $\boldsymbol{a}$ es un punto interior de U_a. Esto es así ya que $\boldsymbol{f}$ es una función continua en $\boldsymbol{a}$ y, por ello, la imagen recíproca de un entorno de $\boldsymbol{b} = \boldsymbol{f}(\boldsymbol{a})$ es un entorno de $\boldsymbol{a}$ y, en particular $U_a = \boldsymbol{f}^{-1}(V_b)$ es, pues, un entorno de $\boldsymbol{a}$.

Como para todo $\boldsymbol{y} \in V_b$ es $(\boldsymbol{f} \circ \boldsymbol{f}^{-1})(\boldsymbol{y}) = \boldsymbol{y}$ y $\boldsymbol{f}^{-1}$ es diferenciable en el punto $\boldsymbol{y}$ y $\boldsymbol{f}$ es diferenciable en el punto $\boldsymbol{x} = \boldsymbol{f}^{-1}(\boldsymbol{y})$, resulta que, de acuerdo con el teorema de diferenciación de la función compuesta (véase [29], 2.º) y como la función identidad $\boldsymbol{y} \mapsto \boldsymbol{i}(\boldsymbol{y})$ es diferenciable y su diferencial es la propia identidad, se verifica:

$$d\boldsymbol{f}(\boldsymbol{x}) \circ d\boldsymbol{f}^{-1}(\boldsymbol{y}) = \boldsymbol{i} \qquad \text{o} \qquad d\boldsymbol{f}^{-1}(\boldsymbol{y}) = [d\boldsymbol{f}(\boldsymbol{x})]^{-1}$$

Como ha resultado que las aplicaciones lineales $d\boldsymbol{f}^{-1}(\boldsymbol{y})$ y $d\boldsymbol{f}(\boldsymbol{x})$ (endomorfismos de $\mathbb{R}^p$) son inversas la una de la otra, también son inversas, una de otra, sus respectivas matrices (jacobianas), es decir, se sabe que:

$$J\boldsymbol{f}^{-1}(\boldsymbol{y}) = [J\boldsymbol{f}(\boldsymbol{x})]^{-1}$$

Las expresiones de las derivadas parciales y de la diferencial de la función inversa, $\boldsymbol{f}^{-1}$, se obtienen derivando parcialmente y diferenciando en la identidad $\boldsymbol{f}(\boldsymbol{f}^{-1}(\boldsymbol{y})) \equiv \boldsymbol{y}$ para $\boldsymbol{y} \in V_b$. Sobre estas expresiones y sobre sus desarrollos en función de sus componentes nos ocupamos a continuación, en el próximo apartado $[38]_1$.

Respecto de la «nota» del final del enunciado, se tiene: si det $[J\boldsymbol{f}(\boldsymbol{a})] \neq 0$ para todo punto $\boldsymbol{a}$ de un cierto abierto $A \subset U$, el teorema es válido si se sustituye U por A y $\boldsymbol{a}$ es cualquier punto de A. El conjunto $B = \boldsymbol{f}(A)$ es abierto pues, para cualquier $\boldsymbol{b} \in B$, el entorno V_b de $\boldsymbol{b}$ está incluido en B, pues $U_a \subset A$.

$[38]_1$ Derivación y diferenciación de la función inversa

Aquí vamos a suponer que, para la siguiente función $\boldsymbol{f} = (f_1, f_2, ..., f_p)$ (usaremos la misma notación que en [38]):

$$\boldsymbol{x} \mapsto \boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x}) \qquad \text{o} \qquad (x_1, x_2, ..., x_p) \mapsto \begin{cases} y_1 = f_1(x_1, x_2, ..., x_p) \\ y_2 = f_2(x_1, x_2, ..., x_p) \\ \dots\dots\dots\dots\dots\dots \\ \dots\dots\dots\dots\dots\dots \\ y_p = f_p(x_1, x_2, ..., x_p) \end{cases}$$

se satisfacen las hipótesis del anterior teorema [38] de la función inversa en torno de $(\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{a}, \boldsymbol{f}(\boldsymbol{a}))$, por lo que existe la función inversa local $\boldsymbol{y} \mapsto \boldsymbol{x} = \boldsymbol{f}^{-1}(\boldsymbol{y})$ en un entorno del punto $\boldsymbol{b} = \boldsymbol{f}(\boldsymbol{a})$, en el que se verifica, pues, que:

$$\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{f}^{-1}(\boldsymbol{y})) \qquad \text{o} \qquad \begin{cases} y_1 = f_1(f_1^{-1}(y_1, ..., y_p), ..., f_p^{-1}(y_1, ..., y_p)) \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ y_p = f_p(f_1^{-1}(y_1, ..., y_p), ..., f_p^{-1}(y_1, ..., y_p)) \end{cases} \qquad [1]$$

(para todo $\boldsymbol{y}$ perteneciente a un cierto entorno del punto $\boldsymbol{b} = \boldsymbol{f}(\boldsymbol{a})$; las componentes de $\boldsymbol{f}^{-1}$ se han denotado por $f_1^{-1}, ..., f_p^{-1}$).

DERIVADAS PARCIALES DE LA FUNCION INVERSA. De acuerdo con la regla de la cadena (véase [29], 1.º), al derivar parcialmente en [1] respecto de la variable y_i (para $i = 1, 2, ..., p$), se obtiene que:

$$\left.\begin{array}{l} 0 = \dfrac{\partial f_1}{\partial x_1}\dfrac{\partial f_1^{-1}}{\partial y_i} + \dfrac{\partial f_1}{\partial x_2}\dfrac{\partial f_2^{-1}}{\partial y_i} + \cdots + \dfrac{\partial f_1}{\partial x_p}\dfrac{\partial f_p^{-1}}{\partial y_i} \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ 1 = \dfrac{\partial f_i}{\partial x_1}\dfrac{\partial f_1^{-1}}{\partial y_i} + \dfrac{\partial f_i}{\partial x_2}\dfrac{\partial f_2^{-1}}{\partial y_i} + \cdots + \dfrac{\partial f_i}{\partial x_p}\dfrac{\partial f_p^{-1}}{\partial y_i} \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ 0 = \dfrac{\partial f_p}{\partial x_1}\dfrac{\partial f_1^{-1}}{\partial y_i} + \dfrac{\partial f_p}{\partial x_2}\dfrac{\partial f_2^{-1}}{\partial y_i} + \cdots + \dfrac{\partial f_p}{\partial x_p}\dfrac{\partial f_p^{-1}}{\partial y_i} \end{array}\right\} \qquad \text{o} \qquad \boldsymbol{u}_i = \sum_{k=1}^{p} \frac{\partial \boldsymbol{f}}{\partial x_k}\frac{\partial f_k^{-1}}{\partial y_i} \qquad [2]$$

siendo $\boldsymbol{u}_i = (0, ..., 1, ..., 0)$, donde todos los elementos son 0, excepto el de lugar i, que es 1. Expresando en forma matricial el anterior sistema [2], se obtiene que:

$$\begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} & \cdots & \dfrac{\partial f_1}{\partial x_p} \\ \cdots & \cdots & \cdots & \cdots \\ \dfrac{\partial f_i}{\partial x_1} & \dfrac{\partial f_i}{\partial x_2} & \cdots & \dfrac{\partial f_i}{\partial x_p} \\ \cdots & \cdots & \cdots & \cdots \\ \dfrac{\partial f_p}{\partial x_1} & \dfrac{\partial f_p}{\partial x_2} & \cdots & \dfrac{\partial f_p}{\partial x_p} \end{bmatrix} \begin{bmatrix} \dfrac{\partial f_1^{-1}}{\partial y_i} \\ \dfrac{\partial f_2^{-1}}{\partial y_i} \\ \vdots \\ \dfrac{\partial f_p^{-1}}{\partial y_i} \end{bmatrix} \quad \text{o} \quad I_i = \left[\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\right]\left[\frac{\partial \boldsymbol{f}^{-1}}{\partial y_i}\right]^t \qquad [2']$$

donde $[\partial \boldsymbol{f}/\partial \boldsymbol{x}]$ denota a la matriz jacobiana de $\boldsymbol{f}$, que también se denota poniendo $J\boldsymbol{f}$, y donde I_i es la matriz columna (traspuesta de $\boldsymbol{u}_i$) que tiene todos sus elementos nulos excepto el de lugar i, que es un 1; el superíndice t significa matriz (columna) traspuesta.

Como las referida matriz jacobiana es regular por hipótesis (pues se supone que el jacobiano det $[J\boldsymbol{f}]$ es no nulo), de [2′] se deduce que

$$\left[\frac{\partial \boldsymbol{f}^{-1}}{\partial y_i}\right]^t = \left[\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\right]^{-1} I_i \quad \text{o} \quad \begin{bmatrix} \dfrac{\partial f_1^{-1}}{\partial y_i} \\ \vdots \\ \dfrac{\partial f_p^{-1}}{\partial y_i} \end{bmatrix} = \left[\frac{\partial(f_1, ..., f_p)}{\partial(x_1, ..., x_p)}\right]^{-1} \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \qquad [3]$$

Esta expresión proporciona, pues, las derivadas parciales de la función inversa $\boldsymbol{f}^{-1} = (f_1^{-1}, ..., f_p^{-1})$ respecto de y_i (una cualquiera de sus p variables $y_1, ..., y_p$).

DIFERENCIAL DE LA FUNCIÓN INVERSA. Según la regla de diferenciación de la función compuesta (véase [29], 2.º), si diferenciamos en las relaciones [1], obtendremos:

$$d\boldsymbol{y} = \sum_{j=1}^{p} \frac{\partial \boldsymbol{f}}{\partial x_j}\, df_i^{-1}$$

esto es:

$$\left\{\begin{array}{l} dy_1 = \dfrac{\partial f_1}{\partial x_1}\, df_1^{-1} + \dfrac{\partial f_1}{\partial x_2}\, df_2^{-1} + \cdots + \dfrac{\partial f_1}{\partial x_p}\, df_p^{-1} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ dy_p = \dfrac{\partial f_p}{\partial x_1}\, df_1^{-1} + \dfrac{\partial f_p}{\partial x_2}\, df_2^{-1} + \cdots + \dfrac{\partial f_p}{\partial x_p}\, df_p^{-1} \end{array}\right. \qquad [4]$$

Estas igualdades, con notación matricial y acudiendo a la matriz jacobiana de f, que denotamos por Jf o $[\partial f/\partial \mathbf{x}]$, se pueden poner en la forma (las matrices con una «t» como supraíndice, que significa «traspuesta de», son matrices columna):

$$\begin{bmatrix} dy_1 \\ \vdots \\ dy_p \end{bmatrix} = Jf \begin{bmatrix} df_1^{-1} \\ \vdots \\ df_p^{-1} \end{bmatrix} \qquad \text{o} \qquad [dy]^t = \left[\frac{\partial f}{\partial \mathbf{x}}\right][df^{-1}]^t \qquad [4']$$

Como la matriz jacobiana $[\partial f/\partial \mathbf{x}]$ es regular (pues se ha supuesto que su determinante es no nulo en el entorno en el que nos movemos), al premultiplicar en [4′] por la matriz $[\partial f/\partial \mathbf{x}]^{-1}$, se llega a que:

$$[df^{-1}]^t = \left[\frac{\partial f}{\partial \mathbf{x}}\right]^{-1}[dy]^t$$

esto es:

$$\begin{bmatrix} df_1^{-1} \\ \vdots \\ df_p^{-1} \end{bmatrix} = \left[\frac{\partial(f_1, \ldots, f_p)}{\partial(x_1, \ldots, x_p)}\right]^{-1} \begin{bmatrix} dy_1 \\ \vdots \\ dy_p \end{bmatrix} \qquad [5]$$

esta expresión [5] es la que proporciona la diferencial $df^{-1} = (df_1^{-1}, \ldots, df_p^{-1})$ de la función inversa $f^{-1} = (f_1^{-1}, \ldots, f_p^{-1})$.

[38]$_2$ Ejercicio

Considérese la función $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ definida mediante:

$$(x, y) \mapsto f(x, y) = (u, v), \qquad \text{siendo} \begin{cases} u = -3x + y^3 \\ v = -3y + x^3 \end{cases} \qquad [1]$$

Analizar si f admite inversa local en torno de un punto (x_0, y_0), en cada uno de los dos casos siguientes:

1.º $(x_0, y_0) = (1, 0)$.
2.º $(x_0, y_0) = (1, 1)$.

En el caso de que la inversa exista y sea derivable, hallar sus derivadas parciales respecto de u en el punto $(u_0, v_0) = f(x_0, y_0)$.

Resolución

La función dada es de clase $\mathscr{C}^{\infty}$ en todo punto $(x_0, y_0) \in \mathbb{R}^2$. Si su jacobiano en (x_0, y_0) es no nulo, entonces es de aplicación el teorema de la función implícita. El jacobiano de f es:

$$\det\,[Jf(x_0, y_0)] = \begin{vmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{vmatrix} = \begin{vmatrix} -3 & 3y^2 \\ 3x^2 & -3 \end{vmatrix} = 9(1 - x^2y^2) \qquad [2]$$

Analicemos cada uno de los dos casos que se citan en el enunciado:

1.º En el punto $(x_0, y_0) = (1, 0)$ el jacobiano de f es no nulo, pues valor det $[Jf(1, 0)] = 9 \neq 0$. Por tanto, aquí es de aplicación el teorema de la función inversa, del que se deduce que f admite inversa local en torno del punto $(1, 0)$, $(u, v) \mapsto (x, y)$, la cual es de clase $\mathscr{C}^1$, al menos, en un cierto entorno de $(1, 0)$.

Las derivadas parciales $x'_u(1, 0)$ e $y'_u(1, 0)$, que se nos piden, se obtienen derivando parcialmente en [1], respecto de u (las variables son u y v; las funciones son x e y); haciéndolo, se obtiene:

$$\begin{cases} 1 = -3x'_u + 3y^2y'_u \\ 0 = -3y'_u + 3x^2x'_u \end{cases} \qquad \text{luego} \begin{cases} 1 = -3x'_u(1,\ 0) + 0 \\ 0 = -3y'_u(1,\ 0) + 3x'_u(1,\ 0) \end{cases}$$

de donde se obtiene que $x'_u(1, 0) = -1/3$ e $y'_u(1, 0) = 1/3$.

2.º En este caso el jacobiano [2] es nulo, det $[Jf(1, 1)] = 9(1 - 1) = 0$, por lo que el teorema de la función inversa no puede aplicarse ahora. Estudiemos, pues, directamente este caso; haciéndolo, vamos a poder comprobar que no existe inversa local de f en torno de $(x, y) = (1, 1)$, pues encontraremos puntos distintos, tan próximos al $(1, 1)$ como se desee, en los que f alcanza el mismo valor.

La restricción de f a la recta $x = y$ es la función

$$(x, x) \mapsto (u, v), \qquad \text{donde} \begin{cases} u = -3x + x^3 \\ v = -3x + x^3 \end{cases}$$

Como la función $x \mapsto -3x + x^3$ presenta un mínimo relativo para $x = 1$, hay valores x' y x'' tan próximos a $x = 1$ como se quiera (siendo $x' < 1 < x''$) tales que $-3x' + x'^3$ y $-3x'' + x''^3$ son iguales, por lo que

$$u(x', x') = u(x'', x'') \qquad \text{y} \qquad v(x', x') = v(x'', x'')$$

para puntos (x', x') y (x'', x'') tan próximos a $(x, y) = (1, 1)$ como se quiera. De ello se desprende, pues, que $f: (x, y) \mapsto (u, v)$ no tiene inversa local cerca de $(1, 1)$.

$[38]_3$ Función inversa de una de clase $\mathscr{C}^r$

Sea $f: U \to \mathbb{R}^p$, $\boldsymbol{x} \mapsto \boldsymbol{y} = f(\boldsymbol{x})$, una función de clase $\mathscr{C}^r$, con $r \geqslant 1$, en un entorno $U \subset \mathbb{R}^p$ de un punto $\boldsymbol{a} \in \mathbb{R}^p$. Si det $[Jf(\boldsymbol{a})] \neq 0$, entonces f admite inversa local en $\boldsymbol{a}$, la cual es de clase $\mathscr{C}^r$ en un entorno del punto $\boldsymbol{b} = f(\boldsymbol{a})$. Es decir, al verificarse las anteriores hipótesis, existen sendos entornos $U_a \subset \mathbb{R}^p$ de $\boldsymbol{a}$ y $V_b \subset \mathbb{R}^p$ de $\boldsymbol{b}$ tales que $f: U_a \to V_b$ es función biyectiva y su inversa $f^{-1}: V_b \to U_a$ (inversa local de f cerca de $\boldsymbol{a}$) es de clase $\mathscr{C}^r$ en V_b. Las derivadas parciales de esta función f^{-1} se obtienen derivando o diferenciando, con la ayuda de la regla de la cadena, en $\boldsymbol{y} \equiv f(f^{-1}(\boldsymbol{y}))$ para $\boldsymbol{y} \in V_b$.

Comprobación

Las anteriores hipótesis se diferencian de las del teorema de la función inversa (véase [38]) en que ahora se exige que f sea de clase $\mathscr{C}^r$, en lugar de pedir que sea, sólo, de clase $\mathscr{C}^1$. Por tanto, ahora se verifica la tesis de aquel teorema y todas sus consecuencias; así pues, existe la citada inversa local $f^{-1}: V_b \to U_a$, que es de clase $\mathscr{C}^1$, al menos, en V_b y sus derivadas parciales están dadas por la fórmula [3] de $[38]_1$, es decir, para todo $y \in V_b$ y si es $x = f^{-1}(y)$, se verifica que:

$$\left[\frac{\partial f^{-1}}{\partial y_i}(y)\right]^t = [Jf(x)]^{-1} I_i = [Jf(f^{-1}(y))]^{-1} I_i \qquad [1]$$

(y_i es la componente i-ésima de y; la inversa de la matriz jacobiana existe, pues ella es regular). Los elementos de la aplicación $y \mapsto Jf(f^{-1}(y))$ son funciones continuas en V_b, ya que f^{-1} es continua y son continuas las derivadas de f; como los elementos de la matriz $[Jf(f^{-1}(y))]^{-1}$ son expresiones racionales (con denominador no nulo) de los elementos de la matriz $Jf(f^{-1}(y))$, resulta que los elementos de la aplicación $y \mapsto [Jf(f^{-1}(y))]^{-1}$ son funciones continuas en V_b y, por ello y de acuerdo con [1], las derivadas parciales $\partial f^{-1}/\partial y_i$ son todas continuas, es decir, f^{-1} es de clase $\mathscr{C}^1$ en V_b. Si es $r = 2$, es decir, si f es de clase $\mathscr{C}^2$ en U_a, de la expresión [1] y razonando como en el párrafo anterior, se obtiene que entonces f^{-1} es de clase $\mathscr{C}^2$ en V_b y, repitiendo el proceso, para $r = 3, \ldots$, se llega a que f^{-1} es de clase $\mathscr{C}^r$ en V_b si f es de clase $\mathscr{C}^r$ en U_a.

$[38]_4$ Observación

En el supuesto de que para la función $f: x \mapsto y = y(x)$ de clase $\mathscr{C}^r$ se verifican las anteriores hipótesis de $[38]_3$ y que, por ello, su inversa local $f^{-1}: y \mapsto x = x(y)$ también es de clase $\mathscr{C}^r$, para obtener las derivadas parciales de f^{-1}, de órdenes $1, 2, \ldots, r$, basta con derivar sucesivamente en los dos miembros de la identidad

$$y = y(x(y)), \qquad \text{o bien } \begin{cases} y_j = y_j(x_1(y_1, \ldots, y_p), \ldots, x_p(y_1, \ldots, y_p)) \\ (p \text{ identidades; para } j = 1, 2, \ldots, p) \end{cases}$$

Al derivar una vez (respecto de la variable y_i, para i fijo, $i = 1, 2, \ldots, p$) se obtiene, según ya sabemos (véase $[38]_1$, [2])

$$\delta_{ij} = \frac{\partial y_j}{\partial x_1}\frac{\partial x_1}{\partial y_i} + \frac{\partial y_j}{\partial x_2}\frac{\partial x_2}{\partial y_i} + \cdots + \frac{\partial y_j}{\partial x_p}\frac{\partial x_p}{\partial y_i} \quad \begin{cases} p \text{ relaciones; } j = 1, \ldots, p \\ \delta_{ij} = 0, \text{ si } i \neq j;\ \delta_{ij} = 1, \text{ si } i = j \end{cases} \qquad [1]$$

de donde, como ya se hizo anteriormente (véase $[38]_1$, [3]), se obtienen las p derivadas parciales $\partial x_h/\partial y_i$ (para $h = 1, \ldots, p$) en función de las p^2 derivadas parciales $\partial y_j/\partial x_h$

(para j, $h = 1, ..., p$) que se suponen conocidas. Si se deriva de nuevo, ahora en [1] y respecto de la variable y_k (para k fijo, $k = 1, 2, ..., p$), se obtiene que:

$$\begin{aligned} 0 = &\left(\frac{\partial^2 y_j}{\partial x_1 \partial x_1}\frac{\partial x_1}{\partial y_k} + \frac{\partial^2 y_j}{\partial x_1 \partial x_2}\frac{\partial x_2}{\partial y_k} + \cdots + \frac{\partial^2 y_j}{\partial x_1 \partial x_p}\frac{\partial x_p}{\partial y_k}\right)\frac{\partial x_1}{\partial y_i} + \frac{\partial y_j}{\partial x_1}\frac{\partial^2 x_1}{\partial y_i \partial y_k} + \\ &+ \left(\frac{\partial^2 y_j}{\partial x_2 \partial x_1}\frac{\partial x_1}{\partial y_k} + \frac{\partial^2 y_j}{\partial x_2 \partial x_2}\frac{\partial x_2}{\partial y_k} + \cdots + \frac{\partial^2 y_j}{\partial x_2 \partial x_p}\frac{\partial x_p}{\partial y_k}\right)\frac{\partial x_2}{\partial y_i} + \frac{\partial y_j}{\partial x_2}\frac{\partial^2 x_2}{\partial y_i \partial y_k} + \qquad [2] \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &+ \left(\frac{\partial^2 y_j}{\partial x_p \partial x_1}\frac{\partial x_1}{\partial y_k} + \frac{\partial^2 y_j}{\partial x_p \partial x_2}\frac{\partial x_2}{\partial y_k} + \cdots + \frac{\partial^2 y_j}{\partial x_p \partial x_p}\frac{\partial x_p}{\partial y_k}\right)\frac{\partial x_p}{\partial y_i} + \frac{\partial y_j}{\partial x_p}\frac{\partial^2 x_p}{\partial y_i \partial y_k} \end{aligned}$$

Suponiendo que son ya conocidas las derivadas parciales primeras (de las $x_1, x_2, ..., x_p$ respecto de y_i y de y_k), las p igualdades [2] (para $j = 1, 2, ..., p$) permiten calcular las p derivadas parciales segundas $\partial^2 x_h/\partial y_i \partial y_k$ (para $h = 1, ..., p$). Nótese que las ecuaciones [2] son lineales, en las p incógnitas $\partial^2 x_h/\partial y_i \partial y_k$, y que el determinante de este sistema es el jacobiano det $[\partial \mathbf{y}/\partial \mathbf{x}]$, que es no nulo, lo que permite, en efecto, despejar las referidas incógnitas.

[38]$_5$ Ejercicio

Se considera la transformación $(x, y) \mapsto (u, v)$, entre puntos de $\mathbb{R}^2$, definida por:

$$(x, y) \mapsto \begin{cases} u = x(y - \mathrm{L}\ y) + 1 \\ v = ye^{y-x} - 2 \end{cases} \qquad \text{(para } y > 0)$$

Compruébese que esta función tiene inversa local $(u, v) \mapsto (x, y)$ en torno del punto $(x, y) = (1, 1)$, que dicha inversa es de clase $\mathscr{C}^\infty$ y hállense sus derivadas parciales x''_{uv} e y''_{uv} en el punto $(u, v) = (2, -1)$, imagen del $(x, y) = (1, 1)$.

Resolución

Es evidente que $u(x, y)$ y $v(x, y)$ son de clase $\mathscr{C}^\infty$ en un entorno de $(x, y) = (1, 1)$. Aquí es de aplicación el teorema de la función implícita, ya que

$$\det \left.\frac{\partial(u, v)}{\partial(x, y)}\right|_{(1,1)} = \left.\begin{vmatrix} u'_x & u'_y \\ v'_x & v'_y \end{vmatrix}\right|_{(1,1)} = \left.\begin{vmatrix} y + \mathrm{L}\ y & x(1 + 1/y) \\ -ye^{y-x} & (1 + y)e^{y-x} \end{vmatrix}\right|_{(1,1)} = \begin{vmatrix} 1 & 2 \\ -1 & 2 \end{vmatrix} \neq 0$$

luego existe la inversa local, que es de clase $\mathscr{C}^\infty$, pues lo son $u(x, y)$ y $v(x, y)$.

Para hallar las derivadas parciales de $x(u, v)$ y de $y(u, v)$, respecto de u y de v, derivamos en las igualdades [1], de lo que se obtiene (se consideran a u y v como variables y a x e y como funciones):

$$\left.\begin{array}{l} 1 = x'_u(y + \mathrm{L}\ y) + x(y'_u + y'_u/y) \\ 0 = y'_u e^{y-x} + y e^{y-x}(y'_u - x'_u) \end{array}\right\} \to \left.\begin{array}{l} 1 = x'_u(1, 1) + 2y'_u(1, 1) \\ 0 = -x'_u(1, 1) + 2y'_u(1, 1) \end{array}\right\} \to \begin{array}{l} x'_u(1, 1) = 1/2 \\ y'_u(1, 1) = 1/4 \end{array}$$

$$\left.\begin{array}{l} 0 = x'_v(y + \mathrm{L}\ y) + x(y'_v + y'_v/y) \\ 1 = y'_v e^{y-x} + y e^{y-x}(y'_v - x'_v) \end{array}\right\} \to \left.\begin{array}{l} 0 = x'_v(1, 1) + 2y'_v(1, 1) \\ 1 = -x'_v(1, 1) + 2y'_v(1, 1) \end{array}\right\} \to \begin{array}{l} x'_v(1, 1) = -1/2 \\ y'_v(1, 1) = 1/4 \end{array}$$

Derivamos ahora en las dos primeras relaciones, de las cuatro precedentes, respecto de v (o en las dos últimas, respecto de u); al hacerlo, se obtiene:

$$\left.\begin{array}{l} 0 = x''_{uv}(y + \mathrm{L}\ y) + x'_u(y'_v + y'_v/y) + x'_v(y'_u + y'_u/y) + x[y''_{uv} + (y''_{uv}y - y'_u y'_v)/y^2] \\ 0 = e^{y-x}[y''_{uv} + y'_u(y'_v - x'_v) + y'_v(y'_u - x'_u) + y(y'_u - x'_u)(y'_v - x'_v) + y(y''_{uv}y - x''_{uv}] \end{array}\right\}$$

y particularizando ahora para $(x, y) = (1, 1)$, se llega a:

$$\left.\begin{array}{r} x''_{uv}(1, 1) + 2y''_{uv}(1, 1) = -1/16 \\ -x''_{uv}(1, 1) + 2y''_{uv}(1, 1) = -1/16 \end{array}\right\} \to \begin{array}{l} x''_{uv}(1, 1) = 0 \\ y''_{uv}(1, 1) = -1/32 \end{array}$$

3.3. DEPENDENCIA FUNCIONAL

Para ilustrar la definición que viene a continuación, empecemos por considerar el caso de las dos funciones, de $\mathbb{R}^2$ en $\mathbb{R}$, siguientes:

$$\begin{array}{l} (x, y) \mapsto u(x, y) = e^{2x-y} \\ (x, y) \mapsto v(x, y) = e^{2y-4x} \end{array} \qquad [1]$$

estas dos funciones se dicen «dependientes» porque existe una cierta relación no nula entre u y v que se anula idénticamente cuando se sustituyen u y v por las expresiones [1]. La tal relación es:

$$F(u, v) = u^2v - 1$$

Obsérvese que así es, en efecto: 1.º, $F(u, v)$ no es nula cuando u y v varían libremente; 2.º, no obstante lo anterior, se verifica que

$$F[u(x, y), v(x, y)] = (e^{2x-y})^2(e^{2y-4x}) - 1 = e^0 - 1 = 0, \qquad \forall\ (x, y) \in \mathbb{R}^2$$

Este hecho conduce a que, al considerar la función $(x, y) \mapsto (u, v)$, cuando (x, y) recorre todo $\mathbb{R}^2$ (conjunto con «dos grados de libertad»), su imagen (u, v) recorre la curva de ecuación $u^2v = 1$ (conjunto con «un grado de libertad»); de modo que la función $(x, y) \mapsto (u, v)$, de $\mathbb{R}^2$ en $\mathbb{R}^2$, transforma «conjuntos bidimensionales» en «conjuntos unidimensionales».

CONCEPTO DE DEPENDENCIA FUNCIONAL

[39]

Sea $C \subset \mathbb{R}^p$ un conjunto abierto y considérense q funciones reales, de clase $\mathscr{C}^1$ en C:

$$\left.\begin{array}{l} f_1 : C \to \mathbb{R}, \quad x \mapsto u_1 = f_1(x) \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ f_q : C \to \mathbb{R}, \quad x \mapsto u_q = f_q(x) \end{array}\right\}$$

1.° Se dice que las funciones $f_1, f_2, ..., f_q$ son dependientes en el conjunto C si existe una función real $F : D \to \mathbb{R}$,

$$(u_1, ..., u_q) \mapsto F(u_1, ..., u_q)$$

de clase $\mathscr{C}^1$ en un conjunto abierto $D \subset \mathbb{R}^q$ que incluya a $f_1(C) \times \cdots \times f_q(C)$, que no se anule en todo un entorno de ningún punto de D y tal que:

$$F[f_1(x), ..., f_q(x)] \equiv 0, \qquad \forall\, x \in C$$

2.° Las funciones $f_1, ..., f_q$ se dicen independientes en C si no son dependientes en C.

3.° Se dice que f_i, una de las funciones dadas, depende de las demás si, siendo $f_1, ..., f_i, ..., f_p$ dependientes, la anterior función F es del tipo

$$F(u_1, ..., u_{i-1}, u_i, u_{i+1}, ..., u_q) = u_i - G(u_1, ..., u_{i-1}, u_{i+1}, ..., u_q)$$

Nota: En lo que sigue, se va a necesitar que las funciones f_i sean de clase $\mathscr{C}^1$, por lo que así se ha supuesto aquí; no obstante este requisito no es necesario en esta definición. En consecuencia con ello, a la función F bastaría con exigirla que fuese continua.

[39]$_1$ Observaciones

- Si unas funciones $f_1, ..., f_q$, todas ellas de $C \subset \mathbb{R}^n$ en $\mathbb{R}$, son dependientes en C, es entonces evidente que dichas funciones son dependientes en cualquier $C' \subset C$, pero pueden no serlo en un $C'' \supset C$. Si dichas funciones son independientes en C, entonces también lo son en todo $C'' \supset C$, pero no han de serlo, necesariamente, en un $C' \subset C$.
- Sean $f_1, ..., f_q, f_{q+1}, ..., f_{q+h}$ unas ciertas funciones, todas ellas de $C \subset \mathbb{R}^n$ en $\mathbb{R}$. Si $f_1, ..., f_q$ son dependientes en C entonces también $f_1, ..., f_q, f_{q+1}, ..., f_{q+h}$ son dependientes en C. En efecto: si es $F : D \to \mathbb{R}$ la función que, por ser $f_1, ..., f_q$ independientes en C, cumple lo señalado en la definición [39], es evidente que, para la función $G : D \times \mathbb{R}^h \to \mathbb{R}$ definida por:

$$G(u_1, ..., u_q, u_{q+1}, ..., u_{q+h}) = F(u_1, ..., u_q)$$

se verifica lo exigido en la definición [39] para garantizar la dependencia de $f_1, ..., f_q, f_{q+1}, ..., f_h$ en C.

- Si una de las funciones $f_1, ..., f_q$, todas ellas de $C \subset \mathbb{R}^n$ en $\mathbb{R}$, es constante en C, entonces dichas funciones son dependientes. En efecto: si es f_i la función constante y si es k el valor que toma esta función en C, entonces la condición [39] de dependencia se verifica para la función F, de $\mathbb{R}^q$ en $\mathbb{R}$, definida por:

$$F(u_1, ..., u_i, ..., u_q) = u_i - k$$

- En Álgebra Lineal se dice que unas funciones $f_1, ..., f_q$, todas ellas de $C \subset \mathbb{R}^p$ en $\mathbb{R}$ son «linealmente dependientes», si existen ciertas constantes $k_1, ..., k_q \in \mathbb{R}$ tales que

$$k_1 f_1(\boldsymbol{x}) + \cdots + k_q f_q(\boldsymbol{x}) = 0, \qquad \forall\, \boldsymbol{x} \in C$$

La dependencia lineal entre funciones es, pues, un caso particular de dependencia funcional, ya que si se cumple la anterior relación entonces también se cumple la condición [39] de dependencia funcional para la función F, de $\mathbb{R}^q$ en $\mathbb{R}$, definida por:

$$F(u_1, ..., u_1) = k_1 u_1 + \cdots + k_q u_q$$

Nótese que la dependencia de funciones acontece, en general, sin necesidad de que dicha dependencia sea lineal. Así por ejemplo, las funciones seno y coseno, $x \mapsto \operatorname{sen} x$ y $x \mapsto \cos x$ para $x \in \mathbb{R}$, no son linealmente dependientes (no son proporcionales), pero son funcionalmente dependientes, puesto que $\operatorname{sen}^2 x + \cos^2 x = 1$ para $x \in \mathbb{R}$, es decir, se cumple la condición [39] de dependencia para la función $F: \mathbb{R}^2 \to \mathbb{R}$, $F(u, v) = u^2 + v^2$.

$[39]_2$ Ejercicio

Se consideran las tres funciones, de $\mathbb{R}^4$ en $\mathbb{R}$, definidas por las relaciones:

$$\boldsymbol{x} = (x_1, x_2, x_3, x_4) \mapsto \begin{cases} u_1 = x_1 + x_2 + x_3 + x_4 \\ u_2 = x_1^2 + x_2^2 + x_3^2 + x_4^2 \\ u_3 = x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4 \end{cases}$$

Compruébese que estas tres funciones $\boldsymbol{x} \mapsto u_i$ son dependientes en $\mathbb{R}^2$, hallando la relación que las liga:

$$F(u_1(\boldsymbol{x}), u_2(\boldsymbol{x}), u_3(\boldsymbol{x})) = 0, \qquad \forall\, \boldsymbol{x} \in \mathbb{R}^4$$

Resolución

Resulta evidente que para $F(u_1, u_2, u_3) = u_1^2 - u_2 - 2u_3$ es (las sumas lo son para $i, j = 1, 2, 3, 4$):

$$F(u_1(\boldsymbol{x}), u_2(\boldsymbol{x}), u_3(\boldsymbol{x})) = (\sum x_i)^2 - (\sum x_i^2) - 2\Big(\sum_{j \neq j} x_i x_j\Big) =$$

$$= \Big(\sum x_i^2 + 2\sum_{i \neq j} x_i x_j\Big) - (\sum x_i^2) - 2\Big(\sum_{i \neq j} x_i x_j\Big) = 0, \qquad \forall\, \boldsymbol{x} \in \mathbb{R}^4$$

CONDICIÓN NECESARIA DE DEPENDENCIA FUNCIONAL

[40]

Sea C un conjunto abierto de $\mathbb{R}^p$ y considérense q funciones reales $f_1, ..., f_q$ de clase $\mathscr{C}^1$ en C, siendo $q \leqslant p$:

$$x = (x_1, ..., x_p) \mapsto \begin{cases} u_1 = f_1(x) = f_1(x_1, ..., x_q, ..., x_p) \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ u_q = f_q(x) = f_q(x_1, ..., x_q, ..., x_p) \end{cases}$$

Si el jacobiano de las funciones $f_1, ..., f_q$ respecto de q de las variables $x_1, ..., x_p^{(*)}$ es distinto de cero en algún punto de C, entonces las citadas funciones $f_1, ..., f_p$ son independientes en C. Dicho de otro modo: si las funciones $f_1, ..., f_p$ son dependientes en C, entonces en todo punto de C se anula el jacobiano de $f_1, ..., f_q$ respecto de q de las variables $x_1, ..., x_p^{(*)}$.

(*) Esto es: existen unas ciertas variables $x_{i_1}, x_{i_2}, ..., x_{i_q}$ tales que el jacobiano de las funciones $f_1, f_2, ..., f_q$ respecto de dichas variables, etc.

Demostración

Obsérvese que las dos propiedades que se acaban de enunciar son equivalentes, pues cada una de ellas es la contrarrecíproca de la otra. Demostremos la primera y hagámoslo por reducción al absurdo, esto es, suponiendo que en cierto punto $\boldsymbol{a} \in C$ es

$$\det\left(\frac{\partial(f_1, ..., f_q)}{\partial(x_1, ..., x_q)}(\boldsymbol{a})\right) \neq 0^{(*)} \qquad [1]$$

y que existe una relación funcional (véase [39]) entre $f_1, ..., f_q$, dada por una función F:

$$\boldsymbol{u} = (u_1, ..., u_q) \mapsto F(u_1, ..., u_q)$$

vamos a buscar una contradicción, lo que nos obligará a rechazar la existencia de la tal función F.

Como las funciones $f_1, ..., f_q$ son de clase $\mathscr{C}^1$ en C y dado que se verifica la desigualdad [1], el teorema de la función implícita (véase [36]) permite asegurar que el sistema de ecuaciones

$$\left.\begin{matrix} f_1(x_1, ..., x_q, ..., x_p) - u_1 = 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ f_q(x_1, ..., x_q, ..., x_p) - u_q = 0 \end{matrix}\right\} \qquad [2]$$

(*) Si las citadas q variables, de entre las $x_1, ..., x_p$ no fuesen las $x_1, ..., x_q$, alternaríamos el orden de las variables hasta llevar a aquéllas a ocupar los lugares 1, ..., q.

define implícitamente a $x_1, ..., x_p$ como funciones de clase $\mathscr{C}^1$ de $u_1, ..., u_q$ y $x_{q+1}, ..., x_p$ en un cierto entorno $U \subset \mathbb{R}^p$ del punto $(b_1, ..., b_q, a_{q+1}, ..., a_p)$, donde $b_1 = f_1(\boldsymbol{a}), ..., b_q = f_q(\boldsymbol{a})$; si a dichas funciones implícitas las llamamos $\varphi_1, ..., \varphi_q$, se verifica, pues, que

$$\begin{cases} f_i(\varphi_1(u_1, ..., u_q, x_{q+1}, ..., x_p), ..., \varphi_q(u_1, ..., u_q, x_{q+1}, ..., x_p), x_{q+1}, ..., x_p) - u_i = 0 \\ \forall\ (u_1, ..., u_q, x_{q+1}, ..., x_p) \in U \qquad (\text{para } i = 1, 2, ..., q) \end{cases} \quad [3]$$

Nótese que, por ser $U \subset \mathbb{R}^p$ un entorno de $(b_1, ..., b_q, a_{q+1}, ..., a_p)$, existen un entorno $U_b \subset \mathbb{R}^q$ de $(b_1, ..., b_q)$ y un entorno $U' \subset \mathbb{R}^{p-q}$ de $(a_{q+1}, ..., a_p)$ tales que $U_b \times U' \subset U$. Consecuentemente, las anteriores relaciones [3] se verifican siempre que $(u_1, ..., u_q) \in U_b$ y que $(x_{q+1}, ..., x_p) \in U'$. Como para la función $(u_1, ..., u_q) \mapsto F(u_1, ..., u_q)$ se verifica que

$$F(f_1(\boldsymbol{x}), ..., f_q(\boldsymbol{x})) \equiv 0, \qquad \forall\ \boldsymbol{x} = (x_1, ..., x_q, x_{q+1}, ..., x_p) \in C \quad [4]$$

esta relación será cierta, en particular, si se toma $(x_{q+1}, ..., x_p) \in U'$ y

$$\left.\begin{array}{l} x_1 = \varphi_1(u_1, ..., u_q, x_{q+1}, ..., x_p) \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ x_q = \varphi_q(u_1, ..., u_q, x_{q+1}, ..., x_p) \end{array}\right\} \quad \text{para} \quad \begin{array}{l} (u_1, ..., u_q) \in U_b \\ (x_{q+1}, ..., x_p) \in U' \end{array} \quad [5]$$

Llevando, pues, [5] a [4] y teniendo en cuenta las relaciones [3], nos encontramos con que

$$F(u_1, ..., u_q) \equiv 0, \qquad \forall\ (u_1, ..., u_q) \in U_b$$

lo que es una contradicción, pues la función F no puede anularse, por la propia definición (véase [39]), en todo un entorno de un punto de su campo de definición.

[40]$_1$ Observación

Si, en las condiciones del teorema anterior, el número q (de funciones) fuese mayor que el número p de incógnitas, entonces las q funciones serían dependientes en C. Para cerciorarse de ello, basta con tener en cuenta que, de ser $q > p$, las funciones f_i (para $i = 1, ..., q$) se pueden considerar como funciones de q variables (definidas en $C \times \mathbb{R}^{q-p}$), sin más que poner

$$f_i(x_1, ..., x_p) = f_i(x_1, ..., x_p) + 0x_{p+1} + \cdots + 0x_q$$

y que el jacobiano

$$\det\left[\frac{\partial(f_1, ..., f_q)}{\partial(x_1, ..., x_q)}\right]$$

es nulo en todo punto del campo de definición, ya que tiene sus $q - p$ últimas filas nulas.

$[40]_2$ Ejercicio

Compruébese que las funciones f y g, de $\mathbb{R}^2$ en $\mathbb{R}$, definidas por

$$f(x, y) = \operatorname{sen}^2 (x + y) \qquad \text{y} \qquad g(x, y) = \cos^2 (x - y)$$

son independientes en todo abierto $C \subset \mathbb{R}^2$.

Resolución

El jacobiano de f y g (respecto de x e y) es:

$$\det \left(\frac{\partial(f, g)}{\partial(x, y)}\right) = -2 \operatorname{sen} (2x + 2y) \operatorname{sen} (2x - 2y),$$

que se anula sólo en los puntos de las rectas $x \pm y = (k/2)\pi$ para $k \in \mathbb{Z}$, por lo que es no nulo en infinidad de puntos de un abierto (cualquiera) $C \subset \mathbb{R}^2$ y, de acuerdo con [40], f y g son entonces independientes en todo abierto $C \subset \mathbb{R}^2$.

CONDICIÓN SUFICIENTE DE DEPENDENCIA FUNCIONAL (TEOREMA DEL RANGO CONSTANTE)

[41]

Sea C un conjunto abierto de $\mathbb{R}^p$ y considérense q funciones reales $f_1, \ldots, f_q$ de clase $\mathscr{C}^1$ en C,

$$x = (x_1, \ldots, x_p) \mapsto f(x) = (u_1, \ldots, u_q) \begin{cases} u_1 = f_1(x) \\ \ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots \\ u_q = f_q(x) \end{cases}$$

tales que la matriz $Jf(x) = [\partial f(x)/\partial x]$, jacobiana de f en x, tiene rango r en todo punto $x \in C$; es decir, tales que: 1.°, todos los menores de orden mayor que r de $Jf(x)$ son nulos en todo $x \in C$; y 2.° para cada $a \in C$, existe un menor de orden r de $Jf(a)$ que es no nulo, o sea

$$\det \left[\frac{\partial(f_{i_1}, \ldots, f_{i_r})}{\partial(x_{j_1}, \ldots, x_{j_r})}(a)\right] \neq 0 \qquad \left(\begin{array}{l}\text{para unos ciertos índices:} \\ (i_1, \ldots, i_r) \subset (1, 2, \ldots, q) \\ (j_1, \ldots, j_r) \subset (1, 2, \ldots, p)\end{array}\right)$$

Entonces, para cada $a \in C$, existe un entorno suyo, U, tal que en él se verifica que: 1.°, las funciones $f_{i_1}, \ldots, f_{i_r}$ son independientes; y 2.°, cada una de las $q - r$ restantes funciones f_i dadas, $i \notin \{i_1, \ldots, i_r\}$, depende funcionalmente de las $f_{i_1}, \ldots, f_{i_r}$.

Demostración

Cambiemos el orden de las funciones $f_1, ..., f_q$ y el orden de las variables $x_1, ..., x_p$ hasta conseguir que $(i_1, ..., i_r)$ y $(j_1, ..., j_r)$ pasen a ser $(1, ..., r)$.

1.° Las funciones $f_1, ..., f_r$ son independientes en C, que es un entorno de $\boldsymbol{a}$, pues para ellas se cumple lo exigido en el teorema anterior (véase [40]).

2.° En esta segunda parte hemos de comprobar que cada una de las funciones $f_{r+1}, ..., f_q$ depende de las $f_1, ..., f_r$ en un cierto entorno de $\boldsymbol{a} = (a_1, ..., a_p)$; vamos a hacerlo sólo con f_{r+1}, pues para las demás se razona de modo análogo. Para ello vamos a comenzar dando los mismos pasos que en la demostración del anterior teorema [40] y, por razones de brevedad, adoptaremos la siguiente notación: 1.°, un punto cualquiera $\boldsymbol{x} = (x_1, ..., x_r, ..., x_p)$ de C se pondrá en la forma:

$$\boldsymbol{x} = (\tilde{\boldsymbol{x}}, \hat{\boldsymbol{x}}), \qquad \text{donde } \tilde{\boldsymbol{x}} = (x_1, ..., x_r) \qquad \text{y} \qquad \hat{\boldsymbol{x}} = (x_{r+1}, ..., x_p)$$

y, en particular, se escribirá $\boldsymbol{a} = (\tilde{\boldsymbol{a}}, \hat{\boldsymbol{a}})$; 2.°, dados $u_1, ..., u_r \in \mathbb{R}$, se pondrá $\tilde{\boldsymbol{u}} = (u_1, ..., u_r)$; 3.°, llamaremos $b_1 = f_1(\boldsymbol{a}), ..., b_r = f_r(\boldsymbol{a})$ y $\tilde{\boldsymbol{b}} = (b_1, ..., b_r)$; 4.°, se llamará $\tilde{\boldsymbol{f}}: C \to \mathbb{R}^r$ a la función cuyas componentes son $f_1, ..., f_r$, es decir, se pone $\tilde{\boldsymbol{f}} = (f_1, ..., f_r)$.

Comencemos considerando el sistema formado por las r ecuaciones:

$$\left.\begin{array}{l} f_1(\tilde{\boldsymbol{x}}, \hat{\boldsymbol{x}}) - u_1 = 0 \\ \cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots \\ f_r(\tilde{\boldsymbol{x}}, \hat{\boldsymbol{x}}) - u_r = 0 \end{array}\right\} \qquad [1]$$

Como la función $\tilde{\boldsymbol{f}} = (f_1, ..., f_r)$ es de clase $\mathscr{C}^1$ en C y como se sabe que el jacobiano de esta función respecto de $\tilde{\boldsymbol{x}}$ es no nulo en $\boldsymbol{a}$, resulta evidente que al anterior sistema [1] le es de aplicación el teorema de la función implícita (véase [36]), del que se desprende que [1] define a r funciones

$$(\tilde{\boldsymbol{u}}, \hat{\boldsymbol{x}}) \mapsto \tilde{\boldsymbol{x}} = \tilde{\boldsymbol{\varphi}}(\tilde{\boldsymbol{u}}, \hat{\boldsymbol{x}}) \left\{\begin{array}{l} x_1 = \varphi_1(\tilde{\boldsymbol{u}}, \hat{\boldsymbol{x}}) \\ \cdots\cdots\cdots \\ \cdots\cdots\cdots \\ x_r = \varphi_r(\tilde{\boldsymbol{u}}, \hat{\boldsymbol{x}}) \end{array}\right. \qquad [2]$$

de clase $\mathscr{C}^1$ en un cierto entorno $U \subset \mathbb{R}^p$ del punto $(\tilde{\boldsymbol{u}}, \hat{\boldsymbol{x}}) = (\tilde{\boldsymbol{b}}, \hat{\boldsymbol{a}})$; estas funciones son, pues, tales que:

$$\left.\begin{array}{l} f_i(\tilde{\boldsymbol{\varphi}}(\tilde{\boldsymbol{u}}, \hat{\boldsymbol{x}}), \hat{\boldsymbol{x}}) - u_i \equiv 0 \\ (\text{para } i = 1, 2, ..., r) \end{array}\right\} \qquad \forall\, (\tilde{\boldsymbol{u}}, \hat{\boldsymbol{x}}) \in U \qquad [3]$$

Si en este sistema [3] se deriva parcialmente respecto de una cualquiera de las coordenadas de $\hat{\boldsymbol{x}} = (x_1, ..., x_p)$, a la que llamaremos x_k, se obtiene que:

$$\sum_{j=1}^{r}\left(\frac{\partial f_i}{\partial x_j}(\tilde{\boldsymbol{\varphi}}(\tilde{\boldsymbol{u}}, \hat{\boldsymbol{x}}), \hat{\boldsymbol{x}})\frac{\partial \varphi_j}{\partial x_k}(\tilde{\boldsymbol{u}}, \hat{\boldsymbol{x}})\right) + \frac{\partial f_i}{\partial x_k}(\tilde{\boldsymbol{\varphi}}(\tilde{\boldsymbol{u}}, \hat{\boldsymbol{x}}), \hat{\boldsymbol{x}}) = 0 \quad (\text{para } i = 1, 2, ..., r)\ \forall\, (\tilde{\boldsymbol{u}}, \hat{\boldsymbol{x}}) \in U \qquad [4]$$

Pues bien, para verificar que la función f_{r+1} depende, en un cierto entorno U_a de $\boldsymbol{a}$, de las funciones $f_1, ..., f_r$, consideraremos la función (que se obtiene de llevar los valores de $x_1, ..., x_r$ definidos por [2] a la expresión de la $r+1$-ésima función dada):

$$(\tilde{\boldsymbol{u}}, \hat{\boldsymbol{x}}) \mapsto f_{r+1}(\tilde{\varphi}(\tilde{\boldsymbol{u}}, \hat{\boldsymbol{x}}), \hat{\boldsymbol{x}}), \qquad \forall\ (\tilde{\boldsymbol{u}}, \hat{\boldsymbol{x}}) \in U \tag{5}$$

y veremos que ella no depende de $\hat{\boldsymbol{x}}$; esto es: si llamamos

$$G(\tilde{\boldsymbol{u}}, \hat{\boldsymbol{x}}) = f_{r+1}(\tilde{\varphi}(\tilde{\boldsymbol{u}}, \hat{\boldsymbol{x}}), \hat{\boldsymbol{x}}), \qquad \forall\ (\tilde{\boldsymbol{u}}, \hat{\boldsymbol{x}}) \in U \tag{6}$$

vamos a comprobar que $G(\tilde{\boldsymbol{u}}, \hat{\boldsymbol{x}})$ no varía, en el entorno dado, al variar $\hat{\boldsymbol{x}}$, y pondremos entonces $G(\tilde{\boldsymbol{u}}, \hat{\boldsymbol{x}}) = G(\tilde{\boldsymbol{u}})$.

De ser ello cierto, el teorema estaría ya probado, pues por la continuidad de $\boldsymbol{x} \mapsto \tilde{\boldsymbol{u}} = \tilde{f}(\boldsymbol{x})$ en $\boldsymbol{x} = \boldsymbol{a}$, existiría entonces un cierto entorno de $\boldsymbol{a}$ tal que, para todo $\boldsymbol{x}$ de él, de [5] se desprendería que:

$$f_{r+1}(\boldsymbol{x}) = G(f_1(\boldsymbol{x}), ..., f_r(\boldsymbol{x}))$$

es decir, la condición de dependencia funcional (véase [39], 3.º) se verificaría (en un cierto entorno de $\boldsymbol{a}$) para la función

$$(u_1, ..., u_r, u_{r+1}) \mapsto u_{r+1} - G(u_1, ..., u_r)$$

Comprobemos que, efectivamente, la función [5] no depende de $\hat{\boldsymbol{x}} = (x_{r+1}, ..., x_p)$ esto es, que son nulas las derivadas parciales de $G(\tilde{\boldsymbol{u}}, \hat{\boldsymbol{x}})$ respecto de todas las variables $(x_{r+1}, ..., x_p)$. Llamando x_k a una cualquiera de dichas variables y derivando respecto de x_k en [6], se obtiene que para $(\tilde{\boldsymbol{u}}, \hat{\boldsymbol{x}}) \in U$ es:

$$\frac{\partial G}{\partial x_k}(\tilde{\boldsymbol{u}}, \hat{\boldsymbol{x}}) = \sum_{j=1}^{r} \left(\frac{\partial f_{r+1}}{\partial x_j}(\tilde{\varphi}(\tilde{\boldsymbol{u}}, \hat{\boldsymbol{x}}), \hat{\boldsymbol{x}}) \frac{\partial \varphi_j}{\partial x_k}(\tilde{\boldsymbol{u}}, \hat{\boldsymbol{x}}) \right) + \frac{\partial f_{r+1}}{\partial x_k}(\tilde{\varphi}(\tilde{\boldsymbol{u}}, \hat{\boldsymbol{x}}), \hat{\boldsymbol{x}}) \tag{7}$$

Esta relación [7] y las relaciones [4] permiten asegurar que, como la función $\boldsymbol{x} \mapsto \tilde{\boldsymbol{u}} = \tilde{f}(\boldsymbol{x})$ es continua en C, existe un cierto entorno de $\boldsymbol{a}$ tal que, para todo $\boldsymbol{x}$ de él, es:

$$\sum_{j=1}^{r} \frac{\partial f_i}{\partial x_j}(\boldsymbol{x}) \frac{\partial \varphi_j}{\partial x_k}(\tilde{f}(\boldsymbol{x}), \hat{\boldsymbol{x}}) + \frac{\partial f_i}{\partial x_k}(\boldsymbol{x}) = \begin{cases} 0, \text{ para } i = 1, 2, ..., r \\ \dfrac{\partial G}{\partial x_k}(\tilde{f}(\boldsymbol{x}), \hat{\boldsymbol{x}}), \text{ si } i = r+1 \end{cases}$$

Es decir, escribiendo las anteriores relaciones de un modo detallado y omitiendo, por razón de brevedad, que las derivadas $\partial f_i / \partial x_j$ lo son en el punto $\boldsymbol{x}$ y que las $\partial\varphi_j/\partial x_k$ y $\partial G/\partial x_k$ lo son en $(\tilde{f}(\boldsymbol{x}), \hat{\boldsymbol{x}})$, se obtiene que, para $\boldsymbol{x}$ en un cierto entorno de $\boldsymbol{a}$, es:

$$\left.\begin{array}{l}
\dfrac{\partial f_1}{\partial x_1}\dfrac{\partial\varphi_1}{\partial x_k}+\cdots+\dfrac{\partial f_1}{\partial x_r}\dfrac{\partial\varphi_r}{\partial x_k}+\dfrac{\partial f_1}{\partial x_k}=0\\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
\dfrac{\partial f_r}{\partial x_1}\dfrac{\partial\varphi_1}{\partial x_k}+\cdots+\dfrac{\partial f_r}{\partial x_r}\dfrac{\partial\varphi_r}{\partial x_k}+\dfrac{\partial f_r}{\partial x_k}=0\\
\dfrac{\partial f_{r+1}}{\partial x_1}\dfrac{\partial\varphi_1}{\partial x_k}+\cdots+\dfrac{\partial f_{r+1}}{\partial x_r}\dfrac{\partial\varphi_r}{\partial x_k}+\dfrac{\partial f_{r+1}}{\partial x_k}=\dfrac{\partial G}{\partial x_k}
\end{array}\right\}\quad\left(\begin{array}{c}\text{para cualquier}\\ k\in\{r+1, ..., p\}\end{array}\right) \qquad [8]$$

Nótese que [8] puede interpretarse como un sistema de $r+1$ ecuaciones lineales en las $r+1$ incógnitas $\partial\varphi_1/\partial x_k$, ..., $\partial\varphi_r/\partial x_k$, 1, que es compatible, pues sabemos que dichas derivadas existen. El determinante de este sistema es nulo (en el entorno de $\boldsymbol{a}$), pues es el jacobiano de $(f_1, ..., f_r, f_{r+1})$ respecto de $(x_1, ..., x_r, x_k)$, o sea, un menor de orden $r+1$ de Jf, y todos los menores de orden mayor que r de Jf son nulos en C. Como el sistema es compatible y su determinante es nulo, son también nulos todos los determinantes que se obtienen al sustituir una columna cualquiera de aquel por la columna de los términos independientes y, en particular, será nulo el determinante que resulta al sustituir la última columna, el cual, al desarrollarle por los elementos de la columna de los términos independientes, se puede poner:

$$\det\left[\frac{\partial(f_1, ..., f_r)}{\partial(x_1, ..., x_r)}\right]\frac{\partial G}{\partial x_k}$$

Como este producto es cero y su primer factor es no nulo, por hipótesis, resulta que $\partial G/\partial x_k = 0$, para todo $\boldsymbol{x}$ del entorno de $\boldsymbol{a}$. Así pues, G no depende de $\hat{\boldsymbol{x}}$, como había que comprobar.

[41]$_1$ Ejercicio

Compruébese que las siguientes funciones f_1, f_2 y f_3, de $\mathbb{R}^3$ en $\mathbb{R}$, son funcionalmente dependientes en $\mathbb{R}^3$ en los dos casos siguientes:

1.° $(x, y, z) \mapsto \begin{cases} f_1(x, y, z) = x + y - z + 1 \\ f_2(x, y, z) = -x + y + z - 1 \\ f_3(x, y, z) = x^2 - y^2 + z^2 - 2xz \end{cases}$

2.° $(x, y, z) \mapsto \begin{cases} f_1(x, y, z) = \operatorname{sen}\,(x + y + z) \\ f_2(x, y, z) = \operatorname{sen}\,(x + y - z) \\ f_3(x, y, z) = \operatorname{sen}\,(x + y)\cos z \end{cases}$

Resolución

1.º La matriz jacobiana de (f_1, f_2, f_3), respecto de (x, y, z), es:

$$\frac{\partial(f_1, f_2, f_3)}{\partial(x, y, z)} = \begin{bmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ 2x - 2z & -2y & 2z - 2x \end{bmatrix}$$

esta matriz tiene rango 2 (sus dos primeras columnas son independientes y la tercera es proporcional a la primera) para cualquiera que sea (x, y, z), luego f_1, f_2 y f_3 son dependientes. Como sus dos primeras filas son linealmente independientes, f_1 y f_2 son funcionalmente independientes y, consecuentemente, f_3 depende de las f_1 y f_2. Según se comprueba fácilmente, la relación funcional entre f_1, f_2 y f_3 es la:

$$(u - 1)(v + 1) + w = 0$$

2.º La matriz jacobiana es ahora:

$$\frac{\partial(f_1, f_2, f_3)}{\partial(x, y, z)} = \begin{bmatrix} \cos(x+y+z) & \cos(x+y+z) & \cos(x+y+z) \\ \cos(x+y-z) & \cos(x+y-z) & -\cos(x+y-z) \\ \cos(x+y)\cos z & \cos(x+y)\cos z & -\operatorname{sen}(x+y)\operatorname{sen} z \end{bmatrix}$$

cuyo rango es 2 o menos, para cualquiera que sea $(x, y, z) \in \mathbb{R}^3$, por lo que f_1, f_2 y f_3 son dependientes en $\mathbb{R}^3$. Según cuesta poco comprobar, la relación funcional entre f_1, f_2 y f_3 es:

$$\operatorname{sen}(x + y + z) + \operatorname{sen}(x + y - z) = \operatorname{sen}(x + y)\cos z$$

3.4. CAMBIOS DE VARIABLES

En ocasiones, para atajar determinados problemas, como la resolución de ecuaciones en derivadas parciales, se recurre a la «técnica de los cambios de variables». Cuando se aborda una tal cuestión, en la que interviene una «relación diferencial», que liga a unas ciertas variables, a una función de ellas y a las derivadas (ordinarias o parciales) de esta respecto de aquellas, puede ser conveniente expresar la referida relación diferencial en términos de otras variables, de otra función y de las correspondientes derivadas. A partir de las expresiones que ligan a las nuevas variables y función con las antiguas, se obtendrán las relaciones entre unas y otras derivadas, lo que llevado a la relación diferencial dada conducirá a la nueva relación diferencial. Se espera que el problema planteado sea, ahora, más fácil de resolver, acudiendo a esta última relación diferencial.

TÉCNICA DEL CAMBIO DE VARIABLE

[42]

Sea dada la siguiente «expresión diferencial» E, en la que z es una función de x e y (de clase $\mathscr{C}^r$, con r «suficientemente» grande), $(x, y) \mapsto z(x, y)$, donde (x, y) recorre cierto abierto de $\mathbb{R}^2$:

$$E = \Phi(x, y, z, z'_x, z'_y, z''_{xx}, z''_{xy}, z''_{yy}, z'''_{xxx}, z'''_{xxy}, ...) \qquad [1]$$

Se desea obtener la forma que adopta E cuando se expresa en términos de unas nuevas variables u y v y de una nueva función $(u, v) \mapsto w(u, v)$, conocidas mediante las siguientes «funciones de cambio» (de clase $\mathscr{C}^r$):

$$u = \varphi(x, y, z), \qquad v = \psi(x, y, z), \qquad w = f(x, y, z) \qquad [2]$$

(se supone que el sistema [2] es localmente invertible y que sus dos primeras relaciones definen implícitamente a x e y). Es decir, se buscan las derivadas $z'_x, z'_y, z''_{xx}, ...$, en función de las derivadas $w'_u, w'_v, w''_{uu}, ...$; al llevar estos resultados a [1] se obtendría la nueva expresión diferencial:

$$E = \Psi(u, v, w, w'_u, w'_v, w''_{uu}, w''_{uv}, w''_{vv}, w'''_{uuu}, w'''_{uuv}, ...) \qquad [3]$$

Para realizar dicho cambio de variable, se puede recurrir a que (suponiendo que se cumplen las hipótesis del teorema de la función implícita) el sistema

$$w = w(u, v), \qquad w = f(x, y, z), \qquad u = \varphi(x, y, z), \qquad v = \psi(x, y, z)$$

define implícitamente a z, u, v y w como funciones de x e y, con lo que existen $z(x, y)$, $u(x, y)$, $v(x, y)$ y $w(x, y)$, de clase $\mathscr{C}^r$, tales que (para todo (x, y) del campo de definición) es

$$\left.\begin{aligned} w(x, y) &\equiv w[u(x, y), v(x, y)] \\ w(x, y) &\equiv f[x, y, z(x, y)] \\ u(x, y) &\equiv \varphi[x, y, z(x, y)] \\ v(x, y) &\equiv \psi[x, y, z(x, y)] \end{aligned}\right\} \quad [S]$$

Derivando o diferenciando en este sistema $[S]$ (en el que las únicas variables son x e y) se obtienen las derivadas buscadas (del modo que se indica a continuación).

Comprobación

Veamos como, en efecto, se pueden obtener las relaciones que ligan a las derivadas de $z(x, y)$ con las derivadas de $w(u, v)$, en función de las derivadas parciales de f, φ y ψ, las cuales son datos del problema. Según se acaba de indicar, esto se puede hacer derivando o diferenciando en $[S]$; veámoslo.

1.° PRIMER PROCEDIMIENTO: DERIVANDO. Si en el sistema [S] derivamos respecto de la variable x (igual se haría con y) se obtiene que:

$$\left.\begin{aligned} w_x &= w_u u_x + w_v v_x \\ w_x &= f_x + f_z z_x \\ u_x &= \varphi_x + \varphi_z z_x \\ v_x &= \psi_x + \psi_z z_x \end{aligned}\right\} \quad [S_x]$$

De este sistema se pueden obtener (si el correspondiente jacobiano es no nulo) z_x, w_x, u_x y v_x en función de w_u y w_v (y de las derivadas de las funciones de cambio f, φ y ψ). En particular, para z_x se obtiene

$$z_x = \frac{f_x - \varphi_x w_u - \psi_x w_v}{-f_z + \varphi_z w_u + \psi_z w_v}$$

Para obtener las derivadas segundas z_{xx}, z_{xy} y z_{yy}, se deriva respecto de x y respecto de y en el sistema [S_x] y en el sistema [S_y], que se obtendría cambiando x por y en el [S_x]. Así, por ejemplo, para obtener z_{xy} (para hallar z_{xx} y z_{yy} se puede proceder de manera análoga), derivando en [S_x] respecto de y, se tiene:

$$\left.\begin{aligned} w_{xy} &= w_{uu}u_x u_y + w_{uv}(u_x v_y + v_x u_y) + w_{vv}v_x v_y + w_u u_{xy} + w_v v_{xy} \\ w_{xy} &= f_{xy} + f_{xz}(z_x + z_y) + f_{zz}z_x z_y + f_z z_{xy} \\ u_{xy} &= \varphi_{xy} + \varphi_{xz}(z_x + z_y) + \varphi_{zz}z_x z_y + \varphi_z z_{xy} \\ v_{xy} &= \psi_{xy} + \psi_{xz}(z_x + z_y) + \psi_{zz}z_x z_y + \psi_z z_{xy} \end{aligned}\right\} \quad [S_{xy}]$$

En este sistema [S_{xy}], las incógnitas son z_{xy}, w_{xy}, u_{xy} y v_{xy}, que se pueden obtener (si el correspondiente jacobiano no es nulo) en función de w_{uu}, w_{uv} y w_{vv} (nótese que z_x, z_y, u_x, u_y, v_x y v_y son ya conocidas: se obtienen de [S_x] y [S_y]); de entre las referidas incógnitas nos interesa especialmente la z_{xy}.

Derivando nuevamente, respecto de x e y en [S_{xx}], [S_{xy}] y [S_{yy}], se obtienen sistemas que permiten calcular las derivadas z_{xxx}, z_{xxy}, z_{xyy} y z_{yyy}. Prosiguiendo de este modo, se calculan las sucesivas derivadas de $z(x, y)$ en función de las de $w(u, v)$.

2.° SEGUNDO PROCEDIMIENTO: DIFERENCIANDO. Con la ayuda de la fórmula de la diferencial de una función compuesta (véase [29], 2.°), vamos ahora a diferenciar las identidades del sistema [S]; al aplicar dicha fórmula hay que tener en cuenta que x e y son las variables independientes, por lo que las diferenciales dx y dy serán los incrementos arbitrarios de x e y, y que dz, dw, du y dv son las diferenciales de las funciones $z(x, y)$, $w(x, y)$, $u(x, y)$ y $v(x, y)$ definidas implícitamente por el sistema. Procediendo como se ha dicho, se obtiene:

$$\left.\begin{aligned} dw &= w_u du + w_v dv \\ dw &= f_x dx + f_y dy + f_z dz \\ du &= \varphi_x dx + \varphi_y dy + \varphi_z dz \\ dv &= \psi_x dx + \psi_y dy + \psi_z dz \end{aligned}\right\} \quad [S']$$

Este sistema lineal permite despejar (si el correspondiente jacobiano es no nulo) a las incógnitas dz, du y dv (también dw, pero ésta carece de interés) en función de dx y dy; al hacerlo se obtiene un resultado como el siguiente, en el que a, b, c, d, e y f denotan a unas ciertas expresiones que dependen de w_u y w_v

$$\begin{aligned} dz &= a(w_u, w_v)\,dx + b(w_u, w_v)\,dy \\ du &= c(w_u, w_v)\,dx + d(w_u, w_v)\,dy \\ dv &= e(w_u, w_v)\,dx + f(w_u, w_v)\,dy \end{aligned} \qquad [1]$$

y, en particular, de la primera de las anteriores igualdades se desprende que las derivadas parciales z_x y z_y son entonces:

$$z_x = a(w_u, w_v) \qquad \text{y} \qquad z_y = b(w_u, w_v)$$

Si queremos obtener ahora las derivadas segundas z_{xx}, z_{xy} y z_{yy}, en función de las derivadas de $w(u, v)$, podemos diferenciar nuevamente, ahora en el sistema $[S']$, de lo que se obtiene (véase [34]; téngase en cuenta que, como x e y son las variables independientes, es $d^2x = o$ y $d^2y = o$):

$$\left.\begin{aligned} d^2w &= w_{uu}du^2 + 2w_{uv}\,du\,dv + w_{vv}\,dv^2 + w_u d^2u + w_v d^2v \\ d^2w &= f_{xx}\,dx^2 + f_{yy}\,dy^2 + f_{zz}\,dz^2 + 2f_{xy}\,dx\,dy + 2f_{xz}\,dx\,dz + 2f_{yz}\,dy\,dz + f_z d^2z \\ d^2u &= \varphi_{xx}\,dx^2 + \varphi_{yy}\,dy^2 + \varphi_{zz}\,dz^2 + 2\varphi_{xy}\,dx\,dy + 2\varphi_{xz}\,dx\,dz + 2\varphi_{yz}\,dy\,dz + \varphi_z d^2z \\ d^2v &= \psi_{xx}\,dx^2 + \psi_{yy}\,dy^2 + \psi_{zz}\,dz^2 + 2\psi_{xy}\,dx\,dy + 2\psi_{xz}\,dx\,dz + 2\psi_{yz}\,dy\,dz + \psi_z d^2z \end{aligned}\right\} \quad [S'']$$

Este sistema, que es lineal en las incógnitas (que son d^2z, d^2u, d^2v y d^2w), permite calcular éstas (si el correspondiente jacobiano es no nulo) en función de dz y dy (nótese que du y dv ya son conocidas, de [1], en función de dx y dy); en particular, de $[S'']$ se obtiene fácilmente para d^2z un resultado del tipo:

$$d^2z = \alpha dx^2 + 2\beta\,dx\,dy + \gamma dy^2 \qquad \left(\begin{array}{c}\text{Donde } \alpha,\ \beta \text{ y } \gamma \text{ denotan} \\ \text{a unas ciertas expresiones de} \\ w_u,\ w_v,\ w_{uu},\ w_{uv} \text{ y } w_{vv}\end{array}\right)$$

es decir, las derivadas parciales z_{xx}, z_{xy} y z_{yy} son entonces:

$$z_{xx} = \alpha, \qquad z_{xy} = \beta, \qquad z_{yy} = \gamma$$

Prosiguiendo de este modo, diferenciando sucesivamente, se obtienen las diferencias sucesivas d^3z, d^4z, etc., de $z(x, y)$, lo que proporciona de modo inmediato las derivadas terceras, cuartas, etc., de $z(x, y)$.

[42]$_1$ Un caso particular: no se cambia la función

En el caso general que se acaba de considerar, se cambian las variables independientes y, también, la función. No obstante, son muy frecuentes los casos en los que sólo es preciso cambiar las variables independientes; entonces se simplifica enormemente el proceso, como pasamos a comprobar:

Se considera, también ahora, la siguiente expresión diferencial (como en [42]):

$$E : \Phi(x, y, z, z_x, z_y, z_{xx}, z_{xy}, z_{yy}, z_{xxx}, z_{xxy}, \ldots)$$

referida a una función $(x, y) \mapsto z$ de clase suficientemente avanzada, y se desea obtener la expresión transformada de E cuando se sustituyen las variables x e y (pero no la función z) por otras u y v, relacionadas con aquéllas de una de las dos maneras siguientes:

1.º $u = \varphi(x, y)$ y $v = \psi(x, y)$.

2.º $x = f(u, v)$ e $y = g(u, v)$.

El problema planteado se reduce, pues, a obtener las derivadas parciales de $z(x, y)$, respecto de x e y, en función de las derivadas parciales de $z(u, v)$, respecto de u y v. Para ello, se procederá como sigue:

1.º Conocidas $u = \varphi(x, y)$ y $v = \psi(x, y)$, resulta que

$$\begin{aligned} z_x &= z_u u_x + z_v v_x = z_u \varphi_x + z_v \psi_x \\ z_y &= z_u u_y + z_v v_y = z_u \varphi_y + z_v \psi_y \end{aligned}$$

Derivando de nuevo (por ejemplo, la primera respecto de y para obtener z''_{xy}; igual se haría para z''_{xx} y z''_{yy}), se obtiene que:

$$z_{xy} = z_{uu}\varphi_x\varphi_y + z_{uv}(\varphi_x\psi_y + \varphi_y\psi_x) + z_{vv}\psi_x\psi_y + z_u\varphi_{xy} + z_v\psi_{xy}$$

2.º Conocidas $x = f(u, v)$ e $y = g(u, v)$, resulta que

$$\begin{aligned} z_u &= z_x x_u + z_y y_u = z_x f_u + z_y g_u \\ z_v &= z_x x_v + z_y y_v = z_x f_v + z_y g_v \end{aligned}$$

De este sistema se pueden despejar z_x y z_y (si el jacobiano de (f, g) respecto de (u, v) es no nulo) en función de z_u y z_v, obteniéndose expresiones del tipo

$$\left.\begin{aligned} z_x &= \alpha z_u + \beta z_v \\ z_y &= \gamma z_u + \delta z_v \end{aligned}\right\} \quad \left(\begin{array}{c}\text{donde } \alpha \text{ y } \beta \text{ son unas ciertas} \\ \text{expresiones de } f_u,\ g_u,\ f_v \text{ y } g_v\end{array}\right)$$

Para obtener las derivadas segundas, de $z(x, y)$, derivaremos en estas últimas relaciones, respecto de x y respecto de y. Así, por ejemplo, derivando en la primera respecto de y se obtiene z_{xy}, que valdrá (se utiliza notación simbólica):

$$\begin{aligned} z_{xy} &= \left(\gamma \frac{\partial}{\partial u} + \delta \frac{\partial}{\partial v}\right)\left(\alpha \frac{\partial}{\partial u} + \beta \frac{\partial}{\partial v}\right) z = \left(\gamma \frac{\partial}{\partial u} + \delta \frac{\partial}{\partial v}\right)(\alpha z_u + \beta z_v) = \\ &= \gamma(\alpha z_{uu} + \beta z_{uv} + \alpha_u z_u + \beta_u z_v) + \delta(\alpha z_{uv} + \beta z_{vv} + \alpha_v z_u + \beta_v z_v) = \\ &= \gamma\alpha z_{uu} + (\gamma\beta + \delta\alpha) z_{uv} + \delta\beta z_{vv} + (\gamma\alpha_u + \delta\alpha_v) z_u + (\gamma\beta_u + \delta\beta_v) z_v \end{aligned}$$

[42]$_2$ Ejercicio

Hallar la transformada de la expresión diferencial

$$E = u\,\frac{\partial^2 z}{\partial u^2} - v\,\frac{\partial^2 z}{\partial u \partial v} + \frac{v}{u}\,\frac{\partial z}{\partial v}$$

(que se refiere a una función $z(u, v)$ de clase $\mathscr{C}^2$) cuando se sustituyen las variables independientes u y v por las x e y, mediante el cambio

$$u = x, \qquad v = y/x$$

Resolución

El sistema

$$\left.\begin{array}{l} z = z(x, y) \\ x = u \\ y = uv \end{array}\right\} \qquad \text{define a} \quad z = z(u, v)$$

Derivando en dicho sistema, respecto de u y v, se obtiene:

$$\begin{aligned} z_u &= z_x x_u + z_y y_u = z_x + v z_y \\ z_v &= z_x x_v + z_y y_v = u z_y \\ z_{uu} &= (z_{xx} + v z_{xy}) + (z_{xy} + v z_{yy})v = z_{xx} + 2v z_{xy} + v^2 z_{yy} \\ z_{uv} &= z_y + (z_{xy} + v z_{yy})u = z_y + u z_{xy} + uv z_{yy} \end{aligned}$$

Por lo que, al llevar estos resultados a E, se obtiene que:

$$\begin{aligned} E &= u(z_{xx} + 2v z_{xy} + v^2 z_{yy}) - v(z_y + u z_{xy} + uv z_{yy}) + \frac{v}{u}\,(u z_y) = \\ &= u z_{xx} + uv z_{xy} = x z_{xx} + y z_{yy} \end{aligned}$$

[42]$_3$ Ejercicio

Sabiendo que $z = z(x, y)$ satisface a la relación (ecuación diferencial):

$$z_{xx} + \frac{1}{2y}\,z_{xy} + \frac{1}{4y^2}\,z_{yy} = \frac{1}{4y^3}\,z_y$$

determinar la relación a la que, por ello, satisface la función $w = w(u, v)$, siendo

$$w = z + x, \qquad u = x, \qquad v = y^2$$

Resolución

Consideremos el siguiente sistema de ecuaciones

$$\left.\begin{aligned} w &= w(u, v) \\ w &= z + x \\ u &= x \\ v &= y^2 \end{aligned}\right\}$$

que define a z, u, v y w como funciones de x e y. Diferenciando dos veces en dicho sistema (téngase en cuenta que x e y son las variables independientes, por lo que $d^2x = o$ y $d^2y = o$), se obtiene:

$$\left.\begin{aligned} dw &= w_u\,du + w_v\,dv \\ dw &= dz + dx \\ du &= dx \\ dv &= 2y\,dy \end{aligned}\right\} \Rightarrow \begin{cases} dz + dx = w_u\,dx + w_v 2y\,dy \;\Rightarrow \\ \Rightarrow\;\; dz = (w_u - 1)\,dx + 2yw_v\,dy \;\Rightarrow \\ \Rightarrow\;\; z_x = w_u - 1 \;\;\text{y}\;\; z_y = 2yw_v \end{cases}$$

$$\left.\begin{aligned} d^2w &= w_{uu}\,du^2 + 2w_{uv}\,du\,dv + w_{vv}\,dv^2 + \\ &\quad + w_u d^2u + w_v d^2v \\ d^2w &= d^2z \\ d^2u &= 0 \\ d^2v &= 2dy^2 \end{aligned}\right\} \Rightarrow \begin{cases} d^2z = w_{uu}\,dx^2 + 2w_{uv}(2y\,dx\,dy) + w_{vv}(2y\,dy)^2 + \\ \quad + w_u \cdot 0 + w_v 2dy^2 = w_{uu}\,dx^2 + \\ \quad + 2(2yw_{uv})\,dx\,dy + (4y^2w_{vv} + 2w_v)\,dy^2 \;\Rightarrow \\ z_{xx} = w_{uu},\; z_{xy} = 2yw_{uv},\; z_{yy} = 4y^2w_{vv} + 2w_v \end{cases}$$

Llevando las anteriores expresiones de z_y, z_{xx}, z_{xy} y z_{yy} a la ecuación diferencial del enunciado, se obtiene:

$$w_{uu} + \frac{1}{2y}(2yw_{uv}) + \frac{1}{4y^2}(4y^2w_{vv} + 2w_v) = \frac{1}{4y^3}(2yw_v), \qquad \text{o sea} \;\; w_{uu} + w_{uv} + w_{vv} = 0$$

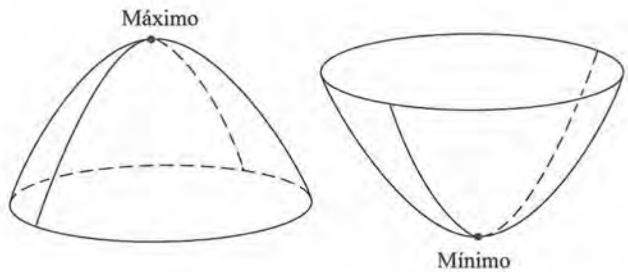
3.5. MÁXIMOS Y MÍNIMOS RELATIVOS

La noción de extremo (máximo o mínimo) relativo o local de una función real de varias variables, de la que en seguida nos ocupamos, es esencialmente la misma que en el caso de una sola variable. Sin embargo, el estudio de los extremos será ahora, que hay $p > 1$ variables, apreciablemente más complejo que cuando era $p = 1$. Si la función f que se estudia es diferenciable, hablaremos aquí, como en el caso $p = 1$, de «puntos estacionarios», que serán aquéllos en los que la diferencial de f es nula; de ser así, también ahora acontece que, de haber algún extremo, éste se presentará en un punto estacionario. Cuando estudiemos una función f que sea de clase $\mathscr{C}^2$, veremos que para que ella tenga un extremo, en un punto estacionario, es suficiente que en él la diferencial segunda de f, que es una forma cuadrática, sea definida positiva (en cuyo caso habrá mínimo) o negativa (caso de máximo), lo que sustituye a las condiciones ya conocidas, para funciones de una sola variable, de derivada segunda positiva o negativa.

❑ EXTREMOS RELATIVOS (O LOCALES)

[43]

EXTREMO RELATIVO. Sea $f: C \to \mathbb{R}$ una función real, definida en un abierto $C \subset \mathbb{R}^p$. Se dice que f tiene un máximo relativo (respectivamente, un mínimo relativo) en sentido estricto en un punto $\boldsymbol{a} \in C$ si se verifica que $f(\boldsymbol{a}) > f(\boldsymbol{x})$ (respectivamente, que $f(\boldsymbol{a}) < f(\boldsymbol{x})$) para todo $\boldsymbol{x}$ de un entorno reducido de $\boldsymbol{a}$.
Si en las definiciones precedentes se sustituyen las relaciones $f(\boldsymbol{a}) > f(\boldsymbol{x})$ y $f(\boldsymbol{a}) < f(\boldsymbol{x})$ por las $f(\boldsymbol{a}) \geqslant f(\boldsymbol{x})$ y $f(\boldsymbol{a}) \leqslant f(\boldsymbol{x})$, respectivamente, se dice entonces que el máximo y el mínimo lo son en sentido amplio.
A los anteriores máximos y mínimos de f en $\boldsymbol{a}$ (donde $\boldsymbol{a}$ es interior del campo de definición de f) se les conoce con el nombre genérico de extremos relativos o extremos locales de f en $\boldsymbol{a}$.



[43]$_1$ Observaciones

1.º Interesa resaltar que, en las anteriores definiciones, el punto $\boldsymbol{a}$ es interior de C; esto es, que si un punto $\boldsymbol{a}$ no es interior del campo de definición de una función f, entonces no tiene sentido hablar sobre si f tiene un extremo relativo en $\boldsymbol{a}$. Así, por ejemplo, la función

$$f: \mathbb{R}_+ \times \mathbb{R}_+, \qquad (x, y) \mapsto f(x, y) = \sqrt{x} + \sqrt{y}$$

tiene en el punto (0, 0) su valor mínimo absoluto:

$$f(0, 0) = 0 < f(x, y), \qquad \forall\, (x, y) \neq (0, 0) \quad \text{con } x \geqslant 0 \text{ e } y \geqslant 0$$

Pero, no obstante, no se dirá que f tiene en (0, 0) un mínimo relativo, ya que en todo entorno de (0, 0) hay puntos en los que no está definida la función.

2.º En la anterior definición, para que una función f alcance un extremo relativo en un punto $\boldsymbol{a}$, no se le exige a f ninguna condición de regularidad; f puede tener un máximo o un mínimo relativo en $\boldsymbol{a}$ sin que sea, ni siquiera, continua en $\boldsymbol{a}$. Considérese, a este respecto, el siguiente ejemplo: la función

$$f: \mathbb{R}^2 \to \mathbb{R}, \qquad (x, y) \mapsto f(x, y) = e[1 - (x^2 + y^2)]$$

(donde e = parte entera de), tiene un máximo relativo en el punto (0, 0), pues $f(0, 0) = 1$ y $f(x, y) < 1$ para todo $(x, y) \in \mathbb{R}^2 - \{(0, 0)\}$. Obsérvese que ello acontece sin necesidad de que f sea continua en (0, 0).

3.° Los extremos relativos en sentido estricto se presentan con muchas más frecuencia que los extremos relativos en sentido amplio. Generalmente, cuando se utilizan los métodos usuales de localización de extremos realtivos de funciones suficientemente regulares, los que mejor suelen caracterizarse son los extremos en sentido estricto.

Para que un extremo relativo, de f en $\boldsymbol{a}$, no lo sea en sentido estricto es necesario que en todo entorno de $\boldsymbol{a}$ haya algún punto $\boldsymbol{x} \neq \boldsymbol{a}$ tal que $f(\boldsymbol{x}) = f(\boldsymbol{a})$. Así ocurre, por ejemplo, con la función

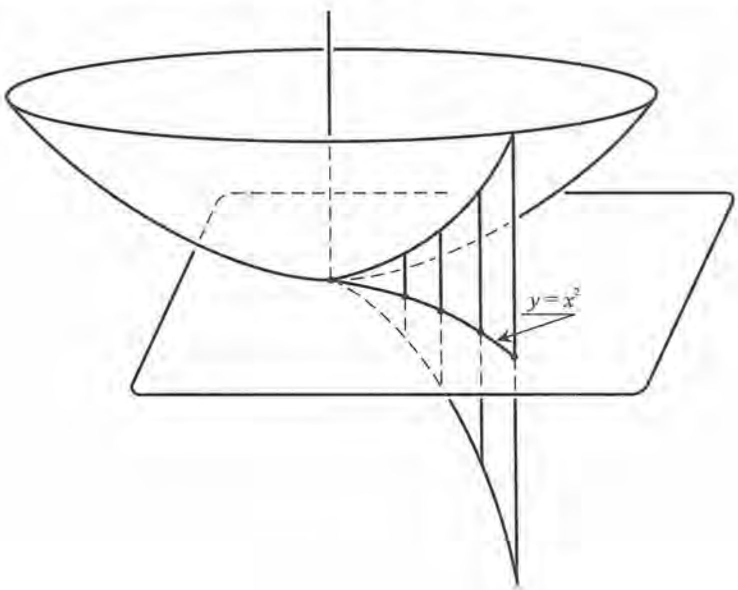
$$f: \mathbb{R}^2 \to \mathbb{R}, \qquad (x, y) \mapsto f(x, y) = (x^2 + 2y^2 - 3)^2$$

que presenta un mínimo relativo en el punto (1, 1) así como en todos los demás puntos de la elipse de ecuación $x^2 + 2y^2 = 3$ (curva continua que tiene infinitos puntos en cualquier entorno de (1, 1)), ya que la función en todos ellos vale 0 y en los demás puntos del plano alcanza valores positivos.

4.° Si la función f tiene un extremo relativo en $\boldsymbol{a}$, entonces es obvio que la restricción f_R, de f a cualquier recta R que pase por $\boldsymbol{a}$, también tiene un extremo relativo en $\boldsymbol{a}$, lo que se expresa diciendo que f tiene extremo relativo en $\boldsymbol{a}$ según cualquier dirección (o recta que pase por $\boldsymbol{a}$). Sin embargo, el recíproco no es cierto, como lo prueba el siguiente ejemplo. La función

$$f: \mathbb{R}^2 \to \mathbb{R}, \qquad (x, y) \mapsto f(x, y) = \begin{cases} x^4, & \text{si } y \neq x^2 \\ -x^4, & \text{si } y = x^2 \end{cases}$$

tiene un mínimo relativo en (0, 0) según toda recta (ya sea la $x = 0$, ya sea $y = mx$ para cualquier $m \in \mathbb{R}$) y, a pesar de ello, es evidente que f no tiene mínimo relativo en (0, 0).



CONDICIÓN NECESARIA DE EXTREMO; PUNTOS ESTACIONARIOS

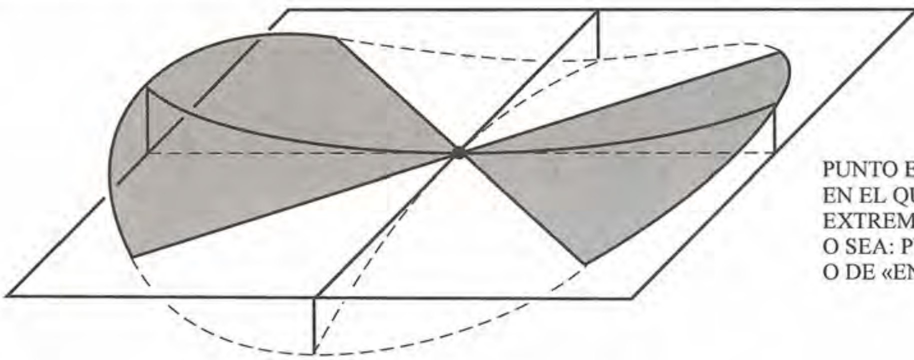
[44]

Considérese una función real, f, definida en un conjunto abierto $C \subset \mathbb{R}^p$,

$$f: C \to \mathbb{R}, \qquad \boldsymbol{x} = (x_1, ..., x_p) \mapsto f(\boldsymbol{x})$$

la cual se supone diferenciable en un cierto punto $\boldsymbol{a} \in C$. Se verifica que:

- Si f tiene un extremo relativo en $\boldsymbol{a}$, entonces la diferencial de f en $\boldsymbol{a}$ es nula, $df(\boldsymbol{a}) = o$; o sea, entonces las derivadas parciales de f en $\boldsymbol{a}$ son todas nulas: $f'_{x_1}(\boldsymbol{a})=0, ..., f'_{x_p}(\boldsymbol{a})=0$. El recíproco es falso, en general.
- Se llaman «puntos estacionarios» (o críticos) de la función diferenciable f a aquellos $\boldsymbol{a} \in C$ tales que $df(\boldsymbol{a}) = o$. Así pues, el anterior resultado se puede enunciar diciendo que: para que f tenga un extremo relativo en $\boldsymbol{a} \in C$ es necesario, pero no suficiente, que $\boldsymbol{a}$ sea un punto estacionario de f.



PUNTO ESTACIONARIO
EN EL QUE NO HAY
EXTREMO RELATIVO;
O SEA: PUNTO «PUERTO»
O DE «ENSILLADURA»

Demostración

1.º En el supuesto de que f es diferenciable en $\boldsymbol{a} \in C$ y sabiendo que f alcanza un extremo relativo en dicho punto $\boldsymbol{a}$, hemos de comprobar que $df(\boldsymbol{a}) = o$, esto es, que $df(\boldsymbol{a})(\boldsymbol{u}) = 0$ para cualquier vector $\boldsymbol{u} \in \mathbb{R}^p$. Para ello, suponiendo que es $\boldsymbol{u} \neq \boldsymbol{o}$ (pues ya sabemos que $df(\boldsymbol{a})(\boldsymbol{o}) = 0$), recurramos a que la restricción de f a la recta $\{\boldsymbol{x} = \boldsymbol{a} + t\boldsymbol{u} / t \in \mathbb{R}\}$ o, más exactamente, a la función

$$t \mapsto \varphi(t) = f(\boldsymbol{a} + t\boldsymbol{u}) \qquad \text{(definida en un entorno de } t = 0\text{)}$$

la cual es derivable en $t = 0$, pues f es diferenciable en $\boldsymbol{x} = \boldsymbol{a}$, y su derivada vale (véase [26]):

$$\varphi'(0) = \lim_{t \to 0} \frac{\varphi(t) - \varphi(0)}{t} = \lim_{t \to 0} \frac{f(\boldsymbol{a} + t\boldsymbol{u}) - f(\boldsymbol{a})}{t} = \frac{\partial f(\boldsymbol{a})}{\partial \boldsymbol{u}} = df(\boldsymbol{a})(\boldsymbol{u})$$

Así las cosas, como f tiene un extremo relativo en $\boldsymbol{a}$, es obvio que φ tiene un extremo relativo en $t = 0$, por lo que se verifica que $\varphi'(0) = 0$, que según lo anterior equivale a $df(\boldsymbol{a})(\boldsymbol{u}) = 0$, como había que verificar.

2.º Comprobemos ahora que la condición $df(\boldsymbol{a}) = o$ no es suficiente para poder garantizar que f tiene un extremo en el punto $\boldsymbol{a}$. Así acontece, en efecto, como pone de manifiesto el siguiente ejemplo. La función

$$f: \mathbb{R}^2 \to \mathbb{R}, \qquad (x, y) \mapsto z = x^2 - y^2$$

tiene diferencial nula en el punto $(x, y) = (0, 0)$, ya que $z'_x(0, 0) = 0$ y $z'_y(0, 0) = 0$, y no obstante carece de extremo en dicho punto, pues

$$z(x, 0) = x^2 > 0 \qquad \text{y} \qquad z(0, y) = -y^2 < 0, \qquad \forall\, x, y \neq 0$$

es decir, en todo entorno de $(0, 0)$ hay puntos en los que z es mayor que $z(0, 0) = 0$ y puntos en los que z es menor que dicho valor.

[44]$_1$ Ejemplo

Consideremos la función, real de dos variables reales,

$$(x, y) \mapsto z = x^2y - x^2 - 2y^2 + 3, \qquad \forall\, (x, y) \in \mathbb{R}^2$$

Los puntos críticos de esta función son las raíces del sistema de ecuaciones

$$\left.\begin{aligned} z'_x(x, y) &= 2xy - 2x = 0 \\ z'_y(x, y) &= x^2 - 4y = 0 \end{aligned}\right\} \qquad \text{o sea, los } \begin{cases} (x_0, y_0) = (0, 0) \\ (x_1, y_1) = (2, 1) \\ (x_2, y_2) = (-2, 1) \end{cases}$$

- En el punto (x_0, y_0) la función dada presenta un máximo relativo estricto, pues se puede poner (operando en coordenadas polares: $x = \rho \cos\theta$ e $y = \rho \operatorname{sen}\theta$):

$$z(0, 0) - z(x, y) = \rho^2(1 + \operatorname{sen}^2\theta - \rho\cos^2\theta \operatorname{sen}\theta)$$

  que es positivo siempre que sea $0 < \rho < 1$.
- En los puntos (x_1, y_1) y (x_2, y_2) la función dada no tiene extremo relativo. Para comprobarlo, nótese que se puede poner, tanto para $(x_i, y_i) = (x_1, y_1)$ como para $(x_i, y_i) = (x_2, y_2)$, que:

$$\Delta z = z(x_i, y_i) - z(x, y) = (x^2 - 2y - 2)(1 - y)$$

  y acontece también que los dos puntos (x_i, y_i) son los de intersección de la parábola $x^2 - 2y - 2 = 0$ con la recta $y - 1 = 0$. Como estas dos curvas dividen al plano en regiones en las que Δz tiene signo constante, el cual es distinto en cada dos regiones contiguas, resulta que en todo entorno de (x_i, y_i) hay puntos (x, y) en los que Δz es positivo y puntos en los que es negativo, por lo que no puede haber extremo relativo en ninguno de los dos puntos (x_i, y_i).

CONDICIÓN SUFICIENTE DE EXTREMO: MÉTODO DE LA DIFERENCIAL SEGUNDA

[45]

Considérese una función real $f\colon C \to \mathbb{R}$, definida en un conjunto abierto $C \subset \mathbb{R}^p$, la cual se supone de clase $\mathscr{C}^2$ en C; sea $\boldsymbol{a} \in C$ un punto estacionario de f, esto es, tal que $df(\boldsymbol{a}) = o$. Entonces, acudiendo a la diferencial segunda $d^2f(\boldsymbol{a})$ (que es una forma cuadrática(*)), se puede asegurar que:

- Si $d^2d(\boldsymbol{a})$ es definida positiva, entonces f tiene en $\boldsymbol{a}$ un mínimo relativo.
- Si $d^2f(\boldsymbol{a})$ es definida negativa, entonces f tiene en $\boldsymbol{a}$ un máximo reativo.
- Si $d^2f(\boldsymbol{a})$ no es definida ni semidefinida, entonces en $\boldsymbol{a}$ no hay un extremo de f. (Si $d^2f(\boldsymbol{a})$ es semidefinida, entonces este método no da información.)

(*) Respecto de las formas cuadráticas y, en concreto, de las formas cuadráticas definidas y semidefinidas, véase $[45]_1$. En particular, la forma cuadrática $d^2f(\boldsymbol{a})$ tiene por matriz a la $[D^2_{ij}f(\boldsymbol{a})]$, que se llama matriz hessiana de f en $\boldsymbol{a}$ (a su determinante se le llama hessiano).

Nota: Si f es de clase $\mathscr{C}^3$ en C y $df(\boldsymbol{a}) = o$, $d^2f(\boldsymbol{a}) = o$ y $d^3f(\boldsymbol{a}) \neq o$, entonces f no tiene un extremo relativo en $\boldsymbol{a}$.

Demostración

Empecemos recordando que $d^2f(\boldsymbol{a})\colon \mathbb{R}^p \to \mathbb{R}$ es la forma cuadrática que a cada incremento de la variable, $\boldsymbol{h} = \boldsymbol{x} - \boldsymbol{a}$, le hace corresponder el valor:

$$d^2f(\boldsymbol{a})(\boldsymbol{h}) = \sum_{i,j=1}^{p} a_{ij}h_ih_j, \qquad \text{donde } \begin{cases} a_{ij} = D^2_{ij}f(\boldsymbol{a}) \\ \boldsymbol{h}(h_1, \ldots, h_p) \end{cases} \tag{1}$$

Conviene resaltar el hecho de que, para cualquiera que sea el número $\alpha \in \mathbb{R}$, es:

$$d^2f(\boldsymbol{a})(\alpha\boldsymbol{h}) = \alpha^2 d^2f(\boldsymbol{a})(\boldsymbol{h}) \tag{2}$$

En lo que sigue, vamos a echar mano del desarrollo limitado de Taylor (véase [33]), que le es de aplicación a la función f, en torno del punto $\boldsymbol{a}$ y con términos hasta el de la diferencial segunda. Dicho desarrollo, como ahora es $df(\boldsymbol{a}) = o$, permite escribir:

$$f(\boldsymbol{a} + \boldsymbol{h}) - f(\boldsymbol{a}) = \frac{1}{2!}\, d^2f(\boldsymbol{a})(\boldsymbol{h}) + o[\|\boldsymbol{h}\|^2] \tag{3}$$

Nótese que, de acuerdo con lo dicho en [2] y para cualquiera que sea $\boldsymbol{h} \neq \boldsymbol{o}$, esta última relación se puede poner en la forma:

$$f(\boldsymbol{a} + \boldsymbol{h}) - f(\boldsymbol{a}) = \frac{\|\boldsymbol{h}\|^2}{2!}\left(d^2f(\boldsymbol{a})\left(\frac{\boldsymbol{h}}{\|\boldsymbol{h}\|}\right) + \frac{o[\|\boldsymbol{h}\|^2]}{\|\boldsymbol{h}\|^2}\right)$$

Por tanto, como el último sumando tiene límite cero para $\boldsymbol{h} \to \boldsymbol{o}$, es evidente que:

$$\text{signo de } [f(\boldsymbol{a} + \boldsymbol{h}) - f(\boldsymbol{a})] = \text{signo de} \left[d^2f(\boldsymbol{a})\left(\frac{\boldsymbol{h}}{\|\boldsymbol{h}\|}\right) + \varepsilon(\boldsymbol{h}) \right] \qquad [4]$$

$$\text{donde } \lim_{\boldsymbol{h} \to \boldsymbol{o}} \varepsilon(\boldsymbol{h}) = 0 \qquad \text{y} \qquad \left\| \frac{\boldsymbol{h}}{\|\boldsymbol{h}\|} \right\| = 1$$

1.° Supongamos, en primer lugar, que la forma cuadrática $d^2f(\boldsymbol{a})$ es definida positiva (en el caso de definida negativa se razona de manera análoga). Consideremos la restricción $f|_E$, de f a la esfera E de centro en el origen y radio unidad; nótese que esta esfera es un conjunto compacto de $\mathbb{R}^p$. Como la función $\boldsymbol{h} \mapsto d^2f(\boldsymbol{a})(\boldsymbol{h})$ (dada por [1]) es evidentemente continua en todo $\mathbb{R}^p$, es entonces continua en E y, como E es compacto, entonces según el teorema de Weierstrass (véase [18]) se puede asegurar que existe cierto $\boldsymbol{h}_0 \in E$ tal que $k = d^2f(\boldsymbol{a})(\boldsymbol{h}_0)$ es el valor mínimo de $f|_E$. Como f es definida positiva, es $d^2f(\boldsymbol{a})(\boldsymbol{h}_0) > 0$, luego podemos resumir diciendo que:

$$\exists\, k > 0 \,/\, \|\boldsymbol{u}\| = 1 \quad \Rightarrow \quad d^2f(\boldsymbol{a})(\boldsymbol{u}) \geqslant k$$

Por ello, acudiendo ahora a la relación [4], como de los dos sumandos del segundo miembro el primero se mantiene mayor o igual que $k > 0$ y el segundo tiende a 0, nos encontramos con que, para todo $\boldsymbol{h}$ de un cierto entorno de $\boldsymbol{a}$, el signo del segundo miembro es «más», luego también lo es el del primer miembro, es decir, f presenta un mínimo relativo en $\boldsymbol{a}$, como había que comprobar.

2.° Vamos a suponer ahora que la forma cuadrática $d^2f(\boldsymbol{a})$ no es ni definida ni semidefinida, es decir, que existen $\boldsymbol{h}_1, \boldsymbol{h}_2 \in \mathbb{R}^p$, no nulos, tales que

$$d^2f(\boldsymbol{a})(\boldsymbol{h}_1) = c_1 > 0 \qquad \text{y} \qquad d^2f(\boldsymbol{a})(\boldsymbol{h}_2) = c_2 < 0$$

Por tanto, llamando $k_1 = c_1/\|\boldsymbol{h}_1\|^2 > 0$ y $k_2 = c_2/\|\boldsymbol{h}_2\|^2 < 0$, de la relación [4] se desprende que, para todo $\lambda \in \mathbb{R}$, es:

$$\text{sig } [f(\boldsymbol{a} + \lambda\boldsymbol{h}_1) - f(\boldsymbol{a})] = \text{sig} \left[d^2f(\boldsymbol{a})\left(\frac{\lambda\boldsymbol{h}_1}{\|\lambda\boldsymbol{h}_1\|}\right) + \varepsilon(\lambda\boldsymbol{h}_1) \right] = \text{sig } [k_1 + \varepsilon_1(\lambda)] \qquad [5]$$

$$\text{sig } [f(\boldsymbol{a} + \lambda\boldsymbol{h}_2) - f(\boldsymbol{a})] = \text{sig} \left[d^2f(\boldsymbol{a})\left(\frac{\lambda\boldsymbol{h}_2}{\|\lambda\boldsymbol{h}_2\|}\right) + \varepsilon(\lambda\boldsymbol{h}_2) \right] = \text{sig } [k_2 + \varepsilon_2(\lambda)] \qquad [5']$$

donde $\varepsilon_1(\lambda) \to 0$ y $\varepsilon_2(\lambda) \to 0$ cuando $\lambda \to 0$. Por tanto, como $k_1 > 0$ y $k_2 < 0$, existe cierto entorno de $\lambda = 0$ en el que [5] es el signo «más» y [5'] es el signo «menos» y, por ello, según las rectas $\{\boldsymbol{x} = \boldsymbol{a} + \lambda\boldsymbol{h}_1/\lambda \in \mathbb{R}\}$ y $\{\boldsymbol{x} = \boldsymbol{a} + \lambda\boldsymbol{h}_2/\lambda \in \mathbb{R}\}$ la función f presenta en $\boldsymbol{a}$ un mínimo y un máximo relativos, respectivamente, lo que permite afirmar que f no tiene un extremo relativo en $\boldsymbol{a}$, como se dice en el enunciado.

3.° Si la forma cuadrática $d^2f(\boldsymbol{a})$ es semidefinida, la función f puede tener extremo relativo en $\boldsymbol{a}$ o puede carecer de él, como se prueba con el siguiente ejemplo. Sea $f: \mathbb{R}^2 \to \mathbb{R}$ la función definida mediante

$$(x, y) \mapsto f(x, y) = x^2 + ky^4$$

Esta función es de clase $\mathscr{C}^2$ en todo $\mathbb{R}^2$ y, en el punto (0, 0), es evidente que $f''_{xx}(0, 0)=2$, $f''_{xy}(0, 0) = 0$ y $f''_{yy}(0, 0) = 0$, con lo que

$$d^2f(0, 0)(\Delta x, \Delta y) = 2(\Delta x)^2$$

que es una forma cuadrática semidefinida positiva. Ahora bien, según el signo de k, se tiene:

- $k > 0 \Rightarrow f(x, y) - f(0, 0) > 0 \Rightarrow$ | f tiene un mínimo relativo en sentido estricto en (0, 0)

- $k < 0 \Rightarrow \left\{\begin{matrix} f(x, 0) - f(0, 0) > 0 \\ f(0, y) - f(0, 0) < 0 \end{matrix}\right\} \Rightarrow$ | f carece de extremo relativo en (0, 0)

- $k = 0 \Rightarrow \left\{\begin{matrix} f(x, y) > f(0, 0), \text{ si } x \neq 0 \\ f(0, y) - f(0, 0) = 0 \end{matrix}\right\} \Rightarrow$ | f tiene un mínimo relativo en sentido amplio en (0, 0)

4.° Supongamos finalmente, para comprobar la «nota», que f es de clase $\mathscr{C}^3$ en C, que $df(\boldsymbol{a}) = o$, que $d^2f(\boldsymbol{a}) = o$ y que $d^3f(\boldsymbol{a}) \neq o$. Nótese, en primer lugar, que esta última relación significa que existe un vector no nulo $\boldsymbol{u} \in \mathbb{R}^p$ tal que $d^3f(\boldsymbol{a})(\boldsymbol{u}) \neq 0$. Consideremos la siguiente función real φ:

$$t \mapsto \varphi(t) = f(\boldsymbol{a} + t\boldsymbol{u}) \qquad \text{(definida en un entorno de } t = 0)$$

Es evidente que esta función es tres veces derivable en $t = 0$ y que:

$$\varphi'(0) = \frac{\partial f(\boldsymbol{a})}{\partial \boldsymbol{u}} = d^2f(\boldsymbol{a})(\boldsymbol{u}) = 0$$

$$\varphi''(0) = \frac{\partial^2 f(\boldsymbol{a})}{\partial \boldsymbol{u}^2} = d^2f(\boldsymbol{a})(\boldsymbol{u}) = 0$$

$$\varphi'''(0) = \frac{\partial^3 f(\boldsymbol{a})}{\partial \boldsymbol{u}^3} = d^3f(\boldsymbol{a})(\boldsymbol{u}) \neq 0$$

luego φ tiene un punto de inflexión en $t = 0$, por lo que $\varphi(t) - \varphi(0)$ tiene distinto signo para $t > 0$ que para $t < 0$, es decir:

$$\text{signo de } [f(\boldsymbol{a} + t\boldsymbol{u}) - f(\boldsymbol{a})] \neq \text{signo de } [f(\boldsymbol{a} - t\boldsymbol{u}) - f(\boldsymbol{a})]$$

(para t en un cierto entorno de $t = 0$), por lo que f no tiene un extremo relativo en $\boldsymbol{a}$.

[45]$_1$ Recordatorios acerca de las formas cuadráticas

Traemos aquí aquellas definiciones y resultados, sobre las formas cuadráticas, que son necesarios para manejarse debidamente en el estudio de los extremos relativos. De cuanto se nos supone conocido a este respecto, del Álgebra Lineal, sólo vamos a ocuparnos de aquello que ahora nos hace falta acerca de las formas cuadráticas, de $\mathbb{R}^p$ en $\mathbb{R}$:

a) *Forma cuadrática.* Una forma cuadrática en $\mathbb{R}^p$ es toda aplicación del tipo:

$$\omega : \mathbb{R}^p \to \mathbb{R}, \qquad \omega(\boldsymbol{x}) = \sum_{i,j=1}^{p} a_{ij}x_i x_j, \qquad \text{siendo } \begin{cases} a_{ij} = a_{ji} \in \mathbb{R} \text{ dados} \\ \boldsymbol{x} = (x_1,\ x_2,\ \ldots,\ x_p) \end{cases}$$

Acudiendo a la matriz $A = [a_{ij}]$ (simétrica de tamaño $p \times p$; donde a_{ij} está situado en la fila i y la columna j), que se llama matriz de ω, y denotando por X a la matriz columna que forman las coordenadas de $\boldsymbol{x}$, se puede poner que

$$\omega(\boldsymbol{x}) = X^t A X = [x_1 \ \cdots \ x_p] \begin{bmatrix} a_{11} & \cdots & a_{1p} \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ a_{p1} & \cdots & a_{pp} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_p \end{bmatrix}$$

b) *Cambio de base; matrices congruentes.* Si en $\mathbb{R}^p$ se cambia la base, con lo que la nueva matriz columna de coordenadas, X', se relaciona con la anterior, X, mediante una relación del tipo $X = QX'$ (donde Q es la «matriz de cambio de coordenadas», de tamaño $p \times p$), la expresión de $\omega(\boldsymbol{x})$ en la nueva base es:

$$\omega(\boldsymbol{x}) = X'^t A' X', \qquad \text{donde } A' = Q^t A Q$$

Se dice que las matrices simétricas A y A' son congruentes si existe Q, matriz regular, tal que $A' = Q^t A Q$.

c) *Diagonalización; rango y signatura.* Existen bases de $\mathbb{R}^p$ en las que la matriz de ω es diagonal. Las matrices diagonales de ω (en las correspondientes bases) tienen todas igual número α de elementos positivos e igual número β de elementos negativos (ley de inercia); se llama signatura de ω (sig ω) al par (α, β). Se llama rango de ω (rang ω) al rango de una cualquiera de las matrices asociadas (en las distintas bases de $\mathbb{R}^p$) a ω, que es el mismo para todas las matrices; se verifica que rang $\omega = \alpha + \beta$.

d) *Relación entre autovalores, rango y signatura.* Una de las matrices diagonales asociadas a ω es aquella cuya diagonal está ocupada por los p autovalores $\lambda_1, \lambda_2, \ldots, \lambda_p$ de A (a cada uno de ellos se le cuenta tantas veces como indique su orden de multiplicidad). Por ello:

$$\text{rang } \omega = (\text{número de autovalores no nulos de } A)$$

$$\text{sig } \omega = (\text{núm. de autovalores positivos de } A,\ \text{núm. de autovalores negativos de } A)$$

A este respecto, para hallar la signatura conviene tener en cuenta que, si la ecuación característica de A es $\det(A - \lambda I) \equiv c_0 + c_1\lambda + \cdots + c_p\lambda^p$, el número de autovalores positivos de A es igual al número de cambios de signo que se presentan en la sucesión $(c_0, c_1, \ldots, c_n)$; si alguno de estos números fuese nulo, $c_i = 0$, se le asignará el signo que se quiera.

e) *Formas definidas y semidefinidas.* Se dice que la forma cuadrática ω es definida positiva (respectivamente, definida negativa) si $\omega(\boldsymbol{x}) > 0$ (respectivamente, si $\omega(\boldsymbol{x}) < 0$), para todo vector no nulo $\boldsymbol{x} \in \mathbb{R}^p$. Se dice que ω es semidefinida positiva (respectivamente, semidefinida negativa) si $\omega(\boldsymbol{x}) \geqslant 0$ (resp., si $\omega(\boldsymbol{x}) \leqslant 0$) para todo $\boldsymbol{x} \in \mathbb{R}^p$ y, además, es $\omega(\boldsymbol{x}_0) \neq 0$ para algún $\boldsymbol{x}_0 \in \mathbb{R}^p$. Se verifica que:

$$\begin{array}{lcll} \omega \text{ es definida positiva} & \Leftrightarrow & \operatorname{sig} \omega = (p, 0) & \\ \omega \text{ es definida negativa} & \Leftrightarrow & \operatorname{sig} \omega = (0, p) & \\ \omega \text{ es semidefinida positiva} & \Leftrightarrow & \operatorname{sig} \omega = (r, 0), & \text{con } r < p \\ \omega \text{ es semidefinida negativa} & \Leftrightarrow & \operatorname{sig} \omega = (0, r), & \text{con } r < p \end{array}$$

f) *Criterio de Sylvester.* Si Δ_k es el determinante de orden k que forman los elementos de las k primeras filas y las k primeras columnas de la matriz A, para $k = 1, 2, ..., p$, entonces se verifica que:

$$[\omega \text{ es definida positiva}] \Leftrightarrow [\Delta_k > 0 \qquad \text{para } k = 1, 2, ..., p]$$

$$[\omega \text{ es definida negativa}] \Leftrightarrow [(-1)^k \Delta^k > 0 \qquad \text{para } k = 1, 2, ..., p]$$

g) *Diagonalización por congruencia mediante operaciones elementales.* Se llaman operaciones elementales, realizadas en líneas paralelas (filas o columnas) de una matriz a las siguientes manipulaciones: cambiar entre sí dos líneas paralelas; multiplicar una línea por un número no nulo; sumarle a una línea cualquier combinación de líneas paralelas a aquélla; o cualquier combinación de las manipulaciones anteriores. Dada una matriz A simétrica, al aplicar a sus filas cualesquiera operaciones elementales y aplicar después a sus columnas las mismas operaciones elementales, se obtiene una matriz simétrica A' que es congruente con A. Procediendo metódicamente, aplicando adecuadas operaciones elementales, a las filas y luego a las columnas de A, es fácil obtener una matriz A' que sea diagonal, con lo que se consigue diagonalizar la forma cuadrática ω que tiene a A por matriz.

[45]$_2$ Ejercicio

Hallar los extremos relativos de la función $u: \mathbb{R}^3 \to \mathbb{R}$ definida por

$$u(x, y, z) = x^2 z + y^2 z + \frac{2}{3} z^3 - 4x - 4y - 10z + 1$$

Resolución

Empecemos hallando los puntos estacionarios de u, entre los que han de estar sus posibles extremos relativos, esto es, resolvamos el sistema de ecuaciones $u'_x = 0$, $u'_y = 0$, $u'_z = 0$:

$$\begin{array}{lccl} u'_x = 2xz - 4 & u'_x = 0 & & x = 2/z \\ u'_y = 2yz - 4 & u'_y = 0 & \Leftrightarrow & y = 2/z \\ u'_z = x^2 + y^2 + 2z^2 - 10 & u'_z = 0 & & x^2 + y^2 + 2z^2 - 10 = 0 \end{array}$$

cuyas soluciones son, obviamente, las cuatro ternas:

$$(x_1, y_1, z_1) = (2, 2, 1), \qquad (x_2, y_2, z_2) = (-2, -2, -1)$$
$$(x_3, y_3, z_3) = (1, 1, 2), \qquad (x_4, y_4, z_4) = (-2, -1, -1)$$

Acudamos al método de la diferencial segunda, d^2u, para lo que obtendremos la matriz hessiana $H(u)$ (matriz de d^2u), que es:

$$H(u) = \begin{bmatrix} u''_{xx} & u''_{xy} & u''_{xz} \\ u''_{yx} & u''_{yy} & u''_{yz} \\ u''_{zx} & u''_{zy} & u''_{zz} \end{bmatrix} = \begin{bmatrix} 2z & 0 & 2x \\ 0 & 2z & 2y \\ 2x & 2y & 4z \end{bmatrix}$$

En cada uno de los cuatro puntos estacionarios de u, la matriz hessiana es entonces la que se indica a continuación, donde también se ha dado una diagonalización de la misma (el signo $\sim$ significa: congruente con), la que nos permite clasificar d^2u (en definida positiva, definida negativa, semidefinida, ni definida ni semidefinida):

$$H(u)_1 = \begin{bmatrix} 2 & 0 & 4 \\ 0 & 2 & 4 \\ 4 & 4 & 4 \end{bmatrix} \sim \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -12 \end{bmatrix} \quad \Rightarrow \quad (d^2u)_1 \text{ no es definida ni es semidefinida}$$

$$H(u)_2 = \begin{bmatrix} -2 & 0 & -4 \\ 0 & -2 & -4 \\ -4 & -4 & -4 \end{bmatrix} \sim \begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 12 \end{bmatrix} \quad \Rightarrow \quad (d^2u)_2 \text{ no es definida ni es semidefinida}$$

$$H(u)_3 = \begin{bmatrix} 4 & 0 & 2 \\ 0 & 4 & 2 \\ 2 & 2 & 8 \end{bmatrix} \sim \begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 6 \end{bmatrix} \quad \Rightarrow \quad (d^2u)_3 \text{ definida positiva}$$

$$H(u)_4 = \begin{bmatrix} -4 & 0 & -2 \\ 0 & -4 & -2 \\ -2 & -2 & -8 \end{bmatrix} \sim \begin{bmatrix} -4 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -6 \end{bmatrix} \quad \Rightarrow \quad (d^2u)_4 \text{ definida negativa}$$

Por tanto, la función u tiene dos extremos relativos, que son:

un mínimo relativo en $(x, y, z) = (1, 1, 2)$, con $u(1, 1, 2) = -53/3$
un máximo relativo en $(x, y, z) = (-1, -1, -2)$, con $u(-1, -1, -2) = 59/3$

[45]$_3$ Caso particular: funciones de dos variables

Para determinar los extremos relativos de una función real de dos variables reales, $(x, y) \mapsto f(x, y)$, de clase $\mathscr{C}^2$ en un conjunto abierto $C \subset \mathbb{R}^2$, se puede proceder del siguiente modo:

1.º Se buscan los puntos estacionarios de f, esto es, los puntos $(a, b) \in C$ tales que $f'_x(a, b) = 0$ y $f'_y(a, b) = 0$; entre los tales puntos (a, b) han de estar los posibles extremos relativos de f.

2.º En el supuesto de que (a, b) es solución del sistema $f'_x(a, b) = 0$, $f'_y(a, b) = 0$ y llamando $\Delta(a, b)$ al determinante hessiano de f en (a, b), es decir, si

$$\Delta(a, b) = \begin{vmatrix} f''_{xx}(a, b) & f''_{xy}(a, b) \\ f''_{yx}(a, b) & f''_{yy}(a, b) \end{vmatrix} = f''_{xx}(a, b) f''_{yy}(a, b) - f''_{xy}(a, b)^2$$

entonces el método de la diferencial segunda conduce a que:

- $[\Delta(a, b) > 0 \quad \text{y} \quad f''_{xx}(a, b) > 0] \quad \Rightarrow \quad \begin{cases} f \text{ tiene un mínimo relativo} \\ \text{estricto en el punto } (a, b) \end{cases}$
- $[\Delta(a, b) > 0 \quad \text{y} \quad f''_{xx}(a, b) < 0] \quad \Rightarrow \quad \begin{cases} f \text{ tiene un máximo relativo} \\ \text{estricto en el punto } (a, b) \end{cases}$
- $[\Delta(a, b) < 0] \quad \Rightarrow \quad \begin{cases} f \text{ no tiene un extremo relativo en } (a, b) \text{ (se dice} \\ \text{que } (a, b) \text{ es punto de ensilladura o puerto de } f) \end{cases}$
- (si $\Delta(a, b) = 0$, el método de la diferencial segunda no da información)

Comprobación

Cuanto se acaba de decir se obtiene particularizando lo recién obtenido en [45], a nuestro caso particular, teniendo en cuenta el criterio de Sylvester (véase $[45]_1$, f); este criterio conduce obviamente a las cuatro conclusiones del punto segundo del anterior enunciado. A este mismo resultado se llega fácilmente si se tiene en cuenta que la diferencial segunda

$$d^2 f(a, b)(\Delta x, \Delta y) = f''_{xx}(a, b)\Delta x^2 + 2f''_{xy}(a, b)\Delta x\, \Delta y + f''_{yy}(a, b)\Delta y^2$$

es definida positiva, definida negativa, semidefinida o ni definida ni semidefinida según que (suponiendo, por ejemplo, que $f''_{xx}(a, b) \neq 0$) la función de segundo grado (en la variable $t = \Delta x/\Delta y$)

$$F(t) = f''_{xx}(a, b)t^2 + 2f''_{xy}(a, b)t + f''_{yy}(a, b)$$

(cuya representación es una parábola de eje «vertical») sea respectivamente: positiva para todo t, negativa para todo t, tenga una raíz doble o tenga dos raíces reales y distintas. Como el discriminante de la ecuación de segundo grado $F(t) = 0$ es $-\Delta(a, b)$, las anteriores posibilidades se corresponden, respectivamente, con $\Delta(a, b) > 0$ y $f''_{xx}(a, b) > 0$, $\Delta(a, b) > 0$ y $f''_{xx}(a, b) < 0$, $\Delta(a, b) = 0$ y $\Delta(a, b) < 0$, lo que confirma lo dicho en el anterior enunciado.

Según ya se ha señalado, en el caso de ser $\Delta(a, b) < 0$ se dice que (a, b) es un punto de ensilladura o puerto de f. Nótese que en estos casos, si el punto variable (x, y) se acerca al (a, b) siguiendo una recta, acontece que para algunas de estas rectas, que pasan por (a, b), f presenta un máximo relativo, en (a, b), y según otras presenta mínimo relativo. Esta es la situación que se presenta con la altura de los puestos de montaña o en una silla de montar, en la zona destinada a ubicar las posaderas; de ahi vienen los nombres de punto puerto o de ensilladura que dijimos anteriormente.

$[45]_4$ Ejercicio

Hallar los extremos relativos de la siguiente función, de $\mathbb{R}^2$ en $\mathbb{R}$:

$$(x, y) \mapsto z = x^4 + y^4 - 2x^2 + 4xy - 2y^2$$

Resolución

Los puntos estacionarios de $z(x, y)$, es decir, las raíces del sistema $z'_x = 0$, $z'_y = 0$, son:

$$\left.\begin{aligned} z'_x &\equiv 4x^3 - 4x + 4y = 0 \\ z'_y &\equiv 4y^3 + 4x - 4y = 0 \end{aligned}\right\} \Rightarrow \left\{\begin{aligned} &(x_0, y_0) = (0, 0) \\ &(x_1, y_1) = (\pm\sqrt{2}, \mp\sqrt{2}) \end{aligned}\right.$$

Para ver si en dichos puntos hay extremo relativo de z, acudiendo a la matriz hessiana

$$H(x, y) = \begin{bmatrix} z''_{xx} & z''_{xy} \\ z''_{yx} & z''_{yy} \end{bmatrix} = \begin{bmatrix} 12x^2 - 4 & 4 \\ 4 & 12y^2 - 4 \end{bmatrix}$$

y a su determinante $\Delta(x, y)$ (hessiano), según lo dicho en $[45]_3$ se tiene:

- $\Delta(x_1, y_1) = 384 > 0$ y $z''_{xx}(x_1, y_1) = 20 > 0 \rightarrow \begin{cases} z \text{ tiene un mínimo relativo en} \\ (x_1, y_1), \text{ que vale } z(x_1, y_1) = 8 \end{cases}$
- $\Delta(x_0, y_0) = 0 \rightarrow$ el método no da información para el punto (x_0, y_0).

En $(x_0, y_0) = (0, 0)$, la función $z(x, y)$ no alcanza un extremo en $(0, 0)$, ya que al estudiar sus restricciones a $x = y$ y a $x = -y$, se obtiene que

$$\begin{aligned} z(x, x) &= 2x^4 && \text{tiene un mínimo relativo en } x = 0 \\ z(x, -x) &= -8x^2 + 2x^4 && \text{tiene un máximo relativo en } x = 0 \end{aligned}$$

3.6. EXTREMOS RELATIVOS CONDICIONADOS

En lo que venimos diciendo hasta aquí, acerca de los extremos relativos de una función real $f: C \rightarrow \mathbb{R}$ (donde $C \subset \mathbb{R}^p$ es un abierto), $\boldsymbol{x} \mapsto f(\boldsymbol{x})$, a las p variables $x_1, x_2, ..., x_p$ (las p componentes de $\boldsymbol{x}$) no se las ha impuesto ninguna ligadura. Dicho de otro modo, las p variables se han podido mover libremente en un conjunto (el C) en el que hay p «grados de libertad»; a dichas p variables no se las ha obligado a satisfacer ninguna relación, no había ataduras entre ellas.

Vamos a ocuparnos ahora del caso en que sí existen unas tales ligaduras: buscaremos los extremos relativos de una función en el supuesto de que sus variables no pueden tomar sus valores libremente, sino que están obligadas a satisfacer a ciertas relaciones funcionales o condiciones de ligadura.

Ejemplo

Se desea hallar el valor mínimo que alcanza la función

$$(x, y) \mapsto z(x, y) = 2x^2 + y^2$$

cuando (x, y) recorre la parábola de ecuación $y = 1 - x^2$.

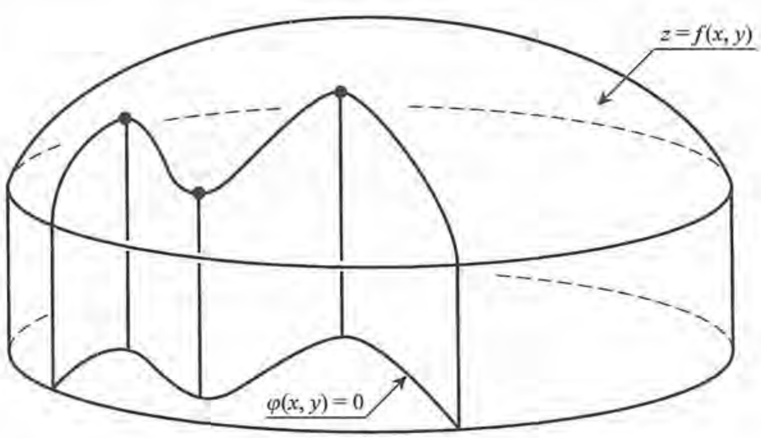
Es evidente que lo que se busca es el valor mínimo de la función

$$f(x) \equiv z(x, 1 - x^2) = 1 + x^4$$

Dicho valor mínimo se alcanza obviamente para $x = 0$ y vale $f(0) = z(0, 1) = 1$. Nótese que este valor no es el mínimo de $z(x, y)$ para $(x, y) \in \mathbb{R}^2$ (cuando entre x e y no existen ligaduras); este último mínimo es claro que vale $z(0, 0) = 0$.

CONCEPTO DE EXTREMO RELATIVO CONDICIONADO

El problema que aquí nos ocupa puede resumirse así (a grandes rasgos y para un caso muy particular): dada una función suficientemente regular $(x, y) \mapsto z = f(x, y)$, real y definida en un cierto abierto de $\mathbb{R}^2$, se desean localizar los valores máximos y mínimos relativos que alcanza z cuando el punto (x, y) recorre una cierta curva $\varphi(x, y) = 0$ (ligadura).



[46]

Sea $C \subset \mathbb{R}^p$ un conjunto abierto. Sea $f: C \to \mathbb{R}$ una función real (de la que se desean hallar los «extremos condicionados»). Sea $\varphi : C \to \mathbb{R}^q$ (con $q < p$) una función, a cuyas componentes llamaremos $\varphi_1, ..., \varphi_q$ (funciones de C en $\mathbb{R}$), y llamemos *ligadura* a la ecuación $\boldsymbol{\varphi}(\boldsymbol{x}) = \boldsymbol{o}$ (o sea, al sistema de ecuaciones $\varphi_1(x) = \cdots = \varphi_q(x) = 0$), donde $\boldsymbol{x} \in C$; se dice que el conjunto $S = \{\boldsymbol{x} \in C / \boldsymbol{\varphi}(\boldsymbol{x}) = \boldsymbol{o}\}$ es la variedad de $\mathbb{R}^p$ que tiene por ecuación a $\boldsymbol{\varphi}(\boldsymbol{x}) = \boldsymbol{o}$. Sea $\boldsymbol{a} \in C$ un punto para el que se cumple la ligadura, $\varphi(\boldsymbol{x}) = \boldsymbol{o}$, o sea, tal que $\boldsymbol{a} \in S$.

Se dice entonces que f tiene en $\boldsymbol{a}$ un «extremo relativo condicionado» por la ligadura $\varphi(\boldsymbol{x}) = \boldsymbol{o}$ si existe un entorno reducido U^* de $\boldsymbol{a}$ tal que se verifica una de las condiciones siguientes:

- $f(\boldsymbol{a}) > f(\boldsymbol{x})$, $\forall\, \boldsymbol{x} \in U^* \cap S$ (en cuyo caso el extremo es un máximo(*)).
- $f(\boldsymbol{a}) < f(\boldsymbol{x})$, $\forall\, \boldsymbol{x} \in U^* \cap S$ (en cuyo caso el extremo es un mínimo(*)).

Esto es, f tiene en $\boldsymbol{a}$ un extremo relativo condicionado por $\boldsymbol{\varphi}(\boldsymbol{x}) = \boldsymbol{o}$ si la restricción $f|_S$ tiene un extremo relativo (ordinario) en $\boldsymbol{a}$.

(*) Estos máximo y mínimo lo son en sentido estricto. Si las anteriores relaciones $f(\boldsymbol{a}) > f(\boldsymbol{x})$ y $f(\boldsymbol{a}) < f(\boldsymbol{x})$ se sustituyen por las $f(\boldsymbol{a}) \geqslant f(\boldsymbol{x})$ y $f(\boldsymbol{a}) \leqslant f(\boldsymbol{x})$, respectivamente, se obtienen los máximos y mínimos en sentido amplio.

[46]₁ Reducción de un extremo condicionado a un extremo ordinario

Además de las anteriores condiciones [46], admitamos también que la ligadura $\boldsymbol{\varphi}(\boldsymbol{x}) = \boldsymbol{o}$, esto es, que el sistema de ecuaciones

$$\left.\begin{array}{c} \varphi_1(x_1, ..., x_q, ..., x_p) = 0 \\ \dots\dots\dots\dots\dots\dots\dots \\ \varphi_q(x_1, ..., x_q, ..., x_p) = 0 \end{array}\right\} \qquad [1]$$

define implícitamente a q de las variables $x_1, ..., x_p$ como funciones implícitas de las restantes $p - q$, en torno del punto $\boldsymbol{a} = (a_1, ..., a_q, ..., a_p)$(*); supongamos que dichas q variables son $x_1, ..., x_q$ (de no ser así, ello se logra reordenando adecuadamente las variables) y llamemos $\psi_1, ..., \psi_q$ a las referidas funciones implícitas que define el sistema [1]:

$$\left.\begin{array}{c} x_1 = \psi_1(x_{q+1}, ..., x_p) \\ \dots\dots\dots\dots\dots\dots \\ x_q = \psi_1(x_{q+1}, ..., x_p) \end{array}\right\} \qquad [2]$$

(*) Para ello es suficiente, según el teorema de la función implícita (véase [36]), que $\boldsymbol{\varphi}$ sea de clase $\mathscr{C}^1$ en un entorno de $\boldsymbol{a}$ y que el jacobiano de $(\varphi_1, ..., \varphi_q)$ respecto de unas ciertas variables $(x_{i_1}, ..., x_{i_q})$ sea no nulo en $\boldsymbol{a}$; esto último equivale a que la diferencial $d\varphi(\boldsymbol{a})$ tenga rango q en el punto $\boldsymbol{a}$.

Vamos a suponer también que estas funciones, no sólo existen, sino que nos son conocidas. Al llevar los valores que se obtienen de [2] a la expresión de $f(\boldsymbol{x})$, se llega a la función F definida, para todo $(x_{q+1}, ..., x_p)$ de un cierto entorno de $(a_{q+1}, ..., a_p)$, mediante:

$$F(x_{q+1}, ..., x_p) = f(\psi_1(x_{q+1}, ..., x_p), ..., \psi_q(x_{q+1}, ..., x_p), x_{q+1}, ..., x_p)$$

y es obvio que la existencia de un extremo relativo de f en $\boldsymbol{a}$ condicionado por la ligadura $\varphi(\boldsymbol{x}) = 0$ equivale a la existencia de un extremo relativo ordinario (sin condiciones) de F en el punto $(a_{q+1}, ..., a_p)$.

Observación

Para que pueda ser de aplicación lo que acabamos de decir, no sólo es necesario que de las ligaduras se puedan «despejar», de un modo efectivo, q de las variables en función de las otras $p - q$, cosa que ocurre en muy pocas ocasiones, sino que las expresiones que así se obtienen deben definir, realmente, a q funciones en un entorno de cada uno de los puntos que hayan de ser estudiados. A este respecto, véase lo que acontece en el siguiente ejemplo.

[46]$_2$ Ejercicio

Considérese la función $f: \mathbb{R}^2 \to \mathbb{R}$ definida por

$$f(x, y) = (x - 1)^2 + y^2$$

que evidentemente tiene su mínimo absoluto (que también lo es relativo) en el punto $(x, y) = (1, 0)$, con $f(1, 0) = 0$. Se desean hallar los extremos relativos de $f(x, y)$ condicionados por la ligadura $x^2 + y^2 = 1$. Como el conjunto $C = \{(x, y) \in \mathbb{R}^2 / x^2 + y^2 = 1\}$ (circunferencia) es un compacto de $\mathbb{R}^2$ y como f es continua en dicho conjunto C, existen los valores máximo y mínimo de f en C (teorema de Weierstrass); estos máximo y mínimo absolutos son, obviamente, extremos relativos de $f(x, y)$ condicionados por $x^2 + y^2 = 1$, por lo que tales extremos han de existir. Pues bien, para buscarlos, procédase del siguiente modo:

1.º En la ligadura, despéjese y^2 en función de x; llévese dicho valor a $f(x, y)$ y hállense los extremos de la función, de la variable x, que así se obtiene. Justifíquese el resultado al que se llega.

2.º Parametrícese la ligadura, poniendo $x = \cos\theta$ e $y = \operatorname{sen}\theta$, llévense dichas expresiones a $f(x, y)$ y hállense los extremos de la función, de la variable $\theta \in \mathbb{R}$, que así se obtiene.

Resolución

1.º Haciendo lo que se pide, obtenemos:

$$x^2 + y^2 = 1 \to y^2 = 1 - x^2; \quad f(x, y(x)) = (x - 1)^2 + (1 - x^2) = 2 - 2x$$

Esta última función carece (aparentemente) de extremos, pues es decreciente para todo x. Lo que realmente ocurre es que el campo de variación de x no es todo $\mathbb{R}$, sino el intervalo $[-1, 1]$ (ya que $y^2 = 1 - x^2$ debe ser no negativo) y, por ello, los valores máximo y mínimo que buscamos son:

$$f(-1, y(-1)) = f(-1, 0) = 4 \qquad \text{y} \qquad f(1, y(1)) = f(1, 0) = 0$$

Nótese que estos valores han quedado «emboscados» (aparentemente no existían) debido a que la expresión $y = \pm\sqrt{1 - x^2}$ (que se obtiene de la ligadura) no es función (ni con el $+$, ni con el $-$) ni en todo un entorno de $x = -1$, ni en todo un entorno de $x = 1$, condición esta que se hace necesaria para que este procedimiento, de obtener los extremos condicionados reduciéndoles a extremos relativos ordinarios, funcione debidamente.

2.° Al hacer $x = \cos\theta$ e $y = \operatorname{sen}\theta$, con lo que se cumple la ligadura, no hay que esperar ningún tipo de anomalía, como la del apartado anterior, ya que estas expresiones definen a x e y como funciones de θ en todo entorno de cualquier $\theta \in \mathbb{R}$. Operando como se indica en el enunciado, se obtiene:

$$f(\cos\theta, \operatorname{sen}\theta) = (\cos\theta - 1)^2 + \operatorname{sen}^2\theta = 2 - 2\cos\theta$$

que alcanza sus valores máximo y mínimo (relativos y absolutos) para $\cos\theta = -1$ y $\cos\theta = 1$, esto es, cuando $\theta = \pi$ y $\theta = 0$, con lo que dichos valores máximo y mínimo son

$$f(\cos\pi, \operatorname{sen}\pi) = f(-1, 0) = 4 \qquad \text{y} \qquad f(\cos 0, \operatorname{sen} 0) = f(1, 0) = 0$$

$[46]_3$ Observación (ilustración del método de Lagrange)

Antes de abordar el «método de los multiplicadores de Lagrange», del que en seguida nos ocuparemos, vamos a hacer algunas consideraciones, no rigurosas, para justificar dicho método.

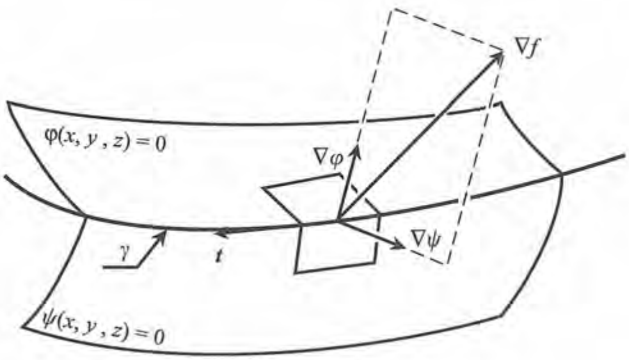
Consideremos el problema de hallar los extremos relativos de $f(x, y, z)$ cuando el punto (x, y, z) recorre la curva γ de ecuaciones $\varphi(x, y, z) = 0$ y $\psi(x, y, z) = 0$. Si en un cierto punto P de la curva hay un tal extremo, esto es, si la variación de f a lo largo de γ alcanza un máximo o un mínimo en P, entonces la variación de f en P según la dirección de la tangente a la curva deberá ser nula (en caso contrario, f crecería o decrecería en P al desplazar (x, y, z) sobre γ), es decir, sería 0 la derivada de f según el vector $\boldsymbol{t}$ tangente a la curva en P, o sea, ∇f tendría que ser perpendicular a $\boldsymbol{t}$. Este vector tangente es perpendicular a los vectores normales, en P, a las superficies $\varphi(x, y, z) = 0$ y $\psi(x, y, z) = 0$, que son los vectores gradiente de $\nabla\varphi$ y $\nabla\psi$ (independientes, salvo que P fuese un punto singular de γ), luego ∇f, por ser perpendicular a $\boldsymbol{t}$, es una combinación lineal de $\nabla\varphi$ y $\nabla\psi$, o sea, existirán ciertos λ, $\mu \in \mathbb{R}$ tales que

$$\nabla f + \lambda\nabla\varphi + \mu\nabla\psi = \boldsymbol{o} \qquad \text{o} \qquad \nabla(f + \lambda\varphi + \mu\psi) = \boldsymbol{o}$$

A la función $g = f + \lambda\varphi + \mu\psi$ se la llama función de Lagrange y se dice que λ y μ son los multiplicadores de Lagrange. Así que, para que en un punto P se presente un extremo relativo de $f(x, y, z)$ condicionado por las ligaduras $\varphi(x, y, z) = 0$ y $\psi(x, y, z) = 0$ es necesario que, para unos ciertos $\lambda, \mu \in \mathbb{R}$, se verifique, en dicho punto, que:

$$\begin{aligned} f_x + \lambda\varphi_x + \mu\psi_x &= 0 \\ f_y + \lambda\varphi_y + \mu\psi_y &= 0 \\ f_z + \lambda\varphi_z + \mu\psi_z &= 0 \end{aligned} \qquad \begin{aligned} \varphi &= 0 \\ \psi &= 0 \end{aligned}$$

(cinco ecuaciones en las cinco incógnitas x, y, z, λ y μ). Resolviendo este sistema se obtienen los «puntos críticos sobre la curva γ», entre los que han de estar los posibles extremos relativos condicionados que andamos buscando. Para comprobar si cada uno de estos puntos críticos es o no un extremo condicionado, se acudirá mas adelante a las derivadas segundas.



CONDICIÓN NECESARIA DE EXTREMO CONDICIONADO: FUNCIÓN DE LAGRANGE

[47]

Sean $f: U \to \mathbb{R}$ y $\varphi: U \to \mathbb{R}^q$ dos funciones de clase $\mathscr{C}^1$ en un entorno $U \subset \mathbb{R}^p$ de un punto $\boldsymbol{a} \in \mathbb{R}^p$; se supone que es $\varphi(\boldsymbol{a}) = \boldsymbol{o}$, que es $q < p$ y que rang $d\varphi(\boldsymbol{a}) = q$(*).

Para que la función $\boldsymbol{x} \mapsto f(\boldsymbol{x})$ tenga un extremo relativo en $\boldsymbol{a}$ condicionado por la ligadura $\varphi(\boldsymbol{x}) = \boldsymbol{o}$ es necesario que existan $\lambda_1, ..., \lambda_q \in \mathbb{R}$ (multiplicadores de Lagrange), únicos, tales que la función $g = f + \lambda_1\varphi_1 + \cdots + \lambda_q\varphi_q$ (donde $\varphi_1, ..., \varphi_q$ son las componentes de φ; a g se la llama función de Lagrange) tenga un punto crítico en $\boldsymbol{a}$, es decir, tales que $dg(\boldsymbol{a}) = \boldsymbol{o}$.

(*) Como la matriz de $d\varphi(\boldsymbol{a})$ es la jacobiana $J\varphi(\boldsymbol{a}) = [\partial\varphi/\partial\boldsymbol{x}](\boldsymbol{a})$, la condición rang $d\varphi(\boldsymbol{a}) = q$ significa que, para algunas variables $x_{i_1}, ..., x_{i_q}$ (que son q de las p componentes de $\boldsymbol{x}$), el jacobiano de $\varphi = (\varphi_1, ..., \varphi_q)$ respecto de $(x_{i_1}, ..., x_{i_q})$ es no nulo. Podemos suponer, sin pérdida de generalidad, que dichas variables son las $x_1, ..., x_q$.

Demostración

Como rang $\varphi(\boldsymbol{a}) = q$, se sabe que (véase la nota al pie del enunciado)

$$\det \left[\frac{\partial(\varphi_1, \ldots, \varphi_q)}{\partial(x_1, \ldots, x_q)} (\boldsymbol{a}) \right] \neq 0 \qquad [1]$$

Por ello y dado que $\boldsymbol{\varphi}$ es de clase $\mathscr{C}^1$ en U, el teorema de la función implícita (véase [36]) permite asegurar que la ecuación $\boldsymbol{\varphi}(\boldsymbol{x}) = \boldsymbol{o}$ (ligadura) define implícitamente a las variables $x_1, \ldots, x_p$ como funciones de clase $\mathscr{C}^1$ de $x_{p+1}, \ldots, x_q$ en torno del punto $\boldsymbol{a} = (a_1, \ldots, a_q, \ldots, a_p)$; sabemos, pues, que:

$$\left.\begin{array}{l} \varphi_1(x_1, \ldots, x_q, \ldots, x_p) = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \varphi_q(x_1, \ldots, x_q, \ldots, x_p) = 0 \end{array}\right\} \quad \text{define a} \quad \left\{\begin{array}{l} x_1 = \psi_1(x_{q+1}, \ldots, x_p) \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ x_q = \psi_q(x_{q+1}, \ldots, x_p) \end{array}\right.$$

de manera que se verifica, para todo $(x_{q+1}, \ldots, x_p)$ de un cierto entorno del punto $\hat{\boldsymbol{a}} = (a_{q+1}, \ldots, a_p)$, que

$$\varphi_i(\psi_1(x_{q+1}, \ldots, x_p), \ldots, \psi_q(x_{q+1}, \ldots, x_p), x_{q+1}, \ldots, x_p) \equiv 0 \qquad [2]$$

$$a_i = \psi_i(a_{q+1}, \ldots, a_p) \qquad (\text{para } i = 1, \ldots, q)$$

(estas funciones ψ_i, de cuya existencia y regularidad, clase $\mathscr{C}^1$, tenemos certeza, no serán conocidas, en general, de modo explícito). En el supuesto de que f tiene en $\boldsymbol{a}$ un extremo relativo condicionado por la ligadura $\varphi(\boldsymbol{x}) = \boldsymbol{o}$ (que es nuestra hipótesis), resulta que la siguiente función (definida en un entorno de $\hat{\boldsymbol{a}}$)

$$F(x_{q+1}, \ldots, x_p) \equiv f(\psi_1(x_{q+1}, \ldots, x_p), \ldots, \psi_q(x_{q+1}, \ldots, x_p), x_{q+1}, \ldots, x_p) \qquad [3]$$

tiene un extremo relativo ordinario (no condicionado) en el punto $\hat{\boldsymbol{a}} = (a_{q+1}, \ldots, a_p)$. De esto último se infiere (véase [44]) que $dF(\hat{\boldsymbol{a}}) = o$, es decir, acudiendo a la regla de diferenciación de una función compuesta (regla de invarianza; véase [29], 2.º), se verifica que: para cualesquiera que sean $dx_{q+1}, \ldots, dx_p$, es:

$$\frac{\partial f(\boldsymbol{a})}{\partial x_1} d\psi_1(\hat{\boldsymbol{a}}) + \cdots + \frac{\partial f(\boldsymbol{a})}{\partial x_q} d\psi_q(\hat{\boldsymbol{a}}) + \frac{\partial f(\boldsymbol{a})}{\partial x_{q+1}} dx_{q+1} + \cdots + \frac{\partial f(\boldsymbol{a})}{\partial x_p} dx_p \equiv 0 \qquad [4]$$

(nótese que, por razones de brevedad, se ha puesto $d\psi_j(\hat{\boldsymbol{a}})$ para representar a $d\psi_j(\hat{\boldsymbol{a}})(dx_{q+1}, \ldots, dx_p)$). Por otra parte, como las funciones, de $x_{q+1}, \ldots, x_p$, que definen los primeros miembros de [2] son nulas en un entorno de $\hat{\boldsymbol{a}}$, se verifica entonces que sus diferenciales en $\hat{\boldsymbol{a}}$ son nulas, lo cual permite poner (de acuerdo con la regla de diferenciación de la función compuesta) que, para $i = 1, ..., q$ y para cualesquiera $dx_{q+1}, ..., dx_p$, es:

$$\frac{\partial \varphi_i(\boldsymbol{a})}{\partial x_1} d\psi_1(\hat{\boldsymbol{a}}) + \cdots + \frac{\partial \varphi_i(\boldsymbol{a})}{\partial x_q} d\psi_q(\hat{\boldsymbol{a}}) + \frac{\partial \varphi_i(\boldsymbol{a})}{\partial x_{q+1}} dx_{q+1} + \cdots + \frac{\partial \varphi_i(\boldsymbol{a})}{\partial x_p} dx_p \equiv 0 \qquad [5]$$

Construyamos ahora la función real g definida, en el entorno U de $\boldsymbol{a}$, mediante:

$$x \mapsto g(\boldsymbol{x}) = f(\boldsymbol{x}) + \lambda_1\varphi_1(\boldsymbol{x}) + \cdots + \lambda_q\varphi_q(\boldsymbol{x}) \qquad [6]$$

y vamos a comprobar que, según se dijo en el enunciado, existen unos únicos $\lambda_1, ..., \lambda_q \in \mathbb{R}$ tales que $dg(\boldsymbol{a}) = o$, o sea, tales que $\partial g(\boldsymbol{a})/\partial x_i = 0$, para $i = 1, ..., q, ..., p$. Para ello, veamos que existen unos únicos $\lambda_1, ..., \lambda_q$ tales que $\partial g(\boldsymbol{a})/\partial x_i = 0$, para $i = 1, ..., q$ y que para dichos $\lambda_1, ..., \lambda_q$ se verifica, además, que $\partial g(\boldsymbol{a})/\partial x_i = 0$ para los restantes valores de i ($i = q + 1, ..., p$). En efecto: las ecuaciones $\partial g(\boldsymbol{a})/\partial x_i = 0$, para $i = 1, ..., q$, son, de acuerdo con [6]:

$$\frac{\partial g(\boldsymbol{a})}{\partial x_i} \equiv \frac{\partial f(\boldsymbol{a})}{\partial x_i} + \lambda_1 \frac{\partial \varphi_1(\boldsymbol{a})}{\partial x_i} + \cdots + \lambda_q \frac{\partial \varphi_q(\boldsymbol{a})}{\partial x_i} = 0 \qquad (\text{para } i = 1, ..., q) \qquad [7]$$

Este sistema, de q ecuaciones lineales en las q incógnitas $\lambda_1, ..., \lambda_q$, tiene solución única, pues el determinante de dicho sistema es el jacobiano en $\boldsymbol{a}$ de $(\varphi_1, ..., \varphi_q)$ respecto de $(x_1, ..., x_q)$ que, según ya se indicó en [1], es no nulo.

Si ahora sumamos miembro a miembro la identidad [4], más la primera de las identidades [5] multiplicada por λ_1, etc., más la q-ésima de las identidades [5] multiplicada por λ_q, obtenemos (a la vista de la definición [6] de g) que:

$$\frac{\partial g(\boldsymbol{a})}{\partial x_1} d\psi_1(\hat{\boldsymbol{a}}) + \cdots + \frac{\partial g(\boldsymbol{a})}{\partial x_q} d\psi_q(\hat{\boldsymbol{a}}) + \frac{\partial g(\boldsymbol{a})}{\partial x_{q+1}} dx_{q+1} + \cdots + \frac{\partial g(\boldsymbol{a})}{\partial x_p} dx_p \equiv 0 \qquad [8]$$

(para cualesquiera $dx_{q+1}, ..., dx_p$)

Si para $(\lambda_1, ..., \lambda_p)$ se toma la solución del sistema [7], la identidad [8] se reduce, pues, a:

$$\frac{\partial g(\boldsymbol{a})}{\partial x_{q+1}} dx_{q+1} + \cdots + \frac{\partial g(\boldsymbol{a})}{\partial x_p} dx_p \equiv 0$$

Como esta última relación se verifica para cualesquiera que sean $dx_{q+1}, ..., dx_p$, resulta entonces que para dichos $\lambda_1, ..., \lambda_p$, se verifica que:

$$\frac{\partial g(\boldsymbol{a})}{\partial x_{q+1}} = 0, ..., \frac{\partial g(\boldsymbol{a})}{\partial x_p} = 0$$

con lo que concluye la demostración.

[47]$_1$ Ejercicio

Hallar los valores máximo y mínimo que alcanza

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

cuando (x, y) recorre la elipse $x^2 + y^2 + xy = 4$.

Resolución

Como f es continua en $\mathbb{R}^2$ y como la elipse es un conjunto compacto de $\mathbb{R}^2$, resulta que, según el teorema de Weierstrass (véase [18], II), existen los máximos y mínimos pedidos. Si acudimos a los multiplicadores de Lagrange (véase [47]), para el problema de hallar extremos relativos de

$$f(x, y), \qquad \text{con la ligadura } \varphi(x, y) \equiv x^2 + y^2 + xy - 4 = 0$$

y determinamos todos los puntos estacionarios de la correspondiente función de Lagrange que verifican a la ligadura, es evidente que los máximos y mínimos pedidos son aquellos dos, de tales puntos estacionarios, en los que f alcanza los valores mayor y menor, respectivamente. Apliquemos, pues, el método de los multiplicadores de Lagrange, para lo que construimos la función

$$g(x, y) = f(x, y) + \lambda\varphi(x, y) = xy + \lambda(x^2 + y^2 + xy - 4)$$

y consideramos el sistema de ecuaciones:

$$\left.\begin{aligned} g'_x(x, y) &\equiv \qquad y + \lambda(2x + y) = 0 \\ g'_y(x, y) &\equiv x^2 + y^2 + xy - 4 = 0 \end{aligned}\right\} \quad ((x, y) \text{ es un punto estacionario de } g)$$

$$\varphi(x, y) \equiv x^2 + y^2 + xy - 4 = 0 \qquad (\text{la ligadura se verifica para } (x, y))$$

Las soluciones de este sistema son, como se comprueba con facilidad:

$$(x_1, y_1; \lambda_1) = (2/\sqrt{3}, 2/\sqrt{3}; -1/3) \qquad (x_2, y_2; \lambda_2) = (-2/\sqrt{3}, -2/\sqrt{3}; -1/3)$$
$$(x_3, y_3; \lambda_3) = (2, -2: 1) \qquad (x_4, y_4; \lambda_4) = (-2, 2: 1)$$

Como los valores de f en los cuatro puntos estacionarios anteriores son

$$f(x_1, y_1) = 4/3, \quad f(x_2, y_2) = 4/3, \quad f(x_3, y_3) = -4, \quad f(x_4, y_4) = -4$$

resulta ya evidente que los valores pedidos son:

máximo: $4/3$, que se alcanza en $(2/\sqrt{3}, 2/\sqrt{3})$ y en $(-2/\sqrt{3}, -2/\sqrt{3})$
mínimo: -4, que se alcanza en $(2, -2)$ y en $(-2, 2)$

CONDICIÓN SUFICIENTE DE EXTREMO CONDICIONADO

[48]

Sean $f: U \to \mathbb{R}$ y $\varphi: U \to \mathbb{R}^q$ dos funciones de clase $\mathscr{C}^2$ en un entorno $U \subset \mathbb{R}^p$ de un punto $\boldsymbol{a} \in \mathbb{R}^p$ tal que $\varphi(\boldsymbol{a}) = \boldsymbol{o}$ y rang $d\varphi(\boldsymbol{a}) = q$, siendo $q < p$. Se supone que $\boldsymbol{a}$ es un punto estacionario de la función de Lagrange $g = f + \lambda_1\varphi_1 + \cdots + \lambda_q\varphi_q$ (donde $\varphi_1, ..., \varphi_q$ son las componentes de $\boldsymbol{\varphi}$ y $\lambda_1, ..., \lambda_q$ son unos ciertos números reales), esto es, se supone que existen $\lambda_1, ..., \lambda_q \in \mathbb{R}$ tales que $dg(\boldsymbol{a}) = o$.

Entonces (si es $\Delta\boldsymbol{x} \in \mathbb{R}^p$ el incremento de la variable $\boldsymbol{x} \in U$, o sea, llamando $\Delta\boldsymbol{x} = \boldsymbol{x} - \boldsymbol{a}$) se verifica que:

1.º Si $d^2g(\boldsymbol{a})(\Delta\boldsymbol{x}) > 0$ para todo $\Delta\boldsymbol{x} \neq \boldsymbol{o}$ tal que $d\varphi(\boldsymbol{a})(\Delta\boldsymbol{x}) = \boldsymbol{o}$ entonces f tiene en $\boldsymbol{a}$ un mínimo relativo condicionado por la ligadura $\varphi(\boldsymbol{x}) = \boldsymbol{o}$.

2.º Si $d^2g(\boldsymbol{a})(\Delta\boldsymbol{x}) < 0$ para todo $\Delta\boldsymbol{x} \neq \boldsymbol{o}$ tal que $d\varphi(\boldsymbol{a})(\Delta\boldsymbol{x}) = \boldsymbol{o}$, entonces f tiene en $\boldsymbol{a}$ un máximo relativo condicionado por la ligadura $\varphi(\boldsymbol{x}) = \boldsymbol{o}$.

3.º Si $d^2g(\boldsymbol{a})(\Delta\boldsymbol{x}_1) > 0$ y $d^2g(\boldsymbol{a})(\Delta\boldsymbol{x}_2) < 0$ para unos ciertos $\Delta\boldsymbol{x}_1$ e $\Delta\boldsymbol{x}_2$ tales que $d\varphi(\boldsymbol{a})(\Delta\boldsymbol{x}_1) = \boldsymbol{o}$ y $d\varphi(\boldsymbol{a})(\Delta\boldsymbol{x}_2) = \boldsymbol{o}$, entonces f no tiene en $\boldsymbol{a}$ un extremo relativo condicionado por la ligadura $\boldsymbol{\varphi}(\boldsymbol{x}) = \boldsymbol{o}$.

Demostración

Para empezar, nótese que la relación $d\varphi(\boldsymbol{a})(\Delta\boldsymbol{x}) = \boldsymbol{o}$, de la que tanto se habla en el enunciado, equivale a la

$$\Delta\boldsymbol{x} \in V, \qquad \text{donde } V \text{ es el núcleo de } d\varphi(\boldsymbol{a}): \mathbb{R}^p \to \mathbb{R}^q \tag{1}$$

Como el rango de $d\varphi(\boldsymbol{a})$ es q, resulta que V tiene dimensión $p - q$. La condición rang $d\varphi(\boldsymbol{a}) = q$ significa que no es nulo el jacobiano de $\boldsymbol{\varphi}$ en $\boldsymbol{a}$ respecto de unas ciertas $x_{i_1}, ..., x_{i_q}$, o sea, de q de sus p variables $x_1, ..., x_p$ (componentes de $\boldsymbol{x}$). Para simplificar la notación, llamaremos $\boldsymbol{y} = (y_1, ..., y_{p-q})$ a las $p - q$ variables x_i distintas de las $x_{i_1}, ..., x_{i_q}$; llamaremos $\boldsymbol{b}$ al $\boldsymbol{y}$ que se obtiene para $\boldsymbol{x} = \boldsymbol{a}$.

Como el jacobiano de φ respecto de $(x_{i_1}, ..., x_{i_q})$ es no nulo en $\boldsymbol{a}$, la ecuación $\varphi(\boldsymbol{x}) = \boldsymbol{o}$ define implícitamente, en torno de $\boldsymbol{a}$, a las variables $x_{i_1}, ..., x_{i_q}$ como funciones de clase $\mathscr{C}^1$ de $\boldsymbol{y}$. Es evidente que cada componente de $\boldsymbol{y}$ es función (función idéntica) de clase $\mathscr{C}^1$ de sí misma, con lo que resulta que existe una función $\boldsymbol{y} \mapsto \boldsymbol{x} = \boldsymbol{\psi}(\boldsymbol{y})$, de clase $\mathscr{C}^1$ en un cierto entorno $U_{\boldsymbol{b}}$ de $\boldsymbol{b}$, tal que $\boldsymbol{\varphi}(\boldsymbol{\psi}(\boldsymbol{y})) \equiv \boldsymbol{o}$ para todo $\boldsymbol{y} \in U_{\boldsymbol{b}}$ y verificando que $\boldsymbol{\psi}(\boldsymbol{b}) = \boldsymbol{a}$. Nótese que $d\boldsymbol{\psi}(\boldsymbol{b})$ (aplicación lineal de $\mathbb{R}^{p-q}$ en $\mathbb{R}^p$) tiene rango $p - q$, es decir, que su imagen $W = \{d\boldsymbol{\psi}(\boldsymbol{b})(\Delta\boldsymbol{y})/\Delta\boldsymbol{y} \in \mathbb{R}^{p-q}\}$ tiene dimensión $p - q$; esto es así, pues entre las componentes de $\boldsymbol{\psi}$ hay $p - q$ de ellas que son la identidad. Nótese también que, por ser $\boldsymbol{\varphi}(\boldsymbol{\psi}(\boldsymbol{y})) \equiv \boldsymbol{o}$, para $\boldsymbol{y} \in U_{\boldsymbol{b}}$, se verifica que $d\varphi(\boldsymbol{a})(d\boldsymbol{\psi}(\boldsymbol{b})(\Delta\boldsymbol{y})) = \boldsymbol{o}$ para todo $\Delta\boldsymbol{y} \in \mathbb{R}^{p-q}$, es decir, que el anterior espacio W está incluido en el núcleo de $d\varphi(\boldsymbol{a})$, al que le habíamos llamado V en [1]. Como V y W tienen ambos dimensión $p - q$ y se verifica que $W \subset V$, de todo ello se concluye que $V = W$, es decir, se verifica que:

$$[\Delta\boldsymbol{x} \in \mathbb{R}^p/d\varphi(\boldsymbol{a})\Delta\boldsymbol{x} = \boldsymbol{o}] \quad \Leftrightarrow \quad [\exists\ \Delta\boldsymbol{y} \in \mathbb{R}^{p-q}/\Delta\boldsymbol{x} = d\boldsymbol{\psi}(\boldsymbol{b})(\Delta\boldsymbol{y})] \tag{2}$$

Por otra parte, es claro que la existencia de un extremo relativo de $\boldsymbol{x} \mapsto f(\boldsymbol{x})$ en el punto $\boldsymbol{a}$ condicionado por la ligadura $\boldsymbol{\varphi}(\boldsymbol{x}) = \boldsymbol{o}$ equivale a la existencia de un extremo relativo (ordinario) de $\boldsymbol{y} \mapsto F(\boldsymbol{y}) = f(\boldsymbol{\psi}(\boldsymbol{y}))$ en el punto $\boldsymbol{b}$. Nótese que como $\boldsymbol{\varphi}(\boldsymbol{\psi}(\boldsymbol{y})) \equiv \boldsymbol{o}$ para todo $\boldsymbol{y}$ del entorno U_b, esto es, como $\varphi_i(\boldsymbol{\psi}(\boldsymbol{y})) = 0$ para $\boldsymbol{y} \in U_b$, entonces la función

$$\boldsymbol{y} \mapsto G(\boldsymbol{y}) = g(\boldsymbol{\psi}(\boldsymbol{y})) = f(\boldsymbol{\psi}(\boldsymbol{y})) + \lambda_1\varphi_1(\boldsymbol{\psi}(\boldsymbol{y})) + \cdots + \lambda_q\varphi_q(\boldsymbol{\psi}(\boldsymbol{y})) \qquad [3]$$

donde g es la función de Lagrange) coincide con la $\boldsymbol{y} \mapsto F(\boldsymbol{y})$ para $\boldsymbol{y} \in U_b$, por lo que se verifica la siguiente equivalencia:

$$\begin{bmatrix} f \text{ tiene un extremo relativo en } \boldsymbol{a} \\ \text{condicionado por la ligadura } \boldsymbol{\varphi}(\boldsymbol{x}) = \boldsymbol{o} \end{bmatrix} \Leftrightarrow \begin{bmatrix} G \text{ tiene en } \boldsymbol{b} \text{ un} \\ \text{extremo relativo} \end{bmatrix} \qquad [4]$$

Vamos a ocuparnos, pues, de las condiciones suficientes para que exista y para que no exista extremo relativo de G en $\boldsymbol{b}$ y ello lo haremos acudiendo (véase [45]) a la diferencial segunda de G en $\boldsymbol{b}$ (nótese que la diferencial primera es nula, pues según [3] es $dG(\boldsymbol{b}) = dg(\boldsymbol{a}) \circ d\boldsymbol{\psi}(\boldsymbol{b}) = o$, ya que $dg(\boldsymbol{a}) = o$ por ser g la función de Lagrange); dicha diferencial segunda es

$$d^2G(\boldsymbol{b}) = d^2g(\boldsymbol{a}) \circ d\boldsymbol{\psi}(\boldsymbol{b}) + dg(\boldsymbol{a}) \circ d^2\boldsymbol{\psi}(\boldsymbol{b})$$

Ahora bien, como $dg(\boldsymbol{a}) = o$, la última relación se puede poner en la forma

$$d^2G(\boldsymbol{b})(\Delta\boldsymbol{y}) = d^2g(\boldsymbol{a})(d\boldsymbol{\psi}(\boldsymbol{b})(\Delta\boldsymbol{y})), \qquad \forall\ \Delta\boldsymbol{y} \in \mathbb{R}^{p-q} \qquad [5]$$

Así pues, los resultados 1.°, 2.° y 3.° del enunciado se deducen fácilmente de [5] y del anterior resultado [2], ya que:

1.° Si $d^2g(\boldsymbol{a})(\Delta\boldsymbol{x}) > 0$ para todo $\Delta\boldsymbol{x} \neq \boldsymbol{o}$ tal que $d\boldsymbol{\varphi}(\boldsymbol{a})(\Delta\boldsymbol{x}) = \boldsymbol{o}$, o sea, si (según [2]) $d^2g(\boldsymbol{a})(d\boldsymbol{\psi}(\boldsymbol{b})(\Delta\boldsymbol{y})) > 0$ para todo $\Delta\boldsymbol{y} \neq \boldsymbol{o}$ de $\mathbb{R}^{p-q}$, esto es, si (según [5]) $d^2G(\boldsymbol{b})$ es definida positiva, entonces (véase [45]) G tiene un mínimo relativo en $\boldsymbol{b}$, es decir, entonces (según [4]) f tiene un extremo relativo en $\boldsymbol{a}$ condicionado por $\boldsymbol{\varphi}(\boldsymbol{x}) = \boldsymbol{o}$.

2.° Para el resultado 2.° del enunciado se razona como en el caso anterior.

3.° Si $d^2g(\boldsymbol{a})(\Delta\boldsymbol{x}_1) > 0$ y $d^2g(\boldsymbol{a})(\Delta\boldsymbol{x}_2) < 0$ para ciertos $\Delta\boldsymbol{x}_1$ y $\Delta\boldsymbol{x}_2$ tales que $d\boldsymbol{\varphi}(\boldsymbol{a})(\Delta\boldsymbol{x}_1) = \boldsymbol{o}$ y $d\boldsymbol{\varphi}(\boldsymbol{a})(\Delta\boldsymbol{x}_2) = \boldsymbol{o}$, como (según [2]) existen $\Delta\boldsymbol{y}_1$, $\Delta\boldsymbol{y}_2 \in \mathbb{R}^{p-q}$ tales que

$$\Delta\boldsymbol{x}_1 = d\boldsymbol{\psi}(\boldsymbol{b})(\Delta\boldsymbol{y}_1) \qquad \text{e} \qquad \Delta\boldsymbol{x}_2 = d\boldsymbol{\psi}(\boldsymbol{b})(\Delta\boldsymbol{y}_2)$$

de [5] se desprende que, para ciertos $\Delta\boldsymbol{y}_1$ y $\Delta\boldsymbol{y}_2$ de $\mathbb{R}^{p-q}$, es

$$d^2G(\boldsymbol{b})(\Delta\boldsymbol{y}_1) > 0 \qquad \text{y} \qquad d^2G(\boldsymbol{b})(\Delta\boldsymbol{y}_2) < 0$$

Por tanto (véase [45]) G no tiene un extremo relativo en $\boldsymbol{b}$, es decir, entonces (según [4]) f no tiene un extremo relativo en $\boldsymbol{a}$ condicionado por $\boldsymbol{\varphi}(\boldsymbol{x}) = \boldsymbol{o}$.

[48]$_1$ Ejercicio

Hallar los máximos y mínimos relativos de

$$f(x, y, z) = z^2 - 2xy$$

cuando (x, y, z) recorre la superficie de ecuación (ligadura):

$$2x^3 + 2y^3 + z^3 = 4$$

Resolución

Construyamos la función de Lagrange, acudiendo a un parámetro (multiplicador de Lagrange) por haber una sola ligadura:

$$g(x, y, z) = f(x, y, z) + \lambda\varphi(x, y, z) = z^2 - 2xy + \lambda(2x^3 + 2y^3 + z^3 - 4)$$

Hallemos ahora los puntos estacionarios de g que cumplen la ligadura, es decir, las soluciones de $dg(x, y, z) = o$ y $\varphi(x, y, z) = 0$:

$$\left.\begin{aligned} g_x(x, y, z) &\equiv -2y + 6\lambda x^2 = 0 \\ g_y(x, y, z) &\equiv -2x + 6\lambda y^2 = 0 \\ g_z(x, y, z) &\equiv 2z + 3\lambda z^2 = 0 \\ \varphi(x, y, z) &\equiv 2x^3 + 2y^3 + z^3 - 4 = 0 \end{aligned}\right\} \quad \text{tiene por raíces a} \left\{\begin{aligned} &(x, y, z; \lambda) = (1, 1, 0; 1/3) \\ &(x, y, z; \lambda) = (-1, -1, 2; -1/3) \end{aligned}\right.$$

Para cada una de estas dos soluciones, vamos a recurrir a las condiciones suficientes de extremo (y de «no extremo») de [48], para lo que empezamos obteniendo

$$d^2g(x, y, z)(dx, dy, dz) = 12\lambda x\, dx^2 + 12\lambda y\, dy^2 + (2 + 6\lambda z)\, dz^2 - 4\, dx\, dy$$

$$d\varphi(x, y, z)(dx, dy, dz) = 6x^2\, dx + 6y^2\, dy + 3z^2\, dz$$

- Para la raíz $(x, y, z; \lambda) = (1, 1, 0; 1/3)$ es:

$$d\varphi(1, 1, 0)(dx, dy, dz) = 6dx + 6dy, \quad \text{que se anula si } dy = -dx \tag{1}$$

$$d^2g(1, 1, 0)(dx, dy, dz) = 4dx^2 + 4dy^2 + 2dz^2 - 4dx\, dy \tag{2}$$

Llevando (2) a (1), se obtiene que

$$d^2g(1, 1, 0)(dx, -dx, dz) = 12\, dx^2 + 2dz^2$$

que es definida positiva (como función de dx y dz), luego f tiene en el punto $(1, 1, 0)$ un mínimo relativo condicionado por $\varphi(x, y, z) = 0$.

- Para la raíz $(x, y, z; \lambda) = (-1, -1, 2; -1/3)$ es:

$$d\varphi(-1, -1, 2)(dx, dy, dz) = 6dx + 6dy + 12dz, \quad \text{que se anula si } dz = -\tfrac{1}{2}(dx + dy) \qquad (3)$$
$$d^2g(-1, -1, 2)(dx, dy, dz) = 4dx^2 + 4dy^2 - 2dz^2 - 4dx\,dy \qquad (4)$$

Llevando (4) a (3), se obtiene que

$$d^2g(-1, -1, 2)\big(dx, dy, -\tfrac{1}{2}(dx + dy)\big) = \tfrac{7}{2}dx^2 + \tfrac{7}{2}dy^2 - 5dx\,dy$$

que es definida positiva (como función de dx y dy), luego f tiene en el punto $(-1, -1, 2)$ otro mínimo relativo condicionado por $\varphi(x, y, z) = 0$.

RESUMEN: MÉTODO DE LOS MULTIPLICADORES DE LAGRANGE

[49]

Sean $f: C \to \mathbb{R}$ y $\varphi: C \to \mathbb{R}^q$ dos funciones de clase $\mathscr{C}^2$ en un abierto $C \subset \mathbb{R}^p$, siendo $q < p$, tal que rang $d\varphi(x) = q$ para $x \in C$. Se quieren hallar los extremos relativos de f condicionados por la ligadura $\varphi(x) = o$; para ello, construyamos la función de Lagrange $g = f + \lambda_1\varphi_1 + \cdots + \lambda_q\varphi_q$ (donde $\lambda_1, ..., \lambda_q \in \mathbb{R}$, que se llaman multiplicadores de Lagrange, y siendo $\varphi_1, ..., \varphi_q$ las componentes de φ) y procedamos del siguiente modo:

- Se resuelve el sistema de ecuaciones $dg(x) = o$ (p ecuaciones) y $\varphi(x) = o$ (q ecuaciones), es decir, se determinan los $x = (x_1, ..., x_p) \in C$ y los $\lambda_x = (\lambda_1, ..., \lambda_q) \in \mathbb{R}^q$ para los que

$$\frac{\partial f(x)}{\partial x_i} + \lambda_1 \frac{\partial \varphi_1(x)}{\partial x_i} + \cdots + \lambda_q \frac{\partial \varphi_q(x)}{\partial x_i} = 0 \ (i = 1, ..., p) \quad \text{y} \quad \varphi_j(x) = 0 \ (j = 1, ..., q) \quad [1]$$

- Para que en un punto $x = a$ haya extremo condicionado es necesario que $x = a$ y $\lambda_x = \lambda_a$ sean solución del anterior sistema [1]. Para analizar si la función f tiene, en un tal punto a (solución de [1]), un extremo condicionado por $\varphi(x) = o$, se determinan el espacio vectorial $V = \{\Delta x \in \mathbb{R}^p / d\varphi(a)(\Delta x) = o\}$, la diferencial segunda $d^2g(a)$ y la restricción $\omega = d^2g(a)|_V$ (que es una forma cuadrática) y se tiene que:

— Si ω es definida positiva, entonces f tiene en a un mínimo relativo condicionado por la ligadura $\varphi(x) = o$.
— Si ω es definida negativa, entonces f tiene en a un máximo relativo condicionado por la ligadura $\varphi(x) = o$.
— Si ω no es definida ni semidefinida entonces f no tiene en a un extremo relativo condicionado por la ligadura $\varphi(x) = o$.
(Si ω es semidefinida, este método no informa sobre si hay extremo o no lo hay.)

Comprobación

Las anteriores conclusiones no son más que un resumen de lo ya obtenido en los dos anteriores apartados [47] y [48].

[49]$_1$ Ejercicio

Hallar los extremos relativos de la función $f: \mathbb{R}^3 \to \mathbb{R}$ definida por

$$f(x, y, z) = xz + yz$$

condicionados por las dos ligaduras siguientes:

$$\begin{aligned} \varphi(x, y, z) &\equiv x^3 - y^3 - 4z - 4 = 0 \\ \psi(x, y, z) &\equiv x + y + z^2 - 3 = 0 \end{aligned}$$

Resolución

Siguiendo el método de los multiplicadores de Lagrange, construyamos la función (con dos parámetros λ y μ):

$$\begin{aligned} g(x, y, z) &= f(x, y, z) + \lambda\varphi(x, y, z) + \mu\psi(x, y, z) = \\ &= (xz + yz) + \lambda(x^3 - y^3 - 4z - 4) + \mu(x + y + z^2 - 3) \end{aligned}$$

y hallemos los puntos estacionarios de g que cumplen las ligaduras:

$$\left.\begin{aligned} g_x(x, y, z) &\equiv z + 3\lambda x^2 + \mu = 0 \\ g_y(x, y, z) &\equiv z - 3\lambda y^2 + \mu = 0 \\ g_z(x, y, z) &\equiv x + y + 4\lambda + 2\mu z = 0 \\ \varphi(x, y, z) &\equiv x^3 - y^3 - 4z - 4 = 0 \\ \psi(x, y, z) &\equiv x + y + z^2 - 3 = 0 \end{aligned}\right\} \quad \begin{array}{l}\text{tiene una única} \\ \text{solución, que es:}\end{array} (x, y, z; \lambda, \mu) = (2, 0, 1; 0, -1)$$

Vayamos ahora a las condiciones suficientes, para lo que hallamos

$$d^2g(2, 0, 1)(dx, dy, dz) = -2dz^2 + 2dx\,dz + 2dy\,dz \tag{1}$$

$$\left.\begin{aligned} d\varphi(2, 0, 1)(dx, dy, dz) &= 12dx - 4dz \\ d\psi(2, 0, 1)(dx, dy, dz) &= dx + dy + 2dz \end{aligned}\right\} \quad \text{se anulan si} \begin{cases} dz = 3dx \\ dy = -7dx \end{cases} \tag{2}$$

Llevando (2) a (1), se obtiene que

$$d^2g(2, 0, 1)(dx, -7dx, 3dx) = -54dx^2$$

que es definida negativa (como función de dx), luego f tiene un máximo relativo en (2, 0, 1) condicionado por las dos ligaduras dadas.

[49]$_2$ Ejercicio

Considérense el círculo cerrado C y la función $f: C \to \mathbb{R}$ siguientes:

$$C = \{(x, y) \in \mathbb{R}^2 / x^2 + y^2 \leqslant 8\}$$
$$(x, y) \mapsto f(x, y) = x^3 + y^3 + 3xy$$

Hallar los extremos absolutos de f en C.

Resolución

Si hallamos los extremos relativos de f en el interior de C (problema que sabemos abordar, pues dicho interior es un abierto), si hallamos los extremos relativos de f en la frontera (que son extremos condicionados por la ligadura $x^2 + y^2 = 8$) y, luego, hallamos el mayor de todos los máximos relativos y el menor de todos los mínimos relativos, dichos mayor y menor(*) son el máximo y el mínimo absolutos buscados, que existen según el teorema de Weierstrass (véase [18]), pues f es continua y C es compacto.

1.º Extremos relativos de f en $\mathring{C}$, interior de C. Procediendo como se indica en [44] y [45], empecemos hallando los puntos estacionarios de f:

$$\left.\begin{array}{l} f_x(x, y) \equiv 3x^2 + 3y = 0 \\ f_y(x, y) \equiv 3y^2 + 3x = 0 \end{array}\right\} \quad \text{tiene por soluciones a} \left\{\begin{array}{l} (x, y) = (0, 0) \\ (x, y) = (-1, -1) \end{array}\right.$$

Para cada una de ellas, hallamos $d^2 f$ y se obtiene

$$d^2 f(0, 0)(dx, dy) = 6dx\,dy \qquad \text{(ni definida ni semidefinida)}$$
$$d^2 f(-1, -1)(dx, dy) = -6dx^2 + 6dx\,dy - 6dy^2 \qquad \text{(definida negativa)}$$

por lo que en $(-1, -1)$ hay un máximo relativo, $f(-1, -1) = 1$, y no hay ningún otro extremo de f en $\mathring{C}$.

2.º Extremos relativos de f condicionados por la ligadura $\varphi(x, y) \equiv x^2 + y^2 = 8$. Recurriendo a la función de Lagrange

$$g(x, y) = f(x, y) + \lambda\varphi(x, y) = x^3 + y^3 + 3xy + \lambda(x^2 + y^2 - 8)$$

y hallando los puntos estacionarios de g que cumplen la ligadura, se tiene:

$$\left.\begin{array}{l} g_x(x, y) \equiv 3x^2 + 3y + 2\lambda x = 0 \\ g_y(x, y) \equiv 3y^2 + 3x + 2\lambda y = 0 \\ \varphi(x, y) \equiv x^2 + y^2 - 8 = 0 \end{array}\right\} \quad \begin{array}{l}\text{cuyas}\\ \text{soluciones}\\ \text{son}\end{array} \left\{\begin{array}{l} (x, y; \lambda) = (2, 2; -9/4) \\ (x, y; \lambda) = (-2, -2; 3/4) \\ (x, y; \lambda) = (-1 + \sqrt{3}, -1 - \sqrt{3}; 9/2) \\ (x, y; \lambda) = (-1 - \sqrt{3}, -1 + \sqrt{3}; 9/2) \end{array}\right.$$

(*) Aunque haremos así las cosas, para ejercitarnos, basta con hallar el mayor y el menor de entre todos los puntos estacionarios (en el interior de C y en la frontera) el mayor de todos ellos es el máximo absoluto de f en C y el menor es el mínimo absoluto.

Para hallar los extremos condicionados acudimos a d^2g, a $d\varphi$, al subespacio V en el que se anula $d\varphi$ y a $d^2g|_V$. Como

$$d^2g(x, y)(dx, dy) = (6x + 2\lambda)\,dx^2 + 6dx\,dy + (6y + 2\lambda)\,dy^2$$
$$d\varphi(x, y)(dx, dy) = 2x\,dx + 2y\,dy$$

resulta fácil obtener que $d^2g|_V$ es definida negativa en (2, 2) y en (−2, −2) y definida positiva en $(-1 + \sqrt{3}, -1 - \sqrt{3})$ y en $(-1 - \sqrt{3}, -1 + \sqrt{3})$, por lo que f alcanza máximos relativos condicionados en (2, 2) y (−2, −2) y mínimos relativos condicionados en $(-1 + \sqrt{3}, -1 - \sqrt{3})$ y $(-1 - \sqrt{3}, -1 + \sqrt{3})$.

3.° Como resultado de todo lo dicho, se puede concluir diciendo que

$$\text{máx } f(C) = \text{máx } \{f(-1, -1), f(2, 2), f(-2, -2)\} = f(2, 2) = 28$$
$$\text{mín } f(C) = \text{mín } \{f(-1 + \sqrt{3}, -1 - \sqrt{3}), f(-1 - \sqrt{3}, -1 + \sqrt{3})\} = -26$$

Ejercicios y problemas

ENUNCIADOS

3.1. Considérese el sistema de ecuaciones

$$\begin{cases} xu + \operatorname{sen} v = \pi/2 \\ yv + \cos u = 0 \end{cases}$$

Analícese si él define implícitamente a dos funciones $u = u(x, y)$ y $v = v(x, y)$ de clase $\mathscr{C}^1$ en un cierto entorno de $(x, y) = (1, 1)$ y tales que $u(1, 1) = \pi/2$, $v(1, 1) = 0$. Hállense expresiones para du y dv.

3.2. Compruébese si existen dos funciones $u(x, y)$ y $v(x, y)$ de clase $\mathscr{C}^2$ definidas localmente por:

$$\begin{cases} xy(u + v) + (x + y)uv = -2 & u(1, 1) = 1 \\ x + y + u + v = 2 & v(1, 1) = -1 \end{cases}$$

Hallar $du(1, 1)$, $dv(1, 1)$, $d^2u(1, 1)$ y $d^2v(1, 1)$.

3.3. Sea $z = e^x \cos y$, donde x e y son las funciones de t definidas implícitamente por las ecuaciones

$$x^3 + e^x = t^2 + t + 1, \qquad yt^2 + y^2t + y - t = 0$$

y tales que $x(0) = 0$ e $y(0) = 1$. Hallar la derivada dz/dt en $t = 0$.

3.4. Sean $x(u, v)$, $y(u, v)$ y $z(u, v)$ las funciones definidas implícitamente por

$$\begin{cases} x^2 + y^2 + 2z^2 - \operatorname{sen}(2u + 3v) = 2 & x(\pi/2, 0) = 1 \\ x^2 + z^2 - y\cos(uv) = 0 & y(\pi/2, 0) = 1 \\ xy + z - \operatorname{sen} u \cos v = 0 & z(\pi/2, 0) = 0 \end{cases}$$

Hallar dx, dy y dz en el punto $(u, v) = (\pi/2, 0)$.

3.5. Si $u(x, y)$ y $v(x, y)$ son funciones de clase $\mathscr{C}^1$ definidas por el sistema de ecuaciones

$$x + y + u + v = a, \qquad x^2 + y^2 + u^2 + v^2 = b^2$$

Obtener expresiones que proporcionen las diferenciales de u y v.

3.6. Supóngase que $(x, y, z) \mapsto f(x, y, z)$ es una función de clase $\mathscr{C}^1$ en un entorno de cierto punto (x, y, z), en el que sus derivadas parciales no son nulas, por lo que la ecuación $f(x, y, z) = 0$ define implícitamente a tres funciones $z = z(x, y)$, $x = x(y, z)$, $y = y(z, x)$, entorno del punto dado. Hallar

$$\frac{\partial z(x, y)}{\partial x}\,\frac{\partial x(y, z)}{\partial y}\,\frac{\partial y(z, x)}{\partial z}$$

3.7. Calcular las derivadas primera y segunda de una función $y = y(x)$ definida implícitamente por la ecuación $f(x, y) = 0$, en los casos:

1.° $f(x, y) = \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} - 1$

2.° $f(x, y) = \mathrm{L}\,\dfrac{\sqrt{x^2 + y^2}}{a} - \operatorname{arctg}\dfrac{y}{x}$

3.8. Sean $u(x, y)$ y $v(x, y)$ funciones definidas implícitamente por el sistema de ecuaciones

$$xy + uv = 1, \qquad xu + yv = 1$$

Hallar $\partial^2 u/\partial x^2$.

3.9. Si $u(x, y)$ y $v(x, y)$ son las funciones definidas implícitamente por el sistema de ecuaciones

$$\begin{cases} x + y + u + v = a \\ x^3 + y^3 + u^3 + v^3 = a^3 \end{cases} \quad \text{con} \quad \begin{cases} u(a, -a) = 0 \\ v(a, -a) = a \end{cases}$$

Hallar las diferenciales primeras y segundas de u y v en el punto $(x, y) = (a, -a)$.

3.10. Si es $z(x, y)$ la función definida implícitamente por

$$z^3 + xz + y = 0, \qquad z(-2, 1) = 1$$

Hallar $dz(-2, 1)$ y $d^2z(-2, 1)$.

3.11. Sean $u(x, y)$ y $v(x, y)$ dos funciones de clase $\mathscr{C}^2$ definidas implícitamente por el sistema de ecuaciones

$$x^2 - xu - v^2 = 0, \qquad y^2 - yv - u^2 = 0$$

Hallar expresiones para du, dv, d^2u, d^2v.

3.12. Sean $u(x, y)$ y $v(x, y)$ las funciones definidas implícitamente por:

$$\begin{cases} x^2 - y^2 + u^2 + 2v^2 = 5 & u(0, 1) = 2 \\ x^2 + y^2 - u^2 - v^2 = -4 & v(0, 1) = -1 \end{cases}$$

Hallar du, dv, d^2u y d^2v en el punto $(x, y) = (0, 1)$.

3.13. Sea $f: \mathbb{R}^2 \to \mathbb{R}$, $(u, v) \mapsto f(u, v)$ una función de clase $\mathscr{C}^1$ tal que la ecuación $f(x^2 - y^2, y^2 - z^2) = 0$ define implícitamente a cierta función $z = z(x, y)$. Calcúlese la expresión:

$$E = yz\frac{\partial z}{\partial x} + zx\frac{\partial z}{\partial y}$$

comprobando que no depende de f.

3.14. Considérese la transformación $(u, v) \mapsto (x, y)$ definida por $x = ue^v$ e $y = ve^u$. Sea $(x, y) \mapsto (u, v)$ la transformación inversa de la anterior tal que $u(1, 0) = 1$ y $v(1, 0) = 0$. Hallar las diferenciales primeras y segundas de $u(x, y)$ y $v(x, y)$ en el punto $(x, y) = (1, 0)$.

3.15. Considérese la transformación $(u, v) \mapsto (x, y)$ definida por $x = u^v$ e $y = v^u$. Sea $(x, y) \mapsto (u, v)$ la transformación inversa de la anterior tal que $u(1, 2) = 1$ y $v(1, 2) = 2$. Hallar $du(1, 2)$ y $dv(1, 2)$.

3.16. Sea $(\rho, \varphi, \theta) \mapsto (x, y, z)$ la transformación que pasa de coordenadas esféricas a cartesianas rectangulares, esto es:

$$\left.\begin{aligned} x &= \rho \operatorname{sen} \varphi \cos \theta \\ y &= \rho \operatorname{sen} \varphi \operatorname{sen} \theta \\ z &= \rho \cos \varphi \end{aligned}\right\}$$

Hallar el jacobiano J de (x, y, z) respecto de (ρ, φ, θ) y las derivadas parciales de las funciones $\rho(x, y, z)$, $\varphi(x, y, z)$ y $\theta(x, y, z)$ que determina la transformación inversa (calcúlense solamente $\partial\rho/\partial y$, $\partial\varphi/\partial z$ y $\partial\theta/\partial x$).

3.17. Sean $u = u(x, y, z)$, $v = v(x, y, z)$ y $w = w(x, y, z)$ tres funciones dadas, de clase $\mathscr{C}^1$ en un entorno de un punto (x_0, y_0, z_0), en el que es no nulo el jacobiano J de (u, v, w) respecto de (x, y, z). Si la inversa de la función $(x, y, z) \mapsto (u, v, w)$, entorno de (x_0, y_0, z_0), es $x = x(u, v, w)$, $y = y(u, v, w)$, $z = z(u, v, w)$, hallar las derivadas parciales de estas funciones en función de las derivadas parciales de las funciones dadas.

3.18. Para x, y, z positivos, considérense las siguientes u, v y w:

$$u = yzf(x), \qquad v = \mathrm{L}\,(x^2y), \qquad w = \mathrm{L}\,z$$

donde f es una función positiva de clase $\mathscr{C}^1$. Hallar la expresión más general de f para la que entre u, v y w existe una relación funcional.

3.19. Hallar las relaciones que deben cumplirse entre a, b, c, d y e para que las siguientes funciones f y g sean dependientes:

$$f(x, y) = e^{x+2y}, \; g(x, y) = ax^2 + 2bxy + cy^2 + dx + ey$$

3.20. Analícese si las tres siguientes funciones f, g y h pueden ser dependientes, en algún conjunto abierto de $\mathbb{R}^3$, para algunos valores no nulos de los parámetros a, b y c:

$$f(x, y, z) = ax + by + cz, \quad g(x, y, z) = x^2 + y^2 + z^2, \quad h(x, y, z) = x^3 + y^3 + z^3$$

3.21. Sea $y = f(x)$ una curva plana de clase $\mathscr{C}^2$. El radio de curvatura de la curva, en el punto $(x, f(x))$ es:

$$\rho(x) = [1 + y'(x)]^{3/2} : y''(x)$$

Hallar la expresión del radio de curvatura cuando la curva se conoce a través de sus ecuaciones paramétricas $x = \varphi(t)$, $y = \psi(z)$.

3.22. La epicicloide es la curva $y = y(x)$ que admite por ecuaciones paramétricas a las:

$$x(t) = (R + r)\cos t - r\cos\left(\frac{R + r}{r}\,t\right)$$
$$y(t) = (R + r)\operatorname{sen} t - r\operatorname{sen}\left(\frac{R + r}{r}\,t\right)$$

Hallar $y'(x)$ e $y''(x)$, en función de t.

3.23. En la siguiente expresión diferencial E, cambiar la función $y(x)$ por la $z(x)$ que se indica ($a, b, c, d \in \mathbb{R}$ dados):

$$E = \frac{y'''}{y'} - \frac{3}{2}\left(\frac{y''}{y'}\right)^2, \qquad z(x) = \frac{ay(x) + b}{cy(x) + d}$$

3.24. Cambiar y por z, siendo $y = \operatorname{tg} z$, en la ecuación diferencial:

$$(1 + y^2)\frac{d^2y}{dx^2} - 2y\left(\frac{dy}{dx}\right)^2 - 2(1 + y^2)\frac{dy}{dx} = 0$$

3.25. En la siguiente ecuación diferencial, cámbiese la función y por la nueva función $z = ye^{-x^2}$ (la x sigue siendo la variable independiente):

$$\frac{d^2y}{dx^2} - 4x\frac{dy}{dx} + (3x^2 - 2)y = 0$$

3.26. En la siguiente ecuación diferencial hágase el siguiente cambio de variables: nueva variable independiente y; nueva función x:

$$\frac{d^2y}{dx^2} + \left(\frac{dy}{dx}\right)^3 = 0$$

3.27. En la siguiente expresión E realizar el cambio de variable $(x, y) \mapsto (u, v)$ que se indica

$$E \equiv z''_{xx} + xz''_{xy} + \left(\frac{1}{2}x^2 - y\right)z''_{yy} \qquad \begin{cases} x = u + v \\ y = uv \end{cases}$$

3.28. Sea $z = f(x, y)$ una función de clase $\mathscr{C}^1$. En las siguientes expresiones diferenciales E, realizar los cambios de variables que se indican:

1.° $E = z_x'^2 + z_y'^2$, cambio $\begin{cases} x = \rho\cos\theta \\ y = \rho\operatorname{sen}\theta \end{cases}$

2.° $E = z_x'^2 + z_y'^2$, cambio $\begin{cases} x = \rho\operatorname{Ch}\theta \\ y = \rho\operatorname{Sh}\theta \end{cases}$

3.29. En la siguiente ecuación en derivadas parciales, cambiar las variables independientes x e y por las $u = 2x - 3y$ y $v = 3x + 2y$:

$$3\frac{\partial z}{\partial x} + z\frac{\partial z}{\partial y} = 0$$

3.30. Sabiendo que $z = z(x, y)$ es una función de clase $\mathscr{C}^2$ que define implícitamente a x como función $x = x(y, z)$, de clase $\mathscr{C}^2$, expresar $\partial^2 z/\partial x\partial y$ en función de las derivadas parciales de $x(y, z)$.

3.31. En la ecuación diferencial $xz'_x + yz'_y = 0$ realizar el cambio de x, y, z por nuevas variables u, v y nueva función w, siendo

$$x = vw \qquad y = wv \qquad z = uv$$

3.32. En la siguiente expresión E, cambiar las variables x e y por las $u = x - y$ y $v = z$ y tomar como nueva función a $w = x + y$

$$E = z''_{xx} + z''_{yy}$$

3.33. Hallar los puntos estacionarios de las siguientes funciones y analizar si en ellos se alcanza un extremo relativo:

1.° $f(x, y) = 2x^2 + y^2 + xy + 5$
2.° $f(x, y) = x^2 + 2y^2 + 3xy - x - 2y + 2$
3.° $f(x, y) = x^2 + 4y^2 - 4xy + 2x - 4y + 3$

3.34. Hallar los extremos relativos de $f(x, y)$, para $\varepsilon = 1$, para $\varepsilon = -1$ y para $\varepsilon = 0$:

$$f(xy) = e^{-x^2 + \varepsilon y^2}$$

3.35. Hallar los máximos y mínimos relativos de la función $f: \mathbb{R}^2 \to \mathbb{R}$ siguiente:

$$f(x, y) = 12x^2 + 12y^2 - x^3y^3 + 5$$

3.36. Hallar los puntos críticos de

$$f(x, y) = (x^2 + y^2)e^{x^2 - y^2}$$

y obtener sus extremos relativos.

3.37. Hallar los puntos críticos y estudiar si en cada uno de ellos hay o no extremo relativo de la función $(x, y) \mapsto z$ siguiente:

$$z = x^8 + y^6 - 2x^4 - 3y^2$$

3.38. Hallar los valores máximos y mínimos relativos que alcanza $f(x, y)$ cuando (x, y) recorre la curva $\varphi(x, y) = 0$ siendo:

$$f(x, y) = x^2 - 4xy + y^2 + 20x + 20y + 10$$
$$\varphi(x, y) = x^2 + y^2 + xy - 12$$

3.39. Hallar los valores máximo y mínimo de

$$f(x, y) = xy + yz + zx + x + y + z$$

cuando (x, y, z) recorre la esfera $x^2 + y^2 + z^2 = 1$.

3.40. Sea $u:\mathbb{R}^3 \to \mathbb{R}$ la función definida por la siguiente expresión, en la que $a > 1$, $b > 1$ y $c > 1$ son dados:

$$u = a^x b^y c^z$$

Hallar los valores máximo y mínimo de u cuando (x, y, z) recorre la esfera de ecuación $x^2 + y^2 + z^2 = r^2$.

3.41. Se considera el elipsoide, de semiejes a, b y c:

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

Se inscribe en el elipsoide una pirámide de base rectangular perpendicular al eje z. Hallar el volumen máximo de una tal pirámide.

3.42. Se consideran todas las pirámides de base cuadrada cuyas aristas tienen, entre las ocho, una longitud constante e igual a k. Hallar, de entre ellas, la pirámide de volumen máximo.

3.43. Hallar las distancias máxima y mínima de un punto genérico de la siguiente curva γ al origen de coordenadas:

$$\gamma \begin{cases} x^2 + 2y^2 + 3z^2 = 1 \\ 2x^2 + 3y^2 + z^2 = 1 \end{cases}$$

3.44. Hallar las distancias máxima y mínima de un punto genérico de la siguiente curva γ al plano $x + y + z = 10$

$$\gamma \begin{cases} x^2 + 2y^2 + 3z^2 = 1 \\ 2x^2 + y - z = 0 \end{cases}$$

3.45. Se considera el elipsoide $(x/a)^2 + (y/b)^2 + (z/c)^2 = 1$. Hallar un punto del elipsoide situado en el primer octante tal que el tetraedro que el plano tangente en él, al elipsoide, determina con los planos coordenados tenga volumen mínimo y hallar éste.

3.46. Sea $f: C \to \mathbb{R}$ la función definida en el círculo $C: x^2 + y^2 \leqslant 4$ mediante $f(x, y) = 4x^2 + 9y^2 - x^2y^2$.

1.° Hallar los extremos relativos de f en $\mathring{C}: x^2 + y^2 < 4$.
2.° Hallar los extremos relativos de f en $\partial C: x^2 + y^2 = 4$.
3.° Hallar el máximo y el mínimo absolutos de f en C.

3.47. Sea $f: C \to \mathbb{R}$ la función definida en el círculo $C: x^2 + y^2 \leqslant 4$ mediante

$$f(x, y) = 3x^2y^2 + 2x^3 + 2y^3$$

Hallar los valores máximo y mínimo de f en C.

SOLUCIONES

3.1. Sí existe, pues el jacobiano vale $2 \neq 0$.

$$du(x, y)(dx, dy) = \frac{-uy\,dx + v \cos v\,dy}{xy + \operatorname{sen} u \cos v}$$

$$dv(x, y)(dx, dy) = \frac{-u \operatorname{sen} u\,dx - xv\,dy}{xy + \operatorname{sen} u \cos v}$$

3.2. El jacobiano es $-4 \neq 0$, luego existen

$$du(1, 1)(dx, dy) = -dx - dy, \quad dv(1, 1)(dx, dy) = 0$$
$$d^2u(1, 1)(dx, dy) = -dx\,dy, \quad d^2v(1, 1)(dx, dy) = dx\,dy$$

3.3. 1

3.4. $dx(\pi/2, 0)(du, dv) = -\frac{1}{3}du - \frac{1}{2}dv$
$dy(\pi/2, 0)(du, dv) = -\frac{2}{3}du - dv$
$dz(\pi/2, 0)(du, dv) = du + \frac{3}{2}dv$

3.5. $$du(x, y)(dx, dy) = \frac{x - v}{v - u}\,dx + \frac{y - v}{v - u}\,dy$$
$$dv(x, y)(dx, dy) = \frac{x - u}{v - u}\,dx + \frac{y - u}{v - u}\,dy$$

3.6. -1

3.7. 1.° $y'(x) = -\dfrac{b^2x}{a^2y}, \quad y''(x) = -\dfrac{b^4x^2 + a^2b^2y^2}{a^4y^3}$

2.° $y'(x) = \dfrac{x + y}{x - y}, \quad y''(x) = \dfrac{2x^2 + 2y^2}{(x - y)^3}$

3.8. $$\frac{\partial^2 u}{\partial x^2} = 2\,\frac{u^2 - y^2}{(1 - 2ux)^3}\,(2u - 3u^2x - xy^2)$$

3.9. $du(a, -a)(dx, dy) = 0$
$dv(a, -a)(dx, dy) = -dx - dy$
$d^2u(a, -a)(dx, dy) = (4/a)\,dx^2$
$d^2v(a, -a)(dx, dy) = -(4/a)\,dx^2$

3.10. $dz(-2, 1)(dx, dy) = -dx - dy$
$d^2z(-2, 1)(dx, dy) = -4dx^2 - 10dx\,dy - 6dy^2$

3.11. $$du = [(2x - u)y\,dx - (2y - v)v\,dy]\,\frac{1}{J}$$
$$dv = [(2y - v)x\,dx - (2x - u)u\,dy]\,\frac{1}{J}$$
$$d^2u = \frac{2}{J}\,[y(dx^2 - dx\,du - dv^2) - 2v(dy^2 - dy\,dv - du^2)]$$
$$d^2v = \frac{2}{J}\,[x(dy^2 - dy\,dv - du^2) - 2u(dx^2 - dx\,du - dv^2)]$$
siendo $J = xy - 4uv$

3.12. $du(0, 1)(dx, dy) = \frac{1}{2}dy; \quad dv(0, 1)(dx, dy) = 0$
$d^2u(0, 1)(dx, dy) = 6dx^2 + \frac{3}{2}dy^2$
$d^2v(0, 1)(dx, dy) = 2dx^2$

3.13. $xf'_u - zz'_x f'_v = 0, \; -yf'_u + (y - zz'_y)f'_v = 0; \; E = xy$

3.14. $du(1, 0)(dx, dy) = dx - dy$
$dv(1, 0)(dx, dy) = dy$
$d^2u(1, 0)(dx, dy) = -2dx\,dy + 2dy^2$
$d^2v(1, 0)(dx, dy) = -dy^2$

3.15. $du(1, 2)(dx, dy) = (1/2)\,dx$
$dv(1, 2)(dx, dy) = -\mathrm{L}\,2dx + dy$

3.16. $J = \rho^2 \operatorname{sen} \varphi; \; \partial\rho/\partial y = \operatorname{sen} \varphi \operatorname{sen} \theta$
$\partial\varphi/\partial z = -\operatorname{sen} \varphi/\rho, \; \partial\theta/\partial x = -\operatorname{sen} \theta/(\rho \operatorname{sen} \varphi)$

3.17. $$\frac{\partial x}{\partial u} = \frac{1}{J}\,\frac{\partial(v, w)}{\partial(y, z)}, \quad \frac{\partial x}{\partial v} = \frac{1}{J}\,\frac{\partial(w, u)}{\partial(y, z)}, \quad \frac{\partial x}{\partial w} = \frac{1}{J}\,\frac{\partial(u, v)}{\partial(y, z)}$$
(análogo resultado para las derivadas de y y de z).

3.18. $$\left|\frac{\partial(u, v, w)}{\partial(x, y, z)}\right| = f'(x) - \frac{2}{x}\,f(x) \equiv 0; \quad \frac{f'(x)}{f(x)} \equiv 2x;$$
$\mathrm{L}\,f(x) \equiv \mathrm{L}\,x^2 + k; \; f(x) = cx^2$

3.19. $c = 2b = 4a$ y $e = 2d$

3.20. No; el jacobiano es $cxy(x-y)+ayz(y-z)+$ $+bxz(z-x)$, que no se anula en ningún abierto de $\mathbb{R}^3$.

3.21. $\rho(t)=\dfrac{[\varphi'(t)^2+\psi'(t)^2]^{3/2}}{[\varphi'(t)\psi''(t)-\varphi''(t)\psi'(t)]}$

3.22. $y'(x)=\operatorname{tg}\left(\dfrac{R+2r}{2r}\,t\right)$

$$y''(x)=\frac{R+2r}{4r(R+r)}\,\frac{1}{\operatorname{sen}\left(t\,\dfrac{R}{2r}\right)\cos^3\left(t\,\dfrac{R+2r}{2r}\right)}$$

3.23. $\dfrac{z'''}{z'}-\dfrac{3}{2}\left(\dfrac{z''}{z'}\right)^2$ (es invariante)

3.24. $\dfrac{d^2z}{dx^2}-2\,\dfrac{dz}{dx}=0$

3.25. $\dfrac{d^2z}{dx^2}-x^2z=0$

3.26. $\dfrac{d^2x}{dy^2}=1$

3.27. $\dfrac{1}{2}\,(z''_{uu}+z''_{vv})$

3.28. 1.º $E=\left(\dfrac{\partial z}{\partial\rho}\right)^2+\dfrac{1}{\rho^2}\left(\dfrac{\partial z}{\partial\theta}\right)^2$

2.º $E=\left(\dfrac{\partial z}{\partial u}\right)^2-\dfrac{1}{u^2}\left(\dfrac{\partial z}{\partial v}\right)^2$

3.29. $\partial z/\partial v=0$

3.30. $z''_{xy}=\dfrac{1}{(x'_z)^3}\,(x'_yx''_{zz}-x'_zx''_{yz})$

3.31. $uw'_u+vw'_v=w$

3.32. $\dfrac{1}{w'^3_v}\,[-2w'^2_vw''_{uu}+4w'_uw''_{uv}-2(1+w'^2_u)w''_{vv}]$

3.33. 1.º (0, 0); mínimo.
2.º (2, −1); puerto.
3.º Todos los de la recta $x-2y+1=0$; mínimos en sentido amplio.

3.34. Punto crítico (0, 0): si $\varepsilon<1$, máximo; si $\varepsilon>1$, puerto; si $\varepsilon=0$ máximo amplio a lo largo de $x=0$.

3.35. En (0, 0) mínimo, $f(0,0)=5$.

3.36. (0, 0), mínimo; (0, 1) y (0, −1), puertos.

3.37. En (1, 1) mínimo; en (1, 0) máximo; en (0, 1) no extremo; en (0, 0) máximo.

3.38. Máximos $f(4,-2)=102$ y $f(-2,4)=102$; mínimos $f(2,2)=82$ y $f(-2,-2)=-78$.

3.39. Puntos críticos $P(1/\sqrt{3},\ 1/\sqrt{3},\ 1/\sqrt{3})$, $Q(-1/\sqrt{3},\ -1/\sqrt{3},\ -1/\sqrt{3})$ y circunferencia $C: x+y+z+1=0,\quad x^2+y^2+z^2=1$. Máximo $f(Q)=7+\sqrt{3}$; mínimo $f(C)=5$.

3.40. Tomando logaritmos: máximo en $(x,y,z)=$ $=(h\,\mathrm{L}\,a,\ h\,\mathrm{L}\,b,\ h\,\mathrm{L}\,c)$ y mínimo en $(x,y,z)=(-h\,\mathrm{L}\,a,\ -h\,\mathrm{L}\,b,\ -h\,\mathrm{L}\,c)$, donde $h=r[(\mathrm{L}\,a)^2+(\mathrm{L}\,b)^2+(\mathrm{L}\,c)^2]^{-1/2}$. Máximo e^{r^2h}; mínimo e^{-r^2h}.

3.41. Vértice en $(0,0,c)$, base en $z=$ constante; vol $=(4/3)xy(c-z)$ con la ligadura $(x/a)^2+$ $+(y/b)^2+(z/c)^2=1$. Tomando logaritmos, $x=\sqrt{3}a/2$, $y=\sqrt{3}b/2$, $z=-c/2$; vol máximo $=(3/2)abc$.

3.42. Aristas de la base: $(5-\sqrt{13})k/12$; aristas laterales: $(\sqrt{13}-2)k/12$.

3.43. Extremos de $x^2+y^2+z^2$ con las dos ligaduras dadas: puntos críticos

$$(x,y,z;\lambda,\mu)=(0,\ \pm\sqrt{2/7},\ \pm\sqrt{1/7};\ -2/7,\ -1/7$$
$$(x,y,z;\lambda,\mu)=(\pm\sqrt{2/5},\ 0,\ \pm\sqrt{1/5};\ -1/5,\ -2/5)$$

Máximo $\sqrt{3/5}$, mínimo $\sqrt{3/7}$.

3.44. Extremos de $(10 - x - y - z)/\sqrt{3}$ con las dos ligaduras dadas. Máximo $\left(8 + 5/\sqrt{29}\right)/\sqrt{3}$; mínimo $\left(8 - 5/\sqrt{29}\right)/\sqrt{3}$.

3.45. Punto de tangencia (x, y, z), volumen $V(x, y, z) = (abc)^2/(6xyz)$; ligadura $(x/a)^2 + (y/b)^2 + (z/c)^2 = 1$. Punto buscado $\left(a/\sqrt{3},\ b/\sqrt{3},\ c/\sqrt{3}\right)$, volumen mínimo $= \left(\sqrt{3}/2\right)abc$.

3.46. 1.° En $(0, 0)$ mínimo. 2.° En $(0, 2)$ y $(0, -2)$ máximo, en $(2, 0)$ y $(-2, 0)$ mínimo. 3.° Máximo $f(0, 2) = f(0, -2) = 36$, mínimo $f(0, 0) = 0$.

3.47. Puntos críticos en $\mathring{C}$: $(0, 0)$ y $(1, -1)$; puntos críticos a lo largo de ∂D: $(0, \pm 2)$, $(\pm 2, 0)$, $\left(\sqrt{2}; \sqrt{2}\right)$, $\left(-\sqrt{2}, -\sqrt{2}\right)$. Máximo $f\left(\sqrt{2}, \sqrt{2}\right) = 12 + 8\sqrt{2}$; mínimo $f(0, -2) = f(-2, 0) = -16$.

El autor ha descubierto, al descifrar un nuevo manuscrito que llegó a sus manos por azar, que la integración es cosa que, viniendo de más antiguo de lo que él pensaba, fue útil a gentes de épocas muy remotas. El autor se goza en ello y quiere hacer partícipe al lector de su gozo, por lo que le ofrece aquí el referido manuscrito.

FAVORECER A LOS PANIAGUADOS PRODUCE DESAFUEROS(*)

En medio del desierto de Anera-Ed Notra-Ek, en lo más profundo de los desfiladeros que, en la cordillera Sadrog Sacor, han horadado los arroyos Ortil-Oidem y Sajomolos, allí por donde se veían obligadas a pasar las caravanas, en estos lugares se encuentran los restos de la, otrora, famosa ciudad de Aertep-Ebrú. Actualmente, de Aertep-Ebrú sólo quedan en pie las ruinas de unos soberbios templos, de unas majestuosas tumbas, del Gran Aljibe, y poco más; todo ello excavado en la roca. Pero esta ciudad tuvo momentos de gran esplendor bajo el dominio de Omaleyos III.

Las aguas que se obtenían de los arroyos Ortil-Oidem y Sajomolos eran las únicas de que se disponía para abastecer a la ciudad. Como quiera que estos dos arroyos eran de muy escaso caudal, hubo que construir el Gran Aljibe y establecer unas rígidas normas para el reparto del agua; de ello se encargaba el Muy Ilustre Regidor de los Asuntos de las Aguas, que en la época que nos ocupa era Nordal, yerno del rey Omaleyos III y protegido de éste. Nordal, que era persona de pocos escrúpulos, se aprovechaba de su cargo amenguando el agua de unos, con disimulo, y aumentando el agua de otros, percibiendo dineros por ello. Los abusos de Nordal fueron tantos que los Comisionados de los tres barrios de la ciudad, Roh-Clem, Rap-Sag y Rasat-Lab, delegaron en Roh-Clem, el de mayor edad y experiencia, para que acudiera al Muy Ilustre Regidor de los Asuntos de la Justicia con el encargo de que pusiera remedio a la situación. Ymed-Etaif, que así se llamaba este regidor, sabiendo que el rey Omaleyos se inclinaba en favor de Nordal y queriendo actuar rectamente, pidió a Roh-Clem que le proporcionara pruebas de lo que decía y le exigió que guardase secreto de tan delicado asunto. Roh-Clem disponía de muchos testimonios y comprobaciones pero, para que sus alegatos fuesen concluyentes, le faltaba un dato importantísimo: la capacidad del Gran Aljibe de la ciudad.

El Gran Aljibe se había excavado en la roca; tenía planta rectangular, de veinte codos de largo por doce codos de ancho; sus paredes laterales eran cuatro despeñaderos verticales. El fondo era irregular y su forma no se alcanzaba a ver, pues el lugar estaba en penumbra y era celosamente guardado por los Servidores del Aljibe, a las órdenes directas de Nordal.

Antes de que comenzara la época de las lluvias, se limpiaba el Gran Aljibe, retirando de sus fondos las piedras, barro y arena que las aguas habían arrastrado. Para facilitar dicha labor, se ponía un enrejado sobre el aljibe. Esta reja la formaban dos hileras perpendiculares de gruesos barrotes de hierro; cada barrote estaba situado a dos codos del siguiente, con lo que el enrejado disponía $10 \times 6 = 60$ cuadrados de dos codos de lado. La limpieza la hacían dos de los servidores del aljibe, provistos de una soga y de una espuerta: uno de ellos se introducía en el aljibe, con la espuerta, por cada uno de los cuadros del enrejado y llenaba la espuerta de sedimentos; el otro izaba la espuerta con la soga, desde arriba.

(*) Esta narración se ha tomado del libro *Los relatos de Gudor Ben Jusá*, de Juan de Burgos, publicado por la Fundación de la Universidad Politécnica de Madrid.

4

El comisionado Roh-Clem sobornó a este último servidor del Aljibe, llamado Rodiart, y consiguió de él que, cada vez que iba a tirar de la soga, hiciera una señal en el punto de ella que quedaba en el borde superior del Aljibe, junto a la reja, con lo que determinó las profundidades medias del aljibe bajo cada uno de los 60 cuadros de la reja; para entendernos mejor, llamemos p_1, p_2, ..., p_{60} a dichas profundidades.

La parte del aljibe situada bajo el cuadro i-ésimo del enrejado (para $i = 1, 2, ..., 60$) tenía un volumen que queda muy bien aproximado (en codos cúbicos) por el número $(2 \times 2)p_i$; por tanto, una muy buena aproximación de la capacidad del aljibe era el número

$$C = (2 \times 2)(p_1 + p_2 + \cdots + p_{60})$$

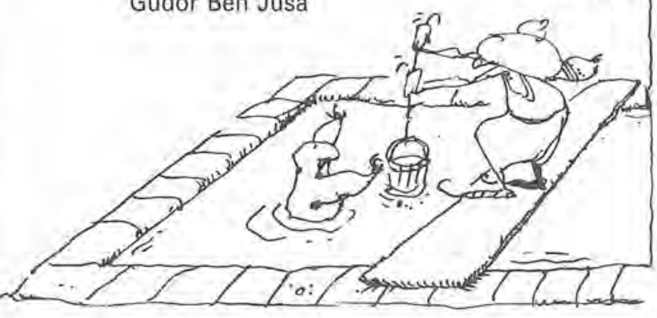
Con los p_i que obtuvo de Rodiart, el Comisionado Roh-Clem dedujo fácilmente el valor C y se lo comunicó secretamente a Ymed-Etaif, el Muy Ilustre Regidor de los Asuntos de la Justicia; a la vista de esta información, Ymed-Etaif ya no tuvo la menor duda: Nordal estafaba con descaro en el tema del agua. Aunque sabía que Nordal era el paniagudo del rey, Ymed-Etaif se armó de valor y expuso los hechos al rey Omaleyos; éste se enfureció, no por lo que pudiera haber hecho Nordal, sino porque se hubiera llegado a saber lo que aquel estaba haciendo. Para guardar las apariencias, Omaleyos III hubo de llamar a Nordal y, delante de Ymed-Etaif, le dijo que se explicara. Nordal, no pudiendo negar la evidencia, admitió que se producía una «merma» en el agua, pero sostuvo que ello era un fenómeno natural que solía ocurrir cuando se mezclan componentes de diversa índole, con características distintas; afirmaba que, al ser tan distintas las aguas de los arroyos Ortil-Oidem y Sajomolos, cosa que era verdad, resultaba que, al juntarse en el Gran Aljibe, se producía una merma, de la que él no era responsable. Apoyó su argumentación haciendo notar que, en una arqueta llena de guijarros de respetable tamaño, se podía añadir gran cantidad de arena sin aumentar por ello el nivel de ocupación de la arqueta. También adujo que si en un balde con agua se añade un cristal de sal, el volumen de la mezcla es notablemente menor que la suma de los volúmenes del agua y de la sal.

Al rey Omaleyos III, las alegaciones de Nordal le parecieron ingeniosísimas y las celebró con grandes risotadas. Ymed-Etaif viendo las cosas perdidas, hizo un último intento para que se hiciera justicia. Se dirigió a Nordal y le preguntó si él creía que, en alguna mezcla, se podían producir «engordes», en lugar de mermas. Nordal dijo, rotundamente, que no. Ymed-Etaif le preguntó, entonces, que cómo era posible que su fortuna fuese notoriamente mayor de lo que se obtendría de sumar lo que él tenía ganado con lo que heredó de sus mayores. El rey manifestó a Ymed-Etaif que lo que había dicho no tenía la menor gracia y le mandó retirarse.

Aquel mismo día, Nordal quedó confirmado en el cargo de Muy Ilustre Regidor de los Asuntos del Agua. También aquel mismo día, Ymed-Etaif fue destituido de su cargo de Muy Ilustre Regidor de los Asuntos de la Justicia.

Al poco tiempo, el comisionado Roh-Clem pidió a Ymed-Etaif que les explicase los motivos de su destitución y de la confirmación de Nordal como Muy Ilustre Regidor de los Asuntos del Agua. Ymed-Etaif, poniendo un gesto de tristeza e ironía, le dijo: conviene convencerse de que «lo que no puede ser, no sólo no puede ser, sino que, además, es imposible».

Gudor Ben Jusá



CAPÍTULO 4

Integrales múltiples (Riemann)

4.1. Integración en intervalos de $\mathbb{R}^p$.—4.2. Clases de funciones integrables en un intervalo.—4.3. Integración en conjuntos medibles.—Ejercicios y problemas.

Según bien conocemos, de estudios anteriores, la integración para funciones de una variable (integrales simples) tiene su origen en el problema de la búsqueda del área de una región plana. La integración para funciones de varias variables (integrales múltiples) tiene una motivación similar: la determinación del volumen de un sólido del espacio.

En el caso de la integral simple, la integración tiene lugar en un intervalo; ahora, que hay más de una variable, el conjunto sobre el que se integra será, en general, mucho más complejo, lo que complica las cosas de un modo apreciable. En el estudio de la integración de funciones de n variables (con $n \geqslant 2$), para escalonar las dificultades, se empieza considerando las integrales sobre un intervalo compacto de $\mathbb{R}^n$; el análisis de tales integrales se puede abordar siguiendo los pasos que se dieron en el caso de integrales simples. Para la integración sobre un conjunto acotado C, que no sea intervalo, se extiende el integrando a un intervalo que incluya a C de modo que sea nulo fuera de C, lo que reduce el problema al de integración sobre un intervalo. Al proceder de este modo, si bien es verdad que se simplifica el conjunto sobre el que se integra, en general se complica el integrando, en el que quedan concentradas todas las posibles «irregularidades» de la integral.

El concepto de integral que se considera es el de Riemann. Con esta integral se puede abordar la gran mayoría de los problemas usuales, teóricos y prácticos. Por otra parte, ello es la base en la que se sustentan las distintas formulaciones que generalizan la noción de integral.

4.1. INTEGRACIÓN EN INTERVALOS DE $\mathbb{R}^p$

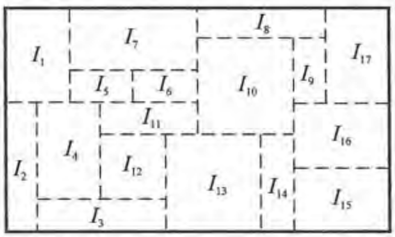
Salvo el último punto, el [60], el resto de lo que se dice en este apartado, en el que se estudian las propiedades generales de las integrales sobre intervalos de $\mathbb{R}^p$, es casi una repetición de lo que, al considerar la integral simple, se estudia en el cálculo integral de una variable. Pudiéramos, pues, haber omitido aquí las demostraciones de los teoremas (cosa que no hemos hechos), sin que de ello se debiera derivar ningún grave inconveniente.

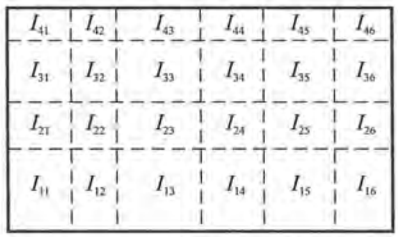
INTERVALOS DE $\mathbb{R}^p$; PARTICIONES DE UN INTERVALO

[50]

- Se llaman *intervalos* compactos de $\mathbb{R}^p$ a los conjuntos $I = [a_1, b_1] \times \cdots \times [a_p, b_p]$, donde $[a_i, b_i]$ es un intervalo compacto de $\mathbb{R}$ para $i = 1, ..., p$. Se llama *medida* de I al número real $\mu(I) = (b_1 - a_1) \cdots (b_p - a_p)$.

- Se dice que un conjunto finito de intervalos compactos, $P = \{I_1, ..., I_n\}$, es una *partición* del intervalo I si $I_1 \cup \cdots \cup I_n = I$ e $\mathring{I}_i \cap \mathring{I}_j = \varnothing$ para $i \neq j$. Se dice que otra partición P' es posterior o más fina que P si todo intervalo de P' está incluido en algún intervalo de P. Se llama *diámetro* de la partición P al mayor de los diámetros de $I_1, ..., I_n$, al cual se le denotará poniendo $|P|$; si P' es posterior a P, entonces $|P'| \leqslant |P|$.
- *Si* $P = \{I_1, ..., I_n\}$ *y* $Q = \{J_1, ..., J_m\}$ son dos particiones de un intervalo I, entonces los conjuntos $I_i \cap J_j$ (para $i = 1, ..., n$ y $j = 1, ..., m$) que no son vacíos forman una partición de I, a la que llamaremos *superposición* de P y Q y la denotaremos poniendo $P * Q$.
- Si $P = \{I_1, ..., I_n\}$ es una partición del intervalo I, entonces la medida de I es la suma de las medidas de las de los intervalos de P, esto es, $\mu(I) = \mu(I_1) + \cdots + \mu(I_n)$.





Comprobación

Vamos a ocuparnos sólo de la última de las afirmaciones anteriores, pues, las demás se verifican trivialmente. Consideremos todos los hiperplanos paralelos a los p hiperplanos coordenados que pasan por los vértices de los intervalos $I_1, ..., I_n$. Estos hiperplanos, al intersecarse, determinan una partición $Q = \{J_1, ..., J_m\}$, que es posterior a P, para la que es evidente que $\mu(I) = \mu(J_1) + \cdots + \mu(J_m)$; también es evidente que, para cada $I_i \in P$, la medida $\mu(I_i)$ es la suma de las medidas de aquellos intervalos de Q que están incluidos en I_i y, como todo intervalo de Q está incluido en algún $I_i \in P$, se concluye que

$$\sum_{i=1}^{n} \mu(I_i) = \sum_{j=1}^{m} \mu(J_j) = \mu(I)$$

$[50]_1$ Observaciones

1. La medida de un intervalo compacto de $\mathbb{R}^2$ es su área; en $\mathbb{R}^3$, la medida de un intervalo es su volumen.

2. La relación «es posterior a», definida entre las particiones de un intervalo compacto, es una relación de orden.
3. El diámetro de un intervalo compacto $I = [a_1, b_1] \times \cdots \times [a_p, b_p]$, de $\mathbb{R}^p$, es la distancia del punto $(a_1, ..., a_p) \in \mathbb{R}^p$ al punto $(b_1, ..., b_p) \in \mathbb{R}^p$, o sea,

$$[(b_1 - a_1)^2 + \cdots + (b_p - a_p)^2]^{1/2}$$

4. Se dice que dos intervalos I y J, de $\mathbb{R}^p$, no se solapan si no tienen puntos interiores comunes (sí pueden tener puntos frontera comunes), es decir, si se cumple una cualquiera de las relaciones:

$$I \cap J \subset F_r(I) \cap F_r(J) \qquad \text{o} \qquad \mathring{I} \cap \mathring{J} = \varnothing$$

5. Si $C = I_1 \cup \cdots \cup I_n$, donde $I_1, ..., I_n$ son intervalos compactos de $\mathbb{R}^p$ tales que cada dos de ellos no se solapan, entonces se dice que C es un «conjunto elemental» y se llama medida de C al número real

$$\mu(C) = \mu(I_1) + \cdots + \mu(I_n)$$

Esta medida $\mu(C)$ es la misma para cualquiera que sea la forma de expresar C como unión de intervalos compactos que no se solapen.
6. Se dice que el intervalo $I = [a_1, b_1] \times \cdots \times [a_p, b_p]$, de $\mathbb{R}^p$, es «degenerado» si es $a_i = b_i$ para algún $i = 1, ..., p$. Cuando nos refiramos a un intervalo, supondremos que éste no es degenerado, salvo que se diga expresamente lo contrario.

SUMAS DE DARBOUX EN UN INTERVALO

[51]

Sea $f : I \to \mathbb{R}$ una función acotada en el intervalo compacto $I \subset \mathbb{R}^p$. Para cada partición $P = \{I_1 ..., I_n\}$ de I, si $m_i = \inf f(I_i)$ y $M_i = \sup f(I_i)$ y $\mu(I_i)$ es la medida de I_i para $i = 1, ..., n$, se llaman sumas inferior y superior (de Darboux) de f correspondientes a la partición P a

$$s(f, p) = \sum_{i=1}^{n} m_i\mu(I_i) \qquad \text{y} \qquad S(f, P) = \sum_{i=1}^{n} M_i\mu(I_i)$$

1.° Llamando $m = \inf f(I)$ y $M = \sup f(I)$ y para cualquiera que sea la partición P de I, se verifica que

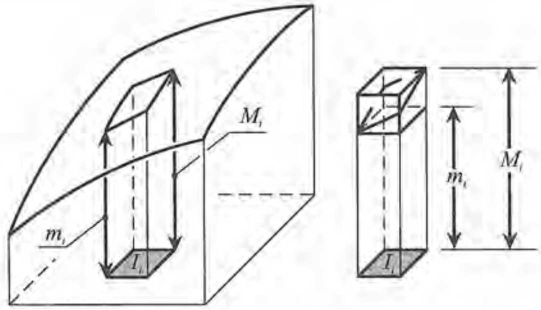
$$m\mu(I) \leqslant s(f, P) \leqslant S(f, P) \leqslant M\mu(I)$$

2.° Si P' es una partición de I posterior a otra partición P, entonces

$$s(f, P') \geqslant s(f, P) \qquad \text{y} \qquad S(f, P') \leqslant S(f, P)$$

3.° Para cualesquiera particiones P_1 y P_2 de I, se verifica que:

$$s(f, P_1) \leqslant S(f, P_2)$$



Demostración

1.° Para cualquiera que sea la partición P, es $m \leqslant m_i \leqslant M_i \leqslant M$ para $i = 1, ..., n$, de donde, como $\mu(I_i) > 0$, se desprende que

$$m\mu(I) = m \sum_{i=1}^{n} \mu(I_i) \leqslant \sum_{i=1}^{n} m_i\mu(I_i) \leqslant \sum_{i=1}^{n} M_i\mu(I_i) \leqslant M \sum_{i=1}^{n} \mu(I_i) = M\mu(I)$$

con lo que $m\mu(I) \leqslant s(f, P) \leqslant S(f, P) \leqslant M\mu(I)$.

2.° (Sólo comprobaremos la primera desigualdad; para la segunda se razona análogamente): si $P' = \{I'_1, ..., I'_{n'}\}$ es una partición posterior a $P = \{I_1, ..., I_n\}$, cada intervalo $I_i \in P$ incluirá a ciertos intervalos de P', que llamaremos $I'_{i1}, ..., I'_{ih_i}$, siendo $h_1 + \cdots + h_n = n'$; llamemos $m'_j = \text{ínf } (I'_j)$ para $j = 1, ..., n'$ y $m'_{ik} = \text{ínf } f(I'_{ik})$ para $k = 1, ..., h_i$, con lo que será $m'_{ik} \geqslant m_i$, ya que $I'_{ik} \subset I_i$, y por tanto (como $m'_{ik} - m_i \geqslant 0$ y $\mu(I'_{ik}) \geqslant 0$):

$$s(f, P') - s(f, P) = \sum_{j=1}^{n'} m'_j\mu(I'_j) - \sum_{i=1}^{n} m_i\mu(I_i) = \sum_{i=1}^{n} \sum_{k=1}^{h_i} m'_{jk}\mu(I'_{ik}) - \sum_{i=1}^{n} m_i\mu(I_i) =$$
$$= \sum_{i=1}^{n} \sum_{k=1}^{h_i} (m'_{ik} - m_i)\mu(I'_{ik}) \geqslant 0$$

3.° Para cualesquiera que sean las particiones P_1 y P_2, recurriendo a la partición $P_1 * P_2$ (superposición de P_1 y P_2, véase [50]), que es posterior a ellas, de acuerdo con lo que se acaba de probar se tiene que:

$$s(f, P_1) \leqslant s(f, P_1 * P_2) \leqslant S(f, P_1 * P_2) \leqslant S(f, P_2)$$

[51]$_1$ Ejercicio

Sea $f: I \to \mathbb{R}$ una función continua en un intervalo compacto I. Pruébese que para cada $\varepsilon > 0$ existe una partición P de I tal que $S(f, P) - s(f, P) < \varepsilon$ (recúrrase a que f es uniformemente continua en I).

Resolución

Como f es continua en I e I es compacto, entonces f es uniformemente continua en I (teorema de Heine; véase [21]) y, por tanto, existe un $\delta > 0$ tal que, para $x, x' \in I$:

$$\|x - x'\| < \delta \quad \Rightarrow \quad |f(x) - f(x')| < \varepsilon/\mu(I) \qquad [1]$$

Sea $P = \{I_1, ..., I_n\}$ una partición cuyo diámetro es menor que δ. Al ser f continua en I_i y como I_i es compacto (para $i = 1, ..., n$) existen $x_i, x_i' \in I_i$ tales que $f(x_i) = m_i$ y $f(x_i') = M_i$ (teorema de Weierstrass; véase [18], II); además es $\|x_i - x_i'\| < \delta$, ya que $x_i, x_i' \in I_i$, con lo que $|f(x_i) - f(x_i')| < \varepsilon/\mu(I)$, es decir $M_i - m_i < \varepsilon/\mu(I)$. De todo ello se obtiene que:

$$S(f, P) - s(f, P) = \sum_{i=1}^{n} (M_i - m_i)\mu(I_i) < \frac{\varepsilon}{\mu(I)} \sum_{i=1}^{n} \mu(I_i) = \varepsilon$$

INTEGRALES INTERIOR Y SUPERIOR EN UN INTERVALO

[52]

Sea $f : I \to \mathbb{R}$ una función acotada en un intervalo compacto I de $\mathbb{R}^p$. Se llaman integral inferior e integral superior de f en I a los siguientes números reales $\underline{\int}_I f$ e $\overline{\int}_I f$, que existen, donde $\mathscr{P}(I)$ denota al conjunto de las particiones de I:

$$\left.\begin{aligned} \underline{\int}_I f &= \sup\ \{s(f, P)/P \in \mathscr{P}(I)\} \\ \overline{\int}_I f &= \inf\ \{S(f, P)/P \in \mathscr{P}(I)\} \end{aligned}\right\} \quad \text{Es} \quad \underline{\int}_I f \leqslant \overline{\int}_I f$$

Para cada $\varepsilon > 0$ existe una partición P de I tal que(*):

$$0 \leqslant \underline{\int}_I f - s(f, P) < \varepsilon \qquad \text{y} \qquad 0 \leqslant S(f, P) - \overline{\int}_I f < \varepsilon \qquad [1]$$

(*) Estas desigualdades se verifican (según se verá en [59]), no solo para una partición P, sino para toda partición de I que tenga diámetro menor que cierto $\delta > 0$, que depende del ε dado.

Demostración

Obsérvese que el conjunto $\{s(f, P)/P \in \mathscr{P}(I)\}$ está acotado superiormente por cualquiera de las sumas superiores $S(f, P)$ (según se desprende de [1]), luego dicho conjunto tiene supremo, es decir, existe la integral inferior $\underline{\int}_I f$ y se verifica que $\underline{\int}_I f \leqslant S(f, P)$ para toda $P \in \mathscr{P}(I)$. De esto último se infiere que existe el ínfimo del conjunto $\{S(f, P)/P \in \mathscr{P}(I)\}$, es decir, que existe $\overline{\int}_I f$ y se verifica que $\underline{\int}_I f \leqslant \overline{\int}_I f$.

Dado $\varepsilon > 0$ (cualquiera), como $\underline{\int}_I f = \sup \{s(f, P)\}$ e $\overline{\int}_I f = \inf \{S(f, P)\}$, existen algunas particiones P_1 y P_2 de I tales que

$$0 \leqslant \underline{\int}_I f - s(P_1) < \varepsilon \qquad \text{y} \qquad 0 \leqslant S(P_2) - \overline{\int}_I f < \varepsilon$$

Llamando $P = P_1 * P_2$ (superposición de P_1 y P_2) y como P es posterior a P_1 y a P_2, resulta que $s(f, P_1) \leqslant s(f, P)$ y $S(f, P_2) \geqslant S(f, P)$, lo que combinado con las anteriores desigualdades conduce a que se verifica la condición [2] del enunciado, pues:

$$0 \leqslant \underline{\int}_I f - s(f, P) \leqslant \underline{\int}_I f - s(f, P_2) < \varepsilon \qquad \text{y} \qquad 0 \leqslant S(f, P) - \overline{\int}_I f \leqslant S(f, P_2) - \overline{\int}_I f < \varepsilon$$

[52]$_1$ Ejercicio

Sean $f : I \to \mathbb{R}$ y $g : I \to \mathbb{R}$ dos funciones acotadas en un intervalo compacto I de $\mathbb{R}^p$; sea $c \in \mathbb{R}$. Pruébese que se verifica:

1.° $$\underline{\int}_I (cf) = \begin{cases} c \underline{\int}_I f, & \text{si } c \geqslant 0 \\ c \overline{\int}_I f, & \text{si } c \leqslant 0 \end{cases} \qquad \text{e} \qquad \overline{\int}_I (cf) = \begin{cases} c \overline{\int}_I f, & \text{si } c \geqslant 0 \\ c \underline{\int}_I f, & \text{si } c \leqslant 0 \end{cases}$$

2.° $$\underline{\int}_I (f+g) \geqslant \underline{\int}_I f + \underline{\int}_I g \qquad \text{e} \qquad \overline{\int}_I (f+g) \leqslant \overline{\int}_I f + \overline{\int}_I g$$

Resolución

Consideramos sólo el caso de las integrales inferiores (de cf y de $f + g$); para las integrales superiores se razona de manera análoga.

1.° Según se comprueba trivialmente, es (en lo que sigue, P es una partición genérica de I):

$$\underline{\int}_I (cf) = \sup \{s(cf, P)\} = \begin{cases} c \sup \{s(f, P)\} = c \underline{\int}_I f, & \text{si es } c \geqslant 0 \\ c \inf \{S(f, P)\} = c \overline{\int}_I f, & \text{si es } c \leqslant 0 \end{cases}$$

2.° Recuérdese que, dados dos conjuntos X e Y y llamando $X + Y = \{x + y / x \in X, y \in Y\}$, se verifica que

$$\sup (X + Y) \leqslant \sup X + \sup Y \qquad \text{e} \qquad \inf (X + Y) \geqslant \inf X + \inf Y$$

Ello nos permite poner, para cualquiera que sea la partición P de I:

$$s(f + g, P) \geqslant s(f, P) + s(g, P)$$

y por tanto, haciendo que P recorra el conjunto de las particiones de I, se tiene

$$\underline{\int}_I (f + g) = \sup \{s(f + g, P)\} \geqslant s(f + g, P) \geqslant s(f, P) + s(g, P) \qquad [1]$$

(la partición P que figura en el último miembro es cualquiera). Sabemos que, para cualquiera que sea $\varepsilon > 0$, existen particiones, P_f y P_g, de I tales que

$$s(P_f, f) \geqslant \int_I f - \frac{\varepsilon}{2} \qquad \text{y} \qquad s(P_g, g) \geqslant \int_I g - \frac{\varepsilon}{2} \tag{2}$$

Si es $P_\varepsilon = P_f * P_g$ (superposición de P_f y P_g), como P_ε es posterior a P_f y a P_g, las desigualdades [2] se verifican si se sustituyen P_f y P_g por P_ε; expresando que así ocurre y sumando dichas desigualdades, se obtiene que

$$s(P_\varepsilon, f) + s(P_\varepsilon, g) \geqslant \int_I f + \int_I g - \varepsilon \tag{3}$$

Como la relación [1] se verifica para cualquier partición P, se verifica en particular para $P = P_\varepsilon$; de las relaciones [1] y [3] se concluye pues, que:

$$\int_I (f+g) \geqslant \int_I f + \int_I g - \varepsilon, \qquad \forall\, \varepsilon > 0$$

y por ser esto cierto para todo $\varepsilon > 0$, se ha de verificar que

$$\int_I (f+g) \geqslant \int_I f + \int_I g$$

FUNCIONES INTEGRABLES EN UN INTERVALO. INTEGRAL

[53]

Sea $f: I \to \mathbb{R}$ una función acotada en un intervalo compacto I de $\mathbb{R}^p$. Se dice que f es integrable (según Riemann) en I si son iguales sus integrales inferior y superior en I y, entonces, ambas se representan poniendo

$$\int_I f, \quad \int_I f\,d\mu, \quad \int_I f(x)\,dx \qquad \text{o} \qquad \int_I \overset{p}{\cdots} \int_I f(x_1, \ldots, x_p)\,dx_1 \ldots dx_p$$

y se dice que este número real es la integral de f en I. Es decir, f es integrable en I si y sólo si existe un único número real, que se denota por $\int_I f$, tal que $s(P) \leqslant \int_I f \leqslant S(P)$ para toda partición P de I.

Condiciones «ε : P» de integrabilidad (de Riemann). Dada una función $f: I \to \mathbb{R}$, acotada en el intervalo compacto I de $\mathbb{R}^p$, se verifica que:

1.ª La función f es integrable en I si, y sólo si, para cada $\varepsilon > 0$ existe una partición P de I tal que $S(f, P) - s(f, P) < \varepsilon$

2.ª La función f es integrable en I y su integral en I es el número A si, y sólo si, para cada $\varepsilon > 0$ existe una partición P de I tal que $A - s(f, P) < \varepsilon$ y $S(f, P) - A < \varepsilon$.

Integrabilidad de las funciones continuas. Si $f: I \to \mathbb{R}$ es una función continua en el intervalo compacto $I \subset \mathbb{R}^p$, entonces f es integrable en I.

Demostración

Nótese que, como $\underline{\int}_I f = \sup \{s(f, P)\}$ e $\overline{\int}_I f = \inf \{S(f, P)\}$, para P recorriendo el conjunto de las particiones de I, resulta que (para $J \in \mathbb{R}$):

$$\underline{\int}_I f \leqslant J \leqslant \overline{\int}_I f \quad \Leftrightarrow \quad \begin{cases} s(f, P) \leqslant J \leqslant S(f, P) \\ \forall P, \text{ partición de } I \end{cases}$$

Así, pues, para que sea $\underline{\int}_I f = \overline{\int}_I f$, es decir, para que la relación $\underline{\int}_I f \leqslant J \leqslant \overline{\int}_I f$ se verifique para un único número real J (que será $J = \underline{\int}_I f = \overline{\int}_I f$), es necesario y suficiente que exista un solo número real J tal que $s(f, p) \leqslant J \leqslant S(f, p)$ para toda partición P de I. Y, aclarado esto, ocupémonos ya de las condiciones «$\varepsilon : P$» de integrabilidad de Riemann y de la integrabilidad de las funciones continuas.

1.º (Primera condición «$\varepsilon : P$»). Si f es integrable en I y dado $\varepsilon > 0$, sabemos que para alguna partición P de I se verifica que (véase [52], [1]):

$$\int_I f - s(f, P) < \frac{\varepsilon}{2} \qquad \text{y} \qquad S(f, p) - \int_I f < \frac{\varepsilon}{2}$$

de donde, sumando, se obtiene que $S(f, P) - s(f, P) < \varepsilon$, luego se cumple entonces la primera condición «$\varepsilon : P$». Supongamos ahora que f no es integrable en I, esto es, que la diferencia $d = \overline{\int}_I f - \underline{\int}_I f$ es positiva, $d > 0$; entonces la primera condición «$\varepsilon : P$» no se verifica al tomar $\varepsilon = d > 0$, ya que para toda partición P de I es:

$$S(f, P) - s(f, P) \geqslant \overline{\int}_I f - \underline{\int}_I f = d$$

2.º (Segunda condición «$\varepsilon : P$»). Si f es integrable en I y dado $\varepsilon > 0$, se sabe (por cumplirse la primera condición «$\varepsilon : P$») que existe una partición P de I tal que $S(f, P) - s(f, p) < \varepsilon$. Para dicha partición, y llamando $A = \int_I f$, se verifica entonces la segunda condición «$\varepsilon : P$», ya que:

$$A - s(f, P) \leqslant S(f, P) - s(f, p) < \varepsilon \qquad \text{y} \qquad S(f, P) - A < S(f, P) - s(f, P) < \varepsilon$$

Recíprocamente, si f no es integrable en I, esto es, si $d = \overline{\int}_I f - \underline{\int}_I f > 0$, entonces la segunda condición «$\varepsilon : P$» no se verifica para $\varepsilon = d/2$, ya que si ella fuese cierta para alguna partición P de I, seria:

$$d = \overline{\int}_I f - \underline{\int}_I f \leqslant S(f, p) - s(f, p) = (S(f, p) - A) + (A - s(f, p)) < \frac{d}{2} + \frac{d}{2} = d,$$

luego habría de ser $d < d$, que evidentemente es falso.

3.º (Integrabilidad de una función continua). Si f es continua en I, entonces se cumple la condición «$\varepsilon : P$» de integrabilidad (para f en I), según ya se vio en el ejercicio $[51]_1$, luego f es integrable en I.

[53]$_1$ Ejercicio

Sea $f: I \to \mathbb{R}$ una función, acotada en el intervalo compacto $I \subset \mathbb{R}^p$, tal que, para cierto intervalo $J \subset I$:

$$f: J \to \mathbb{R} \text{ es integrable en } J$$
$$f(x) = 0 \quad \text{para todo } x \in I - J$$

Pruébese que entonces f es integrable en I y que $\int_I f = \int_J f$

Resolución

Sea P_0 una partición cualquiera de $I - J$(*); para cada partición P_J de J, sea $P_I = P_0 \cup P_J$, que es una partición de I. Como f es nula en $I - J$, es evidente que $s(f, P_I) = s(f, P_J)$ y que $S(f, P_I) = S(f, P_J)$. Por ser f integrable en J, sabemos que $\int_J f$ es el único número real comprendido entre $s(f, P_J)$ y $S(f, P_J)$ para cualquiera que sea la partición P_J, es decir, $\int_J f$ es el único número real comprendido entre $s(f, P_I)$ y $S(f, P_I)$ para todas las particiones $P_I = P_0 \cup P_J$ de I, donde P_J es una partición cualquiera de J. De este último se desprende que $\int_J f$ es el único número real comprendido entre $s(f, P)$ y $S(f, P)$ para cualquiera que sea la partición P de I, lo que significa que f es integrable en I y que $\int_I f = \int_J f$.

[53]$_2$ Ejercicio

Sean $f: I \to \mathbb{R}$ y $g: I \to \mathbb{R}$ dos funciones, acotadas en el intervalo compacto $I \subset \mathbb{R}^p$, tales que, para ciertos puntos dados $a_1, ..., a_n \in I$, es:

$$\begin{cases} g(x) = f(x), & \text{si } \ x \in I - \{a_1, ..., a_n\} \\ g(x) \neq f(x), & \text{si } \ x = a_1, ..., a_n \end{cases}$$

Pruébese que si f es integrable en I, entonces también lo es g y se verifica que $\int_I g = \int_I f$.

Resolución

Si se prueba que la propiedad es cierta en el caso $i = 1$ (o sea si f y g son distintos en un solo punto), aplicando dicho resultado n veces se obtiene la propiedad del enunciado. Supondremos, pues, que hay un punto $a \in I$ tal que $g(x) = f(x)$ para todo $x \in I - \{a\}$ y que $g(a) \neq f(a)$.

Sea dado un $\varepsilon > 0$ cualquiera. Sea I_0 un intervalo de $\mathbb{R}^p$ tal que $a \in I_0 \subset I$ y cuya medida es $\mu(I_0) < \varepsilon/2K$, donde K es una cota de $|f(x)|$ y de $|g(x)|$ para $x \in I$; sea P_0 una partición de I que tiene a I_0 como uno de sus intervalos. Como f es integrable en I, y según la 1.ª condición «$\varepsilon : P$» de integrabilidad, existe una partición P_1 de I tal que $S(f, P_1) - s(f, P_1) < \varepsilon/2$; sea $P' = P_0 * P_1$ (superposición de P_0 y P_1); como P' es posterior a P_1, se verifica que $S(f, P') - s(f, P') < \varepsilon/2$.

(*) Es decir, P_0 es un conjunto finito de intervalos, $P_0 = \{I_1, ..., I_n\}$, tal que $I_1 \cup \cdots \cup I_n = I - J$ e $\mathring{I}_i \cap \mathring{I}_j = \varnothing$ para $i \neq j$.

Para comprobar que g es integrable en I, acudamos a la 2.ª condición «$\varepsilon : P$» de integrabilidad: comprobemos que para g se cumple esta condición si se toma para P la anterior partición P' y siendo $A = \int_I f$ (comprobemos sólo la desigualdad correspondiente a la suma inferior; para la superior se precede de modo análogo). Así ocurre, ya que

$$\left|\int_I f - s(g, P')\right| \leqslant \left|\int_I f - s(f, P')\right| + |s(f, P') - s(g, P')| < |S(f, P') - s(f, P')| +$$

$$+ 2K\mu(I_0) < \frac{\varepsilon}{2} + 2K\frac{\varepsilon}{4K} = \varepsilon$$

$[53]_3$ Ejercicio

Sea f: $[0,1] \times [0,1] \to \mathbb{R}$ la función definida mediante

$$f(x, y) = \begin{cases} 0, & \text{si } x = 0 \text{ o si } y = 0 \\ 0, & \text{si alguno de los } x \text{ o } y \text{ son irracionales} \\ \dfrac{1}{q+s}, & \text{si } x = \dfrac{p}{q} \text{ e } y = \dfrac{r}{s} \text{ (racionales irreducibles)} \end{cases}$$

Compruébese que f es integable en $I = [0,1] \times [0,1]$.

Resolución

Acudamos a la primera condición «$\varepsilon : P$». Sea dado $\varepsilon > 0$ (cualquiera). Llamemos h a un número natural tal que $h > 2/\varepsilon$. El número m de punto $(x, y) \in I$ en los que es $f(x, y) \geqslant 1/h$ es finito, pues dichos puntos son los $(p/q, r/s)$ tales que $q + s \leqslant h$, que son sólo algunos de los que tienen $q \leqslant h$ y $s \leqslant h$ y estos últimos hacen un total que no llega a superar a

$$(1 + 2 + \cdots + h) + (1 + 2 + \cdots + h) = h(h+1);$$

de manera que m es finito, $m \leqslant h(h+1)$. Para cada uno de estos puntos, consideremos un intervalo cerrado I_i (para $i = 1, ..., m$) centrado en él, que tenga longitud menor que $\sqrt[p]{\varepsilon/m}$ y de modo que dichos intervalos no se solapen. Finalmente, consideremos una cualquiera de las particiones P de I a la que pertenezcan los intervalos $I_1, ..., I_m$, que podremos poner

$$P = P_1 \cup P_2, \qquad \text{con } P_1 = \{I_1, ..., I_m\}$$

De lo dicho, resulta evidente que:

$$S(f, P) - s(f, P) = S(f, P) - 0 = S(f, P_1) + S(f, P_2) <$$

$$< m(\sqrt[p]{\varepsilon/m})^p \text{ máx } f(I) + \mu(I)\frac{1}{h} < m\frac{\varepsilon}{m}\cdot\frac{1}{2} + 1\cdot\frac{\varepsilon}{2} = \varepsilon$$

luego para esta partición P, se verifica la primera condición «$\varepsilon : P$» de integrabilidad, como había que verificar.

PROPIEDADES ARITMÉTICAS DE LA INTEGRAL (EN UN INTERVALO)

[54]

Sean $f: I \to \mathbb{R}$ y $g: I \to \mathbb{R}$ dos funciones, acotadas en el intervalo compacto $I \subset \mathbb{R}^p$. Si f y g son integrables en I, se verifica:

1.° *Linealidad de la integral: las funciones* $f+g$ y cf (para cualquier $c \in \mathbb{R}$) son integrables en I y sus integrales son:

$$\int_I (f+g) = \int_I f + \int_I g \qquad \text{e} \qquad \int_I cf = c\int_I f$$

2.° *Integrabilidad del producto:* la función fg es integrable en I.

3.° *Integrabilidad del cociente:* si es $|g(x)| \geqslant k$ para todo $x \in I$ y para un cierto $k > 0$, entonces la función f/g es integrable en I.

4.° *Integrabilidad del valor absoluto:* la función $|f|$ es integrable en I.

CRITERIO LINEAL DE INTEGRABILIDAD. Sean $\varphi: I \to \mathbb{R}$ y $\psi: I \to \mathbb{R}$ dos funciones integrables en I. Sea $F: I \to \mathbb{R}$ una función acotada. Se supone que existen dos constantes $h \geqslant 0$ y $k \geqslant 0$ tales que, para cualquiera que sea el intervalo compacto $J \subset I$, las oscilaciones $\omega(\varphi, J)$, $\omega(\psi, J)$ y $\omega(F, J)$, de φ, ψ y F en I, son tales que $\omega(F, J) \leqslant h\omega(\varphi, J) + k\omega(\psi, J)$ [recuérdese que la oscilación de una función f en J es $\omega(f, J) = \sup f(J) - \inf f(J)$]. Entonces F es integrable en I.

Demostración

Empecemos comprobando el «criterio lineal de integrabilidad», del que aquí haremos luego uso: por cumplirse la desigualdad que figura en el enunciado y para cualquiera que sea la partición $P = \{I_1, ..., I_n\}$ de I, se verifica que:

$$\begin{aligned} S(F, P) - s(F, P) &= \sum_{i=1}^{n} \omega(F, I_i)\,\mu(I_i) \leqslant h \sum_{i=1}^{n} \omega(\varphi, I_i)\,\mu(I_i) + k \sum_{i=1}^{n} \omega(\psi, I_i)\,\mu(I_i) = \\ &= h[S(\varphi, P) - s(\varphi, P)] + k[S(\psi, P) - s(\psi, P)] \end{aligned} \qquad [1]$$

Como φ y ψ son integrales en I y según la primera condición «$\varepsilon : P$» de integrabilidad (véase [53], 1.ª), dado $\varepsilon > 0$ (cualquiera) existen dos particiones P_φ y P_ψ de I tales que

$$S(\varphi, P_\varphi) - s(\varphi, P_\varphi) < \frac{\varepsilon}{2h} \qquad \text{y} \qquad S(\psi, P_\psi) - s(\psi, P_\psi) < \frac{\varepsilon}{2k}$$

Nótese que estas relaciones se verifican también si, en lugar de P_φ y P_ψ, se pone en ellas cualquier partición P posterior a P_φ y P_ψ, como por ejemplo la $P = P_\varphi * P_\psi$ (superposición de P_φ y P_ψ); por ello, de [1] se deduce que, para esta partición P, es:

$$S(F, P) - s(F, P) < h\frac{\varepsilon}{2h} + k\frac{\varepsilon}{2k} = \varepsilon$$

luego F es integrable en I, pues para ella se cumple la primera condición «$\varepsilon : P$».

1.° Esta propiedad se puede obtener a partir del ejercicio [52]$_1$, sin más que tener presente que, ahora, es $\underline{\int}_I f = \overline{\int}_I f$ e $\underline{\int}_I g = \overline{\int}_I g$. No obstante, vamos a dar una demostración directa. Para ello, tendremos en cuenta que, para cualquiera que sea la partición P de I, es(*)

$$s(f+g, P) \geqslant s(f, P) + s(g, P) \qquad \text{y} \qquad S(f+s, P) \leqslant S(f, P) + S(g, P)$$

Recurramos a la segunda condición «$\varepsilon : P$» de integrabilidad (véase [53], 2.ª). Dado $\varepsilon > 0$ (cualquiera) y como f y g son integrables en I, existen particiones P_f y P_g de I tales que

$$\begin{array}{ll} \displaystyle\int_I f - s(f, P_f) < \frac{\varepsilon}{2} & \displaystyle S(f, P_f) - \int_I f < \frac{\varepsilon}{2} \\ \displaystyle\int_I g - s(g, P_g) < \frac{\varepsilon}{2} & \displaystyle S(g, P_g) - \int_I g < \frac{\varepsilon}{2} \end{array} \qquad [1]$$

Nótese que estas relaciones se verifican si, en lugar de P_f y P_g, se pone en ellas una partición P posterior a P_f y a P_g, como por ejemplo la $P = P_f * P_g$ (superposición de P_f y P_g). Por ello, para dicha partición P se verifica que

$$\left(\int_I f + \int_I g\right) - s(f+g, P) \leqslant \left(\int_I f + \int_I g\right) - (s(f, P) + s(g, P)) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

$$S(f+g, P) - \left(\int_I f + \int_I g\right) \leqslant (S(f, P) + S(g, P)) - \left(\int_I f + \int_I g\right) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

luego (según la segunda condición «$\varepsilon : P$») $\int_I f + \int_I g$ es la integral de $f + g$ en I, como había que comprobar.

Para estudiar la integrabilidad de la función cf, se estudiarán por separado la casos $c \geqslant 0$ y $c \leqslant 0$ (sólo consideraremos el $c \leqslant 0$; para el otro se razona igual). si es $c \leqslant 0$, es evidente que $s(cf, P) = cS(f, P)$ y que $S(cf, P) = cs(f, P)$, para toda partición P. Por tanto, dado $\varepsilon > 0$, si es P una partición de I tal que

$$\int_I f - s(f, P) < \frac{\varepsilon}{|c|} \qquad \text{y} \qquad S(f, P) - \int_I f < \frac{\varepsilon}{|c|}$$

(tal partición existe según la segunda condición «$\varepsilon : P$» aplicada a f en I), en el caso $c \leqslant 0$ se obtiene que:

$$c \int_I f - s(cf, P) = c\left(\int_I f - S(f, P)\right) \leqslant (-c)\,\frac{\varepsilon}{|c|} = \varepsilon$$

$$S(cf, P) - c \int_I f = c\left(s(f, P) - \int_I f\right) \leqslant (-c)\,\frac{\varepsilon}{|c|} = \varepsilon$$

luego (según la segunda condición «$\varepsilon : P$») $c\int_I f$ es la integral de cf en I, como habia que probar.

(*) Téngase en cuenta que, dados dos conjuntos X e Y y si es $X + Y = \{x + y / x \in X \text{ e } y \in Y\}$, se verifica que $\sup (X + Y) \leqslant \sup X + \sup Y$ y que $\inf (X + Y) \geqslant \inf X + \inf Y$.

2.º Comprobemos en primer lugar que, si f es integrable en I, entonces también lo es f^2. Para ello, si es K una cota superior de $|f|$, pongamos que, para cualesquiera $x, x' \in I$, es

$$|f(x)^2 - f(x')^2| = |f(x) + f(x')|\,|f(x) - f(x')| \leqslant 2K\,|f(x) - f(x')|$$

Por tanto, para cualquiera que sea el intervalo compacto $J \subset I$, es

$$\omega(f^2, J) \leqslant 2K\omega(f, J)$$

Según el «criterio lineal de integrabilidad» (tomando $F = f^2$, $\varphi = f$, $h = 2K$ y $k = 0$), de esta última relación se desprende la integrabilidad de f^2 en I.

Para comprobar la integrabilidad de fg, recurramos a que

$$fg = \tfrac{1}{2}\left[(f+g)^2 - f^2 - g^2\right]$$

Como f y g son integrables en I, sabemos que también lo son $f + g$, $(f+g)^2$, f^2, g^2 y, por ello, el anterior segundo miembro es integrable en I, es decir, lo es fg.

3.º Es suficiente con que comprobemos que $1/g$ es integrable en I, ya que entonces, por ser f y $1/g$ integrables, sería integrable su producto, que es f/g. Para cualesquiera que sean los puentes $x, x' \in I$, se verifica:

$$\left|\frac{1}{g(x)} - \frac{1}{g(x')}\right| = \frac{|g(x') - g(x)|}{|g(x)g(x')|} \leqslant \frac{1}{c^2}\,|g(x') - g(x)|$$

Así, pues, para cualquiera que sea el intervalo compacto $J \subset I$, es

$$\omega(1/g, J) \leqslant \frac{1}{c^2}\,\omega(g, J)$$

Según el «criterio lineal de integrabilidad» (tomando $F = 1/g$, $\varphi = g$, $h = 1/c^2$ y $k = 0$), de esta última relación se desprende que, como g es integrable en I, también lo es $1/g$.

4.º Para cualesquiera que sean los puntos $x, x' \in I$, se verifica que

$$\big||f(x)| - |f(x')|\big| \leqslant |f(x) - f(x')|$$

Luego, para cualquiera que sea el intervalo compacto $J \subset I$, es

$$\omega(|f|, J) \leqslant \omega\,(f, J)$$

De esto se deduce que, según, el «criterio lineal de integrabilidad» (en el que tomamos $F = |f|$, $\varphi = f$, $h = 1$ y $k = 0$), $|f|$ es integrable en I por serlo f.

[54]$_1$ Ejercicio

Si $f : I \to \mathbb{R}$ es una función positiva, acotada en el intervalo $I \subset \mathbb{R}^p$, que es integrable en I, entonces $\sqrt{f}$ también es integrable en I.

Resolución

Acudimos a la primera condición «$\varepsilon : P$» de integrabilidad (véase [53], 1.ª). Sea dado $\varepsilon > 0$ (cualquiera); llamemos

$$\varepsilon_1 = [\tfrac{1}{2}\,\varepsilon\mu(I)]^2 \qquad \text{y} \qquad \varepsilon_2 = \frac{1}{4}\,\varepsilon^2\,\mu(I)$$

(donde $\mu(I)$ es la medida del intervalo I). Como f es integrable en I y según la primera condición «$\varepsilon : P$», existe una partición $P = \{I_1, ..., I_n\}$ de I tal que $S(f, P) - s(f, P) < \varepsilon_2$. Llamemos $m_i = \inf f(I_i)$ y $M_i = \sup f(I_i)$ para $i = 1, ..., n$ y notemos que:

$$\begin{cases} M_i \leqslant \varepsilon_1 & \Rightarrow \quad \sqrt{M_i} - \sqrt{m_i} < \sqrt{M_i} \leqslant \sqrt{\varepsilon_1} \\ M_i > \varepsilon_1 & \Rightarrow \quad \sqrt{M_i} - \sqrt{m_i} = \dfrac{M_i - m_i}{\sqrt{M_i} + \sqrt{m_i}} < \dfrac{M_i - m_i}{\sqrt{\varepsilon_1}} \end{cases}$$

Por tanto, se tiene que:

$$S(\sqrt{f}, P) - s(\sqrt{f}, P) = \sum_{i=1}^{n} (\sqrt{M_i} - \sqrt{m_i})\mu(I_i) < \sum_{i=1}^{n} \left(\sqrt{\varepsilon_1} + \frac{M_i - m_i}{\sqrt{\varepsilon_1}}\right)\mu(I_i) =$$

$$= \sqrt{\varepsilon_1}\,\mu(I) + \frac{1}{\sqrt{\varepsilon_1}}(S(f, P) - s(f, P)) < \sqrt{\varepsilon_1}\,\mu(I) + \frac{\varepsilon_2}{\sqrt{\varepsilon_1}} = \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

luego para $\sqrt{f}$ se verifica en I la primera condición «$\varepsilon : P$», esto es, $\sqrt{f}$ es integrable en I.

MONOTONÍA DE LA INTEGRAL (EN UN INTERVALO)

[55]

1.° Si $f : I \to \mathbb{R}$ y $g : I \to \mathbb{R}$ son dos funciones, acotadas en el intervalo compacto $I \subset \mathbb{R}^p$, que son integrables en I, entonces se verifica que:

$$(f(x) \leqslant g(x), \quad \forall\, x \in I) \quad \Rightarrow \quad \int_I f \leqslant \int_I g$$

2.° Si $f : I \to \mathbb{R}$ es una función, acotada en el intervalo compacto $I \subset \mathbb{R}^p$, que es integrable en I, entonces se verifica que:

$$\left|\int_I f\right| \leqslant \int_I |f|$$

Demostración

1.º Hay que comprobar que $\int_I g - \int_I f \geqslant 0$, esto es, que $\int_I (g - f) \geqslant 0$. Como $g(\boldsymbol{x}) - f(\boldsymbol{x}) \geqslant 0$ para todo $\boldsymbol{x} \in I$, resulta que $s(g - f, P) \geqslant 0$ para toda partición P de I y de ahí que sea

$$\int_I (g - f) \geqslant s(g - f, P) \geqslant 0$$

2.º Como para todo $\boldsymbol{x} \in I$ es

$$-|f(\boldsymbol{x})| \leqslant f(x) \leqslant |f(x)|$$

de la anterior propiedad de monotonía se deduce que

$$-\int_I |f| \leqslant \int_I f \leqslant \int_I |f|$$

lo que equivale a la relación $|\int_I f| \leqslant \int_I |f|$.

[55]₁ Desigualdad de Cauchy-Schwarz (en un intervalo)

Si $f: I \to \mathbb{R}$ y $g: I \to \mathbb{R}$ son dos funciones, acotadas en el intervalo compacto $I \subset \mathbb{R}^p$, que son integrables en I, entonces se verifica que:

$$\left(\int_I fg\right)^2 \leqslant \left(\int_I f^2\right)\left(\int_I g^2\right)$$

Demostración

Consideremos las dos funciones $\varphi: I \to \mathbb{R}$ definidas por la expresión

$$\varphi(\boldsymbol{x}) = \left(\sqrt{\int_I f^2}\right)g(\boldsymbol{x}) \pm \left(\sqrt{\int_I g^2}\right)f(\boldsymbol{x}), \qquad \forall\, \boldsymbol{x} \in I$$

Obsérvese que, como $\varphi(\boldsymbol{x})^2 \geqslant 0$ para $\boldsymbol{x} \in I$, la propiedad de monotonía (véase [55], 1.º) nos permite asegurar que es $\int_I \varphi \geqslant 0$, o sea, que:

$$\int_I \left\{\left(\int_I f^2\right)g^2 \pm 2\sqrt{\left(\int_I f^2\right)\left(\int_I g^2\right)}fg + \left(\int_I g^2\right)f^2\right\} \geqslant 0$$

operando en el primer miembro de esta desigualdad se obtiene:

$$2\left(\int_I f^2\right)\left(\int_I g^2\right) \pm 2\sqrt{\left(\int_I f^2\right)\left(\int_I g^2\right)}\int_I (fg) \geqslant 0$$

Dividiendo aquí por el número positivo $2\sqrt{(\int_I f^2)(\int_I g^2)}$ se llega trivialmente a la anunciada desigualdad de Cauchy-Schwarz.

$[55]_2$ Ejercicio

Sea $f: I \to \mathbb{R}$ una función acotada en el intervalo compacto $I \subset \mathbb{R}^p$. Pruébese que si es $f(\boldsymbol{x}) \geqslant 0$ para todo $\boldsymbol{x} \in I$, si f es integrable en I, si f es continua en un cierto punto $\boldsymbol{a} \in I$ y es $f(\boldsymbol{a}) > 0$, entonces $\int_I f > 0$.

Resolución

Por ser f continua en $\boldsymbol{a}$, existe un intervalo compacto I_0 (cuya medida $\mu(I_0)$ es, pues, estrictamente positiva) tal que, siendo $\boldsymbol{a} \in I_0 \subset I$, es $f(\boldsymbol{x}) > f(\boldsymbol{a})/2$ para todo $\boldsymbol{x} \in I_0$. Sea $g: I \to \mathbb{R}$ la función definida por:

$$g(\boldsymbol{x}) = f(\boldsymbol{a})/2, \text{ si } \boldsymbol{x} \in I_0; \qquad g(\boldsymbol{x}) = 0, \text{ si } \boldsymbol{x} \notin I_0$$

Como $f(\boldsymbol{x}) \geqslant g(\boldsymbol{x})$ para $\boldsymbol{x} \in I$, de la propiedad de monotonía [55], 1.º se desprende que

$$\int_I f \geqslant \int_I g = \frac{f(\boldsymbol{a})}{2}\,\mu(I_0) > 0$$

TEOREMA DE LA MEDIA (EN UN INTERVALO)

[56]

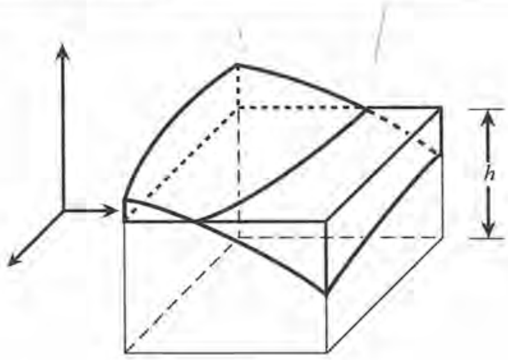
Sea $f: I \to \mathbb{R}$ una función acotada en el intervalo compacto $I \subset \mathbb{R}^p$. Se verifica que:

1.º Si f es integrable en I y si m y M son cotas superior e inferior de f en I, entonces existe $h \in [m, M]$ tal que

$$\int_I f = h\mu(I) \qquad (\mu(I) = \text{medida de } I)$$

2.º Si f es continua en I, entonces existe algún punto $\xi \in I$ tal que

$$\int_I f = f(\xi)\mu(I)$$



Demostración

1.° De acuerdo con la propiedad de monotonía [55],1.ª y por verificarse que $m \leqslant f(\boldsymbol{x}) \leqslant M$ para $\boldsymbol{x} \in I$, se sabe que

$$m\mu(I) = \int_I m\, dx_1 \dots dx_p \leqslant \int_I f(\boldsymbol{x})\, dx_1 \dots dx_p \leqslant \int_I M\, dx_1 \dots dx_p = M\mu(I)$$

Llamando $h = (\int_I f)/\mu(I)$, la anterior relación toma la forma $m \leqslant h \leqslant M$, es decir, el teorema es cierto para dicho valor de h.

2.° En este segundo caso, como f es continua en I, sabemos que f es integrable en I y que alcanza sus valores máximo y mínimo en I (pues I es compacto), con lo que ahora es válido el anterior resultado para $m = \text{mín}\, f(I)$ y $M = \text{máx}\, f(I)$. Además, según la propiedad de Darboux (véase [18]), f alcanza todos los valores compredidos entre las anteriores m y M, luego existe $\xi \in I$ tal que $f(\xi) = \mu$, con lo que $\int_I f = f(\xi)\mu(I)$.

[56]₁ Generalización del teorema de la media

Sean $f: I \to \mathbb{R}$ y $g: I \to \mathbb{R}$ dos funciones acotadas en el intervalo compacto $I \subset \mathbb{R}^p$.

1.° Si f y g son integrables en I y si $m \leqslant f(\boldsymbol{x}) \leqslant M$ y $g(\boldsymbol{x}) \geqslant 0$ para todo $\boldsymbol{x} \in I$, entonces existe $h \in [m, M]$ tal que

$$\int_I fg = h \int_I g$$

2.° Si f es continua en I, si g es integrable en I y si $g(\boldsymbol{x}) \geqslant 0$ para todo $\boldsymbol{x} \in I$, entonces existe algún punto $\xi \in I$ tal que

$$\int_I fg = f(\xi) \int_I g$$

Demostración

1.° Por ser $g(\boldsymbol{x}) \geqslant 0$, de la desigualdad $m \leqslant f(\boldsymbol{x}) \leqslant M$ se deriva la $mg(\boldsymbol{x}) \leqslant f(\boldsymbol{x})g(\boldsymbol{x}) \leqslant \leqslant Mg(\boldsymbol{x})$ para $\boldsymbol{x} \in I$. De esta relación y según la propiedad de monotonía [55], 1.° se desprende que

$$m \int_I g \leqslant \int_I fg \leqslant M \int_I g$$

Dividiendo esta desigualdad por $\int_I g > 0$ (si fuese $\int_I g = 0$, de las anteriores desigualdades se deduciría que $\int_I fg = 0$, con lo que la propiedad a demostrar se verificaría), se llega a que: para $h = (\int_I fg) : (\int_I g)$, es $m \leqslant h \leqslant M$, luego es $\int_I fg = \mu \int_I g$ para un $h \in [m, M]$, como había que comprobar.

2.° Ahora, como se verifican sobradamente las hipótesis del caso anterior para $m = \min f(I)$ y $M = \max f(I)$ (que existen, según el teorema de Weierstrass, pues f continúa en el compacto I), se sabe que $\int_I fg = h \int_I f$ para un $h \in [m, M]$. Además, según la propiedad de Darboux, existe un punto $\xi \in I$ tal que $f(\xi) = h$, con lo que $\int_I fg = f(\xi) \int_I g$.

ADITIVIDAD RESPECTO DEL INTERVALO

[57]

Sea $I \subset \mathbb{R}^p$ un intervalo compacto y supóngase que $\{I_1, ..., I_n\}$ es una partición de I en intervalos compactos(*). Si $f: I \to \mathbb{R}$ es una función acotada en I, entonces: f es integrable en I si, y sólo si, es integrable en todos los intervalos $I_1, ..., I_n$ y se verifica que

$$\int_I f = \int_{I_1} f + \cdots + \int_{I_n} f \qquad [1]$$

(*) $I_1, ..., I_n$ son intervalos compactos de $\mathbb{R}^p$ tales que $I_1 \cup \cdots \cup I_n = I$ e $\mathring{I}_i \cap \mathring{I}_j = \varnothing$ para $i \neq j$.

Demostración

1.° Suponiendo, primero, que f es integrable en I_i para $i = 1, ..., n$, veamos que entonces f es integrable en I y que $\int_I f$ es la suma que figura en el enunciado. Recurrimos para ello a la segunda condición «$\varepsilon : P$» de integrabilidad (véase [53], 2.ª). Dado $\varepsilon > 0$ (cualquiera), como f es integrable en I_i, para cada $i = 1, ..., n$, existe una partición P_i de I_i tal que

$$\int_{I_i} f - s(f, P_i) < \frac{\varepsilon}{n} \qquad \text{y} \qquad S(f, P_i) - \int_{I_i} f < \frac{\varepsilon}{n} \qquad [2]$$

Es evidente que $P = P_1 \cup \cdots \cup P_n$ es una partición de I y que se verifica:

$$s(f, P) = \sum_{i=1}^{n} s(f, P_i) \qquad \text{y} \qquad S(f, P) = \sum_{i=1}^{n} S(f, P_i) \tag{3}$$

Por tanto, de las relaciones [2] y [3] se concluye que, para la partición P, es:

$$\left(\sum_{i=1}^{n} \int_{I_i} f\right) - s(f, P) = \sum_{i=1}^{n} \left(\int_{I_i} f - s(f, P_i)\right) < n\,\frac{\varepsilon}{n} = \varepsilon$$

$$S(f, P) - \left(\sum_{i=1}^{n} \int_{I_i} f\right) = \sum_{i=1}^{n} \left(S(f, P_i) - \int_{I_i} f\right) < n\,\frac{\varepsilon}{n} = \varepsilon$$

lo que, según la segunda condición «$\varepsilon : P$», permite asegurar que f es integrable en I y que cumple la igualdad [1] del enunciado.

2.º Recíprocamente, supongamos ahora que f es integrable en I; de acuerdo entonces con la primera condición «$\varepsilon : P$» de integrabilidad (véase [53],1.ª), dado $\varepsilon > 0$ (cualquiera), existe una partición P_0 de I tal que $S(f, P_0) - s(f, P_0) < \varepsilon$. Sea P la partición que resulta de superponer (véase [50]) P_0 con la partición $\{I_1, ..., I_n\}$ del enunciado; como P es posterior a P_0, se verifica que $S(f, P) - s(f, P) < \varepsilon$. Como P es posterior a $\{I_1, ..., I_n\}$, todo intervalo de P está incluido en uno de los intervalos $I_1, ..., I_n$. Llamemos P_j al conjunto de los intervalos de P que están incluidos en I_j (para $j = 1, ..., n$); es evidente que P_j es una partición de I_j y que

$$s(f, P) = \sum_{j=1}^{n} s(f, P_j) \qquad \text{y} \qquad S(f, P) = \sum_{j=1}^{n} S(f, P_j)$$

Ahora bien, como $S(f, P_j) - s(f, P_j) \geqslant 0$ para todo $j = 1, ..., n$, resulta que para cualquiera que sea $i = 1, ..., n$ es:

$$S(f, P_i) - s(f, P_i) \leqslant \sum_{j=1}^{n} (S(f, P_j) - s(f, P_j)) = \sum_{j=1}^{n} S(f, P_j) - \sum_{j=1}^{n} s(f, P_j) =$$
$$= S(f, P) - s(f, P) < \varepsilon$$

luego se cumple la primera condición «$\varepsilon : P$» de integrabilidad por f en I_i, es decir, existe $\int_{I_i} f$, como había que comprobar (la relación [1] del enunciado también se verifica ahora, pues como existen las integrales $\int_{I_i} f$, es válido lo ya dicho en el apartado 1.º de esta demostración).

[57]$_1$ Observación

Sea $f : I \to \mathbb{R}$ una función acotada en un intervalo compacto $I \subset \mathbb{R}^p$ y considérese otro intervalo compacto I_1 incluido en I. Si f es integrable en I, entonces f es integrable en I_1, como se desprende obviamente de la propiedad anterior (acúdase a una partición P de I que tenga a I_1 como uno de sus intervalos). Si, además, f es no negativa (basta con que sea $f(x) \geqslant 0$ para $x \in I - I_1$), entonces es $\int_{I_1} f \leqslant \int_I f$, pues $\int_I f$ es la suma de $\int_{I_1} f$ con integrales de f en otros subintervalos, en los que f es no negativa, que son por ello no negativas.

SUMAS DE RIEMANN. TEOREMA DE RIEMANN (EN INTERVALOS)

[58]

Sea $f: I \to \mathbb{R}$ una función acotada en el intervalo compacto $I \subset \mathbb{R}^p$.

SUMAS DE RIEMANN. Para cada partición $P = \{I_1, ..., I_n\}$ de I, se llama *familia de puntos intermedios* a cualquiera de los conjuntos $t = \{t_1, ..., t_n\}$, donde $t_i \in I_i$ para $i = 1, ..., n$. Se llama *suma de Riemann* de f, relativa a la partición P y a la familia de puntos t, al número

$$\sigma(f, P, t) = \sum_{i=1}^{n} f(t_i)\mu(I_i) \qquad (\mu(I_i) = \text{medida de } I_i)$$

TEOREMA DE RIEMANN. La función f es integrable en I y su integral es el número A si, y sólo si, se verifica la siguiente condición «$\varepsilon : \delta$»: el número A es tal que, para cada $\varepsilon > 0$, existe un $\delta > 0$ tal que $|A - \sigma(f, P, t)| < \varepsilon$ para toda partición P que tenga diámetro menor que δ y para cualquiera que sea la correspondiente familia t de puntos intermedios: es decir, esquemáticamente:

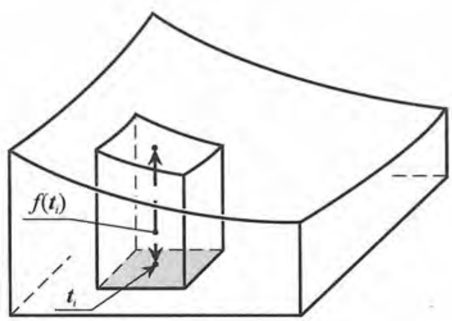
$$\forall\, \varepsilon > 0, \quad \exists\, \delta > 0 \;/\; |P| < \delta \;\Rightarrow\; |A - \sigma(f, P, t)| < \varepsilon, \qquad \forall\, t$$

Para demostrar el anterior teorema, se probarán previamente los siguientes resultados:

LEMA 1.° Dada una partición P del intervalo compacto I, para cada $\varepsilon > 0$ existe un $\delta > 0$ tal que, si Q es una partición de I que tiene diámetro menor que δ, se verifica que la suma de los contenidos de los intervalos de Q que no están incluidos en algún intervalo de P es menor que ε.

LEMA 2.° Dada una partición P del intervalo compacto I y si T es el conjunto de las familias de puntos intermedios correspondientes a la partición P (fija), se verifica que

$$s(f, P) = \inf\,\{\sigma(f, P, t) \,/\, t \in T\} \qquad \text{y} \qquad S(f, P) = \sup\,\{\sigma(f, P, t) \,/\, t \in T\}$$



Demostración

1.° (Lema 1.°). Sea $P = \{I_1, ..., I_n\}$ y considérese una cualquiera de las particiones $Q = \{J_1, ..., J_m\}$ de I que tienen diámetro d menor que la mitad de la menor de las longitudes de las aristas de los intervalos I_i de P; sea δ_0 dicha mitad, con lo que $|Q| < \delta_0$. Para cada intervalo $I_i \in P$:

$$I_i = [a_1, b_1] \times \cdots \times [a_p, b_p]$$

es evidente que los intervalos de Q que tienen sus puntos repartidos en I_i y fuera de I_i están incluidos todos ellos en el conjunto

$$[a_1 - d, b_1 + d] \times \cdots \times [a_p - d, b_p + d] - [a_1 + d, b_1 - d] \times \cdots \times [a_p + d, b_1 - d]$$

y, por tanto, llamando μ_i a la suma de las medidas de dichos intervalos de Q, se verifica que

$$\mu_i < (b_1 - a_1 + 2d) \cdots (b_p - a_p + 2d) - (b_1 - a_1 - 2d) \cdots (b_p - a_p - 2d) = \varphi_i(d)$$

Esta última expresión, a la que hemos llamado $\varphi_i(d)$, considerada como función de la variable d, es continua en el punto $d = 0$, luego existe $\delta_i > 0$ tal que, siempre que sea $d < \delta_i$, se verifica que $\varphi_i(d) < \varepsilon/n$. Por tanto, es $\mu_i < \varepsilon/n$ con tal de tomar $d < \delta_0$ y $d < \delta_i$. Repitiendo n veces este razonamiento, tomando $i = 1, ..., n$, nos encontramos con que, para la suma μ de los contenidos de los intervalos de Q que no están incluidos en algún intervalo de P, es $\mu < \mu_1 + \cdots + \mu_n$ y, por tanto, llamando $\delta = \text{mín}\ \{\delta_0, \delta_1, ..., \delta_n\}$, se verifica que:

$$d = |Q| < \delta \quad \Rightarrow \quad \begin{cases} d < \delta_0,\ \delta_i \\ (i = 1,\ ...,\ n) \end{cases} \quad \Rightarrow \quad \begin{cases} \mu_i < \varepsilon/n \\ (i = 1,\ ...,\ n) \end{cases} \quad \Rightarrow \quad \mu < \mu_1 + \cdots + \mu_n < \varepsilon$$

como había que comprobar.

2.º (Lema 2.º). Sean $P = \{I_1, ..., I_n\}$ (fija) y $\boldsymbol{t} = \{\boldsymbol{t}_1, ..., \boldsymbol{t}_n\}$ ($\boldsymbol{t}$ recorriendo T), donde $\boldsymbol{t}_i \in I_i$ para $i = 1, ..., n$. Como $m_i \leqslant f(\boldsymbol{t}_i) \leqslant M_i$, donde $m_i = \text{ínf}\ f(I_i)$ y $M_i = \sup f(I_i)$, resulta evidente que

$$s(f, P) \leqslant \sigma(f, P, \boldsymbol{t}) \leqslant S(f, P), \qquad \forall\ \boldsymbol{t} \in T$$

luego $s(f, P)$ y $S(f, P)$ son cotas inferior y superior del conjunto de las sumas $\sigma(f, P, \boldsymbol{t})$, cuando $\boldsymbol{t} \in T$. Veamos que $s(f, P)$ es el ínfimo de dicho conjunto (igual se vería que $S(f, P)$ es su supremo), esto es, que dado $\varepsilon > 0$ (cualquiera) existe $\boldsymbol{t}^* \in T$ tal que $\sigma(f, P, \boldsymbol{t}^*) < s(f, P) + \varepsilon$. Así ocurre ya que, como

$$m_i = \text{ínf}\ \{f(\boldsymbol{t}_i)\ /\ \boldsymbol{t}_i \in I_i\}, \qquad \text{para } i = 1, ..., n$$

existe $t_i^* \in I_i$ tal que $f(t_i^*) < m_i + \varepsilon/\mu(I_i)$, pero $i = 1, ..., n$; por tanto, para la familia $t^* = \{t_1^*, ..., t_n^*\}$ es:

$$\sigma(f, P, t^*) = \sum_{i=1}^{n} f(t_i^*)\mu(I_i) < \sum_{i=1}^{n} \left(m_i + \frac{\varepsilon}{\mu(I_i)}\right)\mu(I_i) = s(f, P) + \varepsilon$$

con lo que concluye la demostración.

3.º (La condición «$\varepsilon:\delta$» es necesaria para la integrabilidad). En el supuesto de que f es integrable en I, dado $\varepsilon > 0$ (cualquiera), y según la segunda condición «$\varepsilon:P$» (véase [53], 1.º), se sabe que existe una cierta partición P_0 de I tal que $\int_I f - s(f, P_0) < \varepsilon/2$ y $S(f, P_0) - \int_I f < \varepsilon/2$. De acuerdo con el anterior lema 1.º, existe $\delta > 0$ tal que, para toda partición P de I que tenga diámetro menor que δ, la suma de los contenidos de los intervalos de P que no están incluidos en algún intervalo de P_0 es menor que $\varepsilon/2K$, donde K es una cota superior de $|f|$ en I. Para cualquiera que sea la partición $P = \{I_1, ..., I_n\}$ de I con diámetro menor que δ, situando en primer lugar sus intervalos $I_1, ..., I_h$ que están incluidos en algún intervalo de P_0 y después los $I_{h+1}, ..., I_n$ que no lo están, es evidente que, para cualquiera que sea la familia de puntos intermedios $t = \{t_1, ..., t_n\}$, con $t_i \in I_i$, es:

$$\sigma(f, P, t) = \sum_{i=1}^{h} f(t_i)\mu(I_i) + \sum_{i=h+1}^{n} f(t_i)\mu(I_i) \quad \begin{cases} \leqslant S(f, P_0) + K(\varepsilon/2K) \\ \geqslant s(f, P_0) - K(\varepsilon/2K) \end{cases}$$

Resulta entonces que, si es $|P| < \delta$ y para cualquiera que sea la correspondiente familia t, es:

$$-\varepsilon + \int_I f \leqslant -\frac{\varepsilon}{2} + s(f, P_0) \leqslant \sigma(f, P, t) \leqslant S(f, P_0) + \frac{\varepsilon}{2} < \int_I f + \varepsilon$$

con lo que $|\int_I f - \sigma(f, P, t)| < \varepsilon$, como había que comprobar.

4.º (La condición «$\varepsilon:\delta$» es suficiente para la integrabilidad). Sabiendo que se verifica la anterior condición «$\varepsilon:\delta$», vamos a comprobar la integrabilidad de f en I, viendo que es $A = \int_I f$, para lo que veremos que satisface la segunda condición «$\varepsilon:P$» de integrabilidad (véase [53], 2.ª). Dado $\varepsilon > 0$ (cualquiera), comprobaremos, pues, que existe una partición P de I tal que $A - s(f, P) < \varepsilon$ y $S(f, P) - A < \varepsilon$; solamente verificaremos que se cumple la primera de estas desigualdades (para la segunda se razona de manera análoga). En efecto: por verificarse la condición «$\varepsilon:\delta$», se sabe que existe una partición P de I tal que $|A - \sigma(f, P, t)| < \varepsilon/2$ para cualquier $t \in T$ (se sabe más: esta desigualdad se verifica para toda partición P de I cuyo diámetro sea menor que un cierto $\delta > 0$). Por otra parte, como $s(f, P)$ es el ínfimo del conjunto $\{\sigma(f, P, t) \,/\, t \in T\}$, se sabe que existe una familia $t_0 \in T$ tal que

$$\sigma(f, P, t_0) - \varepsilon/2 < s(f, P)$$

En consecuencia, se puede asegurar que:

$$A - s(f, P) < A - \sigma(f, P, t_0) + \frac{\varepsilon}{2} \leqslant |A - \sigma(f, P, t_0)| + \frac{\varepsilon}{2} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

$[58]_1$ Observación

En lo que acabamos de decir acerca de las sumas de Riemann, en [58], es igualmente válido si los valores $f(t_i)$ se sustituyen por cualesquiera valores $k_i \in [m_i, M_i]$; esto es, si se define: suma de Riemann (de una función acotada $f: I \to \mathbb{R}$; $I \subset \mathbb{R}^p$ intervalo compacto) correspondiente a una partición $P = \{I_1, ..., I_n\}$ de I y a un conjunto de valores $k = \{k_1, ..., k_n\}$, donde $m_i \leqslant k_i \leqslant M_i$ (m_i y M_i son el ínfimo y el supremo de f en I_i) para $i = 1, ..., n$, es el número

$$\sigma(f, P, k) = \sum_{i=1}^{n} k_i \mu(I_i) \qquad (\mu(I_i) = \text{medida de } I_i)$$

Para esta generalización de las sumas de Riemann, no sólo siguen verificándose las propiedades ya estudiadas de dichas sumas, sino que también se cumplen todas las que luego se estudian sobre ellas.

$[58]_2$ Interpretación geométrica de la integral doble

Sea $f: I \to \mathbb{R}$ una función acotada e integrable en el intervalo compacto $I \subset \mathbb{R}^2$. Supongamos que f es positiva, o sea, que $f(x, y) \geqslant 0$ para todo $(x, y) \in I$. Llamemos «conjunto situado bajo la superficie $z = f(x, y)$» a

$$C = \{(x, y, z) \in \mathbb{R}^3 \ / \ (x, y) \in I, \ 0 \leqslant z \leqslant f(x, y)\}$$

Para cada partición $P = \{I_1, ..., I_n\}$ de I y cada familia de puntos intermedios $\boldsymbol{t} = \{\boldsymbol{t}_1, ..., \boldsymbol{t}_n\}$, con $\boldsymbol{t}_i \in I_i$ para $i = 1, ..., n$, consideramos la suma de Riemann

$$\sigma(f, P, \boldsymbol{t}) = \sum_{i=1}^{n} f(\boldsymbol{t}_i)\alpha(I_i), \qquad \text{donde } \alpha(I_i) \text{ es el área de } I_i$$

En ella, el sumando $f(\boldsymbol{t}_i)\alpha(I_i)$ es el volumen de un prisma rectangular cuya base es el intervalo (rectángulo) I_i y cuya altura es un valor intermedio de los que alcanza la función f en los puntos de I_i. Así pues, la suma $\sigma(f, P, \boldsymbol{t})$ puede tomarse como una aproximación del volumen de C; a medida que se toman particiones con mayor número de subintervalos y éstos con menor diámetro, se obtendrá un valor para $\sigma(f, P, \boldsymbol{t})$ que deberá ser cada vez mejor aproximación de dicho volumen. Por ello, y a la vista del teorema de Riemann (véase [58]), lo razonable es acordar que la integral de f en I sea el volumen del conjunto C.

CONDICIÓN «$\varepsilon : \delta$» DE INTEGRABILIDAD (EN UN INTERVALO)

[59]

Si $f: I \to \mathbb{R}$ es una función acotada en el intervalo compacto $I \subset \mathbb{R}^p$, se verifican entonces que:

1.° La función f es integrable en I si, y sólo si, se verifica la siguiente condición (1.ª condición «$\varepsilon : \delta$»): para cada $\varepsilon > 0$ existe un $\delta > 0$ tal que, para toda partición P de I cuyo diámetro sea menor que δ, es $S(f, P) - s(f, P) < \varepsilon$.

2.° La función f es integrable en I y su integral en I es el número A si, y sólo si, se verifica la siguiente condición (2.ª condición «$\varepsilon : \delta$»): para cada $\varepsilon > 0$ existe un $\delta > 0$ tal que, para toda partición P de I cuyo diámetro sea menor que δ, es $A - s(f, P) < \varepsilon$ y $S(f, P) - A < \varepsilon$.

3.° La función f es integrable en I y su integral en I es el número A si, y sólo si, se verifica la siguiente condición (3.ª condición «$\varepsilon : \delta$»): para cada $\varepsilon > 0$ existe un $\delta > 0$ tal que, para toda partición P de I cuyo diámetro sea menor que δ, es $|A - \sigma(f, P, t)| < \varepsilon$ para cualquiera que sea la correspondiente familia t de puntos intermedios.

Demostración

a) La tercera condición «$\varepsilon : \delta$» es, en efecto, condición de integrabilidad, pues ella no es otra cosa que el anterior teorema de Riemann (véase [58]), que ya está demostrado.

b) Comprobemos que la segunda condición «$\varepsilon : \delta$» es equivalente a la tercera condición «$\varepsilon : \delta$», con lo que, como aquella es condición de integrabilidad, también lo será ésta:

- (3.ª $\Rightarrow$ 2.ª) Por verificarse la 3.ª, dado $\varepsilon > 0$, existe $\delta > 0$ tal que

$$|P| < \delta \quad \Rightarrow \quad |A - \sigma(f, P, t)| < \varepsilon/2, \qquad \forall\, t \in T(P)$$

(donde $T(P)$ es el conjunto de las familias de puntos intermedios correspondientes a la partición P). Por otra parte, como $s(f, P)$ es el ínfimo del conjunto $\{\sigma(f, P, t) \,/\, t \in T(P)\}$, existe algún $t_0 \in T(P)$ tal que

$$\sigma(f, P, t_0) - \varepsilon/2 < s(f, P)$$

Por tanto, de lo anterior se desprende que

$$|P| < d \quad \Rightarrow \quad A - s(f, P) < A - \sigma(f, P, t_0) + \frac{\varepsilon}{2} \leqslant |A - \sigma(f, P, t_0)| + \frac{\varepsilon}{2} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

como había que comprobar (la relación $S(f, P) - A < \varepsilon$ se comprueba análogamente).

- (2.ª $\Rightarrow$ 3.ª) Por verificarse la 2.ª, dado $\varepsilon > 0$, existe $\delta > 0$ tal que para toda partición P cuyo diámetro sea menor que δ se verifica que $A - s(f, P) < \varepsilon$ y $S(f, P) - A < \varepsilon$ y, por tanto, es:

$$|P| < \delta \quad \Rightarrow \quad S(f, P) - \varepsilon < A < s(f, P) + \varepsilon \qquad [1]$$

Como sabemos que, para toda partición P y familia $t \in T(P)$, se verifica que

$$-S(f, P) < -\sigma(f, P, t) < -s(f, P) \qquad [2]$$

sumando las desigualdades [1] y [2], concluimos que para toda $t \in T(P)$ es:

$$|P| < \delta \quad \Rightarrow \quad -\varepsilon < A - \sigma(f, P, t) < \varepsilon \quad \Rightarrow \quad |A - \sigma(f, P, t)| < \varepsilon$$

c) Ocupémonos finalmente de la 1.ª condición «$\varepsilon : \delta$».

- Si se verifica la 1.ª condición «$\varepsilon : \delta$», entonces se verifica, evidentemente la 1.ª condición «$\varepsilon : P$» (véase [53], 1.º), por lo que se puede asegurar que f es integrable en I.
- Recíprocamente, si f es integrable en I, entonces ya sabemos que se verifica la 2.ª condición «$\varepsilon : \delta$», por lo que dado $\varepsilon > 0$ existe un $\delta > 0$ tal que

$$|P| < \delta \quad \Rightarrow \quad \int_I f - s(f, P) < \varepsilon/2 \qquad \text{y} \qquad S(f, P) - \int_I f < \varepsilon/2$$

Sumando estas dos desigualdades, se obtiene que

$$|P| < \delta \quad \Rightarrow \quad S(f, P) - s(f, P) < \varepsilon$$

es decir, se verifica la primera condición «$\varepsilon : \delta$».

[59]$_1$ La integral (en un intervalo) como límite de sumas

Sea $f: I \to \mathbb{R}$ una función acotada en el intervalo compacto $I \subset \mathbb{R}^p$. Considérese una sucesión cualquiera, $P_1, \ldots, P_n, \ldots$, de particiones de I cuyos diámetros tienden a cero (cuando $n \to \infty$), esto es, tal que lím $|P_n| = 0$. Se verifica que:

1.º La función f es integrable en I si, y sólo si, acontece que:

$$\lim_{n \to 0} \left(S(f, P_n) - s(f, P_n)\right) = 0$$

Además, si f es integrable en I, su integral vale:

$$\int_I f = \lim_{n \to 0} S(f, P_n) = \lim_{n \to 0} S(f, P_n)$$

2.º Si f es integrable en I y para cualquiera que sea la familia t_n de puntos intermedios correspondiente a la partición P_n, para $n \in \mathbb{N}$, es:

$$\int_I f = \lim_{n \to 0} \sigma(f, P_n, t_n)$$

Demostración

1.º Supongamos primero que f es integrable en I. Dado $\varepsilon > 0$ (cualquiera), se sabe que (véase la 1.ª condición «$\varepsilon:\delta$» de integrabilidad, [59], 1.º) existe $\delta > 0$ tal que $S(f, P) - s(f, P) < \varepsilon$ siempre que sea $|P| < \delta$. Por otra parte, como $|P_n| \to 0$, existe $\nu \in \mathbb{N}$ tal que $n \geqslant \nu \Rightarrow |P_n| < \delta$. De ahí que:

$$n \geqslant \nu \quad \Rightarrow \quad 0 < S(f, P_n) - s(f, P_n) < \varepsilon$$

luego $S(f, P_n) - s(f, P_n)$ tiene límite cero, como había que probar. Recíprocamente, suponiendo ahora que $S(f, P_n) - s(f, P_n)$ tiende a cero, como se verifica que para $n \in \mathbb{N}$ es:

$$0 \leqslant \overline{\int_I} f - \underline{\int_I} f \leqslant S(f, P_n) - s(f, P_n) \to 0$$

ha de ser entonces $\overline{\int}_I f - \underline{\int}_I f = 0$, luego f es integrable en I. En el supuesto de que f es integrable, como se sabe que, para $n \in \mathbb{N}$ es:

$$0 \leqslant \int_I f - s(f, P_n) \leqslant S(f, P_n) - s(f, P_n) \qquad \text{y} \qquad 0 \leqslant S(f, P_n) - \int_I f \leqslant S(f, P_n) - s(f, P_n)$$

tomando límites para $n \to \infty$ y acudiendo a la regla del emparedado, obtenemos que

$$\text{lím}\left(\int_I f - s(f, P_n)\right) = 0 \qquad \text{y} \qquad \text{lím}\left(S(f, P_n) - \int_I f\right) = 0$$

esto es, que $s(f, P_n)$ y $S(f, P_n)$ tienden a la integral, como había que comprobar.

2.º Según se acaba de probar, $s(f, P_n)$ y $S(f, P_n)$ convergen, ambas, hacia $\int_I f$. Como además es

$$s(f, P_n) \leqslant \sigma(f, P_n, t_n) \leqslant S(f, P_n), \qquad \forall\, n \in \mathbb{N}$$

la regla del emparedado permite afirmar que también $\sigma(f, P_n, t_n)$ converge hacia $\int_I f$, como había que comprobar.

INTEGRACIÓN ITERADA (REDUCCIÓN A INTEGRALES SIMPLES)

La integral de una función $f: I \to \mathbb{R}$, en un intervalo compacto $I \subset \mathbb{R}^p$, pongamos $I = [a_1, b_1] \times \cdots \times [a_p, b_p]$, puede expresarse en función de p integrales simples sucesivas, en los intervalos $[a_1, b_1], \ldots, [a_p, b_p]$.

[60]

TEOREMA DE FUBINI. Sea $f: I \to \mathbb{R}$ una función acotada en el intervalo compacto I. Pongamos I en la forma $I = I' \times I''$, donde $I' \subset \mathbb{R}^r$ e $I'' \subset \mathbb{R}^{p-r}$ (con $0 < r < p$); todo $\boldsymbol{x} \in I$ se expresará poniendo $\boldsymbol{x} = (\boldsymbol{x}', \boldsymbol{x}'')$, con $\boldsymbol{x}' \in I'$ y $\boldsymbol{x}'' \in I''$. Para cada $\boldsymbol{x}' \in \boldsymbol{I}'$, sea $f_{x'}: I'' \to \mathbb{R}$ la función definida por $f_{x'}(\boldsymbol{x}'') = f(\boldsymbol{x}', \boldsymbol{x}'')$; llamemos $F_*: I' \to \mathbb{R}$ y $F^*: I' \to \mathbb{R}$ a las funciones definidas por

$$F_*(\boldsymbol{x}') = \underline{\int}_{I''} f_{x'} \qquad \text{y} \qquad F^*(\boldsymbol{x}') = \overline{\int}_{I''} f_{x'}, \qquad \forall\, \boldsymbol{x}' \in I'$$

Si f es integrable en I, entonces F_* y F^* son integrables en I' y se verifica que

$$\int_I f = \int_{I'} F_* \qquad \text{e} \qquad \int_I f = \int_{I'} F^* \tag{*}$$

CÁLCULO DE INTEGRALES MÚLTIPLES. Aplicando reiteradamente el resultado anterior a la función integrable $f: I \to \mathbb{R}$ y si $I \subset \mathbb{R}^p$ es el intervalo compacto $I = [a_1, b_1] \times \cdots \times [a_p, b_p]$, se obtiene que:

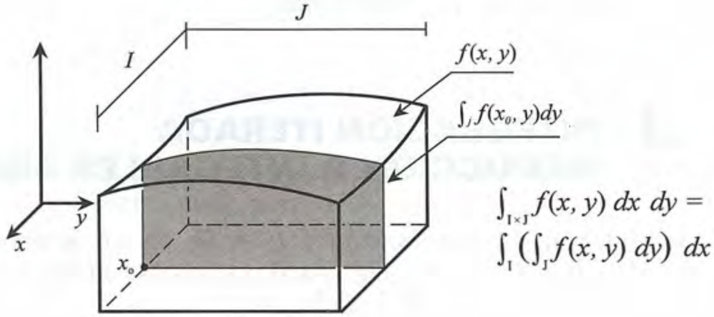
$$\int_I f(x_1, ..., x_p)\, dx_1 \ldots dx_p = \int_{a_1}^{b_1} dx_1 \left\{ \int_{a_2}^{b_2} dx_2 \left(\cdots \left[\int_{a_p}^{b_p} f(x_1, x_2, ..., x_p)\, dx_p \right] \right) \right\} \tag{**}$$

(*) Este resultado se suele expresar poniendo que, si $I = I' \times I''$ y si $\boldsymbol{x} = (\boldsymbol{x}', \boldsymbol{x}'')$, es

$$\int_I f(\boldsymbol{x})\, d\boldsymbol{x} = \int_{I'} \left(\underline{\int}_{I''} f(\boldsymbol{x}', \boldsymbol{x}'')\, d\boldsymbol{x}'' \right) d\boldsymbol{x}' = \int_{I'} \left(\overline{\int}_{I''} f(\boldsymbol{x}', \boldsymbol{x}'')\, d\boldsymbol{x}'' \right) d\boldsymbol{x}'$$

(**) Si el correspondiente integrando no es integrable, el símbolo $\int$ debe sustituirse por cualquiera de los $\underline{\int}$ o $\overline{\int}$. Se denota por $\int_a^b F(t_1, ..., t_i)\, dt_i$ a $\int_a^b F_{t_1 \ldots t_{i-1}}(t_i)\, dt_i$, donde $F_{t_1 \ldots t_{i-1}}$ es la función definida en $[a, b]$ mediante $t_i \mapsto F_{t_1 \ldots t_{i-1}}(t_i) = F(t_1, ..., t_{i-1}, t_i)$. Para $p = 3$ será

$$\int_I f(x, y, z)\, dx\, dy\, dz = \int_{a_1}^{b_1} dx \left\{ \int_{a_2}^{b_2} dy \left(\int_{a_3}^{b_3} f(x, y, z)\, dz \right) \right\}$$



Demostración

Para cualesquiera que sean las particiones $P' = \{I'_1, ..., I'_n\}$ de I' y $P'' = \{I''_1, ..., I''_m\}$ de I'' los $n \cdot m$ intervalos $I'_i \times I''_j$ (para $i = 1, ..., n$ y $j = 1, ..., m$) forman una partición de $I = I' \times I''$, a la que llamaremos $P = \{I_1, ..., I_{nm}\}$. Si llamamos $m_k = \inf f(I_k)$ o $m_{i,j} = \inf f(I'_i \times I''_j)$, es evidente que

$$s(f, P) = \sum_{k=1}^{nm} m_k \mu(I_k) = \sum_{i,j=1}^{n,m} m_{i,j}\mu(I'_i \times I''_j) = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} m_{i,j}\mu(I''_j) \right) \mu(I'_i) \qquad [1]$$

(donde $\mu(I'_i)$ y $\mu(I''_j)$ denotan a las medidas de I'_i e I''_j en cuanto que son intervalos de $\mathbb{R}^r$ y de $\mathbb{R}^{p-r}$, respectivamente). Ahora bien, para cualquiera que sea $\boldsymbol{x}' \in I'_i$ y llamando $m''_j(\boldsymbol{x}') = = \inf f_{x'}(I''_j)$, es evidente que $m_{i,j} \leqslant m''_j(\boldsymbol{x}')$, por lo que para todo $\boldsymbol{x}' \in I'_i$ es:

$$\sum_{j=1}^{m} m_{i,j}\mu(I''_j) \leqslant \sum_{j=1}^{m} m''_j(\boldsymbol{x}')\mu(I''_j) = s(f_{x'}, P'') \leqslant \underline{\int}_{I''} f_{x'} = F_*(\boldsymbol{x}')$$

Tomando ínfimos para $\boldsymbol{x}' \in I'_i$, de esta última desigualdad se deduce que:

$$\sum_{j=1}^{m} m_{i,j}\mu(I''_j) \leqslant \inf F_*(I'_i)$$

Llevando este último resultado a [1], se obtiene:

$$s(f, P) \leqslant \sum_{i=1}^{n} \inf F_*(I'_i)\mu(I'_i) = s(F_*, P')$$

Razonando de manera análoga (considerando los supremos en lugar de los ínfimos), se obtiene que $S(f, P) \geqslant S(F^*, P')$. De estas desigualdades y como $F_* \leqslant F^*$, se desprende que

$$\begin{aligned} s(f, P) \leqslant s(F_*, P') \leqslant S(F_*, P') \leqslant S(F^*, P') \leqslant S(f, P) \\ s(f, P) \leqslant s(F_*, P') \leqslant s(F^*, P') \leqslant S(F^*, P') \leqslant S(f, P) \end{aligned} \qquad [2]$$

Dado $\varepsilon > 0$, como f es integrable en I, según la segunda condición «$\varepsilon : \delta$» de integrabilidad (véase [59], 2.º) existe $\delta > 0$ tal que siempre que sea $|P| < \delta$, se verifica que

$$\int_I f - s(f, P) < \varepsilon \qquad \text{y} \qquad S(f, P) - \int_I f < \varepsilon \qquad [3]$$

Por tanto, si se toma P' de manera que sea $|P'| < \delta/2$, recurriendo a una P'' con $|P''| < \delta/2$ con lo que $|P| < |P'| + |P''| < \delta$, de [2] y [3] se obtiene que:

$$\int_I f - s(F_*, P') \leqslant \int_I f - s(f, P) < \varepsilon, \quad S(F_*, P') - \int_I f \leqslant S(f, P) - \int_I f < \varepsilon$$

$$\int_I f - s(F^*, P') \leqslant \int_I f - s(f, P) < \varepsilon, \quad S(F^*, P') - \int_I f \leqslant S(f, P) - \int_I f < \varepsilon$$

luego la segunda condición «$\varepsilon : \delta$» de integrabilidad permite asegurar que F_* y F^* son inegrables en I' y que $\int_{I'} F^*$ e $\int_{I'} F_*$ son ambas iguales a $\int_I f$, como había que comprobar.

[60]$_1$ Ejercicios

Calcular las siguientes integrales (doble y triple):

$$A = \iint_I x^2 \operatorname{sen}(xy)\,dx\,dy, \qquad I = [0, \pi] \times [0, 1]$$

$$B = \iiint_I \frac{dx\,dy\,dz}{\sqrt{x + y + z + 2}}, \qquad I = [0, 2] \times [0, 1] \times [-1, 4]$$

Resolución

Acudiendo a la integración iterada, se obtiene que:

$$A = \int_0^\pi dx \int_0^1 x^2 \operatorname{sen}(xy)\,dy = \int_0^\pi dx[-x\cos(xy)]_{y=0}^{y=1} = \int_0^\pi (x - x\cos x)\,dx = \\ = \left[\frac{x^2}{2} - \cos x - x \operatorname{sen} x\right]_0^\pi = 2 + \frac{\pi^2}{2}$$

$$B = \int_{-1}^4 dz \int_0^1 dy \int_0^1 \frac{dx}{\sqrt{x + y + z + 2}} = \int_{-1}^4 dz \int_0^1 dy\left[2\sqrt{x + y + z + 2}\right]_{x=0}^{x=2} = \\ = \int_{-1}^4 dz \int_0^1 2\sqrt{y + z + 4}\,dy = \int_{-1}^4 dz\left[\frac{4}{3}(y + z + 4)^{3/2}\right]_{y=0}^{y=1} = \\ = \int_{-1}^4 \frac{4}{3}(z + 5)^{3/2}\,dz = \left[\frac{8}{15}(z + 5)^{5/2}\right]_{-1}^4 = \frac{8}{15}(9^{5/2} - 4^{5/2}) = \frac{1688}{15}$$

[60]$_2$ Generalización de la regla de Barrow

Sea $f: I \to \mathbb{R}$ una función acotada en el intervalo compacto $I = [a, b] \times [c, d]$ de $\mathbb{R}^2$. Si f es integrable en I y si para cada $y \in [c, d]$ la función $[a, b] \to \mathbb{R}$, $x \mapsto F(x, y)$ es una primitiva de la función $[a, b] \to \mathbb{R}$, $x \mapsto f(x, y)$ (esto es, si $\frac{\partial}{\partial x} F(x, y) = f(x, y)$ para todo $(x, y) \in I$), entonces la función $y \mapsto F(b, y) - F(a, y)$ es integrable en $[c, d]$ y se verifica que:

$$\int_a^b \int_c^d f(x, y)\,dx\,dy = \int_c^d (F(b, y) - F(a, y))\,dy$$

Demostración

Sean $P_x = \{x_0, x_1, ..., x_n\}$ y $P_y = \{y_0, y_1, ..., y_m\}$ particiones cualesquiera de $[a, b]$ y de $[c, d]$, sea $y^* = \{y_1^*, ..., y_m^*\}$ un conjunto de puntos intermedios asociado a la partición P_y y llamemos $[P_x, P_y, y^*]$ a la suma

$$[P_x, P_y, y^*] = \sum_{j=1}^{m} \sum_{i=1}^{n} (F(x_i, y_j^*) - F(x_{i-1}, y_j^*))(y_j - y_{j-1}) \qquad [1]$$

Nótese en primer lugar que $[P_x, P_y, y^*]$ se puede poner en la forma

$$\begin{aligned}[P_x, P_y, y^*] &= \sum_{j=1}^{m} \left\{ \sum_{i=1}^{n} (F(x_i, y_j^*) - F(x_{i-1}, y_j^*)) \right\} (y_j - y_{j-1}) = \\ &= \sum_{j=1}^{m} (F(b, y_j^*) - F(a, y_j^*))(y_j - y_{j-1}) = \sigma(G, P_y, y^*)\end{aligned} \qquad [2]$$

que es la suma de Riemann, en $[c, d]$, de la función $y \mapsto G(y) = F(b, y) - F(a, y)$ correspondiente a la partición P_y y a la familia de puntos intermedios y^*.

Por otra parte, para cada $y_j^* \in y^*$, a la función $x \mapsto F(x, y_j^*)$ le es de aplicación en $[x_{i-1}, x_i]$ (para $i = 1, ..., n$) el teorema de los incrementos finitos, del que se obtiene que existe, para cada $j = 1, ..., m$, un punto $x_{ij} \in]x_{i-1}, x_i[$ tal que

$$F(x_i, y_j^*) - F(x_{i-1}, y_j^*) = \frac{\partial}{\partial x} F(x_{ij}^*, y_j^*)(x_i - x_{i-1}) = f(x_{ij}^*, y_j^*)(x_i - x_{i-1})$$

Llevando este resultado a [1], se obtiene que

$$[P_x, P_y, y^*] = \sum_{j=1}^{m} \sum_{i=1}^{n} f(x_{ij}^*, y_j^*)(x_i - x_{i-1})(y_j - y_{j-1}) = \sigma(f, P_{xy}, t^*) \qquad [3]$$

que es la suma de Riemann, en I, de la función f correspondiente a la partición $P_{xy} = \{[x_{i-1}, x_i] \times [y_{j-1}, y_j] \;/\; i = 1, ..., n; j = 1, ..., m\}$ y a la familia de puntos intermedios $t^* = \{(x_{ij}^*, y_j^*) \;/\; i = 1, ..., n; j = 1, ..., m\}$.

Si tomamos ahora las particiones P_x y P_y cada vez más numerosas, haciendo que $|P_x| \to 0$ y $|P_y| \to 0$, como f es integrable en I, de [3] se desprende que $[P_x, P_y, y^*]$ tiende entonces (para cualquiera que sean los y^*) hacia $\int_I f$. Llevando este resultado a [2], nos encontramos con que $\sigma(G, P_y, y^*)$ tiende hacia $\int_I f$ cuando $|P_y| \to 0$ y para cualesquiera que sean los correspondientes conjuntos de puntos intermedios y^*, esto es, G es integrable en $[c, d]$ y su integral es igual a $\int_I f$, como había que comprobar.

4.2. CLASES DE FUNCIONES INTEGRABLES EN UN INTERVALO

Hasta este momento, para averiguar si una función dada $f: I \to \mathbb{R}$ (acotada) es integrable en el intervalo compacto I, disponemos de la propia definición de integral (las integrales superior e inferior son iguales; véase [53]), las condiciones «$\varepsilon : P$» de integrabilidad (véase [53]) y las condiciones «$\varepsilon : \delta$» de integrabilidad (véase [59]). Todos estos criterios son, generalmente, poco prácticos; con ellos es difícil dilucidar si una función dada es o no integrable en un intervalo. Vamos a ocuparnos aquí de un nuevo criterio de integrabilidad, de muy distinta índole que las anteriores, que pone de manifiesto que hay una estrecha relación entre la continuidad y la integrabilidad.

Según ya vimos anteriormente (en [53]), si f es continua en el intervalo compacto I entonces f es integrable en I. Ahora vamos a comprobar que las funciones que son integrables en I son aquellas, y sólo aquellas, que son continuas en «casi todos los puntos» de I; más exactamente, las que son continuas en I excepto, a lo más, en un subconjunto suyo «que tenga medida nula», llamando de este modo a un conjunto que puede ser recubierto por una cantidad numerable de intervalos compactos cuyos contenidos sumen menos que una cantidad positiva que arbitrariamente se prefije.

Antes de probar dicho resultado (debido a Lebesgue), obtendremos otro, de menor alcance, pero suficiente en muchas ocasiones y mucho más fácil de probar, que asegura que una función continua en I excepto en un subconjunto suyo de «contenido nulo» es una función integrable en I, llamando conjunto de «contenido nulo» a un conjunto que puede ser recubierto por una cantidad finita de intervalos compactos cuyos contenidos sumen menos que una cantidad positiva que arbitrariamente se prefije.

❑ CONJUNTOS DE CONTENIDO NULO

[61] Un conjunto $C \subset \mathbb{R}^p$ se dice que tiene *contenido nulo* si, para cada $\varepsilon > 0$, existe un recubrimiento finito de C formado por intervalos compactos cuyas medidas sumen menos que ε, esto es, si existe una familia finita, $I_1, \ldots, I_n$, de intervalos compactos de $\mathbb{R}^p$ tales que (si $\mu(I_i)$ denota la medida de I_i):

$$C \subset I_1 \cup \cdots \cup I_n \qquad \text{y} \qquad \mu(I_1) + \cdots + \mu(I_n) < \varepsilon$$

Propiedades:

1. Un subconjunto de un conjunto de contenido nulo tiene contenido nulo.
2. La unión finita de conjuntos de contenido nulo tiene contenido nulo.
3. Todo conjunto finito tiene contenido nulo.
4. Todo intervalo degenerado de $\mathbb{R}^p$ tiene contenido nulo (en $\mathbb{R}^p$).

Demostración

1. Esta propiedad es evidente (tómese para el subconjunto el mismo recubrimiento que para el conjunto).
2. Si $C_1, ..., C_r$ son conjuntos de contenido nulo y dado $\varepsilon > 0$, entonces para cada $i = 1, ..., r$ existe un recubrimiento finito $\mathscr{R}_i$ de C_i formado por intervalos cerrados cuyas medidas suman menos que ε/r. Por tanto, $\mathscr{R}_1 \cup \cdots \cup \mathscr{R}_r$ es un recubrimiento finito de $C = C_1 \cup \cdots \cup C_r$ formado por intervalos cerrados cuyas medidas suman menos que ε, luego C tiene contenido nulo.
3. Si C es un conjunto finito, $C = \{\boldsymbol{a}_1, ..., \boldsymbol{a}_r\}$ y dado $\varepsilon > 0$, entonces para cada $i = 1, ..., r$ consideramos un intervalo cerrado I_i centrado en $\boldsymbol{a}_i$ y cuyos lados son todos menores que $(\varepsilon/r)^{1/p}$, verificándose, pues, que $\mu(I_i) < \varepsilon/r$, con lo cual $\{I_1, ..., I_r\}$ es un recubrimiento finito de C formado por intervalos cerrados cuyas medidas suman menos que ε, luego C tiene contenido nulo.
4. Un intervalo compacto degenerado de $\mathbb{R}^p$ es un conjunto de la forma $I = I_1 \times \{a\}$, donde I_1 es un intervalo compacto de $\mathbb{R}^{p-1}$ y $a \in \mathbb{R}$; sea μ_1 la medida en $\mathbb{R}^{p-1}$ de I_1. Dado $\varepsilon > 0$, consideremos el intervalo $J = I_1 \times [a - \varepsilon/(3\mu_1), a + \varepsilon/(3\mu_1)]$; es evidente que $\{J\}$ es un recubrimiento de I y que la medida de J es $\mu(J) = (2/3)\varepsilon < \varepsilon$, luego I tiene contenido nulo.

[61]$_1$ Observaciones

1.ª En la anterior definición [61] de conjunto de contenido nulo, los intervalos compactos $I_1, I_2, ..., I_n$ pueden sustituirse por intervalos abiertos y acotados.

2.ª En la anterior definición de conjunto de contenido nulo (véase [61]), los intervalos compactos $I_1, ..., I_n$ se pueden tomar «equiláteros» (con sus lados de igual longitud) y todos ellos de igual tamaño.

Comprobación

1.ª Para comprobar esta propiedad, veamos previamente que se verifica lo que sigue: dado un intervalo compacto cualquiera

$$I = [a_1, b_1] \times \cdots \times [a_p, b_p]$$

para cada $\delta > 0$ existe un $\gamma > 0$ tal que, llamando

$$I_\gamma =]a_1 - \gamma, b_1 + \gamma[\times \cdots \times]a_p - \gamma, b_p + \gamma[$$

(nótese que $I \subset I_\gamma$ para todo $\gamma > 0$, con lo que $\mu(I) < \mu(I_\gamma)$), las medidas de I e I_γ son tales que

$$\mu(I_\gamma) < \mu(I) + \delta$$

Así ocurre, en efecto, pues ello se desprende obviamente de los siguientes hechos:

$$\mu(I) = \prod_{i=1}^{n} (b_i - a_i), \qquad \mu(I_\gamma) = \prod_{i=1}^{n} (b_i - a_i + 2\gamma)$$

la función $\gamma \mapsto \mu(I_\gamma)$ es continua en $\gamma = 0$ y $\mu(I_0) = \mu(I)$.

Visto esto, comprobemos ya la propiedad 1.ª. Si C es un conjunto de contenido nulo (utilizando en la definición intervalos compactos) y dado $\varepsilon > 0$, existe una familia finita de intervalos compactos, $I_1, ..., I_n$, que recubren C y cuyas medidas suman menos de $\varepsilon/2$. Según la proposición anterior, existen ciertos intervalos abiertos y acotados $I'_1, ..., I'_n$ tales que $I_i \subset I'_i$ y $\mu(I'_i) < \mu(I_i) + \varepsilon/(2n)$ (para todo $i = 1, ..., n$), con lo que $\{I'_1, ..., I'_n\}$ es un recubrimiento finito de C formado por intervalos abiertos y acotados y tal que

$$\mu(I'_1) + \cdots + \mu(I'_n) < \mu(I_1) + \frac{\varepsilon}{2n} + \cdots + \mu(I_n) + \frac{\varepsilon}{2n} < \frac{\varepsilon}{2} + n\,\frac{\varepsilon}{2n} = \varepsilon$$

como había que comprobar. El recíproco es evidente: si $\{J_1, ..., J_n\}$ es recubrimiento finito de C mediante intervalos abiertos y acotados cuyas medidas suman menos que un $\varepsilon > 0$ dado, es evidente que $\{\bar{J}_1, ..., \bar{J}_n\}$ (donde $\bar{J}_i$ es la adherencia de J_i) es recubrimiento finito de C mediante intervalos compactos cuyas medidas (iguales a las de J_i) suman menos que ε.

2.ª Para comprobar esta propiedad, veamos previamente que se verifica lo que sigue: si J es un intervalo compacto de $\mathbb{R}^p$ que incluye a otro intervalo compacto I, pongamos $I = [a_1, b_1] \times \cdots \times [a_p, b_p]$, entonces para cada $\delta > 0$ existe un $\gamma > 0$ tal que, para toda partición P de J que tenga diámetro menor que γ, llamando $J_1, ..., J_m$ a los intervalos de P que tienen algún punto de I, se verifica que

$$\mu(J_1) + \cdots + \mu(J_m) < \mu(I) + \delta$$

Así ocurre, en efecto: Llamando I_γ al intervalo

$$I_\gamma = [a_1 - \gamma, b_1 + \gamma] \times \cdots \times [a_p - \gamma, b_p + \gamma]$$

y si $|P| < \gamma$, entonces los intervalos $J_1, ..., J_n$ de P que tienen algún punto de I están incluidos en I_γ, con lo que $\mu(J_1) + \cdots + \mu(J_m) \leqslant \mu(I_\gamma)$. Como la función $\gamma \mapsto \mu(I_\gamma)$ es evidentemente continua en $\gamma = 0$ y como $\mu(I_0) = \mu(I)$, resulta que, dado $\delta > 0$, existe un $\gamma > 0$ tal que siempre que sea $|P| < \gamma$ es $\mu(J_1) + \cdots + \mu(J_m) < \mu(I) + \delta$.

Probemos ya la propiedad 2.ª. Si C es un conjunto de contenido nulo, dado $\varepsilon > 0$, existe una familia finita de intervalos compactos $I_1, ..., I_n$ tales que recubren C y cuyas medidas suman menos de $\varepsilon/2$. Apliquemos ahora el resultado anterior a cada uno de los intervalos $I_1, ..., I_n$ (esto es, tomando $I = I_i$), para lo que acudimos a un intervalo compacto J que incluya a todos los I_i, el cual puede definirse equilátero, por lo que admite, pues, particiones formadas por intervalos equiláteros todos de igual tamaño. Así que, tomando $\delta = \varepsilon/(2n)$, existen los correspondientes γ_i (uno por cada intervalo I_i);

si es $\gamma = \text{mín}\{\gamma_1, ..., \gamma_n\}$, entonces para toda partición P de J en intervalos equiláteros de igual tamaño de lado menor que $\gamma/\sqrt{p}$, con lo que $|P| < \gamma$, si son $I'_1, ..., I'_m$ los intervalos de P que contienen algún punto de algunos de los intervalos $I_1, ..., I_n$, se verifica que

$$\mu(I'_1) + \cdots + \mu(I'_m) \leqslant \mu(I_1) + \cdots + \mu(I_n) + n\,\frac{\varepsilon}{2n} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

lo que prueba nuestra propiedad, ya que los intervalos $I'_1, ..., I'_m$ son equiláteros de igual tamaño y constituyen un recubrimiento de C, pues recubren $I_1 \cup \cdots \cup I_n$ y $\{I_1, ..., I_n\}$ es recubrimiento de C.

$[61]_2$ Las funciones de clase $\mathscr{C}^1$ conservan el contenido nulo

Sea $\varphi : A \to \mathbb{R}^p$, donde $A \subset \mathbb{R}^p$ es abierto, una transformación de clase $\mathscr{C}^1$ en A. Si C es un conjunto de contenido nulo tal que $\overline{C} \subset A$, entonces $\boldsymbol{\varphi}(C)$ es también un conjunto de contenido nulo.

Demostración

Como C ha de estar acotado, también lo está $\overline{C}$, el cual, por ser cerrado, es entonces compacto. Como además es $\overline{C} \subset A$ y A es abierto, resulta que (véase $[08]_1$) la distancia ρ de $\overline{C}$ a la frontera de A es positiva (no nula). Considérese la unión, para $\boldsymbol{x}$ recorriendo $\overline{C}$, de las bolas abiertas $B(\boldsymbol{x}, \rho/2)$; es evidente que esta unión, a la que llamaremos B, es un conjunto abierto, acotado, y tal que $\overline{C} \subset B \subset \overline{B} \subset A$.

Por ser $\boldsymbol{\varphi}$ de clase $\mathscr{C}^1$ en A, los gradientes $\nabla\varphi_1, ..., \nabla\varphi_p$, de las p componentes de la función $\boldsymbol{\varphi}$, son continuas en A, lo son en $\overline{B}$ y, como este conjunto $\overline{B}$ es compacto, están todas ellas acotadas en $\overline{B}$ y, por ello, en B; así pues, existe $K > 0$ tal que $\|\nabla\varphi_i(\boldsymbol{x})\| < K$ para $\boldsymbol{x} \in \overline{B}$ e $i = 1, ..., p$.

Dado $\varepsilon > 0$ (cualquiera), como C tiene contenido nulo, este conjunto puede cubrirse con un número finito de intervalos compactos $I_1, ..., I_n$ equiláteros de igual tamaño (véase [61], 2.°) y cuyas medidas suman menos de ε; nótese que estos intervalos pueden quedar todos incluidos en B, lo cual así se supondrá, con tal de tomar sus diámetros menores que $\rho/2$ (para lo que bastará con tomarlos de modo que sus medidas sumen menos que un valor suficientemente pequeño). Como los intervalos I_j (para $j = 1, ..., n$) son conjuntos convexos incluidos en B, en cada uno de ellos le es aplicable a las funciones φ_i (para $i = 1, ..., p$) el teorema del valor medio (véase [27]), el cual permite poner que (en lo que sigue: ξ es un cierto punto del segmento $]\boldsymbol{x}, \boldsymbol{x}'[$; el signo $\leqslant$ es consecuencia de la desigualdad de Cauchy-Schwarz del producto escalar, $|\boldsymbol{u} \cdot \boldsymbol{v}| \leqslant \|\boldsymbol{u}\|\,\|\boldsymbol{v}\|$):

$$|\varphi_i(\boldsymbol{x}) - \varphi_i(\boldsymbol{x}')| = |d\varphi_i(\xi)(\boldsymbol{x} - \boldsymbol{x}')| = |\nabla\varphi_i(\xi) \cdot (\boldsymbol{x} - \boldsymbol{x}')| \leqslant \|\nabla\varphi_i(\xi)\|\,\|\boldsymbol{x} - \boldsymbol{x}'\| < K\|\boldsymbol{x} - \boldsymbol{x}'\|$$

$$(\text{para } \boldsymbol{x}, \boldsymbol{x}' \in J_j,\ j = 1, ..., n;\ i = 1, ..., p)$$

Dado que $\|x - x'\|$ es menor o igual que el diámetro de I_j, se verificará que

$$\|x - x'\| \leqslant \sqrt{p}\,\sqrt[p]{\mu(I_j)} < \sqrt{p}\,\sqrt[p]{\varepsilon/n}, \qquad \text{luego } |\varphi_i(x) - \varphi_i(x')| < K\sqrt{p}\,\sqrt[p]{\varepsilon/n}$$
$$(\text{para } x, x' \in I_j,\ j = 1, \ldots, n;\ i = 1, \ldots, p)$$

De la última desigualdad se desprende que $\varphi(I_j)$ está incluido en un intervalo compacto de contenido $(K\sqrt{p})^p(\varepsilon/n)$ y, consecuentemente, $\varphi(I_1 \cup \cdots \cup I_n)$ está incluido en n intervalos compactos cuyos contenidos suman no más de $(K\sqrt{p})^p\varepsilon$. Como estos intervalos cubren a $\varphi(C)$, de lo últimamente dicho se infiere que $\varphi(C)$ tiene contenido nulo.

[61]$_3$ La gráfica de una función integrable tiene contenido nulo

Sea $\varphi : J \to \mathbb{R}$ una función acotada en el intervalo compacto $J \subset \mathbb{R}^{p-1}$. Si φ es integrable (o, en particular, continua) en J, entonces la gráfica de φ, es decir, el siguiente conjunto de $\mathbb{R}^p$:

$$G = \{(x, y) \in \mathbb{R}^p \ / \ x \in J,\ y = \varphi(x)\}$$

tiene contenido nulo en $\mathbb{R}^p$.

Demostración

Dado $\varepsilon > 0$, como φ es integrable en J, existe una partición $\{J_1, \ldots, J_n\}$ de J en intervalos compactos de $\mathbb{R}^{p-1}$ tal que

$$\sum_{i=1}^{n} (M_i - m_i)\mu(J_i) < \varepsilon, \qquad \text{donde } \begin{cases} m_i = \inf\ \varphi(J_i) \\ M_i = \sup\ \varphi(J_i) \end{cases}$$

Llamando $I_i = J_i \times [m_i, M_i]$, para $i = 1, \ldots, n$, es evidente que $\{I_1, \ldots, I_n\}$ es un recubrimiento finito de G mediante intervalos compactos (de $\mathbb{R}^p$) para el que se cumple que

$$\mu(I_1) + \cdots + \mu(I_n) = \sum_{i=1}^{n} (M_i - m_i)\mu(J_i) < \varepsilon$$

lo cual prueba que el conjunto G tiene contenido nulo.

[61]$_4$ Ejercicio

Sea $C \subset \mathbb{R}^p$ el conjunto que forman los puntos del intervalo $I = [0, 1]^p \subset \mathbb{R}^p$ que tienen todas sus coordenadas racionales; esto es, $C = [0, 1]^p \cap \mathbb{Q}^p$. Pruébese que C no tiene contenido nulo.

Resolución

Sea $\{I_1, ..., I_n\}$ un recubrimiento finito de C mediante intervalos compactos; nótese que el conjunto $J = I_1 \cup \cdots \cup I_n$ es cerrado, con lo que $C \subset J$. Cualquier punto $\boldsymbol{a} \in I$ es, obviamente, de acumulación de C, con lo que $\boldsymbol{a} \in \bar{C}$, luego $\boldsymbol{a} \in J$; consecuentemente, se ha de verificar que $I \subset J$. Por ello, se puede asegurar que

$$\mu(I) \leqslant \mu(J) \leqslant \mu(I_1) + \cdots + \mu(I_n)$$

con lo que esta suma de contenidos no se puede conseguir menor que cualquier $\varepsilon > 0$ (basta tomar $\varepsilon < \mu(I)$), por lo que C no tiene contenido nulo.

CONTINUIDAD SALVO EN CONJUNTO DE CONTENIDO NULO IMPLICA INTEGRABILIDAD

Se considera aquí una clase de funciones integrables que es suficientemente general como para abarcar a la mayoría de las funciones que aparecen en la integración práctica. Nos referimos a las funciones que son discontinuas en, sólo, un conjunto de contenido nulo.

[62]

Sea $f: I \to \mathbb{R}$ una función acotada en un intervalo compacto $I \subset \mathbb{R}^p$. Si f es continua en $I - C$, donde $C \subset I$ es un conjunto de contenido nulo, entonces f es integrable en I.

Demostración

Para empezar, notemos que, dado un intervalo compacto $I_0 \subset I$ y llamando

$$I = [a_1, b_1] \times \cdots \times [a_p, b_p] \qquad \text{e} \qquad I_0 = [a_1^0, b_1^0] \times \cdots \times [a_p^0, b_p^0]$$

donde $a_i \leqslant a_i^0 < b_i^0 \leqslant b_i$ para $i = 1, ..., p$, los «hiperplanos» $x_i = a_i$, $x_i = a_i^0$, $x_i = b_i^0$ y $x_i = b_i$ (para $i = 1, ..., p$) dividen a I en 3^p intervalos, que determinan una partición de I (si algunos de los 3^p intervalos fueran «degenerados», ellos se excluirían de la partición); a dicha partición la denotaremos por $P(I_0)$.

La integrabilidad de f en I la probaremos acudiendo a la primera condición «$\varepsilon : P$» (véase [53]). Sea dado un $\varepsilon > 0$ cualquiera. Como C es un conjunto de contenido nulo, existe un recubrimiento finito $\{I_1, ..., I_n\}$ de C formado por intervalos cerrados cuyas medidas suman menos que $\varepsilon/(2^{p+2}H)$, donde H es una cota superior de $|f|$ en I. Sea I_i' el intervalo compacto concéntrico con I_i y cuyos lados tienen doble longitud que los de éstos (para $i = 1, ..., n$), con lo que $\mu(I_i') = 2^p\mu(I_i)$; la familia $\{I_1', ..., I_n'\}$ es un recubrimiento finito de C formado por intervalos compactos cuyas medidas suman menos que $2^p[\varepsilon/(2^{p+2}H)]$, esto es, que $\varepsilon/(4H)$. Nótese que se puede suponer que $I_i' \subset I$ (para $i = 1, ..., n$), pues de no ser así, el anterior recubrimiento podría sustituirse por el $\{I_1' \cap I, ..., I_n' \cap I\}$.

Sea $P(I_i')$ la partición de I que determina el intervalo I_i' (para $i = 1, ..., n$), de la que hablamos al comienzo (en el primer párrafo de esta demostración), y consideremos la partición

$$P_0 = P(I_1') * \cdots * P(I_n')$$

(superposición de las $P(I_i')$; véase [50]). Los intervalos que constituyen P_0 los clasificaremos en dos: llamaremos $J_1, ..., J_h$ a aquellos que están incluidos en alguno de los intervalos I_i' (para $i = 1, ..., n$) y llamaremos $K_1, ..., K_k$ a aquellos que no están incluidos en algún I_i', con lo que no se solapan con ningún I_i' (o sea, $\overset{\circ}{I_i'} \cap \overset{\circ}{K_j} = \varnothing$ para $i = 1, ..., n$ y $j = 2, ..., k$; nótese que entonces $I_i \cap K_j = \varnothing$, luego $C \cap K_j = \varnothing$).

Como f es continua en K_j, y, por ello, integrable en K_j (para $j = 1, ..., k$), sabemos (según la primera condición «$\varepsilon : P$») que existe una partición P_j de K_j tal que $S(f, P_j) - s(f, P_j) < \varepsilon/(2k)$.

Pues bien, sea P la siguiente partición de I:

$$P = P_J \cup P_1 \cup \cdots \cup P_k, \qquad \text{donde } P_J = \{J_1, ..., J_h\}$$

Para esta partición se verifica que:

$$S(f, P) - s(f, P) = S(f, P_J) - s(f, P_J) + \sum_{j=1}^{k} (S(f, P_j) - s(f, P_j)) <$$
$$< 2H \sum_{i=1}^{n} \mu(I_i') + k \frac{\varepsilon}{2k} < 2H \frac{\varepsilon}{4H} + \frac{\varepsilon}{2} = \varepsilon$$

luego, para esta partición P, la función f verifica en I a la primera condición «$\varepsilon : P$», esto es, f es integrable en I.

[62]$_1$ Corolario

Sea $f: I \to \mathbb{R}$ una función acotada en un intervalo compacto $I = [a_1, b_1] \times \cdots \times [a_p, b_p]$ de $\mathbb{R}^p$. Para cada $i = 1, ..., p$, se llama I_i a la proyección de I sobre el plano coordenado i-ésimo, es decir,

$$I_i = [a_1, b_1] \times \cdots \times [a_{i-1}, b_{i-1}] \times [a_{i+1}, b_{i+1}] \times \cdots \times [a_p, b_p]$$

Para $i = 1, ..., p$, sean $\varphi_i, ..., \psi_i$ funciones (en cantidad finita) de I_i en $[a_i, b_i]$, que se suponen integrables (o, en particular, continuas) en I_i.

Si f es continua en $I - C$, donde C es la unión de las gráficas de todas las funciones $\varphi_i, ..., \psi_i$ (para $i = 1, ..., p$), entonces f es integrable en I.

Demostración

Las gráficas de las funciones $\varphi_i, ..., \psi_i$ son conjuntos de contenido nulo (véase [61]$_3$) y, por tanto, la unión C de todas estas gráficas también es un conjunto de contenido nulo (véase [61], 2). Estamos, pues, en las condiciones de la anterior propiedad [62], que nos dice que f es integrable en I.

$[62]_2$ Ejercicio

Sean $f: I \to \mathbb{R}$ y $g: I \to \mathbb{R}$ dos funciones acotadas en el intervalo compacto I de $\mathbb{R}^p$. Si f y g coinciden en $I - C$, donde $C \subset I$ es un conjunto de contenido nulo, y si g es integrable en I, pruébese que entonces f es integrable en I.

Comprobación

Esta propiedad se prueba de igual modo que la [62]. Tanto es así, que sólo hay que cambiar, allí, la frase «Como f es continua en K_j y, por ello, integrable en K_j» por «Como f es integrable en K_j, pues los g y f y g coinciden en K_j».

$[62]_3$ Observación

Las funciones acotadas $f: I \to \mathbb{R}$ ($I \subset \mathbb{R}^p$ intervalo compacto) que son discontinuas, sólo, en un conjunto de contenido nulo de I no son las únicas funciones integrables en I (es decir, en general no se verifica la propiedad recíproca de la [62]), sino que la clase de las funciones integrables en I es más amplia que la considerada en [62]. En breve veremos que las funciones integrables en I son las acotadas que son discontinuas, a lo sumo, en un conjunto de «medida» nula de I. De estos conjuntos nos ocupamos a continuación.

Para confirmar la anterior afirmación, considérese el siguiente ejemplo. Sea $f: I \to \mathbb{R}$, donde $I = [0, 1] \times [0, 1]$, la función definida por:

$$f(x, y) = \begin{cases} 0, & \text{si } x = 0 \quad \text{o si} \quad y = 0 \\ 0, & \text{si alguno de los } x \text{ o } y \text{ son irracionales} \\ \dfrac{1}{q+s}, & \text{si } x = \dfrac{p}{q} \text{ e } y = \dfrac{r}{s} \text{ (racionales irreducibles)} \end{cases}$$

Esta función es discontinua en todo punto $(x_0, y_0) \in I \cap \mathbb{Q}^2$, ya que $f(x_0, y_0) \neq 0$ y en todo entorno de (x_0, y_0) hay puntos (x, y) con alguna de sus coordenadas irracionales en los que, por ello, es $f(x, y) = 0$. Por tanto, el conjunto de los puntos de discontinuidad de f es $I \cap \mathbb{Q}^2$, al menos(*), que no es un conjunto de contenido nulo, como se vio en el ejercicio $[61]_4$. Ello no impide que f sea integrable en I, como se comprobó en $[53]_3$.

CONJUNTOS DE MEDIDA NULA

Según se acaba de anunciar, para caracterizar a integrabilidad relacionándola con la continuidad y obtener la clase de las funciones integrables, se precisa de los conjuntos de medida nula.

(*) Si $(x_0, y_0) \in I - \mathbb{Q}^2$, entonces f es continua en (x_0, y_0), pues $f(x_0, y_0) = 0$ y f tiene límite 0 en (x_0, y_0). Para comprobar esto último, dado $\varepsilon > 0$, acudamos a un $h \in \mathbb{N}$ tal que $h > 1/\varepsilon$, y notemos que los $(p/q, r/s)$ tales que $q + s \leqslant h$ es finito, con lo que existe un entorno de (x_0, y_0) en el que no hay ninguno de ellos, luego para (x, y) en dicho entorno es $|f(x, y) - 0| < 1/h < \varepsilon$.

[63]

Un conjunto $C \subset \mathbb{R}^p$ se dice que tiene *medida nula* si, para cada $\varepsilon > 0$, existe un recubrimiento numerable (que, en particular, puede ser finito) de C formado por intervalos compactos cuyas medidas suman menos que ε, esto es, si existe una sucesión numerable, $I_1, ..., I_n, ...$, de intervalos compactos de $\mathbb{R}^p$ tales que

$$C = \bigcup_{n=1}^{\infty} I_n \qquad \text{y} \qquad \sum_{n=1}^{\infty} \mu(I_n) < \varepsilon$$

Propiedades:

1. Un subconjunto de un conjunto de medida nula tiene medida nula.
2. La unión numerable de conjuntos de medida nula tiene medida nula.
3. Todo conjunto numerable tiene medida nula.
4. Todo conjunto de contenido nulo tiene medida nula (el recíproco es falso).
5. Todo conjunto de medida nula y compacto tiene contenido nulo.

Demostración

1. Esta propiedad es evidente (tómese para el subconjunto al mismo recubrimiento que para el conjunto).
2. Si $C_1, ..., C_n, ...$, son conjuntos de medida nula y dado $\varepsilon > 0$, entonces para cada $n \in \mathbb{N}$ existe un recubrimiento numerable $\mathscr{R}_n$ de C_n formado por intervalos cerrados cuyas medidas suman menos que $\varepsilon/2^n$. Por tantop, $\mathscr{R}_1 \cup \cdots \cup \mathscr{R}_n \cup \cdots$ es un recubrimiento numerable de $C = C_1 \cup \cdots \cup C_n \cup \cdots$ formado por intervalos cerrados cuyos contenidos suman menos que

$$\frac{\varepsilon}{2} + \frac{\varepsilon}{2^2} + \cdots + \frac{\varepsilon}{2^n} + \cdots = \varepsilon$$

(obsérvese que los contenidos son números no negativos, por lo que el orden en el que se tomen los anteriores intervalos cerrados puede ser alterado sin que ello tenga ningún efecto; no se olvide que las series de términos no negativos son conmutativas y asociativas), luego C tiene medida nula.
3. El conjunto que forma un solo punto es, obviamente, de medida nula. Un conjunto numerable es, pues, la unión numerable de conjuntos de medida nula, por lo que, según la propiedad anterior, aquel tiene medida nula.
4. Esta propiedad es evidente. El recíproco es falso, como prueba el siguiente ejemplo. El conjunto $C = [0, 1]^p \cap \mathbb{Q}^p$, de puntos de $\mathbb{R}^p$, no tiene contenido nulo, según se vio en $[61]_4$, pero tiene medida nula, pues es un conjunto numerable.
5. Sea C un conjunto de medida nula y compacto. Dado $\varepsilon > 0$ (cualquiera), y como C tiene medida nula, se sabe que C admite un recubrimiento numerable, $\{I_1, ..., I_n, ...\}$, formado por intervalos compactos cuyas medidas suman menos que $\varepsilon/2^p$. Sea I_n' el intervalo compacto que tiene el mismo centro que I_n y cuyos lados miden doble de los de I_n, con lo que $\mu(I_n') = 2^p\mu(I_n)$ y, por ello, la suma de los contenidos de los intervalos I_n' (para $n \in \mathbb{N}$) es menor que ε.

Para mostrar que C tiene contenido nulo, bastará con que comprobemos la existencia de cierto $h \in \mathbb{N}$ tal que $C \subset I'_1 \cup \cdots \cup I'_h$. Esto lo vamos a hacer por reducción al absurdo, es decir, suponiendo que para cada $n \in \mathbb{N}$ existe algún punto $\boldsymbol{a}_n \in C$ tal que $\boldsymbol{a}_n \notin I'_1 \cup \cdots \cup I'_n$, se va a llegar a una contradicción: Como C es compacto, la sucesión $\boldsymbol{a}_1, ..., \boldsymbol{a}_n, ...$, formada por puntos de C, ha de tener algún límite de oscilación $\boldsymbol{a} \in C$ (véase [08], II), es decir, admitiría alguna subsucesión $\boldsymbol{a}_{n_1}, ..., \boldsymbol{a}_{n_i}, ...$, convergente hacia $\boldsymbol{a} \in C$. Como $\{I_1, ..., I_n, ...\}$ es un recubrimiento de C, existe $k \in \mathbb{N}$ tal que $\boldsymbol{a} \in I_k$, luego $\boldsymbol{a}$ es un punto interior de I'_k, es decir, $\mathring{I}'_k$ es un entorno (abierto) de $\boldsymbol{a}$. Ahora bien, los índices n_i son a partir de uno de ellos mayores que k, luego a partir de ese (como se ha supuesto que $\boldsymbol{a}_n \notin I'_1 \cup \cdots \cup I'_n$, con lo que $\boldsymbol{a}_n \notin I'_s$ para todo $s \leqslant n$) se verifica, pues, que $\boldsymbol{a}_{n_i} \notin \mathring{I}'_k$, lo que entra en contradicción con el hecho de ser $\boldsymbol{a}$ el límite de $\boldsymbol{a}_{n_1}, ..., \boldsymbol{a}_{n_i}, ...$, ya que $\mathring{I}'_k$ resultó ser entorno de $\boldsymbol{a}$.

[63]$_1$ Observación

En la definición de conjunto de medida nula (véase [63]), los intervalos compactos $I_1, ..., I_n, ...$, pueden sustituirse por intervalos abiertos y acotados

Comprobación

Se puede razonar aquí siguiendo el mismo camino que en [61]$_1$, introduciendo las oportunas adaptaciones. Basta ahora con sustituir la familia finita $I_1, ..., I_n$ por la familia numerable $I_1, ..., I_n, ...$, y sustituir la exigencia $\mu(I'_i) < \mu(I_i) + \varepsilon/(2n)$ para $i = 1, ..., n$, por la $\mu(I'_i) < \mu(I_i) + \varepsilon/2^i$ para $i \in \mathbb{N}$, con lo que

$$\sum_{i=1}^{\infty} \mu(I'_i) < \sum_{i=1}^{\infty} \mu(I_i) + \sum_{i=1}^{\infty} \frac{\varepsilon}{2^i} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

[63]$_2$ Ejercicio

Sea $J \subset \mathbb{R}^{p-1}$ un intervalo cualquiera y sea $\varphi : J \to \mathbb{R}$ una función integrable (o, en particular, continua) en todo intervalo compacto incluido en J. Pruébese que la gráfica G de φ,

$$G = \{(\boldsymbol{x}, y) \in \mathbb{R}^p \ / \ \boldsymbol{x} \in J, \ y = f(\boldsymbol{x})\}$$

tiene medida nula en $\mathbb{R}^p$.

Resolución

Sea $I_1, ..., I_n, ...$, una sucesión creciente de intervalos compactos de $\mathbb{R}^{p-1}$ cuyo límite es J (tal sucesión es evidente que existe), esto es, tal que

$$I_n \subset J, \quad \forall\, n \in \mathbb{N} \qquad \text{y} \qquad \bigcup_{n=1}^{\infty} I_n = J$$

Si llamamos G_n a la gráfica de la restricción de φ a I_n, es evidente que $G = G_1 \cup \cdots \cup G_n \cup \cdots$. Los conjuntos G_n son todos ellos de contenido nulo (como se vio en $[61]_3$), luego son de medida nula. Por ello, G ha resultado ser la unión numerable de conjuntos de medida nula, luego G tiene también entonces medida nula.

$[63]_3$ Ejercicio

Sea $f: I \to \mathbb{R}$ una función acotada en un intervalo compacto $I \subset \mathbb{R}^p$. Si $f(\boldsymbol{x}) \geqslant 0$ para todo $\boldsymbol{x} \in I$, si f es integrable en I y si $\int_I f = 0$, pruébese que entonces los puntos de I en los que f no se anula forman un conjunto de medida nula.

Resolución

Acudamos al conjunto $N_h = \{\boldsymbol{x} \in I \,/\, f(\boldsymbol{x}) > 1/h\}$, donde $h \in \mathbb{N}$ es fijo, y comprobemos que este conjunto tiene contenido nulo. Dado $\varepsilon > 0$ (cualquiera), como $\int_I f = 0$, existe una partición P tal que $S(f, P) < \varepsilon/h$. Llamando $J_1, ..., J_n$ a aquellos intervalos de P en los que hay puntos de N_h, se puede poner:

$$\frac{\varepsilon}{h} > S(f, P) = \sum_{J \in P} \sup f(J)\mu(J) \geqslant \sum_{i=1}^{n} \sup f(J_i)\mu(J_i) > \frac{1}{h} \sum_{i=1}^{n} \mu(J_i)$$

por lo que, para la familia finita de intervalos compactos $\{J_1, ..., J_n\}$ se verifica que:

$$N_h \supset J_1 \cup \cdots \cup J_n \qquad \text{y} \qquad \mu(J_1) + \cdots + \mu(J_n) < \varepsilon$$

es decir, el conjunto N_h tiene contenido nulo (véase [61]).

Resulta evidente que se verifica que

$$\{\boldsymbol{x} \in I \,/\, f(\boldsymbol{x}) \neq 0\} = N_1 \cup \cdots \cup N_h \cup \cdots$$

y como todos los N_h (para $h \in \mathbb{N}$) tienen contenido nulo, resulta que su unión tiene contenido nulo (véase [63],2), como había que comprobar.

❑ OSCILACIÓN DE UNA FUNCIÓN EN UN PUNTO

Para poder caracterizar a las funciones integrables en un intervalo compacto I, que es aquí nuestro objetivo, vamos a precisar del concepto de oscilación de una función en un punto. La integrabilidad en I resultará equivalente a que exista continuidad en «casi todo» punto de I, más exactamente, a que los puntos de discontinuidad formen un conjunto de medida nula. El interés que aquí tiene la oscilación en un punto radica en el hecho de que la discontinuidad en un punto equivale a oscilación no nula en dicho punto; la discontinuidad en el punto será tanto más «notable», estará tanto más lejos de la continuidad, será tanto más «patológica» cuanto mayor sea la oscilación en él.

[64]

Sea $f: C \to \mathbb{R}$ una función acotada en un conjunto $C \subset \mathbb{R}^p$ y considérese un punto $a \in C$. Para cada $\rho > 0$, representemos por $\omega(f, a, \rho)$ a la oscilación de f en el conjunto de los puntos de C que están en la bola $B(a, \rho)$, esto es

$$\omega(f, a, \rho) = \sup \{f(x) / x \in C, \|x - a\| < \rho\} - \inf \{f(x) / x \in C, \|x - a\| < \rho\}$$

La función $\rho \mapsto \omega(f, a, \rho)$, de la variable real $\rho > 0$, es creciente y acotada, por lo que tiene límite cuando $\rho \to 0$, al que representaremos por $\omega(f, a)$:

$$\omega(f, a) = \lim_{\rho \to 0} \omega(f, a, \rho)$$

y recibe el nombre de *oscilación* de la función f en el punto a.

TEOREMA. La función $f: C \to \mathbb{R}$, acotada en el conjunto $C \subset \mathbb{R}^p$, es continua en un punto $a \in C$ si, y sólo si, es nula la oscilación de f en a.

Demostración

Antes de abordar la demostración del teorema, hagamos alguna observación acerca de la definición de oscilación. Nótese en primer lugar que para todo $\rho > 0$ es $\omega(f, a, \rho) \geqslant 0$, pues el supremo de un conjunto es mayor o igual que su ínfimo. Nótese también que, cuando ρ decrece, con él decrece $\omega(f, a, \rho)$, pues entonces $B(a, \rho)$ se reduce y, por ello, el supremo y el ínfimo de f en $C \cap B(a, \rho)$ disminuye y aumenta, respectivamente. Como la función $\rho \mapsto \omega(f, a, \rho)$ es creciente y no negativa para $\rho > 0$, entonces tiene límite cuando $\rho \to 0$ y dicho límite es no negativo, es decir, existe $\omega(\rho, a)$ el cual es no negativo.

1.° Supongamos primero que f es continua en a. Dado $\varepsilon > 0$ (cualquiera), existe un $\delta > 0$ tal que

$$x \in C \cap B(a, \delta) \quad \Rightarrow \quad f(a) - \frac{\varepsilon}{3} < f(x) < f(a) + \frac{\varepsilon}{3} \quad \Rightarrow$$

$$\Rightarrow \quad f(a) - \frac{\varepsilon}{3} \leqslant m_\delta \leqslant M_\delta \leqslant f(a) + \frac{\varepsilon}{3} \quad \Rightarrow \quad M_\delta - m_\delta \leqslant 2\,\frac{\varepsilon}{3} < \varepsilon$$

donde m_δ y M_δ representan al ínfimo y al supremo de f en $C \cap B(a, \delta)$. Por tanto,

$$0 < \rho < \delta \quad \Rightarrow \quad [m_\rho \geqslant m_\delta,\ M_\rho \leqslant M_\delta] \quad \Rightarrow \quad \omega(f, a, \rho) = M_\rho - m_\rho \leqslant M_\delta - m_\delta < \varepsilon$$

lo que prueba que $\omega(f, a, \rho) \to 0$ cuando $\rho \to 0$, es decir, que $\omega(f, a) = 0$.

2.° Supongamos ahora que $\omega(f, a) = 0$, con lo que, dado $\varepsilon > 0$ (cualquiera), existe un $\delta > 0$ tal que $\omega(f, a, \delta) < \varepsilon$, o sea, tal que $M_\delta - m_\delta < \varepsilon$. Por tanto, se verifica que:

$$x \in C \cap B(a, \delta) \quad \Rightarrow \quad |f(x) - f(a)| \leqslant M_\delta - m_\delta < \varepsilon$$

luego f es continua en a.

[64]$_1$ Ejemplo

Considérese la función $f: \mathbb{R}^2 \to \mathbb{R}$ definida por

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

Por muy pequeño que sea $\rho > 0$, la oscilación de la función f en la bola $B(\boldsymbol{o}, \rho) =$ $= \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < \rho^2\}$ es

$$\omega(f, \boldsymbol{o}, \rho) = \sup f(B(\boldsymbol{o}, \rho)) - \inf f(B(\boldsymbol{o}, \rho)) = 1 - (-1) = 2$$

Obsérvese que esta función no es continua en el punto $\boldsymbol{o} = (0, 0)$.

[64]$_2$ Dos propiedades de la oscilación en un punto

Dada una función $f: I \to \mathbb{R}$, acotada en el intervalo compacto $I \subset \mathbb{R}^p$, se verifica que:

1.° Para cualquiera que sea $k > 0$, el conjunto C_k que forman los puntos de I en los que f tiene oscilación mayor o igual que k, esto es, $C_k = \{\boldsymbol{x} \in I \mid \omega(f, \boldsymbol{x}) \geqslant k\}$, es un conjunto cerrado(*).

2.° Si la oscilación de f en cualquiera de los puntos de I es menor que un cierto número $r > 0$, esto es, si $\omega(f, \boldsymbol{x}) < r$ para todo $\boldsymbol{x} \in I$, entonces existe una partición P de I tal que la oscilación de f en cada uno de los intervalos de P es menor que r y, por tanto, $S(f, P) - s(f, P) < r\mu(I)$.

Demostración

1.° Hemos de comprobar que, si $\boldsymbol{a}$ es un punto de acumulación de C_k, entonces $\boldsymbol{a} \in C_k$ (nótese que un tal punto $\boldsymbol{a}$ debe pertenecer a I, pues I es cerrado y $C_k \subset I$). Así ocurre, en efecto:

Por ser $\boldsymbol{a}$ de acumulación de C_k y para cualquiera que sea $\rho > 0$, se sabe que existe algún punto $\boldsymbol{b} \in B(\boldsymbol{a}, \rho) \cap C_k$; como $\boldsymbol{b} \in B(\boldsymbol{a}, \rho)$, existe $\rho' > 0$ tal que $B(\boldsymbol{b}, \rho') \subset B(\boldsymbol{a}, \rho)$, con lo que

$$B(\boldsymbol{b}, \rho') \cap I \subset B(\boldsymbol{a}, \rho) \cap I, \qquad \text{luego } \omega(f, \boldsymbol{b}, \rho') \leqslant \omega(f, \boldsymbol{a}, \rho)$$

Pero como, por ser $\boldsymbol{b} \in C_k$, es $\omega(f, \boldsymbol{b}, \rho') \geqslant k$, de la anterior desigualdad se desprende que $\omega(f, \boldsymbol{a}, \rho) \geqslant k$, y ello para todo $\rho > 0$. De este resultado se deduce, tomando límites para $\rho \to 0$, que $\omega(f, \boldsymbol{a}) \geqslant k$, es decir, que $\boldsymbol{a} \in C_k$, como había que comprobar.

(*) Este teorema también se verifica si se sustituye el intervalo compacto I por un conjunto cerrado $C \subset \mathbb{R}^n$, lo que se demuestra como aquí (sin más que sustituir I por C).

2.º Consideremos las particiones $P_1, P_2, ..., P_n, ...$, de I que se obtienen de dividir todos los lados de I en $2^1, 2^2, ..., 2^n, ...$, partes iguales. Para alguna de estas particiones ha de verificarse el teorema, como comprobaremos por reducción al absurdo. Esto es, suponiendo que, en cada una de las particiones P_n, hay algún subintervalo I_n tal que, llamando $\omega(f, I_n)$ a la oscilación de f en él, se verifica que $\omega(f, I_n) \geqslant r$, vamos a llegar a una contradicción:

Llamando $\boldsymbol{x}_n \in I$ al centro de I_n, la sucesión $(\boldsymbol{x}_n)$ tiene algún límite de oscilación $\boldsymbol{a}$, que pertenece a I. Es evidente que, para cualquiera que sea $\rho > 0$, el entorno $B(\boldsymbol{a}, \rho)$ de $\boldsymbol{a}$ contiene a alguno de los intervalos I_n, por lo que

$$\omega(f, \boldsymbol{a}, \rho) \geqslant \omega(f, I_n) \geqslant r, \qquad \forall \rho > 0$$

Tomando ahora límites para $\rho \to 0$, nos encontramos con que deberá ser $\omega(f, \boldsymbol{a}) \geqslant r$, lo que es falso, pues como $\boldsymbol{a} \in I$, sabemos que $\omega(f, \boldsymbol{a}) < r$. De esta contradicción se infiere la veracidad del teorema.

CARACTERIZACIÓN DE LEBESGUE DE LA INTEGRABILIDAD (EN UN INTERVALO)

[65]

TEOREMA DE LEBESGUE (DE INTEGRABILIDAD RIEMANN). Sea $f: I \to \mathbb{R}$ una función acotada en el intervalo compacto $I \subset \mathbb{R}^p$. La función f es integrable en I si, y sólo si, los puntos de I en los que f es discontinua forman un conjunto de medida nula.

Demostración

En lo que sigue, usaremos la misma notación que en los apartados anteriores:

$$\omega(f, \boldsymbol{a}) = \text{oscilación de } f \text{ en el punto } \boldsymbol{a} \in I$$
$$C_k = \{\boldsymbol{x} \in I \,/\, \omega(f, \boldsymbol{x}) \geqslant k\}, \qquad \text{para cada } k > 0$$

1.º Primeramente probemos que, si f es integrable en I, entonces el conjunto $C = \{\boldsymbol{x} \in I \,/\, f$ es discontinua en $\boldsymbol{x}\}$ tiene medida nula. Para ello, nótese que (según el teorema de [64]):

$$C = \{\boldsymbol{x} \in I \,/\, \omega(f, \boldsymbol{x}) > 0\} = C_1 \cup C_{1/2} \cup \cdots \cup C_{1/n} \cup \cdots$$

por lo que bastará con probar (véase [63], 2 y 4) que $C_{1/n}$ tiene contenido nulo para todo $n \in \mathbb{N}$. Elegido un $n \in \mathbb{N}$ y dado $\varepsilon > 0$ (cualesquiera), como f es integrable en I, existe una partición P_n de I tal que $S(f, P_n) - s(f, P_n) < \varepsilon/(2n)$. Los intervalos de P_n los dividiremos en dos clases: llamemos $J_1, ..., J_r$ a los intervalos de P_n que contienen en su interior puntos de $C_{1/n}$ y llamemos $K_1, ..., K_s$ a los intervalos de P_n que no tienen en su interior puntos de $C_{1/n}$, con lo que si en alguno de estos intervalos hay puntos de $C_{1/n}$, dichos puntos estarán en su frontera. Ello permite poner

$$\frac{\varepsilon}{2n} > S(f, P_n) - s(f, P_n) = \sum_{i=1}^{r} \omega(f, J_i)\mu(J_i) + \sum_{j=1}^{s} \omega(f, K_j)\mu(K_j) \geqslant$$
$$\geqslant \sum_{i=1}^{r} \omega(f, J_i)\mu(J_i) \geqslant \sum_{i=1}^{r} \omega(f, \boldsymbol{x}_i)\mu(J_i) > \frac{1}{n} \sum_{i=1}^{r} \mu(J_i)$$

donde $\boldsymbol{x}_i$ es un punto de $C_{1/n}$ situado en el interior de J_i (nótese que existe $\rho > 0$ tal que $B(\boldsymbol{x}_i, \rho) \subset J_i$, por lo que $\omega(f, J_i) \geqslant \omega(f, \boldsymbol{x}_i, \rho) \geqslant \omega(f, \boldsymbol{x}_i)$). De la anterior relación se desprende que

$$\frac{\varepsilon}{2n} > \frac{1}{n} \sum_{i=1}^{r} \mu(J_i), \qquad \text{luego } \sum_{i=1}^{r} \mu(J_i) < \frac{\varepsilon}{2}$$

Por otra parte, el conjunto de los puntos de $C_{1/n}$ que no están en el interior de alguno de los J_i está incluido en las fronteras de los intervalos K_j; como estas fronteras forman un conjunto de contenido nulo (pues son una cantidad finita de intervalos degenerados), existe un número finito de intervalos compactos $L_1, ..., L_q$ que le recubren y cuyas medidas suman menos de $\varepsilon/2$. De todo lo dicho resulta que $\{J_1, ..., J_r, L_1, ..., L_q\}$ es un recubrimiento finito de $C_{1/n}$ mediante intervalos compactos cuyas medidas suman menos que $(\varepsilon/2) + (\varepsilon/2) = \varepsilon$, luego $C_{1/n}$ tiene contenido nulo, como había que comprobar.

2.º Probemos ahora que, si el conjunto C de los puntos de discontinuidad de f tiene medida nula, entonces f es integrable en I, es decir, que (según la primera condición «$\varepsilon : P$» de integrabilidad; véase [53], 1.ª) para cada $\varepsilon > 0$ existe una partición P de I tal que $S(f, P) - s(f, P) < \varepsilon$. Para ello, acudamos al conjunto

$$C_\delta = \{\boldsymbol{x} \in I \ / \ \omega(f, \boldsymbol{x}) \geqslant \delta\}, \qquad \text{donde } \delta = \frac{\varepsilon}{2H + \mu(I)}$$

(se llama H a una cota superior de $|f|$ en I). Este conjunto C_δ es cerrado (según $[64]_2$, 1.º) y acotado (pues $C_\delta \subset I$), esto es, C_δ es compacto. Por otra parte, como $C_\delta \subset C$ (véase el teorema de [64]) y C es de medida nula, C_δ es de medida nula; por ello, al ser C_δ compacto y de medida nula, se sabe que (véase [63], 5) C_δ tiene contenido nulo, luego existe una familia finita de intervalos abiertos y acotados (véase $[63]_1$), $A_1, ..., A_n$, tal que:

$$C_\delta \subset A_1 \cup \cdots \cup A_n \qquad \text{y} \qquad \mu(A_1) + \cdots + \mu(A_n) < \delta$$

Las caras de todos los intervalos A_i (para $i = 1, ..., n$), mejor dicho, los «hiperplanos» en los que están situadas dichas caras, determinan una partición P_0 de I, la cual está formada por dos clases de intervalos cerrados, que llamaremos $P' = \{J_1, ..., J_r\}$ y $P'' = \{K_1, ..., K_s\}$, donde: 1.º, cada uno de los intervalos de P' está contenido en alguno de los intervalos cerrados $\bar{A}_i$ (adherencia de A_i); y 2.º, cada uno de los intervalos de P'' tiene intersección vacía con A_i para $i = 1, ..., n$, por lo que en cualquiera de dichos intervalos de P'' no hay ningún punto de C_δ. Nótese que para P' se verifica que:

$$S(f, P') - s(f, P') = \sum_{j=1}^{r} \omega(f, J_j)\mu(J_j) \leqslant 2H \sum_{j=1}^{r} \mu(J_j) < 2H\delta$$

Ocupémonos ahora de cada uno de los intervalos $K_k \in P''$ (para $k = 1, ..., s$). Como en K_k no hay puntos de C_δ, para todo $x \in K_k$ se verifica que $\omega(f, x) < \delta$. De esto último se infiere (según se probó en $[64]_2$, 2.°) que existe una partición P_k del intervalo K_k tal que

$$S(f, P_k) - s(f, P_k) < \delta\mu(K_k), \qquad \text{para } k = 1, ..., s$$

Obsérvese que $P = P' \cup P_1 \cup \cdots \cup P_s$ es una partición de I y que para ella se verifica que:

$$S(f, P) - s(f, P) = S(f, P') - s(f, P') + \sum_{k=1}^{s} (S(f, P_k) - s(f, P_k)) <$$

$$< 2H\delta + \delta \sum_{k=1}^{s} \mu(K_k) < 2H\delta + \delta\mu(I) = \varepsilon$$

con lo que concluye la demostración.

$[65]_1$ Observación

Acudiendo al teorema de Lebesgue [65] se pueden probar, como simples corolarios, las propiedades relativas a la integrabilidad de funciones, algunas de las cuales ya han sido demostradas trabajosamente con anterioridad en [54], que aseguran la integrabilidad de ciertas funciones obtenidas combinando funciones integrables (sólo se asegura la integrabilidad; nada se dice aquí sobre el valor de la correspondiente integral). Así, por ejemplo:

Dadas dos funciones $f: I \to \mathbb{R}$ y $g: I \to \mathbb{R}$, acotadas en el intervalo compacto I de $\mathbb{R}^p$, se verifica que:

1. Si f es integrable en I y si J es un intervalo compacto incluido en I, entonces f es integrable en J.
2. Si f y g son integrables en I, entonces también son integrables en I: 1.°, la función $hf + kg$, para cualesquiera que sean $h, k \in \mathbb{R}$; 2.°, la función fg, y 3.°, la función f/g, si $|g(x)| \geqslant$ k para todo $x \in I$ y para un cierto $k > 0$.
3. Si f es integrable en I, entonces $|f|$ es integrable en I.
4. Si f y g coinciden salvo en un conjunto numerable, entonces una de ellas es integrable si, y sólo si, lo es la otra.
5. Si f y g tienen los mismos puntos de discontinuidad, entonces una de ellas es integrable si, y sólo si, lo es la otra.
6. Si f es integrable en I y si $\varphi: [m, M] \to \mathbb{R}$ es una función continua en $[m, M]$, donde $m = \inf f(I)$ y $M = \sup f(I)$, entonces la función $\varphi \circ f: I \to \mathbb{R}$ es integrable en I.

Comprobación

Llamemos $D(F)$ al conjunto de los puntos de discontinuidad de una función $F: I \to \mathbb{R}$; para indicar que F es integrable en I, pondremos $F \in \mathscr{I}(I)$.

1. Como $f \in \mathscr{I}(I)$, $D(f)$ tiene medida nula, luego el conjunto de las discontinuidades de f en $J \subset I$, por estar incluido en $D(f)$, también es de medida nula, por lo que $f \in \mathscr{I}(J)$.
2. Como $D(hf + kg)$, $D(fg)$ y $D(f/g)$ están incluidos en $D(f) \cap D(g)$ y como los conjuntos $D(f)$ y $D(g)$ tienen medida nula, pues $f, g \in \mathscr{I}(I)$, resulta que los primeros conjuntos tienen medida nula, luego $hf + kg$, fg y $f/g \in \mathscr{I}(I)$.

3. Como $D(|f|) \subset D(f)$ y $D(f)$ tiene medida nula, pues $f \in \mathscr{I}(I)$, resulta que $D(|f|)$ tiene medida nula, luego $|f| \in \mathscr{I}(I)$.
4. $D(f)$ y $D(g)$ difieren en un conjunto numerable y, por ello, de medida nula. Por ello, $D(f)$ es de medida nula si, y sólo si, los $D(g)$ (véase [63], 2), es decir, $f \in \mathscr{I}(I) \Leftrightarrow g \in \mathscr{I}(I)$.
5. $f \in \mathscr{I}(I)$ si, y sólo si, $D(f)$ tiene medida nula y, como $D(f) = D(g)$, esto equivale a $g \in \mathscr{I}(I)$.
6. Como $D(\varphi \circ f) \subset D(f)$ y $D(f)$ es de medida nula, pues $f \in \mathscr{I}(I)$, resulta que $D(\varphi \circ f)$ tiene medida nula, luego $\varphi \circ f \in \mathscr{I}(I)$.

4.3. INTEGRACIÓN EN CONJUNTOS MEDIBLES

Vamos a ocuparnos ahora de la integración de funciones acotadas sobre conjuntos acotados de una clase más general que los intervalos compactos, que son los únicos que hemos considerado hasta el momento. Ello se hará con facilidad aproyándonos en lo ya dicho sobre la integración en intervalos compactos. La existencia de estas nuevas integrales dependerá, pues, de las funciones que se integran y de los conjuntos sobre los que se integra. Estos van a ser los «conjuntos medibles» (según Jordan), de los que nos ocuparemos en breve; según se verá, tales conjuntos no son otros que los que tienen frontera de medida nula. Los integrados, acotados, serán funciones continuas salvo en conjuntos de medida nula.

❑ CONJUNTOS MEDIBLES (SEGÚN JORDAN)

[66]

Un conjunto $C \subset \mathbb{R}^p$ se dice que es *medible* (según Jordan) si está acotado y existe la integral $\int_I \chi_C$, donde: 1.°, $\chi_C : \mathbb{R}^p \to \mathbb{R}$ es la «función característica» de C (definida por $\chi_C(x) = 1$, si $x \in C$; $\chi_C(x) = 0$, si $x \notin C$), y 2.°, I es cualquier intervalo compacto que contenga a C (esta definición no depende del I que elija). En tal caso, a dicha integral se la llama *medida* de C y la representaremos por $\mu(C)$; esto es:

$$\mu(C) = \int_I \chi_C \qquad \text{donde } \chi_C(x) = \begin{cases} 1, & \text{si } x \in C \\ 0, & \text{si } x \notin C \end{cases}$$

CARACTERIZACIÓN. Un conjunto $C \subset \mathbb{R}^p$ es medible si, y sólo si, está acotado y su frontera, $F_r(C)$, es un conjunto de contenido nulo.

CONJUNTOS SIMPLES. Se llaman conjuntos simples de $\mathbb{R}^p$ a los conjuntos de la forma

$$S(\varphi, \psi) = \{(x, y) \in \mathbb{R}^p \,/\, x \in J,\ \varphi(x) \leqslant y \leqslant \psi(x)\}$$

donde J es un intervalo compacto de $\mathbb{R}^{p-1}$ y $\varphi : J \to \mathbb{R}$ y $\psi : J \to \mathbb{R}$ son las funciones acotadas en J y tales que $\varphi(x) \leqslant \psi(x)$ para todo $x \in J$.

Si φ y ψ son integrables (o, en particular, continuas) en J, entonces el conjunto simple $S(\varphi, \psi)$ es un conjunto medible.

Demostraciones

1.° Empecemos comprobando que las anteriores definiciones, de conjunto medible y de medida, no dependen del intervalo compacto I que se tome para incluir en él al conjunto C. Si I' fuese otro intervalo cerrado tal que $C \subset I'$, acudiendo al intervalo compacto $J = I \cap I'$, como $C \subset J$ y, por ello, χ_C es nula en $I - J$ y en $I' - J$, de la aditividad de la integral respecto del intervalo (véase [57]) se deduce que χ_C es integrable en I si, y sólo si, lo es en I' y se verifica que

$$\int_I \chi_C = \int_J \chi_C = \int_{I'} \chi_C$$

2.° Los puntos de discontinuidad de la función $\chi_C : \mathbb{R}^p \to \mathbb{R}$ son los de la frontera de C y sólo ellos. En efecto: 1.°, si $\boldsymbol{x} \in F_r(C)$, entonces en todo entorno de $\boldsymbol{x}$ hay puntos en los que χ_C vale 1 y puntos en los que χ_C vale 0, luego χ_C es discontinua en $\boldsymbol{x}$; 2.°, si $\boldsymbol{x} \notin F_r(C)$, entonces $\boldsymbol{x}$ es interior en C o es exterior a C, luego existe un entorno U de $\boldsymbol{x}$ tal que $\chi_C(U) = 1$, o $\chi_C(U) = 0$, respectivamente, luego χ_C es continua en $\boldsymbol{x}$.

3.° Caracterización. El conjunto C es medible si, y sólo si, está acotado y $\chi_C : \mathbb{R}^p \to \mathbb{R}$ es integrable en I. Esto último equivale (según el teorema de Lebesgue, véase [65]) a que C esté acotado y los puntos de discontinuidad de χ_C formen un conjunto de medida nula, es decir, a que C esté acotado y $F_r(C)$ tenga medida nula. Ahora bien, como se sabe que, para cualquier conjunto $C \subset \mathbb{R}^p$, su frontera $F_r(C)$ es un conjunto cerrado, resulta que la condición C acotado y $F_r(C)$ de medida nula equivale (véase [63], 4 y 5) a la condición C está acotado y $F_r(C)$ tiene contenido nulo, como había que comprobar.

4.° Conjuntos simples. Si las funciones acotadas $\varphi : J \to \mathbb{R}$ y $\psi : J \to \mathbb{R}$ son integrables en el intervalo compacto $J \subset \mathbb{R}^{p-1}$, entonces sus gráficas

$$G_\varphi = \{(\boldsymbol{x}, y) \in \mathbb{R}^p \;/\; \boldsymbol{x} \in J,\ y = \varphi(\boldsymbol{x})\} \qquad \text{y} \qquad G_\psi = \{(\boldsymbol{x}, y) \in \mathbb{R}^p \;/\; \boldsymbol{x} \in J,\ y = \psi(\boldsymbol{x})\}$$

son conjuntos de contenido nulo (véase $[61]_3$).

La frontera del conjunto simple $S(\varphi, \psi)$ está compuesta por las gráficas G_φ y G_ψ y por 2^{p-1} conjuntos acotados «degenerados», incluidos $F_r(J) \times \mathbb{R}$ (incluidos en 2^{p-1} planos, paralelos a planos coordenados); como todos estos conjuntos, tanto los G_φ y G_ψ como los 2^{p-1} conjuntos degenerados, tienen contenido nulo, entonces la frontera de $S(\varphi, \psi)$ tiene contenido nulo y, como $S(\varphi, \psi)$ está acotado, resulta (según se acaba de probar) que $S(\varphi, \psi)$ es un conjunto medible.

$[66]_1$ Observaciones

Comencemos recordando que (véase [66]) un conjunto acotado $C \subset \mathbb{R}^p$ se dice medible (según Jordan) si su función característica $\chi_C : \mathbb{R}^p \to \mathbb{R}$ es integrable en un intervalo compacto I que incluye a C y la medida de C es, entonces, $\mu(C) = \int_I \chi_C$ (que no depende de I).

Si $P = \{I_1, ..., I_n\}$ es una partición cualquiera de I, las sumas superior e inferior (véase [51]) $s(\chi_C, P)$ y $S(\chi_C, P)$, de la función $\chi_C : I \to \mathbb{R}$ correspondientes a la partición P, a las cuales denotaremos también por $\underline{\mu}_C(P)$ y $\bar{\mu}_C(P)$, son

$$\underline{\mu}_C(P) = s(\chi_C, P) = \sum_{\substack{I_i \in P \\ I_i \subset C}} \mu(I_i) \qquad \text{y} \qquad \bar{\mu}_C(P) = S(\chi_C, P) = \sum_{\substack{I_i \in P \\ I_i \cap C \neq \varnothing}} \mu(I_i)$$

1. De acuerdo con la primera condición «$\varepsilon : \delta$» de integrabilidad en un intervalo (véase [59], 1.º), el conjunto acotado $C \subset \mathbb{R}^p$ es medible si, y sólo si, para cada $\varepsilon > 0$ existe un $\delta > 0$ tal que, si P es una partición de un intervalo compacto que incluya a C, entonces:

$$|P| < \delta \quad \Rightarrow \quad \bar{\mu}_C(P) - \underline{\mu}_C(P) < \varepsilon$$

2. De acuerdo con la segunda condición «$\varepsilon : \delta$» de integrabilidad en un intervalo (véase [59], 2.º), el conjunto acotado $C \subset \mathbb{R}^p$ es medible y su medida en $\mu(C)$ si, y sólo si, para cada $\varepsilon > 0$ existe un $\delta > 0$ tal que, si P es una partición de un intervalo compacto que incluya a C, entonces:

$$|P| < \delta \quad \Rightarrow \quad \mu(C) - \varepsilon < \underline{\mu}_C(P) \leqslant \bar{\mu}_C(P) < \mu(C) + \varepsilon$$

3. De acuerdo con lo ya dicho sobre «la integral como límite de sumas» (véase $[59]_1$, 1.º), si $P_1, ..., P_n, ...,$ es una sucesión de particiones, de un intervalo compacto que incluya al conjunto acotado $C \subset \mathbb{R}^p$, cuyos diámetros tienden a cero (cuando $n \to \infty$), esto es, si $\lim |P_n| = 0$, entonces C es medible si, y sólo si, acontece que

$$\lim_{n\to\infty} (\bar{\mu}_C(P_n) - \underline{\mu}_C(P_n)) = 0$$

y, en tal caso, la medida de C vale:

$$\mu(C) = \lim_{n\to\infty} \underline{\mu}_C(P_n) = \lim_{n\to\infty} \bar{\mu}_C(P_n)$$

4. Si el conjunto acotado $C \subset \mathbb{R}^p$ es medible, para cada $\varepsilon > 0$ existe una familia finita $J_1, ..., J_m$ de intervalos compactos de $\mathbb{R}^p$ tales que

$$C \subset J_1 \cup \cdots \cup J_m \qquad \text{y} \qquad \mu(J_1) + \cdots + \mu(J_m) < \mu(C) + \varepsilon$$

Así ocurre ya que, como $\mu(C) = \int_I \chi_C$, existe una partición P de I tal que $0 < \bar{\mu}_C(P) - \mu(C) < \varepsilon$ y, de acuerdo con la expresión que definía a $\bar{\mu}_C(P)$, si $J_1, ..., J_m$ son los intervalos de P que contiene algún punto de C, es

$$\bar{\mu}_C(P) = \mu(J_1) + \cdots + \mu(J_m)$$

$[66]_2$ Primeras propiedades de los conjuntos medibles

Los conjuntos medibles de $\mathbb{R}^p$ y sus medidas satisfacen a las siguientes propiedades:

1.º Para todo conjunto medible C es $\mu(C) \geqslant 0$.
2.º Si C y D son conjuntos medibles tales que $C \subset D$, entonces $\mu(C) \leqslant \mu(D)$.
3.º Si C y D son conjuntos medibles, entonces también son medibles $C \cup D$, $C \cap D$ y $C - D$ y se verifica que:

$$\mu(C \cup D) = \mu(C) + \mu(D) - \mu(C \cap D) \qquad \text{y} \qquad \mu(C - D) = \mu(C) - \mu(C \cap D)$$

4.° Todo intervalo compacto $I = [a_1, b_1] \times \cdots \times [a_p, b_p]$ de $\mathbb{R}^p$ es un conjunto medible y su medida es el valor $\mu(I) = (b_1 - a_1) \cdots (b_p - a_p)$ (es decir, para los intervalos compactos la medida que ahora se define coincide con la que se dio en [50], al comienzo del capítulo).

Demostración

1.° Como $\chi_C(\boldsymbol{x}) \geqslant 0$ para todo $\boldsymbol{x} \in \mathbb{R}^p$, de la propiedad de monotonía (véase [55]) se deduce que $\mu(C) = \int_I \chi_C \geqslant \int_I 0 = 0$ ($I \subset \mathbb{R}^p$ intervalo compacto; $C \subset I$).

2.° Como $\chi_C(\boldsymbol{x}) \leqslant \chi_D(\boldsymbol{x})$ para todo $\boldsymbol{x} \in \mathbb{R}^p$, de la propiedad de monotonía se deduce que $\mu(C) = \int_I \chi_C \leqslant \int_I \chi_D = \mu(D)$ ($I \subset \mathbb{R}^p$ intervalo compacto; $C \subset D \subset I$).

3.° Como C y D son medibles, entonces $F_r(C)$ y $F_r(D)$ son acotados y tienen contenido nulo, luego $F_r(C) \cup F_r(D)$ es acotado y tiene contenido nulo (véase [61], 2); ahora bien, como en el conjunto $F_r(C) \cup F_r(D)$ están incluidos los $F_r(C \cup D)$, $F_r(C \cap D)$ y $F_r(C - D)$, resulta que estos tres conjuntos están acotados y tienen contenido nulo (véase [61], 1), es decir, $C \cup D$, $C \cap D$ y $C - D$ son medibles: existen $\mu(C \cup D)$, $\mu(C \cap D)$ y $\mu(C - D)$. Como $\chi_{C \cup D} = \chi_C + \chi_D - \chi_{C \cap D}$, integrando (en un intervalo compacto que incluya a $C \cup D$) se obtiene que $\mu(C \cup D) = \mu(C) + \mu(D) - \mu(C \cap D)$. Como $(C - D) \cup (C \cap D) = C$ y $(C - D) \cap (C \cap D) = \varnothing$, aplicando el anterior resultado a $C - D$ y $C \cap D$ (en lugar de los C y D), se obtiene que $\mu(C) = \mu(C - D) + \mu(C \cap D) + 0$, como había que comprobar.

4.° Si $I = [a_1, b_1] \times \cdots \times [a_p, b_p]$, su medida se puede expresar (acudiendo a la integración iterada; véase [60]):

$$\mu(I) = \int_I \chi_I = \int_{a_1}^{b_1} dx_1 \left\{ \cdots \left[\int_{a_{p-1}}^{b_{p-1}} dx_{p-1} \left(\int_{a_p}^{b_p} 1\, dx_p \right) \right] \right\} =$$

$$= \int_{a_1}^{b_1} dx_1 \left\{ \cdots \left[\int_{a_{p-1}}^{b_{p-1}} dx_{p-1} (b_p - a_p) \right] \right\} = \int_{a_1}^{b_1} dx_1 \{ \cdots [(b_{p-1} - a_{p-1})(b_p - a_p)] \} = \cdots$$

$$= \int_{a_1}^{b_1} (b_2 - a_2) \cdots (b_p - a_p)\, dx_1 = (b_1 - a_1)(b_2 - a_2) \cdots (b_p - a_p)$$

[66]$_3$ Las funciones $\mathscr{C}^1$-biyectivas transforman medibles en medibles

Sea $\varphi : A \to \mathbb{R}^p$ una transformación, definida en un abierto $A \subset \mathbb{R}^p$, que es de clase $\mathscr{C}^1$ en A e inyectiva en A. Si $C \subset \mathbb{R}^p$ es un conjunto medible tal que $C \subset A$ y si el jacobiano $\det J\varphi(\boldsymbol{x})$ es no nulo en todo punto $\boldsymbol{x} \in \mathring{C}$, entonces se verifica que $F_r(\varphi(C)) = \varphi(F_r(C))$ y el conjunto $\varphi(C)$ es también medible.

Nota: La condición $\det J\varphi(\boldsymbol{x}) \neq 0$ para todo $\boldsymbol{x} \in \mathring{C}$ se utiliza, sólo, para garantizar (véase el teorema de la función inversa, [38]) que para cada $\boldsymbol{x} \in \mathring{C}$ existen sendos entornos (abiertos) U y V de $\boldsymbol{x}$ y $\varphi(\boldsymbol{x})$ tales que $\varphi : U \to V$ es biyectiva. Si φ es inyectiva, esto último está garantizado si $\varphi(\mathring{C})$ es abierto, pues entonces $U = \mathring{C}$ y $V = \varphi(\mathring{C})$. La condición «$\varphi$ es inyectiva en A» sólo se precisa para comprobar que es $F_r(\varphi(C)) \supset \varphi(F_r(C))$, pero no es necesaria para garantizar que $\varphi(C)$ es medible.

Demostración

1.º Veamos primero que $F_r(\varphi(C)) \subset \varphi(F_r(C))$. El conjunto $\overline{C}$ es compacto (es cerrado y está acotado, pues lo está C, ya que es medible) luego también es compacto $\varphi(\overline{C})$, pues φ es continua en $\overline{C}$. Como $\varphi(C) \subset \varphi(\overline{C})$ y este último conjunto es compacto, resulta que $F_r(\varphi(C)) \subset \varphi(\overline{C}) = \varphi(\mathring{C} \cup F_r(C))$, por lo que, para cada $y \in F_r(\varphi(C))$, existe un $x \in \mathring{C} \cup F_r(C)$ tal que $y = \varphi(x)$ y, entonces, para que se cumpla la relación a demostrar, es suficiente con que $x \notin \mathring{C}$. Así ocurre, ya que si fuese $x \in \mathring{C}$ y como $\det J\varphi(x) \neq 0$, según el teorema de la función inversa (véase [38]) existirían sendos entornos $U \subset \mathring{C}$ y V de x y de y tales que $\varphi(U) = V$, luego y sería un punto interior de $\varphi(\mathring{C}) \subset \varphi(C)$, lo que no es posible, pues y es de la frontera de $\varphi(C)$.

2.º Veamos ahora que $\varphi(C)$ es medible. Como $\varphi(C)$ está acotado, pues $\varphi(C) \subset \varphi(\overline{C})$ y este último conjunto está acotado por ser compacto, sólo hay que comprobar que $F_r(\varphi(C))$ tiene contenido nulo (véase [66]). Como ya sabemos (por lo dicho en 1.º) que $F_r(\varphi(C)) \subset \varphi(F_r(C))$, será suficiente con probar que $\varphi(F_r(C))$ tiene contenido nulo (véase [61], 1) y esto es cierto, pues $F_r(C)$ tiene contenido nulo (por ser C medible) y φ (según $[61]_2$) conserva la propiedad «tener contenido nulo».

3.º Veamos finalmente que $\varphi(F_r(C)) \subset F_r(\varphi(C))$, esto es, que si $x \in F_r(C)$ entonces $\varphi(x) \in F_r(\varphi(C))$. Como $F_r(C) \subset A$ y A es abierto, por ser $x \in F_r(C)$ se puede asegurar que existen dos sucesiones (x_n) y (x'_n) con límite x de puntos de C y de $A - C$, respectivamente. Dado que φ es continua en x, se sabe que las sucesiones $(\varphi(x_n))$ y $(\varphi(x'_n))$ tienen límite $\varphi(x)$. Por otra parte, como φ es inyectiva en A, es $\varphi(x_n) \in \varphi(C)$ y $\varphi(x'_n) \notin \varphi(C)$ para $n \in \mathbb{N}$, luego $\varphi(x)$ es límite de una sucesión de puntos de $\varphi(C)$ y límite de una sucesión de puntos que no son de $\varphi(C)$, es decir, $\varphi(x)$ es un punto frontera de $\varphi(C)$, como había que comprobar.

$[66]_4$ Ejercicio

Sabiendo que se verifican las hipótesis del teorema anterior, compruébese que:

1.º $\varphi(\mathring{C}) = \overset{\circ}{\widehat{\varphi(C)}}$.

2.º $\varphi(\overline{C}) = \overline{\varphi(C)}$.

Resolución

Recuérdese que, para una aplicación inyectiva $F: X \to Y$ y si $X_1 \subset X$ y $X_2 \subset X$, es $F(X_1 - X_2) = F(X_1) - F(X_2)$. Por ello y como $F_r(\varphi(C)) = \varphi(F_r(C))$, se tiene:

1.º $\overset{\circ}{\widehat{\varphi(C)}} = \varphi(C) - F_r(\varphi(C)) = \varphi(C) - \varphi(F_r(C)) = \varphi(C - F_r(C)) = \varphi(\mathring{C})$.

2.º $\overline{\varphi(C)} = \varphi(C) \cup F_r(\varphi(C)) = \varphi(C) \cup \varphi(F_r(C)) = \varphi(C \cup \varphi(F_r(C)) = \varphi(\overline{C})$.

$[66]_5$ Observación

Sea C un conjunto acotado, de puntos de $\mathbb{R}^p$; sea $\chi_C : \mathbb{R}^p \to \mathbb{R}$ la función característica de C (esto es: $\chi_C(x) = 1$, si $x \in C$; $\chi_C(x) = 0$, si $x \notin C$). Se llaman «medida interior» y «medida exterior» de C a las siguientes integrales inferior y superior:

$$\underline{\mu}(C) = \underline{\int_I} \chi_C \qquad \text{y} \qquad \overline{\mu}(C) = \overline{\int_I} \chi_C$$

($I \subset \mathbb{R}^p$ es un intervalo compacto cualquiera que incluye a C; los $\underline{\mu}(C)$ y $\bar{\mu}(C)$ no dependen de I). Si las medidas interior y exterior de C son iguales, entonces ambas se representan por $\mu(C)$, a este número se le llama medida de C y se dice que C es medible. Para hallar las medidas interior y exterior de un conjunto acotado C, acudimos a las sumas inferior y superior de Darboux, $s(\chi_C, P_n)$ y $S(\chi_C, P_n)$, de la función característica χ_C, correspondiente a una sucesión de particiones $P_1, ..., P_n, ...$ de I cuyos diámetros tienden a cero, con lo que

$$\underline{\mu}(C) = \lim_{n\to\infty} s(\chi_C, P_n) \qquad \text{y} \qquad \bar{\mu}(C) = \lim_{n\to\infty} S(\chi_C, \mathrm{P}_n)$$

Para cada intervalo J de P_n, el ínfimo $m_J = \inf \chi_C(J)$ y el supremo $M_J = \sup \chi_C(J)$ son: 1.°, si $J \subset C$, entonces $m_J = M_J = 1$; 2.°, si $J \cap C = \varnothing$, entonces $m_J = M_J = 0$, y 3.°, si $J \not\subset C$ y $J \cap C \neq \varnothing$, entonces $m_J = 0$ y $M_J = 1$, con lo que resulta que

$$\underline{\mu}(C) = \lim_{n\to\infty} \sum_{J \in P'_n} \mu(J), \qquad P'_n = \{J \in P_n \,/\, J \subset C\}$$

$$\bar{\mu}(C) = \lim_{n\to\infty} \sum_{J \in P''_n} \mu(J), \qquad P'_n = \{J \in P_n \,/\, J \cap C \neq \varnothing\}$$

Así pues, la medida interior (respectivamente, exterior) de un conjunto acotado C resulta ser el límite de las sumas de las medidas de los intervalos de una partición P_n, que están incluidos en (respectivamente, que tienen algún punto de) C, cuando la partición P_n se hace cada vez más numerosa de modo que su diámetro tienda a cero.

$[66]_6$ Aserto

Si $C \subset \mathbb{R}^p$ es un conjunto medible y dado $\varepsilon > 0$ (cualquiera), existen dos conjuntos medibles C' y C'' tales que:

1.° $C' \subset C \subset C''$.
2.° $\mu(C) - \mu(C') < \varepsilon$ y $\mu(C'') - \mu(C) < \varepsilon$.
3.° C' y C'' son: o ambos compactos, o ambos abiertos, o uno de ellos (cualquiera) compacto y el otro abierto.

Demostración

Acudiendo a un intervalo compacto I que incluya a C, sabemos que (véase [66]) $\mu(C) = \int_I \chi_C$. De acuerdo con la caracterización «$\varepsilon : P$» de la integral (véase [53]), podemos asegurar que hay una partición $P = \{I_1, ..., I_n\}$ de I tal que $\mu(C) - s(\chi, P) < \varepsilon$ y $S(\chi, P) - \mu(C) < \varepsilon$. Llamemos C_i y C_e a los siguientes conjuntos compactos:

- C_i es la unión de los intervalos de P que están incluidos en C (estos intervalos de P son aquellos en los que el ínfimo de χ vale 1; en los demás, dicho ínfimo es 0); nótese, pues, que $s(\chi, P) = \mu(C_i)$.
- C_e es la unión de los intervalos de P que contienen algún punto de C (estos intervalos de P son aquellos en los que el supremo de χ vale 1; en los demás, dicho supremo es 0); nótese, pues que $S(\chi, P) = \mu(C_e)$.

Para estos dos conjuntos compactos se verifican entonces las condiciones del enunciado ($C' = C_i$ y $C'' = C_e$). El conjunto C_i puede sustituirse por su interior $\mathring{C}_i$, que es abierto, pues $\mathring{C}_i \subset C_i \subset C$ y $\mu(\mathring{C}_i) = \mu(C_i)$, con lo que para $\mathring{C}_i$ se verifica lo exigido en el enunciado ($C' = \mathring{C}_i$). Aplicando lo ya probado al conjunto $\bar{C} = I - C$, se sabe que existe un intervalo compacto $C_1 \subset I$ tal que $C_1 \subset \bar{C}$ y $\mu(\bar{C}) - \mu(C_1) < \varepsilon$; para el conjunto $C'' = I - C_1$, que es abierto, se verifica, pues, que $C \subset C''$ y

$$\mu(C'') - \mu(C) = [\mu(I) - \mu(C_1)] - [\mu(I) - \mu(\bar{C})] = \mu(\bar{C}) - \mu(C_1) < \varepsilon$$

es decir, la propiedad a demostrar también se verifica para un C'' abierto.

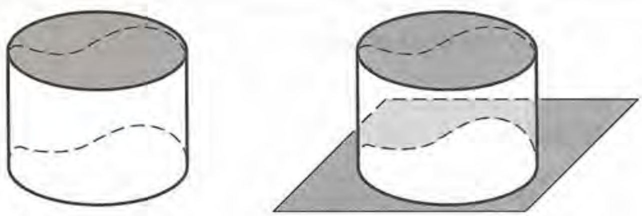
INTEGRABILIDAD EN UN CONJUNTO MEDIBLE. INTEGRAL

[67]

Sea $f: C \to \mathbb{R}$ una función acotada en un conjunto medible $C \subset \mathbb{R}^p$. Se dice que f es integrable en C si existe la integral $\int_I f_C$, donde: 1.º, $f_C: \mathbb{R}^p \to \mathbb{R}$ es la función definida por $f_C(\boldsymbol{x}) = f(\boldsymbol{x})$, si $\boldsymbol{x} \in C$, y $f_C(\boldsymbol{x}) = 0$, si $\boldsymbol{x} \notin C$, y 2.º, $I \subset \mathbb{R}^p$ es cualquier intervalo compacto que contenga a C (esta definición no depende del I que se elija). En tal caso, se dice que dicha integral es la integral de f en C y se la representa por $\int_C f$, $\int_C f\,d\mu$ o $\int_C f(\boldsymbol{x})\,d\boldsymbol{x}$; esto es:

$$\int_C f = \int_I f_C, \qquad \text{donde } f_C(\boldsymbol{x}) = \begin{cases} f(\boldsymbol{x}), & \text{si } \boldsymbol{x} \in C \\ 0, & \text{si } \boldsymbol{x} \notin C \end{cases}$$

CARACTERIZACIÓN DE LA INTEGRABILIDAD. Para que una función $f: C \to \mathbb{R}$, acotada en un conjunto medible $C \subset \mathbb{R}^p$, sea integrable en C es necesario y suficiente que el conjunto de los puntos de discontinuidad de f tenga medida nula.



Demostración

1.º Comprobemos en primer lugar de las anteriores definiciones que no dependen del intervalo compacto I, con $C \subset I$, que se tome. Supongamos, en efecto, que se sustituye I por otro intervalo compacto I', con $C \subset I'$. Si es J el intervalo compacto $J = I \cap I'$, como f_C se anula en $I - J$ y en $I' - J$, pues $C \subset J$, acudiendo a la propiedad de aditividad respecto del intervalo (véase [57]) se concluye que f_C es integrable en I si, y sólo si, lo es en I' y se verifica que $\int_I f_C = \int_J f_C = \int_{I'} f_C$.

2.° Caracterización. Para que exista $\int_C f$, es decir, para que f_C sea integrable en I, es necesario y suficiente (según la caracterización de Lebesgue; véase [65]) que el conjunto, D_C, de los puntos de discontinuidad de f_C tenga medida nula. Los puntos de D_C o son de discontinuidad de f o pertenecen a la frontera de C, verificándose evidentemente que $D \subset D_C$ y $D_C \subset D \cup F_r(C)$, donde se ha llamado D al conjunto de los puntos de discontinuidad de f. Como $F_r(C)$ tiene contenido nulo (pues C es medible), de las dos inclusiones anteriores se desprende que D_C tiene medida nula si, y sólo si, D tiene medida nula, lo que prueba, finalmente, la caracterización de la integrabilidad.

[67]$_1$ Expresión integral de la medida de un conjunto

A la vista de las definiciones de conjunto medible (véase [66]) y de integral en un conjunto medible (véase [67]), se puede decir que: un conjunto acotado $C \subset \mathbb{R}^p$ es medible si, y sólo si, la función unidad (o sea, la $f: C \to \mathbb{R}$ tal que $f(x) = 1$ para todo $x \in C$) es integrable en C y, entonces, la medida de C es el número

$$\mu(C) = \int_C 1$$

[67]$_2$ Integral en un conjunto medible y sumas de Riemann

Supongamos que la función $f: C \to \mathbb{R}$, acotada en el conjunto medible C, es integrable en C; esto es, que existe $\int_C f = \int_I f_C$, donde I es un intervalo compacto que incluye a C y $f_C: \mathbb{R}^p \to \mathbb{R}$ es la función definida por $f_C(\boldsymbol{x}) = f(\boldsymbol{x})$, si $\boldsymbol{x} \in C$, y $f_C(\boldsymbol{x}) = 0$, si $\boldsymbol{x} \notin C$. A la integral $\int_I f_C$ se le puede aplicar el teorema de Riemann [58], lo que permite aproximarla mediante las oportunas sumas de Riemann; en concreto, dado $\varepsilon > 0$ (cualquiera), existe $\delta > 0$ tal que, siempre que una partición $P = \{I_1, ..., I_h\}$ de I tenga diámetro menor que δ y para cualquiera que sea correspondiente familia $\boldsymbol{x} = \{\boldsymbol{x}_1, ..., \boldsymbol{x}_n\}$ de puntos intermedios, se verifica que:

$$\left| \int_C f - \sum_{I_i \in P} f_C(\boldsymbol{x}_i)\mu(I_i) \right| < \varepsilon \qquad [1]$$

Si para cada I_i que no esté totalmente incluido en C tomamos $\boldsymbol{x}_i$ en $I - C$, como para tales $\boldsymbol{x}_i$ es $f_C(\boldsymbol{x}_i) = 0$ y para los restantes $\boldsymbol{x}_i$ es $f_C(\boldsymbol{x}_i) = f(\boldsymbol{x}_i)$, nos encontramos con que la relación [1] toma la forma

$$\left| \int_C f - \sum_{I_i \subset C} f(\boldsymbol{x}_i)\mu(I_i) \right| < \varepsilon \qquad [2]$$

Si para cada I_i que tenga algún punto en C se toma $\boldsymbol{x}_i$ en C, como para tales $\boldsymbol{x}_i$ es $f_C(\boldsymbol{x}_i) = = f(\boldsymbol{x}_i)$ y para los restantes $\boldsymbol{x}_i$ es $f_C(\boldsymbol{x}_i) = 0$, nos encontramos con que la relación [1] toma la forma

$$\left| \int_C f - \sum_{I_i \cap C \neq \varnothing} f(\boldsymbol{x}_i)\mu(I_i) \right| < \varepsilon \qquad [3]$$

Si se considera ahora una sucesión cualquiera $P_1, ..., P_n, ...,$ de partición de I cuyos diámetros tiendan a cero (cuando $n \to \infty$), entonces acudiendo a que $\int_I f_C$ es el límite de las correspondientes sucesiones de sumas de Riemann (véase $[59]_4$) y de acuerdo con lo que acabamos de obtener en [2] y [3], podemos asegurar que:

$$\int_C f = \lim_{n\to\infty} \sum_{I_i \in P'_n} f(\boldsymbol{x}_i)\mu(I_i) = \lim_{n\to\infty} \sum_{I_i \in P''_n} f(\boldsymbol{x}_i)\mu(I_i)$$

$$P'_n = \{I_i \in P_n \,/\, I_i \subset C\} \qquad P''_n = \{I_i \in P_n \,/\, I_i \cap C \neq \varnothing\}$$

PROPIEDADES BÁSICAS DE LA INTEGRAL (EN UN CONJUNTO MEDIBLE)

[68]

1. PROPIEDADES ARITMÉTICAS. Sean $f: C \to \mathbb{R}$ y $g: C \to \mathbb{R}$ dos funciones acotadas en un conjunto medible $C \subset \mathbb{R}^p$. Si f y g son integrables en C, entonces también son integrables en C las funciones $f + g$, cf (para todo $c \in \mathbb{R}$), $|f|$, fg y f/g, esta última si $|g(\boldsymbol{x})| \geqslant k$ para todo $\boldsymbol{x} \in C$ y para un cierto $k > 0$. Además, las integrales de $f + g$ y cf son:

$$\int_C (f+g) = \int_C f + \int_C g \qquad \text{y} \qquad \int_C cf = c\int_C f$$

2. MONOTONÍA DE LA INTEGRAL. Sean $f: C \to \mathbb{R}$ y $g: C \to \mathbb{R}$ dos funciones acotadas en un conjunto medible $C \subset \mathbb{R}^p$. Si f y g son integrables en C, se verifica que

$$(f(\boldsymbol{x}) \leqslant g(\boldsymbol{x}), \quad \forall\, \boldsymbol{x} \in C) \;\Rightarrow\; \int_C f \leqslant \int_C g$$

Y en particular (si para f se satisfacen las anteriores exigencias) es:

$$\left|\int_C f\right| \leqslant \int_C |f|$$

3. TEOREMA DE LA MEDIA. Sea $f: C \to \mathbb{R}$ una función acotada en un conjunto medible $C \subset \mathbb{R}^p$. Si f es integrable en C y si M y m son cotas superior e inferior de f en C, entonces $m\mu(C) \leqslant \int_C f \leqslant M\mu(C)$. Así pues, existe $h \in [m, M]$ tal que:

$$\int_C f = h\mu(C) \qquad (\mu(C) = \text{medida de } C)$$

Y, en particular, si f es continua en C y C es (además de tener medida no nula) compacto y conexo, entonces existe algún punto $\xi \in C$ tal que

$$\int_C f = f(\xi)\mu(C)$$

Demostración

1.° Propiedades aritméticas. Como f y g son integrables en C, los conjuntos D_f y D_g de discontinuidad de f y de g tienen medida nula (véase [67]). Los conjuntos de los puntos de discontinuidad de $f + g$, cf, $|f|$, fg y f/g (esta última cumpliendo la exigencia del enunciado) están incluidas en $D_f \cup D_g$ y, como éste tiene medida nula (véase [63], 1 y 2), también lo tienen aquéllos, luego las anteriores funciones son integrables en C. Además, acudiendo a las funciones $f_C : \mathbb{R}^p \to \mathbb{R}$ y $g_C : \mathbb{R}^p \to \mathbb{R}$ (que coinciden con f y g en C y son nulas en $\mathbb{R}^p - C$) y a un intervalo compacto I que contenga a C, de acuerdo con la linealidad de las integrales en intervalos (véase [54]) se puede poner:

$$\int_C (f + g) = \int_I (f + g)_C = \int_I (f_C + g_C) = \int_I f_C + \int_I g_C = \int_C f + \int_I g$$

$$\int_C cf = \int_I (cf)_C = \int_I cf_C = c \int_I f_C = c \int_C f$$

2.° Monotonía de la integral. Acudiendo, como antes, a f_C, g_C e I y de acuerdo con la propiedad de monotonía de las integrales en intervalos, como $f_C(\boldsymbol{x}) \leqslant g_C(\boldsymbol{x})$ para todo $\boldsymbol{x} \in I$, se verifica que:

$$\int_C f = \int_I f_C \leqslant \int_I g_C = \int_C g$$

Aplicando este último resultado a f y $|f|$, como $-|f(\boldsymbol{x})| \leqslant f(\boldsymbol{x}) \leqslant |f(\boldsymbol{x})|$ para todo $\boldsymbol{x} \in C$, se concluye que

$$-\int_C |f| = \int_C -|f| \leqslant \int_C f \leqslant \int_C |f|, \qquad \text{luego } \left|\int_C f\right| \leqslant \int_C f$$

3.° Teorema de la media. Como $m \leqslant f(\boldsymbol{x}) \leqslant M$ para todo $\boldsymbol{x} \in C$, de la propiedad de monotonía se desprende que

$$\int_C m \leqslant \int_C f \leqslant \int_C M, \qquad \text{o sea } \; m\mu(C) \leqslant \int_C f \leqslant M\mu(C)$$

Si es $\mu(C) = 0$, la relación a demostrar se verifica para cualquier $h \in \mathbb{R}$. Si es $\mu(C) \neq 0$, para $h = (\int_C f) : \mu(C)$ se verifica que $m \leqslant h \leqslant M$ y $\int_C f = h\mu(C)$. Si, además, f es continua en C y C es conexo y compacto, entonces: 1.°, según el teorema de Weierstrass (véase [18], II), existen $m = \text{mín } f(C)$ y $M = \text{máx } f(C)$; 2.°, como m y M son valores de f en C y como $m \leqslant h \leqslant M$, según la propiedad de Darboux (véase [19]) existe $\xi \in C$ tal que $f(\xi) = h$, por lo que $\int_C f = f(\xi)\mu(C)$.

[68]₁ Corolario (integración en un conjunto de medida nula)

Si $C \subset \mathbb{R}^p$ es un conjunto de medida nula, y si $f: C \to \mathbb{R}$ es una función acotada en C, entonces f es integrable en C y es $\int_C f = 0$.

Comprobación

El conjunto de los puntos de discontinuidad de f, como está incluido en C y éste tiene medida nula, es también de medida nula, luego f es integrable en C (véase [67]). Como f está acotada en C, existen cotas, m y M, inferior y superior de f en C, con lo que (según [68],3) es $m\mu(C) \leqslant \int_C f \leqslant M\mu(C)$, luego como $\mu(C) = 0$, es $0 \leqslant \int_C f \leqslant 0$, esto es, $\int_C f = 0$.

$[68]_2$ Ejercicio

Sea $C \subset \mathbb{R}^p$ un conjunto medible; sea $f: C \to \mathbb{R}$ una función acotada y llamemos $m = \inf f(C)$ y $M = \sup f(C)$; considérese una función $\varphi: [m, M] \to \mathbb{R}$. Si f es integrable en C y si φ es continua en $[m, M]$, pruébese que entonces la función compuesta $\varphi \circ f: C \to \mathbb{R}$ es integrable en I.

Resolución

Como f es integrable en C, el conjunto $D(f)$ de los puntos de discontinuidad de f tiene medida nula (véase [67]). Como el conjunto $D(\varphi \circ f)$, de los puntos de discontinuidad de $\varphi \circ f$, está incluido en $D(f)$ y éste tiene medida nula, $D(\varphi \circ f)$ tiene medida nula, luego $\varphi \circ f$ es integrable en C.

$[68]_3$ Desigualdad de Cauchy-Schwarz

Sean $f: C \to \mathbb{R}$ y $g: C \to \mathbb{R}$ dos funciones acotadas en un conjunto medible $C \subset \mathbb{R}^p$. Si f y g son integrables en C, entonces se verifica que

$$\left(\int_C fg\right)^2 \leqslant \left(\int_C f^2\right)\left(\int_C g^2\right)$$

Comprobación

Acudiendo a las funciones $f_C: \mathbb{R}^p \to \mathbb{R}$ y $g_C: \mathbb{R}^p \to \mathbb{R}$ (que coinciden con f y g en C y son nulas en $\mathbb{R}^p - C$) y a un intervalo compacto I que contenga a C, de acuerdo con la desigualdad de Cauchy-Schwarz para intervalos (véase $[55]_1$) se puede poner:

$$\left(\int_C fg\right)^2 = \left(\int_I (fg)_C\right)^2 = \left(\int_I f_C g_C\right)^2 \leqslant \left(\int_I f_C^2\right)\left(\int_I f_C^2\right) =$$
$$= \left(\int_I (f^2)_C\right)\left(\int_I (g^2)_C\right) = \left(\int_C f^2\right)\left(\int_C g^2\right)$$

$[68]_4$ Generalización del teorema de la media

Sean $f: C \to \mathbb{R}$ y $g: C \to \mathbb{R}$ dos funciones acotadas en un conjunto medible $C \subset \mathbb{R}^p$.

1.º Si f y g son integrables en C y si $m \leqslant f(x) \leqslant M$ y $g(x) \geqslant 0$ para todo $x \in C$, entonces existe $h \in [m, M]$ tal que

$$\int_C fg = h \int_C g$$

2.° Si C es compacto y conexo, si f es continua en C, si g es integrable en C y si $g(\boldsymbol{x}) \geqslant 0$ para todo $\boldsymbol{x} \in C$, entonces existe algún punto $\xi \in C$ tal que

$$\int_C fg = f(\xi) \int_C g$$

Demostración

1.° Por ser $m \leqslant f(\boldsymbol{x}) \leqslant M$ y $g(\boldsymbol{x}) \geqslant 0$, es $mg(\boldsymbol{x}) \leqslant Mg(\boldsymbol{x})$ para todo $\boldsymbol{x} \in C$ y, por la monotonía de la integral (véase [68], 2), se verifica que $m \int_C f \leqslant \int_C fg \leqslant M \int_I g$. Dividiendo esta desigualdad por $\int_C g > 0$ (si fuese $\int_C g = 0$, la igualdad a demostrar se verificaría trivialmente para cualquier $h \in \mathbb{R}$), se concluye que para $h = (\int_C fg) : (\int_C g)$ se verifica que $m \leqslant h \leqslant M$ e $\int_C fg = h \int_C g$.

2.° Si, además, f es continua en C y C es conexo y compacto, entonces existen $m = \text{mín} f(C)$ y $M = \text{máx } f(C)$ (aplíquese el teorema de Weierstrass; [18]) y existe algún punto $\xi \in C$ tal que $f(\xi) = h$ (aplíquese la propiedad de Darboux; [19]), por lo que se verifica que $\int_C fg = f(\xi) \int_C g$.

ADITIVIDAD DE LA INTEGRAL (EN CONJUNTOS MEDIBLES)

[69]

1. Sea $f: C \to \mathbb{R}$ una función acotada en un conjunto medible $C \subset \mathbb{R}^p$; sea C_0 un subconjunto medible de C. Si f es integrable en C, entonces f es integrable en C_0.
2. Sean C_1 y C_2 dos conjuntos medibles de $\mathbb{R}^p$; sea $f: C_1 \cup C_2 \to \mathbb{R}$ una función acotada en $C_1 \cup C_2$. Si f es integrable en C_1 y en C_2, entonces f es integrable en $C_1 \cup C_2$ y en $C_1 \cap C_2$, y se verifica que

$$\int_{C_1 \cup C_2} f = \int_{C_1} f + \int_{C_2} f - \int_{C_1 \cap C_2} f$$

(si $C_1 \cap C_2 = \varnothing$, entonces existe $\int_{C_1 \cup C_2} f$ y es $\int_{C_1 \cup C_2} f = \int_{C_1} f + \int_{C_2} f$)

3. Sea $f: C \to \mathbb{R}$ una función acotada sobre un conjunto medible $C \subset \mathbb{R}^p$; sea $N \subset C$ un conjunto de medida nula. Si f es integrable en C, entonces f es integrable en $C - N$ y se verifica que

$$\int_{C-N} f = \int_C f \qquad (N \text{ es de medida nula})$$

4. Sean C_1 y C_2 dos conjuntos medibles de $\mathbb{R}^p$; sea $f: C_1 \cup C_2 \to \mathbb{R}$ una función acotada en $C_1 \cup C_2$. Si f es integrable en C_1 y en C_2, entonces:

$$\mathring{C}_1 \cap \mathring{C}_2 = \varnothing \quad \Rightarrow \quad \int_{C_1 \cup C_2} f = \int_{C_1} f + \int_{C_2} f$$

Demostración

En lo que sigue, llamaremos $D(F, A)$ al conjunto de los puntos de discontinuidad de una función $F: A \to \mathbb{R}$ y acudiremos a la caracterización de la integridad [67]: la función acotada F es integrable en el conjunto medible A si, y sólo si, $D(F, A)$ tiene medida nula.

1. Como f es integrable en C, el conjunto $D(f, C)$ tiene medida nula; dado que $C_0 \subset C$, se verifica que $D(f, C_0) \subset D(f, C)$, por lo que $D(f, C_0)$ tiene medida nula, luego f es integrable en C_0.
2. Como f es integrable en C_1 y en C_2, entonces $D(f, C_1)$ y $D(f, C_2)$ son de medida nula; ahora bien, como $D(f, C_1 \cup C_2)$ y $D(f, C_1 \cap C_2)$ están incluidos en la unión de $D(f, C_1)$ y $D(f, C_2)$ y dicha unión tiene medida nula (véase [63]), resulta que $D(f, C_1 \cup C_2)$ y $D(f, C_1 \cap C_2)$ también tienen medida nula, por lo que f es integrable en $C_1 \cup C_2$ y en $C_1 \cap C_2$. Llamemos (como en [67]) $F_A: \mathbb{R}^p \to \mathbb{R}$ a la función (asociada a una función dada $F: A \to \mathbb{R}$) definida por $F_A(x) = 1$, si $x \in A$ y $F_A(x) = 0$, si $x \notin A$; es evidente que

$$f_{C_1 \cup C_2}(x) = f_{C_1}(x) + f_{C_2}(x) - f_{C_1 \cap C_2}(x), \qquad \forall\, x \in \mathbb{R}^p$$

(los anteriores primero y segundo miembro son iguales en todos los casos, esto es, para x recorriendo $C_1 - C_2$, $C_2 - C_1$, $C_1 \cap C_2$ y $\mathbb{R}^p - C_1 \cup C_2$; en efecto, para cada uno de ellos se tiene $f(x) = f(x) + 0 - 0$, $f(x) = 0 + f(x) - 0$, $f(x) = f(x) + f(x) - f(x)$ y $0 = 0 + 0 - 0$, respectivamente). Por ello, acudiendo a un intervalo compacto I tal que $C_1 \cup C_2 \subset I$ y según la linealidad de la integral en intervalos (véase [54], 1.º), resulta que:

$$\int_{C_1 \cup C_2} f = \int_I f_{C_1 \cup C_2} = \int_I f_{C_1} + \int_I f_{C_2} - \int_I f_{C_1 \cap C_2} = \int_{C_1} f + \int_{C_2} f - \int_{C_1 \cap C_2} f$$

3. De acuerdo con la anterior propiedad 1, existe $\int_{C-N} f$. Como N tiene medida nula, es $\int_N f = 0$ (véase $[68]_1$). Dado que $C = (C - N) \cup N$ y que $(C - N) \cap N = \varnothing$, de la propiedad anterior se desprende que

$$\int_C f = \int_{C-N} f + \int_N f = \int_{C-N} f + 0 = \int_{C-N} f$$

4. Como $\mathring{C}_1 \cap \mathring{C}_2 = \varnothing$, se verifica que $C_1 \cap C_2 \subset F_r(C_1) \cup F_r(C_2)$. Debido a que C_1 y C_2 son medibles, se sabe que $F_r(C_1)$ y $F_r(C_2)$ tienen contenido nulo (véase [66]), luego tienen medida nula; por tanto, también tiene medida nula el conjunto $C_1 \cap C_2$ (véase [63], 1 y 2), por lo que $\int_{C_1 \cap C_2} f = 0$. De la anterior propiedad 2 se deduce, pues, la propiedad de ahora.

$[69]_1$ Observaciones

1.º Sean $f: C \to \mathbb{R}$ y $g: C \to \mathbb{R}$ dos funciones acotadas en un conjunto medible $C \subset \mathbb{R}^p$. Si $f(x) = g(x)$ para todo $x \in C - N$, donde N es un conjunto de medida nula, y f es integrable en C, entonces también g es integrable en C y se verifica que $\int_C f = \int_C g$.

2.º Sea $C \subset \mathbb{R}^p$ un conjunto medible; sea $N \subset \mathbb{R}^p$ un conjunto de contenido nulo; sea $f: C \cup N \to \mathbb{R}$ una función acotada. Si f es integrable en C, entonces f es integrable en $C \cup N$ (que es medible), y se verifica que

$$\int_{C \cup N} f = \int_C f \qquad (N \text{ es de contenido nulo})$$

3.º Sea $f: C \to \mathbb{R}$ una función acotada en un conjunto medible $C \subset \mathbb{R}^p$; supongamos que f se extiende a $\bar{C}$ manteniéndose acotada. Si f es integrable en C, entonces también es integrable en $\mathring{C}$, en $\bar{C}$ y en todo C' verificando que $\mathring{C} \subset C' \subset \bar{C}$ (que son conjuntos medibles), siendo además:

$$\int_{\mathring{C}} f = \int_C f = \int_{C'} f = \int_{\bar{C}} f$$

Demostración

1.º De acuerdo con la anterior propiedad [69], 3, se tiene que:

$$\int_C f = \int_{C-N} f = \int_{C-N} g = \int_C g$$

2.º La frontera de $C \cup N$ está incluida en $F_r(C) \cup N$ y, como este conjunto tiene contenido nulo ($F_r(C)$ tiene contenido nulo por ser C medible); resulta que también es de contenido nulo la frontera de $C \cup N$. Este conjunto $C \cup N$, como además de acotado por serlo C y N, es, pues, medible. El conjunto $D(C \cup N)$ de los puntos de discontinuidad de $f: C \cup N \to \mathbb{R}$ está includo en $D(C) \cup N$, donde $D(C)$ es el conjunto de los puntos de discontinuidad de $f: C \to \mathbb{R}$, y como estos dos conjuntos tienen medida nula ($D(C)$ tiene medida nula por ser $f: C \to \mathbb{R}$ integrable en C) también $D(C \cup N)$ tiene medida nula, luego f es integrable en $C \cup N$ (que es medible). Sabiendo ya que existe $\int_{C \cup N} f$, como $\int_N f = 0$ e $\int_{C \cap N} f = 0$ (pues N y $C \cap N$ son de medida nula), de la propiedad [69],2 se desprende que:

$$\int_{C \cup N} f = \int_C f + \int_N f - \int_{C \cap N} f = \int_C f + 0 - 0 = \int_C f$$

3.º Como $\bar{C} = C \cup N$, donde $N = F_r(C)$ es un conjunto de contenido nulo (pues C es medible), es de aplicación el resultado anterior del que se desprende, pues, que $\bar{C}$ es medible, que existe $\int_{\bar{C}} f$ y que es $\int_{\bar{C}} f = \int_C f$. Como $C' = \bar{C} - N$ para cierto $N \subset F_r(C)$, como $\bar{C}$ es medible y como N tiene contenido nulo (por tenerlo $F_r(C)$), de la propiedad [69], 3 se deduce que f es integrable en C' (que es medible) y que $\int_{C'} f = \int_C f$. Finalmente, para $\mathring{C}$ es válido, en particular, lo que se acaba de obtener para cualquier C' que verifique a $\mathring{C} \subset C' \subset \bar{C}$.

INTEGRACIÓN ITERADA (EN CONJUNTOS MEDIBLES)

[70]

I. TEOREMA DE FUBINI. Sea $f: C \to \mathbb{R}$ una función acotada en un conjunto medible $C \subset \mathbb{R}^p$. Para cada $\boldsymbol{x} = (x_1, ..., x_r, x_{r+1}, ..., x_p) \in \mathbb{R}^p$, sean $\boldsymbol{x}' = (x_1, ..., x_r)$ y $\boldsymbol{x}'' = (x_{r+1}, ..., x_p)$, con lo que $\boldsymbol{x} = (\boldsymbol{x}', \boldsymbol{x}'')$; llamaremos $C' \subset \mathbb{R}^r$ a la proyección de C sobre el subespacio de $\mathbb{R}^r$ que engendrán sus r primeros ejes; para cada $\boldsymbol{x}' \in C'$, llamaremos $C''(\boldsymbol{x}') = \{\boldsymbol{x}'' \in \mathbb{R}^{p-r} / \boldsymbol{x} = (\boldsymbol{x}', \boldsymbol{x}'') \in C\}$. Si f es integrable en C y los conjuntos C' y $C''(\boldsymbol{x}')$ son medibles (este último para todo $\boldsymbol{x}' \in C'$), entonces se verifica que (además de existir las siguientes integrales en C') es:

$$\int_C f(\boldsymbol{x})\, d\boldsymbol{x} = \int_{C'} \left(\underline{\int}_{C''(\boldsymbol{x}')} f(\boldsymbol{x}', \boldsymbol{x}'')\, d\boldsymbol{x}'' \right) d\boldsymbol{x}' = \int_{C'} \left(\overline{\int}_{C''(\boldsymbol{x}')} f(\boldsymbol{x}', \boldsymbol{x}'')\, d\boldsymbol{x}'' \right) d\boldsymbol{x}'$$

II. CASO DE INTEGRACIÓN EN CONJUNTOS SIMPLES. Sea S el conjunto simple (véase [66]) de $\mathbb{R}^p$ que está definido por las funciones acotadas $\varphi : C \to \mathbb{R}$ y $\psi : C \to \mathbb{R}$ (siendo $\varphi(\boldsymbol{x}) \leqslant \psi(\boldsymbol{x})$ para todo $\boldsymbol{x} \in C$ y donde $C \subset \mathbb{R}^{p-1}$ es un conjunto medible); esto es:

$$S = \{(\boldsymbol{x}, y) \in \mathbb{R}^p \,/\, \boldsymbol{x} \in C,\ \varphi(\boldsymbol{x}) \leqslant y \leqslant \psi(\boldsymbol{x})\}$$

Sea $f: S \to \mathbb{R}$ una función acotada en S. Si φ y ψ son integrables (o, en particular, continuas) en C, con lo que S es medible, y si f es integrable (o, en particular, continua) en S, entonces se verifica que

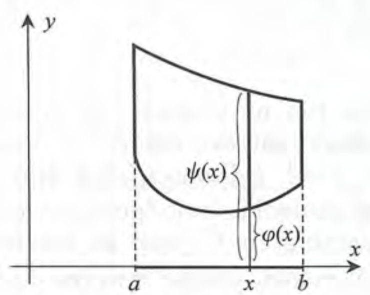
$$\int_S f \equiv \int_S f(\boldsymbol{x}, y)\, d\boldsymbol{x}\, dy = \int_C \left(\int_{\varphi(\boldsymbol{x})}^{\psi(\boldsymbol{x})} f(\boldsymbol{x}, y)\, dy \right) d\boldsymbol{x}$$

Nota: En los casos $p = 2$ y $p = 3$, la fórmula anterior toma la forma:

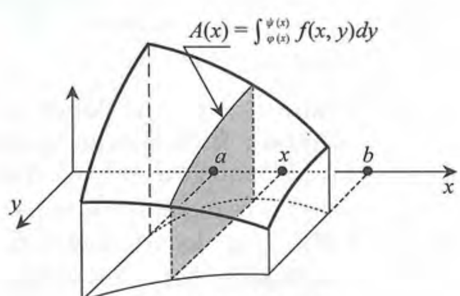
$$\int_S f(x, y)\, dx\, dy = \int_a^b \left(\int_{\varphi(x)}^{\psi(x)} f(x, y)\, dy \right) dx, \qquad \text{donde } S = \{(x, y) \in \mathbb{R}^2 \,/\, a \leqslant x \leqslant b,\ \varphi(x) \leqslant y \leqslant \psi(x)\}$$

$$\int_S f(x, y, z)\, dx\, dy\, dz = \int_C \left(\int_{\varphi(x, y)}^{\psi(x, y)} f(x, y, z)\, dz \right) dx\, dy,$$

donde $S = \{(x, y, z) \in \mathbb{R}^3 \,/\, (x, y) \in C,\ \varphi(x, y) \leqslant z \leqslant \psi(x, y)\}$



S: $a \leqslant x \leqslant b$, $\varphi(x) \leqslant y \leqslant \psi(x)$



VOLUMEN $= \int_a^b A(x)\, dx$

Demostración

I. Este resultado es una consecuencia inmediata del teorema de Fubini para intervalos (véase [60]): acudiendo a cualquier intervalo compacto $I \subset \mathbb{R}^p$ que incluya a C, al aplicar el referido teorema [60] y con la notación que allí se usa (véase la nota al pie), se tiene:

$$\int_C f(\boldsymbol{x})\,d\boldsymbol{x} = \int_I f(\boldsymbol{x})\,d\boldsymbol{x} = \int_{I'} \left(\int_{I''} f(\boldsymbol{x}', \boldsymbol{x}'')\,d\boldsymbol{x}''\right) d\boldsymbol{x}' = \int_{C'} \left(\int_{C''(\boldsymbol{x}')} f(\boldsymbol{x}', \boldsymbol{x}'')\,d\boldsymbol{x}''\right) d\boldsymbol{x}'$$

(otro tanto se puede hacer para el caso de la integral superior).

II. Sean, respectivamente, α y β cotas inferior y superior de φ y ψ en C y sea J un intervalo compacto que contenga a C, con lo que S está incluido en el intervalo compacto $I = [\alpha, \beta] \times J$. Sea $f_S : \mathbb{R}^p \to \mathbb{R}$ la función definida por $f_S(\boldsymbol{x}) = f(\boldsymbol{x})$, si $\boldsymbol{x} \in S$, y $f_S(\boldsymbol{x}) = 0$, si $\boldsymbol{x} \notin S$. A la función $f_S : I \to \mathbb{R}$ le es de aplicación el teorema de Fubini (véase [60] y sustitúyanse f, I' e I'' por f_S, J y $[\alpha, \beta]$, respectivamente), del que se desprende que

$$\int_S f = \int_J \left(\int_\alpha^\beta f_S(\boldsymbol{x}, y)\,dy\right) d\boldsymbol{x} \qquad [1]$$

Ahora bien, para cada $\boldsymbol{x} \in J$, el integrando de $\int_\alpha^\beta f_S(\boldsymbol{x}, y)\,dy$ es la función

$$[\alpha, \beta] \to \mathbb{R},\ y \mapsto f_S(\boldsymbol{x}, y) = \begin{cases} f(\boldsymbol{x}, y)\chi_C(\boldsymbol{x}), & \text{si } y \in [\varphi(\boldsymbol{x}), \psi(\boldsymbol{x})] \\ 0, & \text{si } y \notin [\varphi(\boldsymbol{x}), \psi(\boldsymbol{x})] \end{cases}$$

(donde $\chi_C : J \to \mathbb{R}$ es la función característica del conjunto C), con lo que

$$\int_\alpha^\beta f_S(\boldsymbol{x}, y)\,dy = \int_{\varphi(\boldsymbol{x})}^{\psi(\boldsymbol{x})} f(\boldsymbol{x}, y)\chi_C(\boldsymbol{x})\,dy = \chi_C(\boldsymbol{x}) \int_{\varphi(\boldsymbol{x})}^{\psi(\boldsymbol{x})} f(\boldsymbol{x}, y)\,dy$$

Llevando este resultado a [1] se obtiene finalmente que:

$$\int_S f = \int_J \chi_C(\boldsymbol{x})\left(\int_{\varphi(\boldsymbol{x})}^{\psi(\boldsymbol{x})} f(\boldsymbol{x}, y)\,dy\right) d\boldsymbol{x} = \int_C \left(\int_{\varphi(\boldsymbol{x})}^{\psi(\boldsymbol{x})} f(\boldsymbol{x}, y)\,dy\right) d\boldsymbol{x}$$

como había que comprobar.

[70]₁ Observaciones sobre el cálculo de integrales

1. Supongamos que se desea calcular la integral doble $\int_S f(x, y)\,dx\,dy$, de una función $f : S \to \mathbb{R}$, integrable en el conjunto medible S. Si S es un conjunto simple, como el recién descrito en la «nota» de [70], entonces la fórmula que allí se da reduce el cálculo de la integral doble al de dos integrales simples. Pero, en lugar de ello, también podría ocurrir que S fuera un conjunto simple respecto de la otra variable:

$$S = \{(x, y) \in \mathbb{R}^2 \;/\; c \leqslant y \leqslant d,\ \alpha(y) \leqslant x \leqslant \beta(y)\}$$

(donde $\alpha:[c, d] \to \mathbb{R}$ y $\beta:[c, d] \to \mathbb{R}$ son funciones integrables y tales que $\alpha(y) \leqslant \beta(y)$ para $c \leqslant y \leqslant d$); en tal caso, la reducción a dos integrales simples viene dada por

$$\int_S f(x, y)\, dx\, dy = \int_c^d \left(\int_{\alpha(y)}^{\beta(y)} f(x, y)\, dx \right) dy$$

Para calcular dicha integral $\int_S f$ por este procedimiento iterativo, se mirará si C es un conjunto simple, hallando las funciones φ y ψ o α y β, según el caso. En general, aunque el conjunto C no sea simple, puede descomponerse en dos o más conjuntos simples; de ser así, se expresará $\int_S f$ como suma de integrales que pueden hallarse por integración iterada.

2. En el caso de una integral triple $\int_S f(x, y, z)\, dx\, dy\, dz$, en donde S es un conjunto simple respecto de la variable z, como el que se indica en la anterior «nota» de [70], entonces la fórmula que allí figura reduce el cálculo de dicha integral triple al de una integral simple y, después, de una integral doble; si el conjunto sobre el que integra en esta última es, a su vez, un conjunto simple, ella se reducirá a dos integrales simples. Así, por ejemplo, si se cumplen las condiciones de regularidad exigidas en [70], II, es:

$$\int_S f(x, y, z)\, dx\, dy\, dz = \int_a^b \left\{ \int_{\alpha(x)}^{\beta(x)} \left(\int_{\varphi(x, y)}^{\psi(x, y)} f(x, y, z)\, dz \right) dy \right\} dx$$

siendo

$$S = \{(x, y, z) \in \mathbb{R}^3 \ / \ a \leqslant x \leqslant b,\ \alpha(x) \leqslant y \leqslant \beta(x),\ \varphi(x, y) \leqslant z \leqslant \psi(x, y)\}$$

Para aplicar la integración iterada al cálculo de una integral triple $\int_S f$, habrá que analizar si S es un conjunto simple respecto de alguna de las tres variables, a la que llamaremos z, localizando las correspondientes funciones φ y ψ y el conjunto C; nótese que este conjunto es, entonces, la proyección de S sobre el plano $z = 0$, paralelamente al «eje de la z». En general, S no será un conjunto simple; no obstante, es usual que S se pueda descomponer en conjuntos simples, lo que permite expresar la integral (acudiendo a la propiedad aditiva) como suma de integrales que pueden calcularse recurriendo a la integración iterada.

[70]$_2$ Ejercicios

Calcular las siguientes integrales dobles en los conjuntos S que se indican:

1. $I = \iint_S (2\sqrt{x} - 3y^2)\, dx\, dy, \qquad S = \{(x, y) \in \mathbb{R}^2 \ / \ 0 \leqslant x \leqslant 1,\ x^2 \leqslant y \leqslant \sqrt[4]{x}\}.$

2. $J = \iint_S xy\, dx\, dy$, S es el conjunto acotado del plano que está limitado por los tres arcos de curva: 1.ª, $y = 9 - x^2$ para $0 \leqslant x \leqslant 3$; 2.ª, $y = 9(1 - x^2)$ para $0 \leqslant x \leqslant 1$; 3.ª, $y = (x - 1)(x - 2)^2(x - 3)$ para $1 \leqslant x \leqslant 3$.

Resolución

$$I = \int_0^1 dx \int_{x^2}^{x^{1/4}} (2\sqrt{x} - 3y^2)\, dy = \int_0^1 dx[2\sqrt{x}y - y^3]_{y=x^2}^{y=x^{1/4}} = \int_0^1 (x^{3/4} - 2x^{5/2} + x^6)\, dx =$$

$$= \left[\frac{4}{7}x^{7/4} - \frac{4}{7}x^{7/2} + \frac{1}{7}x^7\right]_0^1 = \frac{1}{7}$$

Para calcular J, dividiremos S en dos subconjuntos S_1 y S_2 mediante la recta $y = 0$ (llamamos S_1 al subconjunto situado en el primer cuadrante); con ello, se tiene:

$$J_1 = \iint_{S_1} xy\, dx\, dy = \int_0^9 dy \int_{\sqrt{1-y/9}}^{-\sqrt{9-y}} xy\, dx = \int_0^9 dy\left[\frac{1}{2}x^2y\right]_{x=\sqrt{1-y/9}}^{x=\sqrt{9-y}} =$$

$$= \int_0^9 \left(4y - \frac{4}{9}y^2\right) dy = \left[2y^2 - \frac{4}{27}y^3\right]_0^9 = 54$$

$$J_2 = \iint_{S_2} xy\, dx\, dy = \int_1^3 dx \int_{(x-1)(x-2)^2(x-3)}^{0} xy\, dy = \int_1^3 dx\left[\frac{1}{2}xy^2\right]_{y=(x-1)(x-2)^2(x-3)}^{y=0} =$$

$$= \int_1^3 -\frac{1}{2}(x^9 - 16x^8 + 110x^7 - 424x^6 + 1001x^5 -$$

$$- 1480x^4 + 1336x^3 - 672x^2 + 144x)\, dx = 4{,}282\ldots$$

$$J = J_1 + J_2 = 58{,}282\ldots$$

[70]$_3$ Ejercicio

Calcular la siguiente integral triple:

$$I = \iiint_S z\, dx\, dy\, dz,$$

$$\text{donde } S = \left\{(x, y, z) \in \mathbb{R}^3 \;/\; x \geqslant 0,\, y \geqslant 0,\, z \geqslant 0,\, \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leqslant 1\right\}$$

Resolución

$$I = \int_0^a dx \int_0^{y(x)} dy \int_0^{z(x,\,y)} z\, dz, \qquad \text{donde} \begin{cases} y(x) = b(1 - x^2/a^2)^{1/2} \\ z(x,\, y) = c(1 - x^2/a^2 - y^2/b^2)^{1/2} \end{cases}$$

$$I = \int_0^a dx \int_0^{y(x)} dy\left[\frac{1}{2}z^2\right]_0^{z(x,\,y)} = \int_0^a dx \int_0^{y(x)} \frac{c^2}{2}\left(1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}\right) dy =$$

$$= \int_0^a dx\left[\frac{c^2}{2}\left(y - \frac{x^2}{a^2}y - \frac{y^3}{3b^2}\right)\right]_0^{y(x)} = \int_0^a \frac{c^2b}{3a^3}(a^2 - x^2)^{3/2} dx =$$

$$= \frac{c^2b}{3a^3}\left[\frac{x(a^2 - x^2)^{3/2}}{4} + \frac{3a^2x(a^2 - x^2)^{1/2}}{8} + \frac{3}{8}a^2 \operatorname{arcsen}\frac{x}{a}\right]_0^a = \frac{1}{16}\pi abc^2$$

$[70]_4$ Fórmulas integrales de áreas y volúmenes

- Según sabemos (véase $[67]_1$), la medida $\mu(C)$, de un conjunto medible $C \subset \mathbb{R}^p$, viene dada por la integral $\mu(C) = \int_C 1$. En particular, si el conjunto medible C es plano ($p = 2$) o si es tridimensional ($p = 3$), entonces su medida, que se llama área o volumen, respectivamente, vendrá dada por:

$$[\text{Área de } C \subset \mathbb{R}^2 \text{ (medible)}] = \alpha(C) = \iint_C dx\,dy \qquad [1]$$

$$[\text{Volumen de } C \subset \mathbb{R}^3 \text{ (medible)}] = \upsilon(C) = \iiint_C dx\,dy\,dz \qquad [2]$$

- Si el conjunto medible $C \subset \mathbb{R}^2$ es un conjunto simple, acudiendo a la integración iterada (véase [70],II) de la anterior expresión [1] del área de C, $\alpha(C)$, se desprende que

$$\alpha(C) = \iint_C dx\,dy = \int_a^b dx\left(\int_{\varphi(x)}^{\psi(x)} dy\right) = \int_a^b (\psi(x) - \varphi(x))\,dx$$
$$\text{donde } C = \{(x, y) \in \mathbb{R}^2 \;/\; a \leqslant x \leqslant b,\ \varphi(x) \leqslant y \leqslant \psi(x)\}$$

Este resultado también es de utilidad en el caso en que el conjunto C, sin ser un conjunto simple, puede descomponerse en dos o más conjuntos que sí lo sean.

- Si C es un conjunto medible tridimensional, $C \subset \mathbb{R}^3$, que es simple respecto de una de las coordenadas, la z, la integral [2] que da su volumen se puede expresar recurriendo a la integración iterada:

$$\upsilon(C) = \iiint_C dx\,dy\,dz = \iint_{C_0} dx\,dy\left(\int_{\alpha(x,y)}^{\beta(x,y)} dz\right) = \iint_{C_0} (\beta(x, y) - \alpha(x, y))\,dx\,dy \qquad [3]$$
$$\text{donde } C = \{(x, y, z) \in \mathbb{R}^3 \;/\; (x, y) \in C_0,\ \alpha(x, y) \leqslant z \leqslant \beta(x, y)\}$$

También ahora, si C no es simple, pero se puede descomponer en varios conjuntos simples, el volumen de C se puede expresar como la suma de las correspondientes integrales, extendidas a los subconjuntos.

- Un caso que aparece muy frecuentemente en las aplicaciones es el que se obtiene para $\alpha = 0$ en la expresión [3]. Se trata, pues, del volumen del conjunto $C \subset \mathbb{R}^3$ limitado por: 1.º, el plano $z = 0$; 2.º, el cilindro de generatrices paralelas al «eje de la z» y que tiene por directriz al contorno (frontera) de cierto conjunto medible C_0 de dicho plano, y 3.º, la superficie de ecuación $z = \beta(x, y) \geqslant 0$. Dicho volumen viene, entonces, dado por la fórmula

$$\upsilon(C) = \iint_{C_0} \beta(x, y)\,dx\,dy$$

De lo dicho resulta, entonces, que la integral doble de una función positiva β en un conjunto medible C_0 se puede interpretar geométricamente como un volumen. Así pues, a veces, los volúmenes pueden expresarse también (además de hacerlo mediante integrales triples) acudiendo a una integral doble. Algo análogo ocurre con las áreas: aun cuando las áreas quedan definidas por integrales dobles, en ocasiones pueden expresarse, como ya es conocido, recurriendo a integrales simples.

- Dado un conjunto medible $C \subset \mathbb{R}^3$, recurriendo a un intervalo compacto que le contenga, pongamos $C \subset [\alpha, \beta] \times J$, donde $J \subset \mathbb{R}^2$ es un intervalo compacto, la integración iterada permite poner la igualdad [2] en la forma

$$v(C) = \int_\alpha^\beta dz\left(\iint_J f_C(x, y, z)\,dx\,dy\right) = \int_\alpha^\beta dz\left(\iint_{C(z)} f(x, y, z)\,dx\,dy\right)$$

donde $f_C(x, y, z) = \begin{cases} f(x, y, z), & \text{si } (x, y, z) \in C \\ 0, & \text{si } (x, y, z) \notin C \end{cases}$; $C(z) = C \cap \{(x, y, z) \,/\, (x, y) \in \mathbb{R}^2\}$

Para cada $z \in [\alpha, \beta]$ fijo, la integral $\int_{C(z)} f(x, y, z)\,dx\,dy$ es el área del conjunto plano $C(z)$, que es la sección del conjunto C por el correspondiente plano perpendicular al «eje de la z», por lo que se puede poner

$$v(C) = \int_\alpha^\beta a(z)\,dz, \text{ donde } a(z) = \text{área de la sección de } C \text{ por el plano de «cota» } z \quad [4]$$

que expresa el volumen como una integral simple (nótese que $a(z)$ se calculará, en general, mediante una integral doble).

- Principio de Cavalieri. Si dos conjuntos medibles, C_1 y C_2, tienen la misma altura y, al cortarlos por planos «horizontales» (perpendiculares del «eje de la z»), las secciones en uno y otro tienen la misma área, entonces los volúmenes de C_1 y de C_2 son iguales. Este resultado es consecuencia inmediata de [4] pues, al aplicar esta fórmula a C_1 y a C_2, las $a(z)$ son iguales en ambos.

[70]$_5$ Ejercicio

Hallar el área de la región plana R situada (en ejes oxy usuales): «debajo» de la parábola $y = 4x - x^2$, «encima» del eje $y = 0$ y «a la derecha» de la recta $y = 6 - 3x$.

Resolución

Llamando R_1 y R_2 a los subconjuntos de R que quedan a la derecha y a la izquierda de la recta $x = 2$, se obtiene:

$$a(R) = a(R_1) + a(R_2) = \iint_{R_1} dx\,dy + \iint_{R_2} dx\,dy = \int_1^2 dx \int_{6-3x}^{4x-x^2} dy + \int_2^4 dx \int_0^{4x-x^2} dy =$$

$$= \int_1^2 (4x - x^2 - 6 + 3x)\,dx + \int_2^4 (4x - x^2)\,dx = \left[\frac{7}{2}x^2 - \frac{x^3}{3} - 6x\right]_1^2 + \left[2x^2 - \frac{x^3}{3}\right]_2^4 = \frac{15}{2}$$

[70]$_6$ Ejercicio

Hallar el volumen del cuerpo sólido S que (en ejes $oxyz$ usuales) está situado «por debajo» del paraboloide $z = 1 - x^2 - y^2$ y «por encima» del plano $z = 1 - y$.

Resolución

El plano y el paraboloide se cortan a lo largo de una elipse cuya proyección sobre el plano $z = 0$ es la circunferencia $x^2 + y^2 - y = 0$. El volumen pedido es, pues:

$$v(S) = \int_0^1 dy \int_{-\sqrt{y-x^2}}^{\sqrt{y-x^2}} [(1 - x^2 - y^2) - (1 - y)]\, dx = \int_0^1 dy \left[(y - y^2)x - \frac{x^3}{3} \right]_{x=-\sqrt{y-y^2}}^{x=\sqrt{y-y^2}} =$$

$$= \frac{4}{3} \int_0^1 (y - y^2)^{3/2} dy \overset{(*)}{=} \frac{1}{12} \int_{-\pi/2}^{\pi/2} \cos^4 \theta \, d\theta \overset{(**)}{=} \frac{\pi}{32}$$

(*) Haciendo el cambio de variable $2y - 1 = \operatorname{sen} \theta$.
(**) Según la fórmula de Wallis, es $\int_0^{\pi/2} \cos^4 \theta \, d\theta = 3\pi/16$.

[70]$_7$ Ejercicio (medida de una bola de $\mathbb{R}^p$)

Dado $r > 0$, considérese la bola $B_p = \{x \in \mathbb{R}^p \,/\, \|x\| \leqslant r\}$. Se sabe que para $p = 1, 2$ y 3, la medida (longitud, área, volumen) de esta bola es $\mu(B_1) = 2r$, $\mu(B_2) = \pi r^2$ y $\mu(B_3) = (4\pi/3)r^3$. Pruébese que para $p \in \mathbb{N}$, es $\mu(B_p) = c_p r^p$, para un cierto $c_p \in \mathbb{R}$ independiente de r, que se pide.

Resolución

Acudiendo al método de inducción, suponiendo que es $\mu(B_{p-1}) = c_{p-1} r^{p-1}$, hemos de probar que es $\mu(B_p) = c_p r^p$, para cierto $c_p \in \mathbb{R}$ que no depende de r. Para cada $z \in [-r, r]$, consideremos el «hiperplano» de ecuación $x_p = z$ (se llama x_p a la coordenada p-ésima de x) secciona a la bola B_p según una bola de dimensión $p - 1$ y radio $(r^2 - z^2)^{1/2}$, a la que llamaremos $B(z)$. De acuerdo con la hipótesis de inducción, la medida de esta sección $B(z)$ es:

$$\mu(B(z)) = c_{p-1}(r^2 - z^2)^{(p-1)/2}$$

Echando mano de la integración iterada, la medida $\mu(B_p)$ se puede poner:

$$\mu(B_p) = \int_{B_p} 1 = \int_{-r}^{r} dz \left(\int_{B(z)} 1 \right) = \int_{-r}^{r} c_{p-1}(r^2 - z^2)^{(p-1)/2} dz \overset{(*)}{=} 2c_{p-1} r^p \int_0^{\pi/2} \cos^p \theta \, d\theta$$

(*) Cambio de variable $z = r \operatorname{sen} \theta$.

Por tanto, la propiedad a demostrar es efectivamente cierta: basta con tomar

$$c_p = 2c_{p-1}I_p, \qquad \text{donde } I_p = \int_0^{\pi/2} \cos^p \theta \, d\theta$$

Acudiendo a las expresiones de estas integrales, que vienen dadas por una fórmula de Wallis(*), resulta evidente que (según que p sea par, $p = 2n$, o que p sea impar, $p = 2n + 1$) se tiene:

$$\mu(B_{2n}) = \frac{\pi^n}{n!} r^{2n} \qquad \text{y} \qquad \mu(B_{2n+1}) = \frac{2^{2n+1}\pi^n n!}{(2n+1)!} r^{2n+1}$$

INTEGRACIÓN POR CAMBIO DE VARIABLE

Vamos aquí a generalizar, al caso de las integrales múltiples, la regla de sustitución o fórmula de cambio de variable de las integrales simples, esto es, aquella que dice que

$$\int_{\varphi(\alpha)}^{\varphi(\beta)} f(x)\, dx = \int_\alpha^\beta f(\varphi(t)) \frac{d\varphi(t)}{dt}\, dt$$

(donde $\varphi : [\alpha, \beta] \to \mathbb{R}$ es una función de clase $\mathscr{C}^1$ en $[\alpha, \beta]$ y f, función real definida en el intervalo $\varphi([\alpha, \beta])$, es continua en este intervalo). Esta generalización no es fácil; así como, para el caso de una variable, la fórmula anterior se prueba sin ninguna dificultad, la demostración de la correspondiente regla de sustitución para integrales múltiples es una tarea delicada y laboriosa.

[71] Sean $A \subset \mathbb{R}^p$ un conjunto abierto y $C \subset \mathbb{R}^p$ un conjunto medible tales que $\bar{C} \subset A$. Sea $\varphi : A \to \mathbb{R}^p$ una transformación inyectiva en A, de clase $\mathscr{C}^1$ en A y tal que su jacobiano no se anula en ningún punto de $\mathring{C}$ (o sea, $\det J\varphi(x) \neq 0$, $\forall\, x \in \mathring{C}$). Sea $f : \bar{D} \to \mathbb{R}$ una función continua en el conjunto $\bar{D} = \varphi(\bar{C})$. Se verifica entonces que $\int_D f = \int_C (f \circ \varphi)|\det J\varphi|$, es decir:

$$\int_D f(y)\, dy = \int_{\varphi^{-1}(D)} f(\varphi(x))|\det J\varphi(x)|\, dx \qquad [1]$$

La fórmula [1], de integración por sustitución, en la que φ es la función de cambio de variable, se puede poner:

$$\int_D f = \int_C f^*, \qquad \text{donde} \begin{cases} D = \varphi(C) \\ f^* = (f \circ \varphi)|\det J\varphi| \end{cases}$$

Más adelante (véase $[71]_5$) se enuncian, sin demostración, otros teoremas de cambio de variable; se trata de refinamientos del que aquí se considera.

(*) $\displaystyle\int_0^{\pi/2} \cos^{2n} \theta \, d\theta = \frac{(2n-1)(2n-3)\cdots 3\cdot 1}{2n(2n-2)\cdots 4\cdot 2}\cdot\frac{\pi}{2}$ e $\displaystyle\int_0^{\pi/2} \cos^{2n+1} \theta\, d\theta = \frac{2n(2n-2)\cdots 4\cdot 2}{(2n+1)(2n-1)\cdots 3\cdot 1}$

Demostración

Empecemos notando que, según lo obtenido en $[66]_3$ y $[66]_4$, es $D = \varphi(C)$ o $C = \varphi^{-1}(D)$, $F_r(D) = \varphi(F_r(C))$ o $F_r(C) = \varphi^{-1}(F_r(D))$, $\mathring{D} = \varphi(\mathring{C})$ o $\mathring{C} = \varphi^{-1}(\mathring{D})$ y D es un conjunto medible, por lo que existe $\int_D f$; nótese también que existe $\int_C (f \circ \varphi)|\det J\varphi|$ ya que, además de ser C medible, su integrando es una función continua en $\bar{C}$. Conviene también observar que, de acuerdo con $[69]_1$, 3.º y como $\mathring{C} = \varphi^{-1}(\mathring{D})$, es:

$$\int_D f = \int_{\mathring{D}} f \qquad \text{e} \qquad \int_{\varphi^{-1}(D)} f^* = \int_{\varphi^{-1}(\mathring{D})} f^* \qquad (\text{se llama } f^* = (f \circ \varphi)|\det J\varphi|)$$

por lo que para demostrar [1] se puede sustituir D por $\mathring{D}$ o, si se prefiere, puede suponerse que D es abierto.

Para probar este teorema, empezaremos considerando casos particulares y, apoyándonos para cada uno de ellos, en lo probado en el anterior, mostraremos finalmente que es cierto el caso general del enunciado:

PRIMER ASERTO. Si la fórmula [1], de cambio de variable, se verifica cuando el conjunto C se sustituye por cualquier intervalo compacto I incluido en $\mathring{C}$ (esto es, si se verifica que $\int_{\varphi(I)} f = \int_I f^*$, para todo intervalo compacto $I \subset \mathring{C}$), entonces la referida fórmula de cambio de variable se verifica en el caso general (es decir, tal y como está enunciada en [71]).

Demostración. Según se acaba de decir anteriormente, bastará con mostrar que la susodicha fórmula [1] se verifica cuando en ella, en lugar de poner D y $\varphi^{-1}(D) = C$, se pone $\mathring{D}$ y $\varphi^{-1}(\mathring{D}) = \mathring{C}$. Se probará, pues, esto último. A lo largo de la demostración, para simplificar la notación, la función $(f \circ \varphi)|\det J\varphi|$ se representa por f^*, como ya se indica en el enunciado.

Para empezar, nótese que, como f es continua en $\bar{D}$ y f^* es continua en $\bar{C}$, ambas funciones están acotadas en sus respectivos conjuntos, por lo que admiten una cota común en ellos; sea $c > 0$ una tal cota. De acuerdo con lo dicho en $[66]_6$ y dado $\varepsilon > 0$ (cualquiera), como $\mathring{D} = \varphi(\mathring{C})$ es un conjunto medible, existe un conjunto compacto y medible $K \subset \mathring{D}$ tal que

$$\mu(\mathring{D}) - \mu(K) < \frac{\varepsilon}{2c}$$

Como el conjunto $H = \varphi^{-1}(K)$ es compacto (véase [18], I), como $\mathring{C}$ es abierto y como $H \subset \mathring{C}$, se sabe que (véase $[08]_1$) la distancia d del conjunto H a la frontera de $\mathring{C}$ es estrictamente positiva, $d > 0$.

Sea I un intervalo compacto que incluya a $\mathring{C}$ y consideremos una partición $P = \{I_1, ..., I_n\}$ de I cuyo diámetro sea: 1.º, menor que la distancia d, y 2.º, tal que la suma de las medidas de los intervalos de P que contienen algún punto de $F_r(\mathring{C})$ sea menor que $\varepsilon/(2c)$ (ello es posible, pues $F_r(\mathring{C})$ tiene contenido nulo por ser $\mathring{C}$ medible). Si es H_P el compacto que resulta de unir los intervalos de P que están incluidos en $\mathring{C}$ y como $|P| < d$, resulta que $H \subset H_P \subset \mathring{C}$, por lo que, llamando $K_P = \varphi(H_P)$, se tiene:

$$\mu(\mathring{D}) - \mu(K_P) = \mu(\mathring{D}) - \mu(\varphi(H_P)) \leqslant \mu(\mathring{D}) - \mu(\varphi(H)) = \mu(\mathring{D}) - \mu(K) < \frac{\varepsilon}{2c}$$

Además, de acuerdo con la segunda condición exigida a $|P|$, se verifica también que $\mu(\overset{\circ}{C}) - \mu(H_P) < \varepsilon/(2c)$.

Si es J uno cualquiera de los intervalos que constituyen H_P, como $J \subset C_0$ y de acuerdo con la hipótesis de nuestro aserto, es $\int_{\varphi(J)} f = \int_J f^*$. Sumando todas estas igualdades, obtenidas al variar J, y de acuerdo con la aditividad de la integral, se obtiene que $\int_{K_P} f = \int_{H_P} f^*$.

Acudiendo ahora a la cota $c > 0$, de f en $\bar{D}$ y de f^* en $\bar{C}$, al teorema de la media (véase [68], 3), a las propiedades de los conjuntos medibles (véase $[66]_2$) y a la aditividad de la integral (véase [69]), se puede poner (para ciertos valores $\xi \in \overset{\circ}{D} - K_P$ y $\zeta \in \overset{\circ}{C} - H_P$):

$$\left|\int_{\overset{\circ}{D}} f - \int_{K_P} f\right| = \left|\int_{\overset{\circ}{D} - K_P} f\right| = |f(\xi)\mu(\overset{\circ}{D} - K_P)| \leqslant c(\mu(\overset{\circ}{D}) - \mu(K_P)) < \frac{\varepsilon}{2}$$

$$\left|\int_{\overset{\circ}{C}} f^* - \int_{H_P} f^*\right| = \left|\int_{\overset{\circ}{C} - H_P} f\right| = |f(\zeta)\mu(\overset{\circ}{C} - H_P)| \leqslant c(\mu(\overset{\circ}{C}) - \mu(H_P)) < \frac{\varepsilon}{2}$$

Como, según ya obtuvimos, es $\int_{K_P} f = \int_{H_P} f^*$, de las dos desigualdades precedentes se desprende que:

$$\left|\int_{\overset{\circ}{D}} f - \int_{\overset{\circ}{C}} f^*\right| \leqslant \left|\int_{\overset{\circ}{D}} f - \int_{K_P} f\right| + \left|\int_{\overset{\circ}{C}} f^* - \int_{H_P} f^*\right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} < \varepsilon$$

y como este resultado se verifica para todo $\varepsilon > 0$, resulta que $\int_{\overset{\circ}{D}} f = \int_{\overset{\circ}{C}} f^*$, como había que comprobar.

SEGUNDO ASERTO. La fórmula [1] de cambio de variable se verifica en el caso particular de que la función $\varphi : A \to \mathbb{R}^p$, $\boldsymbol{x} \mapsto \boldsymbol{y} = \varphi(\boldsymbol{x})$, es del siguiente tipo: para cierto $i_0 \in \{1, ..., p\}$ y cierta función real $\varphi_0 : A \to \mathbb{R}$, es:

$$\varphi(\boldsymbol{x}) = (y_1, ..., y_p), \qquad \text{donde} \begin{cases} y_i = x_i, & \text{si } i \neq i_0 \\ y_{i_0} = \varphi_0(\boldsymbol{x}) \end{cases}$$

Demostración. De acuerdo con lo ya obtenido en el anterior «primer aserto», bastará ahora con demostrar que este «segundo aserto» se verifica cuando se sustituyen los conjuntos C y $D = \varphi(C)$ por los I y $\varphi(I)$, donde I es cualquier intervalo compacto incluido en C. Hemos, pues, de comprobar que, para un tal intervalo compacto I, es $\int_{\varphi(I)} f = \int_I f^*$, en el caso particular de la función φ ya citada.

Cambiando, si es preciso, el orden de las coordenadas $y_1, ..., y_p$ de $\boldsymbol{y}$, podemos suponer que es $i_0 = 1$. De acuerdo con ello, nótese que en nuestro caso particular, para el jacobiano $\det J\varphi$ se tiene:

$$|\det J\varphi(\boldsymbol{x})| = \left|\det \begin{bmatrix} \partial\varphi_0/\partial x_1 & \partial\varphi_0/\partial x_2 & \cdots & \partial\varphi_0/\partial x_p \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}(\boldsymbol{x})\right| = \left|\frac{\partial\varphi_0(\boldsymbol{x})}{\partial x_1}\right| = |D_1\varphi_0(\boldsymbol{x})|$$

Acudiendo a la fórmula de integración iterada (véase [60]), como ahora se manejan funciones continuas, por lo que existen todas las integrales que en dicha fórmula aparecen, se obtiene que

$$\left|\int_I f^*\right| = \int_I (f\circ\varphi)|D_1\varphi_0| = \int_{I''}\left(\int_{I'}(f\circ\varphi)|D_1\varphi_0|\right)$$

donde I' e I'' son las proyecciones de I sobre el primer eje coordenado y sobre el subespacio de $\mathbb{R}^p$ que engendran los otros $p-1$ ejes, respectivamente. Para cada $(x_2, ..., x_p)\in I''$ y de acuerdo con la fórmula de cambio de variable para las integrales simples, al aplicar el cambio $x_1 \mapsto y_1 = \varphi_0(x_1, x_2, ..., x_p)$ a la integral sobre I', se obtiene que

$$\int_{I'}(f\circ\varphi)|D_1\varphi_0| = \int_{I'} f(\varphi_1(\boldsymbol{x}), x_2, ..., x_p)\left|\frac{\partial\varphi_0(\boldsymbol{x})}{\partial x_1}\right| dx_1 = \int_{\varphi_0(I')} f(y_1, x_2, ..., x_p)\, dy_1 \quad (*)$$

Llevando este resultado a la anterior igualdad y como $\varphi = (\varphi_0, i, ..., i)$, donde i denota a la aplicación idéntica, resulta claro que

$$\int_I f^* = \int_{I''} dx_2, ..., dx_p\left(\int_{\varphi_0(I')} f(y_1, x_2, ..., x_p)\, dy_1\right) = \int_{\varphi(I)} f(y_1, y_2, ..., y_p)\, d\boldsymbol{y} = \int_{\varphi(I)} f$$

TERCER ASERTO. La fórmula [1], de cambio de variable, se verifica si se sustituye el conjunto C por cualquier intervalo compacto I incluido en $\mathring{C}$.

Demostración. Vamos a proceder por inducción sobre la dimensión p del espacio: como la propiedad se verifica, según ya se sabe, para el caso de integrales simples ($p = 1$), debemos comprobar que, en el supuesto de que ella sea cierta siempre que la dimensión sea menor o igual que $p - 1$, entonces también lo es si la dimensión es p. Veamos que así ocurre, para lo que empezaremos comprobando que la fórmula de cambio de variable es, entonces, válida en todo intervalo compacto de algún cierto entorno U de cada uno de los puntos $\boldsymbol{a}\in I$ y, después, deduciremos de ello que dicha fórmula se verifica en I.

Como se sabe que $\det J\boldsymbol{\varphi}(\boldsymbol{x}) \neq 0$ para $\boldsymbol{x}\in\mathring{C}$, resulta que para todo $\boldsymbol{x}\in\mathring{C}$ es

$$\operatorname{rang}\begin{bmatrix}\partial\varphi_1/\partial x_1 & \cdots & \partial\varphi_1/\partial x_p\\ \cdots & \cdots & \cdots\\ \cdots & \cdots & \cdots\\ \partial\varphi_{p-1}/\partial x_1 & \cdots & \partial\varphi_{p-1}/\partial x_p\end{bmatrix}(\boldsymbol{x}) = p - 1$$

(*) Nótese que (para $x_2, ..., x_p$ fijos) la función de cambio de variable $x_1 \mapsto y_1 = \varphi_0(\boldsymbol{x})$ es monótona, pues $\partial\varphi_0/\partial x_1$ tiene signo constante (por tenerlo $\det J\boldsymbol{\varphi}$, pues es continuo y no se anula), con lo que al prescindir del valor absoluto en $|\partial\varphi_0/\partial x_1|$ sólo se alteraría el signo del integrando y esta alteración del signo se contrarresta con la correspondiente alteración de los límites de integración (extremos del intervalo I_2). Es decir, la fórmula de sustitución se aplica enunciada así: si $x \mapsto y = \varphi(x)$ es de clase $\mathscr{C}^1$ y monótona en un intervalo compacto I y si $y \mapsto F(y)$ es continua en el intervalo $\varphi(I)$, entonces

$$\int_{\varphi(I)} F(y)\, dy = \int_I F(\varphi(x))\varphi'(x)\, dx$$

(si $I = [a, b]$, $\varphi(a) = \alpha$ y $\varphi(b) = \beta$, entonces: para φ creciente es $\varphi(I) = [\alpha, \beta]$ y para φ decreciente $\varphi(I) = [\beta, \alpha]$).

Por tanto, fijado cualquier $\boldsymbol{a} \in I$, para $\boldsymbol{x} = \boldsymbol{a}$ la anterior matriz admite una submatriz de tamaño $(p-1) \times (p-1)$ que es regular; sin pérdida de generalidad, podemos suponer que esta matriz es la que resulta de suprimir la última columna de aquélla. Como $\boldsymbol{\varphi}$ es de clase $\mathscr{C}^1$ en $\mathring{C}$ y $\boldsymbol{a} \in \mathring{C}$, existe, pues, un cierto entorno $U_{\boldsymbol{a}} \subset \mathring{C}$ de $\boldsymbol{a}$ tal que para todo $\boldsymbol{x} \in U_{\boldsymbol{a}}$ es

$$\det \begin{bmatrix} \partial\varphi_1/\partial x_1 & \cdots & \partial\varphi_1/\partial x_{p-1} \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ \partial\varphi_{p-1}/\partial x_1 & \cdots & \partial\varphi_{p-1}/\partial x_{p-1} \end{bmatrix} \neq 0$$

Este resultado permite aplicar el teorema de la función inversa (véase [38]) en torno del punto $\boldsymbol{a}$ a la función $U_{\boldsymbol{a}} \to \mathbb{R}^p$, $\boldsymbol{x} \mapsto \boldsymbol{z} = \boldsymbol{\varphi}_0(\boldsymbol{x})$ definida por

$$\boldsymbol{z} = \boldsymbol{\varphi}_0(\boldsymbol{x}) = \big(\varphi_1(\boldsymbol{x}), \ldots, \varphi_{p-1}(\boldsymbol{x}), x_p\big)$$

(esta función auxiliar $\boldsymbol{\varphi}_0 : U \to \mathbb{R}^p$ nada tiene que ver con la función $\varphi_0 : A \to \mathbb{R}$ que se consideró en el anterior «segundo aserto») la cual admite, pues, función inversa local de clase $\mathscr{C}^1$, $\boldsymbol{z} \mapsto \boldsymbol{x} = \boldsymbol{\psi}_0(\boldsymbol{z})$, definida en un cierto entorno V de $\boldsymbol{\varphi}_0(\boldsymbol{a})$ y cuya imagen $U \subset U_{\boldsymbol{a}}$ es un entorno de $\boldsymbol{a}$, que es del tipo:

$$\boldsymbol{x} = \boldsymbol{\psi}_0(\boldsymbol{z}) = \big(\psi_1(\boldsymbol{z}), \ldots, \psi_{p-1}(\boldsymbol{z}), z_p\big)$$

Si llamamos $\boldsymbol{\varphi}^* : V \to \mathbb{R}^p$, $\boldsymbol{z} \mapsto \boldsymbol{y} = \boldsymbol{\varphi}^*(\boldsymbol{z})$, a la función definida mediante

$$\boldsymbol{y} = \boldsymbol{\varphi}^*(\boldsymbol{z}) = \big(z_1, \ldots, z_{p-1}, \varphi_p(\psi_1(\boldsymbol{z}), \ldots, \psi_{p-1}(\boldsymbol{z}), z_p)\big)$$

que es de clase $\mathscr{C}^1$ en V, resulta evidente que $\boldsymbol{\varphi} = \boldsymbol{\varphi}^* \circ \boldsymbol{\varphi}_0$ en U:

$$\boldsymbol{\varphi}^*(\boldsymbol{\varphi}_0(\boldsymbol{x})) = \boldsymbol{\varphi}(\boldsymbol{x}), \qquad \forall\, \boldsymbol{x} \in U \text{ (entorno de } \boldsymbol{a}\text{)}$$

Consideremos ahora uno cualquiera de los intervalos compactos incluidos en el entorno U de $\boldsymbol{a}$; sea J dicho intervalo. Llamemos J' y J'' a las proyecciones de J sobre el p-ésimo eje coordenado de $\mathbb{R}^p$ y sobre el subespacio de $\mathbb{R}^p$ que engendra los $p-1$ ejes restantes, respectivamente, con lo que $J = J' \times J''$. Para cada $x_p \in J'$, sea

$$K(x_p) = \{(\varphi_1(\boldsymbol{x}), \ldots, \varphi_{p-1}(\boldsymbol{x})) \in \mathbb{R}^{p-1} \,/\, \boldsymbol{x} \in \{x_p\} \times J''\}$$

Llamemos $g : V \to \mathbb{R}$, $\boldsymbol{z} \mapsto g(\boldsymbol{z})$ a la función definida por

$$g(\boldsymbol{z}) = f(\boldsymbol{\varphi}^*(\boldsymbol{z}))|\det J\boldsymbol{\varphi}^*(\boldsymbol{z})|$$

De acuerdo con nuestra hipótesis de inducción, aplicada, para cada $x_p \in J'$, en el intervalo J'' a la función g (en realidad se aplica al intervalo $J'' \times \{x_p\}$) y tomando como función de cambio de variables a $\boldsymbol{\varphi}_0$, es:

$$\int_{K(x_p)} g(\boldsymbol{z})\, dz_1, \ldots, dz_{p-1} = \int_{J''} g(\boldsymbol{\varphi}_0(\boldsymbol{x}))|\det J\boldsymbol{\varphi}_0(\boldsymbol{x})|\, dx_1, \ldots, dx_{p-1}$$

Recurriendo ahora a la fórmula de integración iterada (véase [70]; téngase en cuenta que $K(x_p)$, J' y J'' son medibles y que los integrandos son funciones continuas), se obtiene:

$$\int_{\varphi_0(J)} g(z)\,dz = \int_{J'}\left(\int_{K(z_p)} g(z)\,dz_1, \ldots, dz_{p-1}\right)dz_p =$$

$$= \int_{J'}\left(\int_{J''} g(\varphi_0(x))|\det J\varphi_0(x)|\,dx_1, \ldots, dx_{p-1}\right)dx_p = \int_J g(\varphi_0(x))|\det J\varphi_0(x)|\,dx$$

Por otra parte, según los teoremas de derivación de una función compuesta (véase [29], 2.ª), como sabemos que $\varphi = \varphi^* \circ \varphi_0$, se verifica que:

$$J\varphi(x) = J\varphi^*(\varphi_0(z)) \cdot J\varphi_0(x)$$

por lo que, teniendo en cuenta la expresión con la que definimos a $g(z)$, la última igualdad entre integrales se puede poner en la forma:

$$\int_{\varphi_0(J)} g(z)\,dz = \int_J f(\varphi(x))|\det J\varphi(x)|\,dx = \int_J f^*(x)\,dx$$

Si aplicamos ahora el resultado obtenido en el «segundo aserto», tomando, para lo que allí son C y φ, el conjunto $\varphi_0(J)$ y la función φ^* (nótese que φ^* sólo altera una de las coordenadas de la variable), se obtiene que:

$$\int_{\varphi^*(\varphi_0(J))} f(y)\,dy = \int_{\varphi_0(J)} f(\varphi^*(z))|\det J\varphi^*(z)|\,dz = \int_{\varphi_0(J)} g(z)\,dz$$

De ésta y de la anterior igualdad y como $\varphi^* \circ \varphi_0 = \varphi$, se desprende que

$$\int_{\varphi(J)} f(y)\,dy = \int_J f^*(x)\,dx, \qquad \forall\ J \subset U \ (J \text{ intervalo compacto})$$

Así pues, se ha probado que cada punto $a \in I$ admite un entorno $U = B_a$, al radio del cual le llamaremos $\varepsilon_a > 0$, tal que $\int_{\varphi(J)} f = \int_J f^*$, para todo intervalo compacto $J \subset B_a$. Como la familia que forman las bolas abiertas $B'_a = B(a, \varepsilon_a/2)$, para a recorriendo I, es un recubrimiento abierto de I, el cual es un conjunto compacto, entonces, según el teorema de Heine-Borel (véase [08]), resulta que hay un número finito de dichas bolas que, sólo ellas, cubren I; sean éstas las correspondientes a los puntos $a_1, \ldots, a_n$ de I y llamemos $\varepsilon = \text{mín}\,\{\varepsilon_{a_1}/2, \ldots, \varepsilon_{a_n}/2\}$. Sea $P = \{J_1, \ldots, J_m\}$ una partición de I que tenga diámetro menor que ε. Se verifica que cualquier intervalo $J_i \in P$ está totalmente incluido en alguna de las bolas B_{a_j}, para $j = 1, \ldots, n$; así ocurre, en efecto, ya que: tomando un punto cualquiera $x_i \in J_i$ (por ejemplo, su centro), como $\{B'_{a_1}, \ldots, B'_{a_n}\}$ es un recubrimiento de I, es $x_i \in B'_{a_j}$ para un cierto $j = 1, \ldots, n$, luego para cualquier punto $x'_i \in J_i$ es:

$$\|x'_i - a_j\| \leqslant \|x'_i - x_i\| + \|x_i - a_j\| < |I_i| + \varepsilon_{a_j}/2 \leqslant \varepsilon + \varepsilon_{a_j}/2 \leqslant \varepsilon_{a_j}$$

luego $x'_i \in B_{a_j}$ para todo $x'_i \in J_i$, o sea, $J_i \subset B_{a_j}$.

Por tanto, para cualquiera de los intervalos $J_i \in P$ se verifica que $\int_{\varphi(J_i)} f = \int_{J_i} f^*$ y, por la aditividad de la integral, se concluye que

$$\int_{\varphi(I)} f = \sum_{i=1}^{n} \int_{\varphi(J_i)} f = \sum_{i=1}^{n} \int_{J_i} f^* = \int_I f^*$$

con lo que concluye la demostración del «tercer aserto».

CUARTO ASERTO. La fórmula [1], del cambio de variable, se verifica en los términos en los que figura en el enunciado.

Demostración. Según el «primer aserto», para que se verifique la fórmula [1] del enunciado es suficiente con que se verifique lo que se afirma en el «tercer aserto», que ha resultado ser cierto.

[71]$_1$ Ejercicio

Calcular la integral $I = \iint_D f$, donde D y f son (para un cierto $a > 0$):

$$D = \{(x, y) \in \mathbb{R}^2 \;/\; x^{2/3} + y^{2/3} \leqslant a^{2/3}\}$$
$$f(x, y) = (a^{2/3} - x^{2/3} - y^{2/3})^2$$

Indicación: realizar el cambio de variable $\varphi : (u, v) \mapsto (x, y)$ dado por

$$x = u \cos^3 v, \qquad y = u \operatorname{sen}^3 v$$

Resolución

Según se comprueba fácilmente, se cumplen las condiciones del anterior teorema [71]; en particular, el jacobiano de φ es:

$$J\varphi(u, v) = \begin{vmatrix} \cos^3 v & -3u \cos^2 v \operatorname{sen} v \\ \operatorname{sen}^3 v & 3u \operatorname{sen}^2 v \cos v \end{vmatrix} = 3u \operatorname{sen}^2 v \cos^2 v$$

La fórmula de cambio de variable permite poner

$$I = \iint_C (a^{2/3} - u^{2/3})^2 3u \operatorname{sen}^2 v \cos^2 v \, du \, dv$$
$$\text{donde } C = \varphi^{-1}(D) = \{(u, v) \in \mathbb{R}^2 \;/\; 0 \leqslant u \leqslant a,\ 0 \leqslant v \leqslant 2\pi\}$$

Integrando por iteración, se obtiene que:

$$I = 3 \int_0^a (a^{2/3} - u^{2/3})^2 u \, du \left(4 \int_0^{\pi/2} \operatorname{sen}^2 v \cos^2 v \, dv \right) = \frac{3\pi}{4} \int_0^a u(a^{2/3} - u^{2/3})^2 \, du = \frac{3\pi a^{10/3}}{80}$$

$[71]_2$ Cambio a coordenadas polares (en el plano)

Considérese la aplicación $\varphi:\mathbb{R}^2 \to \mathbb{R}^2$, $(\rho, \theta) \mapsto (x, y)$, definida mediante

$$x = \rho \cos\theta, \qquad y = \rho \operatorname{sen}\theta$$

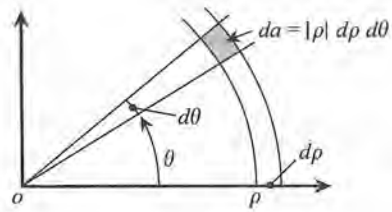
que es de clase $\mathscr{C}^\infty$ en $\mathbb{R}^2$ y cuyo jacobiano es $\det J\varphi(\rho, \theta) = \rho$. Sean C y D dos conjuntos medibles de $\mathbb{R}^2$ tales que $\varphi: C \to D$ es una biyección. Sea $f: D \to \mathbb{R}$ una función integrable en D. Si se cumplen las condiciones de regularidad exigidas en el teorema de cambio de variable [71], entonces

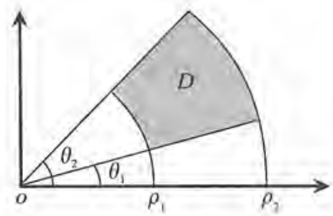
$$\iint_D f(x, y)\, dx\, dy = \iint_C f(\rho\cos\theta, \rho \operatorname{sen}\theta)|\rho|\, d\rho\, d\theta$$

Suele ser frecuente la integración en conjuntos como los siguientes, en donde θ_1, θ_2, ρ_1 y ρ_2 son unos ciertos valores fijos:

$$C = \{(\rho, \theta) \in \mathbb{R}^2 \,/\, \rho_1 \leqslant \rho \leqslant \rho_2,\ \theta_1 \leqslant \theta \leqslant \theta_2\}$$
$$D = \varphi(C) = \{(x, y) \in \mathbb{R}^2 \,/\, \rho_1^2 \leqslant x^2 + y^2 \leqslant \rho_2^2,\ \theta_1 \leqslant \operatorname{arctg}(x/y) \leqslant \theta_2\}$$

(D es, en ejes (x, y), el segmento de corona circular comprendido entre las circunferencias de radio ρ_1 y ρ_2 y las semirrectas de inclinación θ_1 y θ_2; el conjunto C, en ejes cartesianos ρ, θ, es el rectángulo comprendido entre las rectas (paralelas a los ejes) $\rho = \rho_1$, $\rho = \rho_2$, $\theta = \theta_1$ y $\theta = \theta_2$).





Ejemplo

Para calcular la integral doble

$$I = \iint_D \frac{dx\, dy}{(1 + x^2 + y^2)^{3/2}}, \qquad \text{donde } D = \{(x, y) \in \mathbb{R}^2 \,/\, 0 \leqslant x \leqslant 1,\ 0 \leqslant y \leqslant x\}$$

cambiando a coordenadas polares ($x = \rho\cos\theta$, $y = \rho \operatorname{sen}\theta$), se obtiene:

$$I = \iint_C \frac{\rho}{(1 + \rho^2)^{3/2}}\, d\rho\, d\theta, \qquad \text{donde } C = \{(\rho, \theta) \in \mathbb{R}^2 \,/\, 0 \leqslant \theta \leqslant \pi/4,\ 0 \leqslant \rho \leqslant \sec\theta\}$$

integrando por iteración, para I se obtiene:

$$I = \int_0^{\pi/4} d\theta \int_0^{\sec\theta} \frac{\rho}{(1+\rho^2)^{3/2}}\, d\rho = \int_0^{\pi/4} [(1+\rho^2)^{-1/2}]_{\rho=0}^{\rho=\sec\theta}\, d\theta =$$

$$= \int_0^{\pi/4} \left(1 - \frac{\cos\theta}{\sqrt{2-\operatorname{sen}^2\theta}}\right) d\theta = \left[\theta - \operatorname{arcsen}\left(\frac{\sqrt{2}}{2}\operatorname{sen}\theta\right)\right]_0^{\pi/4} = \frac{\pi}{4} - \frac{\pi}{6} = \frac{\pi}{12}$$

$[71]_3$ Cambio a coordenadas cilíndricas (en $\mathbb{R}^3$)

Considérese la aplicación $\varphi: \mathbb{R}^3 \to \mathbb{R}^3$, $(\rho, \theta, \zeta) \mapsto (x, y, z)$, definida por:

$$x = \rho\cos\theta, \qquad y = \rho\operatorname{sen}\theta, \qquad z = \zeta$$

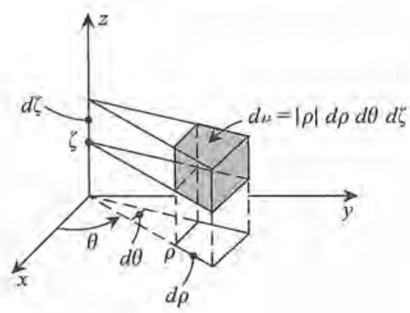
que es de clase $\mathscr{C}^\infty$ en $\mathbb{R}^3$ y cuyo jacobiano es det $J\varphi(\rho, \theta, \zeta) = \rho$. Sean C y D dos conjuntos medibles de $\mathbb{R}^3$ tales que $\varphi: C \to D$ es una biyección. Sea $f: D \to \mathbb{R}$ una función integrable en D. Si se cumplen las condiciones de regularidad exigidas en el teorema de cambio de variable [71], entonces

$$\iiint_D f(x, y, z)\, dx\, dy\, dz = \iiint_C f(\rho\cos\theta, \rho\operatorname{sen}\theta, \zeta)|\rho|\, d\rho\, d\theta\, d\zeta$$

Suele ser frecuente la integración en conjuntos como los siguientes, en donde $\rho_1, \rho_2, \theta_1, \theta_2$, ζ_1 y ζ_2 son unos ciertos valores fijos:

$$C = \{(\rho, \theta, \zeta) \in \mathbb{R}^3 \;/\; \rho_1 \leqslant \rho \leqslant \rho_2, \quad \theta_1 \leqslant \theta \leqslant \theta_2, \quad \zeta_1 \leqslant \zeta \leqslant \zeta_2\}$$
$$D = \varphi(C) = \{(x, y, z) \in \mathbb{R}^3 \;/\; \rho_1^2 \leqslant x^2 + y^2 \leqslant \rho_2^2, \quad \theta_1 \leqslant \operatorname{arctg}(x/y) \leqslant \theta_2, \quad \zeta_1 \leqslant z \leqslant \zeta_2\}$$

(D es, en ejes (x, y, z), el conjunto situado entre los cilindros de radios ρ_1 y ρ_2 y eje $x = y = 0$, entre los semiplanos de longitudes (o acimutes) θ_1 y θ_2 y entre los planos de cotas ζ_1 y ζ_2; el conjunto C, en ejes cartesianos $(\rho\theta\zeta)$, es el paralelepípedo comprendido entre los planos (paralelos a los de coordenadas) $\rho = \rho_1$, $\rho = \rho_2$, $\theta = \theta_1$, $\theta = \theta_2$, $\zeta = \zeta_1$ y $\zeta = \zeta_2$).



Ejemplo

Para calcular la integral triple

$$I = \iiint_D \frac{yz}{x}\, dx\, dy\, dz,$$

$$\text{donde } D = \{(x, y, z) \in \mathbb{R}^3 \;/\; y > 0,\ z > 0,\ x^2 + y^2 < ax,\ x^2 + y^2 + z^2 < a^2\}$$

cambiando a coordenadas cilíndricas ($x = \rho \cos\theta$, $y = \rho \operatorname{sen}\theta$, $z = \zeta$), se obtiene que:

$$I = \iiint_C \frac{\operatorname{sen}\theta}{\cos\theta}\, \zeta\rho\, d\rho\, d\theta\, d\zeta,$$

$$\text{donde } C = \{(\rho, \theta, \zeta) \in \mathbb{R}^3 \;/\; 0 < \theta < \pi/2,\ 0 < \rho < a\cos\theta,\ 0 < \zeta < \sqrt{a^2 - \rho^2}\}$$

integrando por iteración, para I se obtiene:

$$I = \int_0^{\pi/2} \frac{\operatorname{sen}\theta}{\cos\theta}\, d\theta \int_0^{a\cos\theta} \rho\, d\rho \int_0^{\sqrt{a - \rho^2}} \zeta\, d\zeta = \int_0^{\pi/2} \frac{\operatorname{sen}\theta}{\cos\theta}\, d\theta \int_0^{a\cos\theta} \frac{1}{2}\, \rho(a^2 - \rho^2)\, d\rho =$$
$$= \frac{1}{8}\, a^4 \int_0^{\pi/2} (2\cos\theta \operatorname{sen}\theta - \cos^3\theta \operatorname{sen}\theta) = \frac{3a^4}{32}$$

$[71]_4$ Cambio a coordenadas esféricas (en $\mathbb{R}^3$)

Considérese la aplicación $\boldsymbol{\varphi} : \mathbb{R}^3 \to \mathbb{R}^3$, $(x, y, z) \mapsto (\rho, \theta, \varphi)$, definida mediante

$$x = \rho \cos\varphi \cos\theta, \qquad y = \rho \cos\varphi \operatorname{sen}\theta, \qquad z = \rho \operatorname{sen}\varphi \quad (*)$$

que es de clase $\mathscr{C}^\infty$ en $\mathbb{R}^3$ y cuyo jacobiano es $\det J\boldsymbol{\varphi}(\rho, \theta, \varphi) = \rho^2 \cos\varphi$. Sean C y D dos conjuntos medibles de $\mathbb{R}^3$ tales que $\boldsymbol{\varphi} : C \to D$ es una biyección. Sea $f : D \to \mathbb{R}$ una función integrable en D. Si se cumplen las condiciones de regularidad exigidas en el torema de cambio de variable [71], entonces:

$$\iiint_D f(x, y, z)\, dx\, dy\, dz = \iiint_C f(\rho\cos\varphi\cos\theta,\ \rho\cos\varphi\operatorname{sen}\theta,\ \rho\operatorname{sen}\varphi)\rho^2|\cos\varphi|\, d\rho\, d\theta\, d\varphi$$

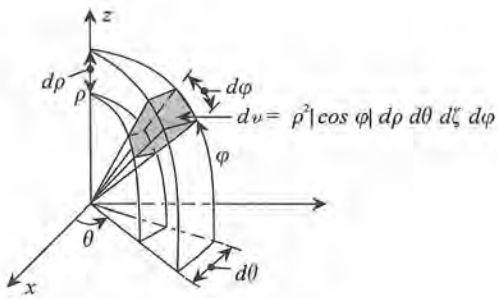
(*) En esto de las coordenadas esféricas hay quienes prefieren utilizar, en lugar del ángulo φ (latitud), el ángulo $\psi = \pi/2 - \varphi$ (colatitud); cuando así se hace, las fórmulas de cambio son:

$$x = \rho \operatorname{sen}\psi \cos\theta, \qquad y = \rho \operatorname{sen}\psi \operatorname{sen}\theta, \qquad z = \rho\cos\psi$$

Suele ser frecuente la integración en conjuntos como los siguientes, en donde ρ_1, ρ_2, θ_1, θ_2, φ_1 y φ_2 son unos ciertos valores fijos:

$$C = \{(\rho, \theta, \varphi) \in \mathbb{R}^2 \ / \ \rho_1 \leqslant \rho \leqslant \rho_2, \ \theta_1 \leqslant \theta \leqslant \theta_2, \ \varphi_1 \leqslant \varphi \leqslant \varphi_2\}$$
$$D = \varphi(D) = \{(x, y, z) \in \mathbb{R}^2 \ / \ \rho_1^2 \leqslant x^2 + y^2 + z^2 \leqslant \rho_2^2, \ \theta_1 \leqslant \operatorname{arctg}(y/x) \leqslant \theta_2,$$
$$\varphi_1 \leqslant \operatorname{arctg} z/\sqrt{x^2 + y^2} \leqslant \varphi_2\}$$

(D es, en ejes (xyz), el conjunto situado entre las esferas de radio ρ_1 y ρ_2 y centro en el origen, entre los semiplanos de longitudes (acimutes) θ_1 y θ_2 y entre los conos de centro el origen, eje vertical y semiángulos en el vértice $\pi/2 - \varphi_1$ y $\pi/2 - \varphi_2$ (colatitudes); el conjunto C, en ejes rectangulares ($\rho\theta\varphi$), es el paralelepípedo comprendido entre los planos (paralelos a los de coordenadas) $\rho = \rho_1$, $\rho = \rho_2$, $\theta = \theta_1$, $\theta = \theta_2$, $\varphi = \varphi_1$ y $\varphi = \varphi_2$.



Ejemplo

Para calcular el volumen υ del sólido S de $\mathbb{R}^3$ encerrado en el interior de la superficie de ecuación (en ejes cartesianos rectangulares xyz):

$$(x^2 + y^2 + z^2)^2 = 2z(x^2 + y^2)$$

cambiando a coordenadas esféricas, en las que la ecuación de la superficie es $\rho =$ $= 2 \operatorname{sen} \varphi \cos^2 \varphi$ y, por tanto, el sólido S es:

$$S = \{(\rho, \theta, \varphi) \in \mathbb{R}^2 \ / \ 0 \leqslant \theta \leqslant 2\pi, \ 0 \leqslant \varphi \leqslant \pi/2, \ 0 \leqslant \rho \leqslant 2 \operatorname{sen} \varphi \cos^2 \varphi\}$$

el volumen pedido se puede poner (recuérdese que el jacobiano de la función de cambio es $\rho^2 \cos \varphi$)

$$\upsilon = \iiint_S \rho^2 \cos\varphi \, d\rho \, d\theta \, d\varphi = \int_0^{2\pi} d\theta \int_0^{\pi/2} \cos\varphi \, d\varphi \int_0^{2 \operatorname{sen}\varphi \cos^2\varphi} \rho^2 \, d\rho =$$
$$= \int_0^{2\pi} d\theta \int_0^{\pi/2} \frac{8}{3} \operatorname{sen}^3 \varphi \cos^7 \varphi \, d\varphi = \int_0^{2\pi} d\theta \left[\frac{8}{3} \left(-\frac{1}{8} \cos^8 \varphi + \frac{1}{10} \cos^{10} \varphi \right) \right]_{\varphi = 0}^{\varphi = \pi/2} = \int_0^{2\pi} \frac{d\theta}{15} = \frac{2\pi}{15}$$

$[71]_5$ Observaciones (otros teoremas de cambio de variable)

El anterior teorema [71], de cambio de variable en integrales múltiples, admite multitud de variaciones y refinamientos. Aun cuando no entraremos a demostrar ni los unos ni los otros, sí vamos a enunciar algunos teoremas, que vienen a ser algo así como un resumen de todo ello, en los que se alivia la condición de no anulación del jacobiano:

UN SEGUNDO TEOREMA DE CAMBIO DE VARIABLE (en el que se permite que el jacobiano se anule en un conjunto de contenido nulo):

> Sean $A \subset \mathbb{R}^p$ un conjunto abierto y sea $\varphi : A \to \mathbb{R}^p$ una transformación inyectiva en A, de clase $\mathscr{C}^1$ en A y tal que $\{x \in A \ / \ \det J\varphi(x) = 0\}$ es un conjunto de contenido nulo. Se verifica entonces que:
>
> 1.° Si $C \subset \mathbb{R}^p$ es un conjunto medible y tal que $\bar{C} \subset A$, entonces el conjunto $D = \varphi(C)$ es medible y la medida de D es $\mu(D) = \int_C |\det J\varphi|$.
>
> 2.° Si, además, $f : D \to \mathbb{R}$ es una función integrable en D, entonces la función $(f \circ \varphi)$ es integrable en C y es $\int_D f = \int_C (f \circ \varphi)|\det J\varphi|$.

A este respecto, bueno es recordar el siguiente teorema de Sard: si $A \subset \mathbb{R}^p$ es un conjunto abierto y si $\varphi : A \to \mathbb{R}^p$ es una transformación de clase $\mathscr{C}^1$ en A, entonces el conjunto $\{x \in A \ / \ \det J\varphi(x) = 0\}$ tiene medida nula.

UN TERCER TEOREMA DE CAMBIO DE VARIABLE (en el que la no anulación del jacobiano sólo se precisa en el interior del conjunto de integración):

> Sea $A \subset \mathbb{R}^p$ un conjunto abierto y sea $\varphi : A \to \mathbb{R}^p$ una transformación de clase $\mathscr{C}^1$ en A. Se verifica entonces que:
>
> 1.° Si $C \subset \mathbb{R}^p$ es un conjunto medible tal que $\bar{C} \subset A$, si la restricción de φ a $\mathring{C}$ es inyectiva y si $\det J\varphi(x) \neq 0$ para todo $x \in \mathring{C}$, entonces el conjunto $D = \varphi(C)$ es medible y la medida de D es $\mu(D) = \int_C |\det J\varphi|$.
>
> 2.° Si, además, $f : D \to \mathbb{R}$ es una función integrable en D, entonces la función $(f \circ \varphi)$ es integrable en C y es $\int_D f = \int_C (f \circ \varphi)|\det J\varphi|$.

Una variante de este teorema se obtiene al sustituir la exigencia de jacobiano no nulo en $\mathring{C}$ por un requisito similar: pedir que φ sea $\mathscr{C}^1$-inversible(*) en $\mathring{C}$. Concretamente:

> Sea $A \subset \mathbb{R}^p$ un conjunto abierto y sea $\varphi : A \to \mathbb{R}^p$ una transformación de clase $\mathscr{C}^1$ en A. Se verifica entonces que:
>
> 1.° Si $C \subset \mathbb{R}^p$ es un conjunto medible tal que $\bar{C} \subset A$ y si la restricción de φ a $\mathring{C}$ es $\mathscr{C}^1$-invertible(*), en $\mathring{C}$, entonces el conjunto $D = \varphi(C)$ es medible y la medida de D es $\mu(D) = \int_C |\det J\varphi|$.

(*) Si $U \subset \mathbb{R}^p$ es un conjunto abierto, se dice que una transformación $\varphi : U \to \mathbb{R}^p$ es $\mathscr{C}^1$-invertible en U si $V = \varphi(U)$ es abierto, si $\varphi : U \to V$ es una biyección, si $\varphi : U \to \mathbb{R}^p$ es de clase $\mathscr{C}^1$ en U y si $\varphi^{-1} : V \to \mathbb{R}^p$ es de clase $\mathscr{C}^1$ en V.

2.° Si, además, $f: D \to \mathbb{R}$ es una función integrable en D, entonces la función $(f \circ \varphi)$ es integrable en C y es $\int_D f = \int_C (f \circ \varphi)|\det J\varphi|$.

El interés de este «tercer teorema», en sus dos versiones, radica en que la no anulación del jacobiano o, lo que viene a ser equivalente, la condición de ser $\mathscr{C}^1$-invertible, sólo se le exige a la función de cambio, φ, en el interior del conjunto de integración, lo cual es una gran ventaja en el momento de las aplicaciones, en las que no es infrecuente que el referido jacobiano se anule en la misma frontera del referido conjunto de integración.

[71]$_6$ Interpretación geométrica del jacobiano

Destaquemos, del teorema del cambio de variable, aquello que se refiere a la medida del conjunto imagen, esto es, que (con la formulación del que hemos llamado «un tercer teorema del cambio de variable» en [71]$_5$): dada una transformación $\varphi: A \to \mathbb{R}^p$ de clase $\mathscr{C}^1$, en un abierto $A \subset \mathbb{R}^p$, y dado un conjunto medible $C \subset \mathbb{R}^p$, con $\bar{C} \subset A$, si la restricción de φ a $\mathring{C}$ es inyectiva y tal que $\det J\varphi(x) \neq 0$ para todo $x \in \mathring{C}$, entonces $D = \varphi(C)$ es también un conjunto medible y su medida es igual a:

$$\mu(D) = \int_{\varphi^{-1}(D)} |J\varphi(x)|\, dx$$

Esta fórmula expresada en la función de $C = \varphi^{-1}(D)$, en lugar de hacerlo en términos de $D = \varphi(C)$, no es otra que la:

$$\mu(\varphi(C)) = \int_C |J\varphi(x)|\, dx$$

Si a esta última integral le aplicamos el teorema de la media (no se olvide que el integrando es una función continua; véase [68]), podemos afirmar que existe un cierto punto $\xi \in C$ tal que:

$$\mu(\varphi(C)) = \mu(C)|\det J\varphi(\xi)|$$

Así pues, el jacobiano de φ (mejor dicho, el valor del jacobiano de φ en un cierto punto intermedio de C) es el «factor de dilatación» o coeficiente local de proporcionalidad que permite pasar de la medida de un conjunto C a la medida de su conjunto imagen $\varphi(C)$.

❑ ALGO ACERCA DEL ÁREA DE UNA SUPERFICIE

Aquí vamos a limitarnos a hacer algunas consideraciones someras, faltas de solidez, para justificar las fórmulas que permiten calcular el área de una superficie de $\mathbb{R}^3$. Cuanto ahora decimos son, pues, solamente unas notas de carácter práctico.

[72]

Consideremos la superficie que forman los puntos $(x, y, z) \in \mathbb{R}^3$ imagen de una función vectorial $\boldsymbol{S}: A \to \mathbb{R}$, $(u, v) \mapsto (x, y, z) = \boldsymbol{S}(u, v)$, donde $\boldsymbol{S}$ es de clase $\mathscr{C}^1$ en un conjunto abierto $A \subset \mathbb{R}^2$. Si $C \subset A$ es un conjunto medible, el área de la porción de superficie $\boldsymbol{S}(C)$ viene dado por ($\wedge$ denota al producto vectorial):

$$\mathscr{a}(\boldsymbol{S}(C)) = \iint_C \left\| \frac{\partial \boldsymbol{S}(u, v)}{\partial u} \wedge \frac{\partial \boldsymbol{S}(u, v)}{\partial v} \right\| du\, dv$$

---

1. Acudiendo a los siguientes productos escalares E, F y G (que se llaman coeficientes de la primera forma fundamental), se puede poner

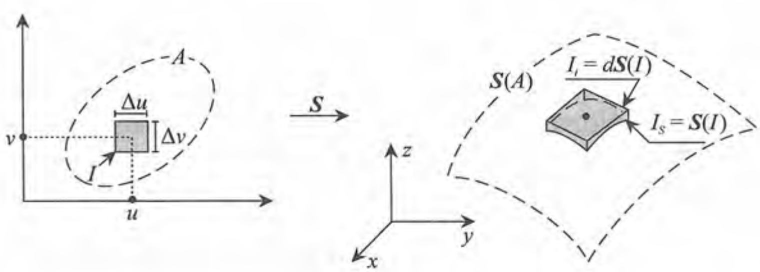
$$\mathscr{a}(\boldsymbol{S}(C)) = \iint_C \sqrt{EG - F^2} \qquad \text{donde } E = \frac{\partial \boldsymbol{S}}{\partial u} \cdot \frac{\partial \boldsymbol{S}}{\partial u}, \quad F = \frac{\partial \boldsymbol{S}}{\partial u} \cdot \frac{\partial \boldsymbol{S}}{\partial v} \quad \text{y} \quad G = \frac{\partial \boldsymbol{S}}{\partial v} \cdot \frac{\partial \boldsymbol{S}}{\partial v}$$

2. Si la porción de superficie es la de ecuación explícita $z = f(x, y)$, para $(x, y) \in C$, su área es:

$$\mathscr{a} = \iint_C \sqrt{1 + f_x'(x, y)^2 + f_y'(x, y)^2}\, dx\, dy$$

3. Cuando se considera la superficie de revolución engendrada, al girar alrededor del «eje de la x», por la curva de clase $\mathscr{C}^1$ (del plano xy) de ecuación $y = f(x) \geqslant 0$ para $a \leqslant x \leqslant b$, el área es ahora

$$\mathscr{a} = 2\pi \int_a^b f(x)\sqrt{1 + f'(x)^2}\, dx$$



Justificación

Consideremos un «pequeño» intervalo I, del plano auxiliar uv, que esté incluido en C; sea $(u, v) \in C$ el centro de dicho intervalo y sean Δu e Δv las longitudes de sus lados. En correspondencia con este intervalo I, disponemos de la «pequeña» porción de superficie $I_s = \boldsymbol{S}(I)$ y del correspondiente paralelogramo (también «pequeño») I_t en el plano tangente, que no es otro que el $I_t = d\boldsymbol{S}(u, v)(I)$, donde $d\boldsymbol{S}(u, v)$ denota a la diferencial de $\boldsymbol{S}$ en el punto (u, v); nótese que I_t es el paralelogramo que tiene su centro en el punto $\boldsymbol{S}(u, v)$ y cuyos lados son paralelos a los vectores

$$d\boldsymbol{S}(u, v)(\Delta u, 0) = \frac{\partial \boldsymbol{S}(u, v)}{\partial u}\, \Delta u \qquad \text{y} \qquad d\boldsymbol{S}(u, v)(0, \Delta v) = \frac{\partial \boldsymbol{S}(u, v)}{\partial v}\, \Delta v$$

y que, por consiguiente, el área de este paralelogramo I_t es la norma del producto vectorial de estos dos vectores, es decir,

$$\alpha(I_t) = \left\| \frac{\partial S(u, v)}{\partial u} \wedge \frac{\partial S(u, v)}{\partial v} \right\| \Delta u \, \Delta v \qquad [1]$$

Es razonable suponer que si P es una partición, de un intervalo compacto de $\mathbb{R}^2$ que incluya a C, el área de la porción de superficie $S(C)$ sea el límite de las sumas $\Sigma\alpha(I_s)$, para $I \in P$ e $I \subset C$, cuando $|P| \to 0$. Si se sustituye cada pequeña porción de superficie I_s por el correspondiente paralelogramo I_t (que es una especie de «pequeña escama» plana tangente a I_s en su punto medio), cabe esperar que $\Sigma\alpha(I_s)$ y $\Sigma\alpha(I_t)$ tengan el mismo límite cuando $|P| \to 0$. Ahora bien, el límite de esta última suma es, a la vista de la expresión [1], la integral del enunciado, por lo que se conviene en acordar que el valor de la referida integral sea el área de la porción de superficie $S(C)$.

- Para comprobar la primera de las notas que figuran al pie del enunciado, si llamamos $\boldsymbol{a} = \partial S/\partial u$ y $\boldsymbol{b} = \partial S/\partial v$, bastará con que veamos que $\|\boldsymbol{a} \wedge \boldsymbol{b}\|^2$ y $\|\boldsymbol{a}\|^2\|\boldsymbol{b}\|^2 - (\boldsymbol{a} \cdot \boldsymbol{b})^2$ son iguales. Esto se comprueba fácilmente acudiendo al ángulo $\theta = \text{áng}\,(\boldsymbol{a}, \boldsymbol{b})$, ya que

$$\|\boldsymbol{a} \wedge \boldsymbol{b}\|^2 = \|\boldsymbol{a}\|^2\|\boldsymbol{b}\|^2 \operatorname{sen}^2 \theta = \|\boldsymbol{a}\|^2\|\boldsymbol{b}\|^2(1 - \cos^2 \theta) =$$
$$= \|\boldsymbol{a}\|^2\|\boldsymbol{b}\|^2\left(1 - \frac{(\boldsymbol{a} \cdot \boldsymbol{b})^2}{\|\boldsymbol{a}\|^2\|\boldsymbol{b}\|^2}\right) = \|a\|^2\|b\|^2 - (\boldsymbol{a} \cdot \boldsymbol{b})^2$$

- Cuando la superficie S se conoce a través de su ecuación explicita $z = f(x, y)$ (según se indica en la nota 2 del enunciado), como entonces $x = u$, $y = v$, $z = f(u, v)$ son ecuaciones paramétricas de la superficie, ahora será $\partial S/\partial u = (1, 0, f'_x)$ y $\partial S/\partial v = (0, 1, f'_y)$, con lo que

$$E = 1 + (f'_x)^2, \quad F = f'_x f'_y, \quad G = 1 + (f'_y)^2, \quad \sqrt{EG - F^2} = \sqrt{1 + (f'_x)^2 + (f'_y)^2}$$

por lo que es válida la fórmula que se da en dicha nota 2.
- El caso de la superficie de revolución se comprueba fácilmente, ya que ella admite por ecuaciones paramétricas a las

$$(x, \theta) \mapsto S(x, \theta) = (x, y, z) \begin{cases} x = x \\ y = \varphi(x) \cos \theta \\ z = \varphi(x) \operatorname{sen} \theta \end{cases} \qquad \begin{pmatrix} a \leqslant x \leqslant b \\ 0 \leqslant \theta \leqslant 2\pi \end{pmatrix}$$

con lo que $\partial S/\partial x = (1, \varphi'(x) \cos \theta, \varphi'(x) \operatorname{sen} \theta)$ y $\partial S/\partial(\theta) = (0, -\varphi(x) \operatorname{sen} \theta, \varphi(x) \cos \theta)$, luego

$$E = 1 + \varphi'(x)^2, \quad F = 0, \quad G = \varphi(x)^2, \quad \sqrt{EG - F^2} = \varphi(x)\sqrt{1 + \varphi'(x)^2}$$

de donde se desprende trivialmente la fórmula del enunciado (nota 3).

[72]$_1$ Ejercicio

Hallar el área de la porción de la superficie de la Tierra (de radio r) comprendida entre los meridianos de longitudes θ_1 y θ_2 y los paralelos de latitudes φ_1 y φ_2. Hacer una estimación de la superficie de la península Ibérica, suponiendo que ésta ocupa la porción de superficie terrestre comprendida entre los meridianos $\theta_1 = 0°\ O$ y $\theta_2 = 9°\ O$ y los paralelos $\varphi_1 = 37°\ N$ y $\varphi_2 = 43°\ N$ (tómese $r = 6.380$ km).

Resolución

La porción de superficie terrestre de la que hay que obtener el área admite las siguientes ecuaciones paramétricas:

$$(\theta, \varphi) \mapsto S(\theta, \varphi) = (x, y, z) \begin{cases} x = r\cos\varphi\cos\theta \\ y = r\cos\varphi\,\mathrm{sen}\,\theta \\ z = r\,\mathrm{sen}\,\varphi \end{cases} \qquad \begin{pmatrix} \theta_1 \leqslant \theta \leqslant \theta_2 \\ \varphi_1 \leqslant \varphi \leqslant \varphi_2 \end{pmatrix}$$

Por tanto, $\partial S/\partial\theta$, $\partial S/\partial\varphi$, E, F y G valen ahora:

$$\frac{\partial S}{\partial \theta} = (-r\cos\varphi\,\mathrm{sen}\,\theta,\ r\cos\varphi\cos\theta,\ 0) \qquad E = r^2\cos^2\varphi$$
$$F = 0$$
$$\frac{\partial S}{\partial \varphi} = (-r\,\mathrm{sen}\,\varphi\cos\theta,\ -r\,\mathrm{sen}\,\varphi\,\mathrm{sen}\,\theta,\ r\cos\ \varphi) \qquad G = r^2$$

y, por ello, el área pedida será

$$\alpha = \int_{\theta_1}^{\theta_2} d\theta \int_{\varphi_1}^{\varphi_2} \sqrt{r^4\cos^2\varphi}\, d\varphi = \int_{\theta_1}^{\theta_2} r^2(\mathrm{sen}\,\varphi_2 - \mathrm{sen}\,\varphi_1)\, d\theta = r^2(\theta_2 - \theta_1)(\mathrm{sen}\,\varphi_2 - \mathrm{sen}\,\varphi_1)$$

La aproximación para el área de la península Ibérica será, pues,

$$\alpha = 6.380^2\,\frac{\pi(9-0)}{180}\,(\mathrm{sen}\ 43° - \mathrm{sen}\ 37°) = 512.679\ \mathrm{km}^2$$

(es una mala aproximación, con error de un 12%, pues también lo son θ_1, θ_2, φ_1 y φ_2).

Ejercicios y problemas

ENUNCIADOS

4.1. Sean C y D dos conjuntos medibles de $\mathbb{R}^p$. Pruébese que también es medible el conjunto

$$C - D = \{x \in \mathbb{R}^p \;/\; x \in C, x \notin D\}$$

4.2. Sea $C \subset \mathbb{R}^p$ un conjunto acotado y tal que el conjunto C^a de sus puntos de acumulación es finito. Pruébese que C tiene contenido nulo.

4.3. Sea $C \subset \mathbb{R}^p$ un conjunto acotado y tal que el conjunto C^a de sus puntos de acumulación tiene contenido nulo. Pruébese que C tiene contenido nulo.

4.4. Sea $I \subset \mathbb{R}^p$ un intervalo compacto y sean f y g dos funciones acotadas de I en $\mathbb{R}$. Pruébese que

$$\underline{\int_I} (f+g) \geqslant \underline{\int_I} f + \underline{\int_I} g \quad \text{e} \quad \overline{\int_I} (f+g) \leqslant \overline{\int_I} f + \overline{\int_I} g$$

4.5. Sea $I \subset \mathbb{R}^p$ un intervalo compacto. Una función acotada $\varphi: I \to \mathbb{R}$ se dice «escalonada» si existe una partición $P = \{I_1, ..., I_n\}$ de I tal que φ es constante en $\overset{\circ}{I_i}$, para todo $I_i \in P$. Pruébese que:

1.º Si φ es escalonada en I, entonces es integrable en I y se verifica que:

$$\int_I \varphi = c_1\mu(I_1) + \cdots + C_n\mu(I_n)$$

(donde c_i es el valor, constante, que toma φ en todo punto de $\overset{\circ}{I_i}$).

2.º Para cualquiera que sea la función acotada $f: I \to \mathbb{R}$, es

$$\underline{\int_I} f = \sup\left\{\int_I \varphi \;/\; \varphi \leqslant f, \varphi \text{ escalonada en } I\right\}$$

$$\overline{\int_I} f = \inf\left\{\int_I \varphi \;/\; \varphi \geqslant f, \varphi \text{ escalonada en } I\right\}$$

4.6. Sea $f: I \to \mathbb{R}$ la función definida, en $I = [0, 1] \times \times [0, 2]$, mediante: $f(x, y) = 0$, si $x \in \mathbb{Q}$; $f(x, y) = = y - 1$, si $x \notin \mathbb{Q}$. Hallar $\varphi(y) = \underline{\int_0^2} f(x, y)\,dx$ y $\psi(y) = \overline{\int_0^2} f(x, y)\,dx$. Comprobar que f no es integrable en I. Comprobar que, no obstante, existe la integral iterada $\int_0^1 dx \int_0^2 f(x, y)\,dy$.

4.7. Sean $f: [a, b] \to \mathbb{R}$ y $g: [a, b] \to \mathbb{R}$ dos funciones integrables en sus respectivos intervalos; sea ϕ la función real definida en $I = [a, b] \times [c, d]$ por $\phi(x, y) = f(x)g(y)$. Comprobar que ϕ es integrable en I y que $\int_I \phi = (\int_a^b f)(\int_c^d g)$.

4.8. Sea $I \subset \mathbb{R}^p$ un intervalo compacto; sea $f: I \to \mathbb{R}$ una función acotada que es integrable en I. Si $\int_I f = 0$, si $f(x) \geqslant 0$, para todo $x \in J$ y si f es continua en un punto $a \in I$, pruébese que $f(a) = 0$.

4.9. Sea $I \subset \mathbb{R}^p$ un intervalo compacto y sea $f: I \to \mathbb{R}$ una función acotada. Si $f(x) = 0$ para todo $x \in I - N$, donde N es un conjunto de medida nula, pruébese que f es integrable en I y que $\int_I f = 0$.

4.10. Sea $I \subset \mathbb{R}^p$ un intervalo compacto y sea $f: I \to \mathbb{R}$ una función acotada. Si $f(x) \geqslant 0$, para todo $x \in I$, si f es integrable en I y si $\int_I f = 0$, pruébese que entonces $C = \{x \in I \;/\; f(x) = 0\}$ es un conjunto de medida nula (indicación: acudir a los conjuntos $C_n = \{x \in I \;/\; f(x) \geqslant 1/n\}$).

4.11. Sea f la función real definida en $[0, 2] \times [0, 1]$ mediante las expresiones: $f(x, y) = 0$, si $x \in \mathbb{Q}$; $f(x, y) = x - 1$, si $x \notin \mathbb{Q}$. Recurriendo a las sumas superiores e inferiores, hacer $\underline{\int_I} f$ e $\overline{\int_I} f$.

4.12. Sea $C \subset \mathbb{R}$ un conjunto de medida nula; sea $D = C \times \mathbb{R}$. Pruébese que $D \subset \mathbb{R}^2$ tiene medida nula.

4.13. Calcular la integral doble

$$I = \iint_C \frac{(x+y)^2}{(x+2y+1)(y+1)}\,dx\,dy,$$

donde $C = [0, 2] \times [0, 1]$

4.14. Calcular la siguiente integral doble

$$I = \iint_C e^{(y-x)/(y+x)}\,dx\,dy,$$

donde $C = \{(x, y) \in \mathbb{R}^2 \,/\, x \geqslant 0, y \geqslant 0, x + y \leqslant 1\}$

4.15. Calcular la siguiente integral doble

$$I = \iint_C \frac{3xy(x^2+y^2) - x^3 - y^3}{(x^2+y^2)^{3/2}}\,dx\,dy$$

donde $C = \{(x, y) \in \mathbb{R}^2 \,/\, x^2 + y^2 \leqslant 1\}$

4.16. Calcular la siguiente integral doble en los dos casos que se indican:

$$I = \iint_C \frac{1}{(x^2+y^2+1)^2}\,dx\,dy, \qquad \text{siendo } C\text{:}$$

1.º El bucle de la curva $(x^2 + y^2)^2 = x^2 - y^2$ (lemniscata).

2.º El triángulo de vértices (0, 0), (2, 0) y $(1, \sqrt{3})$.

4.17. Calcular la siguiente integral triple

$$I = \iiint_C |xyz|\,dx\,dy\,dz$$

$$\text{donde } C = \left\{(x, y, z) \in \mathbb{R}^3 \;\middle/\; \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leqslant 1\right\}$$

4.18. Hallar el volumen del sólido

$$S = \{(x, y, z) \in \mathbb{R}^3 \,/\, x^2 + y^2 \leqslant 1,\ x^2 + z^2 \leqslant 1\}$$

4.19. Calcular la siguiente integral triple:

$$I = \iiint_C z\,dx\,dy\,dz$$

donde $C = \{(x, y, z) \in \mathbb{R}^3 \,/\, x^2 + y^2 + z^2 \leqslant 1,\ x^2 + y^2 - z^2 \leqslant 0,\ z \geqslant 0\}$

4.20. Calcular la integral de $xy + yz + zx$ extendida a la región del primer octante situada «debajo» del plano $x + y + z = 1$.

4.21. Calcular la siguiente integral triple

$$I = \iiint_C \frac{dx\,dy\,dz}{x^2 + y^2 + (z-3)^2}$$

siendo $C = \{(x, y, z) \in \mathbb{R}^3 \,/\, x^2 + y^2 + z^2 \leqslant 1\}$

4.22. Calcular el valor de

$$I = \int_2^4 dx \int_{4/x}^{(20-4x)/(8-x)} (4y - y^2)\,dy$$

4.23. Sea $f\colon C \to \mathbb{R}^2$ una función de clase $\mathscr{C}^1$ en el conjunto medible $C \subset \mathbb{R}^2$; y considérese la integral $I = \iint_C [(f_x')^2 + (f_y')^2]$. Realícese en I el cambio «inversión de centro el origen y potencia 1», esto es $(x, y) \mapsto (u, v)$ con $u = x/(x^2 + y^2)$ y $v = y/(x^2 + y^2)$ (se supone que $(0, 0) \notin C$).

4.24. Calcular la siguiente integral doble:

$$I = \iint_C \frac{dx\,dy}{\sqrt{(x^2 + y^2 + c^2)^2 - 4c^2x^2}}$$

$$\text{donde } C = \left\{(x, y) \in \mathbb{R}^2 \;\middle/\; \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1\right\}$$

$$\text{y} \quad a^2 = b^2 + c^2$$

(Indicación: cambio de variables $x = c \cos u \operatorname{Ch} v$ e $y = c \operatorname{sen} v \operatorname{Sh} v$).

4.25. Calcular la siguiente integral triple

$$I = \iiint_C \frac{xz^3}{y^2 - 1}\,dx\,dy\,dz$$

donde $C = \{(x, y, z) \in \mathbb{R}^3 \,/\, x \geqslant 0, y \geqslant 0, z \geqslant 0, x^2 + y^2 + z^2 \leqslant 1\}$

4.26. Calcular la siguiente integral doble

$$I = \iint_C xe^{-x^2/y}\,dx\,dy,$$

donde $C = \{(x, y) \in \mathbb{R}^2 \,/\, x \geqslant 0, 1 \leqslant y \leqslant 2, y > x^2\}$

4.27. Calcular la siguiente integral triple:

$$I = \iiint_C \frac{dx\,dy\,dz}{(1 + x + y + z)^3}$$

donde $C = \{(x, y, z) \in \mathbb{R}^3 \,/\, x \geqslant 0, y \geqslant 0, z \geqslant 0, x + y + z \leqslant 1\}$

4.28. Si es C el conjunto, del plano xy, situado en el interior de la curva de ecuación polar $\rho = 1 + \cos\theta$ (cardioide), hallar $I = \iint_C x\,dx\,dy$.

4.29. Calcular la integral doble $I = \iint_C x^2y^2\,dx\,dy$, siendo C el conjunto acotado situado entre las rectas $y = x$ e $y = 4x$ y las hipérbolas $xy = 1$ y $xy = 2$.

4.30. Calcular la integral doble

$$I = \iint_C |x - \operatorname{sen} y|\,dx\,dy$$

donde $C = \{(x, y) \in \mathbb{R}^2 / 0 \leqslant x \leqslant 1,\ 0 \leqslant y \leqslant \pi\}$

4.31. Calcular $\iint_I f$, siendo $I = [0, 1] \times [0, 1]$ y $f: I \to \mathbb{R}$ la función definida por $f(x, y) = x$, si $x > y$; $f(x, y) = y^2$, si $x \leqslant y$.

4.32. Calcular la integral triple de x^2yz^3 extendida al recinto limitado entre los planos $y = 0$ y $z = 0$ y las dos superficies cilíndricas $y^2 = ax - x^2$ y $z^2 = 4ax$.

4.33. Hallar el área A de la región del plano que forman los puntos interiores a la curva $\rho = a(1 + \cos\theta)$ (cardioide; ρ y θ coordenadas polares) y exteriores al círculo $\rho = a$.

4.34. Calcular el volumen del sólido

$$\{(x, y, z) \in \mathbb{R}^3 / \left[\sqrt{x^2 + y^2} - 1\right]^2 / a^2 + z^2 / b^2 \leqslant 1\}$$

4.35. Sea ABC un triángulo, cuyos lados miden a, b y c; sea h la altura correspondiente al lado AB. Se hace girar el triángulo alrededor de su lado AB. Hallar la superficie y el volumen del sólido que se obtiene al girar.

4.36. Hallar el volumen del sólido

$$\left\{(x, y, z) \in \mathbb{R}^3 \Big/ \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leqslant 1, \alpha x + \beta y + \gamma z \geqslant 1\right\}$$

4.37. Se considera la superficie S de ecuaciones paramétricas $x = \rho\cos\theta$, $y = \rho\operatorname{sen}\theta$, $z = f(\theta)$, donde f es una función de clase $\mathscr{C}^1$. Sea V el volumen de la porción de S limitada entre los planos $\theta = \theta_0$ y $\theta = \theta_1$ y el cilindro de eje vertical $(x^2 + y^2)^{1/2} = f'(\theta)$. Si A el área de la proyección sobre $z = 0$ de dicha porción de superficie, hallar V/A.

4.38. Obtener una expresión para el área de la porción de la superficie de ecuación cilíndrica $z = F(\rho, \theta)$ que se obtiene cuando (ρ, θ) recorre un conjunto medible C.

4.39. Hallar el área A de la porción de la superficie $z^2 = 2xy$ que se proyecta en el plano xy en el rectángulo R de lados $x = 0$, $x = 2$, $y = 0$ e $y = 1$.

4.40. Hallar el área A de la porción de la esfera $x^2 + y^2 + z^2 = 25$ situada por encima del plano $z = 0$ y en el interior del cilindro $2x^2 + y^2 = 25$.

4.41. Hallar el área A de la porción del paraboloide $x^2 + y^2 = az^2$ situado en el interior del cilindro $4(x^2 + y^2)^2 = a^2(x^2 - y^2)$.

4.42. Hallar el volumen V del sólido formado por los puntos interiores al cilindro $x^2 + y^2 = 2x$ que están situados sobre el plano $z = 0$ y bajo el paraboloide $z = x^2 + y^2$.

4.43. Hallar el área A de la porción del helicoide $z = \operatorname{arctg}(y/x)$ que está situada en el interior del cilindro $x^2 + y^2 = 1$.

4.44. Hallar el área A de la porción del cono $x^2 + y^2 = az^2$ situada por encima del plano $z = 0$ y en el interior del cilindro $x^2 + y^2 = by$.

4.45. Hallar el volumen V del sólido encerrado entre los dos cilindros $z = x^2$ y $z = 4 - y^2$.

SOLUCIONES

4.1. Con notación de $[66]_1$: $\bar{\mu}_C(P) \geqslant \bar{\mu}_{C-D}(P) + \underline{\mu}_D(P)$, $\bar{\mu}_C(P) \leqslant \underline{\mu}_{C-D}(P) + \bar{\mu}_C(P)$, luego $\bar{\mu}_C(P) - \bar{\mu}_D(P) \leqslant \underline{\mu}_{C-D}(P) \leqslant \bar{\mu}_{C-D}(P) \leqslant \bar{\mu}_C(P) - \underline{\mu}_D(P)$; hacer que $|P| \to 0$.

4.2. Si C^a tiene h puntos, con centro en cada uno tómese un intervalo de lado $l < \sqrt[p]{\varepsilon/(2h)}$.

4.3. $\{I_1, ..., I_n\}$ recubrimiento compacto de C^a, $\mu(I_2) + \cdots + \mu(I_n) < \varepsilon/2$. La parte de C no cubierta por los I_i es finita, luego de contenido nulo.

4.4. $\forall\ P$, $s(f, P) + s(g, P) \leqslant s(f + g, P)$. $\forall\ \varepsilon > 0$, $\exists\ P/\underline{\int}_I f + \underline{\int}_I g - \varepsilon \leqslant s(f, P) + s(g, P) \leqslant s(f + g, P) \leqslant \underline{\int}_I (f + g)$.

4.5. 1.° φ es discontinua en conjunto de contenido nulo; aditividad de la integral.
2.° Para cada $P = \{I_1, ..., I_n\}$, sea $\varphi_P : I \to \mathbb{R}$ tal que $\varphi(\mathring{I}_i) = \inf f(I_i)$, luego $\int_I \varphi_P = s(f, P)$.

4.6. $\varphi(y) = y - 1$ si $0 \leqslant y \leqslant 1$, $\varphi(y) = 0$ si $1 \leqslant y \leqslant 2$; $\psi(y) = 0$ si $0 \leqslant y \leqslant 1$, $\psi(y) = y - 1$ si $1 \leqslant y \leqslant 2$; $\int_0^2 \varphi = -1/2$, $\int_0^2 \psi = 1/2$. Si existiese $\int_I f$, estas dos últimas integrales serían iguales.
$\int_0^1 dx \int_0^2 f(xy)\, dy = \int_0^1 0\, dx = 0$.

4.7. $(x, y) \mapsto F(x, y) = f(x)$ y $(x, y) \mapsto G(x, y) = g(y)$ integrables, pues $D(F) = D(f) \times [c, d]$ y $D(G) = [a, b] \times D(g)$ tienen medida nula (D = conjunto de puntos de discontinuidad); $\phi = F \cdot G$ integrable.

4.8. $f(\boldsymbol{a}) = h > 0$; I_0 intervalo compacto tal que $\boldsymbol{a} \in I_0 \subset I$ y $f(\boldsymbol{x}) > h/2$ para $\boldsymbol{x} \in I_0$; $\int_I f \geqslant \int_{I_0} f \geqslant h\mu(I_0)/2 > 0$.

4.9. Sea $g: I \to \mathbb{R}$ la función nula; g integrable, $g = f$ salvo en conjunto de medida nula, luego $\exists \int_I f = \int_I g = 0$.

4.10. $C = \bigcup_{n \in \mathbb{N}} C_n$; $(\mu(C_n) = 0,\ \forall\ n) \Rightarrow \mu(C) = 0$. Fijado n y $\forall\ \varepsilon > 0$, $\exists\ P/0 \leqslant S(f, P) \leqslant \varepsilon/n$; es $M(f, I_i) \geqslant 1/n$; la familia $\{I_i \in P / I_i \cap C_n \neq \varnothing\}$ recubre C_n y es tal que sus medidas suman menos que $nS(f, P) \leqslant \varepsilon$.

4.11. $\underline{\int}_I f = -1/2$, $\overline{\int}_I f = 1/2$.

4.12. $D = \bigcup_{n \in \mathbb{N}} D_n$ con $D_n = C \times [-n, n]$; D_n es de medida nula: $\forall\ \varepsilon$, $\exists\ I_1, ..., I_m, .../C \subset \bigcup_{m \in \mathbb{N}} I_m$, $\sum_{m=1}^{\infty} \mu(I_m) < \varepsilon/2n$, acudir a los $J_m = I_m \times [-n, n]$.

4.13. Descomponer en fracciones simples el integrando: $1/(x + 2y + 1) + (x - 1)/(y + 1)^2$; integración iterada: $I = (5/2)$ L 5 − 3 L 3.

4.14. Cambio de variable $y - x = u$, $y + x = v$ e integración iterada; $I = (1/2)$ Sh 1.

4.15. $I = 0$.

4.16. Cambiando a coordenadas polares (ρ, θ):

$$1.° \quad I = \int_{-\pi/4}^{\pi/4} d\theta \int_0^{\sqrt{\cos 2\theta}} \frac{\rho\, d\rho}{(1 + \rho^2)^2} = \frac{\pi}{4} - \frac{1}{2}$$

$$2.° \quad I = \int_0^{\pi/3} d\theta \int_0^{\sqrt{3}/\cos(\theta - \pi/6)} \frac{\rho\, d\rho}{(1 + \rho^2)^2} = \frac{\sqrt{3}}{2} \operatorname{arctg} \frac{1}{2}$$

4.17. Cambiar de $(x/a, y/b, z/c)$ a esféricas; $I = a^2 b^2 c^2/6$.

4.18. $$16 \int_0^1 dz \int_z^1 dy \int_0^{\sqrt{1 - y^2}} dx = 16/3$$

4.19. Coordenadas esféricas:

$$I = 4 \int_0^{\pi/2} d\theta \int_{\pi/4}^{\pi/2} \operatorname{sen} \varphi \cos \varphi\, d\varphi \int_0^1 \rho^3\, d\rho = \frac{\pi}{8}$$

4.20. $$4 \int_0^1 x\, dx \int_0^{1-x} y\, dy \int_0^{1-x-y} dz = \frac{1}{30}$$

4.21. Coordenadas cilíndricas:

$$I = 4 \int_{-1}^1 dz \int_0^{\sqrt{1 - z^2}} \frac{\rho\, d\rho}{\rho^2 + (z - 3)^2} \int_0^{\pi/2} d\theta = 2\pi(1 - (4/3) \text{ L } 2)$$

4.22. Cambiar el orden de integración;

$$I = \int_1^2 (4y - y^2)\, dy \int_{4/y}^{(20 - 8y)/(4 - y)} dx = \frac{4}{3}$$

4.23. $I = \iint_D [(f'_u)^2 + (f'_v)^2]$, con $D = \{(u, v) / (x, y) \in C\}$

4.24. $I = 4 \iint_D du\, dv$, D es el rectángulo de lados $u = 0$, $u = \pi/2$, $v = 0$, $v = \operatorname{Arg\,Ch}(a/c)$; $I = \pi[\mathrm{L}\,(a + b) - \mathrm{L}\,(a - b)]$

4.25. $I = \displaystyle\int_0^1 \frac{dy}{y^2 - 1} \int_0^{\sqrt{1-y^2}} z^3\, dz \int_0^{\sqrt{1-y^2-z^2}} x\, dx = -\frac{1}{45}$

4.26. $I = \displaystyle\int_1^2 \frac{1}{2}\, y\, dy \int_0^{\sqrt{y}} (2x/y) e^{-x^2/y}\, dx = 3(e - 1)/(4e)$

4.27. $I = \displaystyle\int_0^1 dz \int_0^{1-z} dy \int_0^{1-y-z} dz/(1 + x + y + z)^3 = (\mathrm{L}\,2 - 5/8)/2$

4.28. Coordenadas polares:

$$I = 2 \int_0^{\pi} \cos\theta\, d\theta \int_0^{1 + \cos\theta} \rho^2\, d\rho = \frac{5\pi}{4}$$

4.29. Coordenadas polares:

$$I = (1/4) \int_{\pi/4}^{\operatorname{arctg} 4} \operatorname{sen}^2 2\theta\, d\theta \int_{\sqrt{2/\operatorname{sen} 2\theta}}^{\sqrt{4/\operatorname{sen} 2\theta}} \rho^5\, d\rho = (7/3)\,\mathrm{L}\,2$$

4.30. $\varphi(y) = \displaystyle\int_0^1 |x - \operatorname{sen}\, y|\, dx = \operatorname{sen}^2 y - \operatorname{sen} y + \frac{1}{2}$; $I = \displaystyle\int_0^{\pi} \varphi(y)\, dy = \pi$

4.31. $\varphi(y) = \displaystyle\int_0^1 f(x, y)\, dx = y^3 - \frac{y^2}{2} + \frac{1}{2}$; $I = \displaystyle\int_0^1 \varphi(y)\, dy = -1$

4.32. $\displaystyle\int_0^a x^2\, dx \int_0^{\sqrt{ax - x^2}} y\, dy \int_0^{\sqrt{4ax}} z^3\, dz = \frac{a^9}{21}$

4.33. $A = 2 \displaystyle\int_0^{\pi/2} d\theta \int_a^{a(1 + \cos\theta)} \rho\, d\rho = \frac{a^2(\pi + 8)}{4}$

4.34. Coordenadas cilíndricas; $2\pi^2 ab$.

4.35. Superficie $\pi ch^2/3$; volumen $\pi h(a + b)$.

4.36. Cambio $x = ax'$, $y = by'$, $z = cz'$; volumen $= = \pi abc(2h^3 - 3h^2 + 1)/(3h^3)$, donde:

$$h = \sqrt{(a\alpha)^2 + (b\beta)^2 + (c\gamma)^2}$$

4.37. $V = \displaystyle\int_{\theta_0}^{\theta_1} d\theta \int_0^{f'(\theta)} \sqrt{\rho^2 + f'(\theta)^2}\, d\rho = \frac{1}{2}\left[\sqrt{2} + \mathrm{L}\left(1 + \sqrt{2}\right)\right] \int_{\theta_1}^{\theta_2} f'(\theta)\, d\theta = \left[\sqrt{2} + \mathrm{L}\left(1 + \sqrt{2}\right)\right] \mathrm{A}$

4.38. $\displaystyle\iint_C \sqrt{(F)^2 + (F)^2 (F'_\rho)^2 + (F'_\theta)^2}$

4.39. $A = \displaystyle\iint_R \sqrt{1 + (z'_x)^2 + (z'_y)^2} = \left(1/\sqrt{2}\right) \iint_R \left(\sqrt{x/y} + \sqrt{y/x}\right) dx\, dy = 4$

4.40. Coordenadas polares:

$$A = 5 \int_0^{2\pi} d\theta \int_0^{5/\sqrt{1 + \cos^2\theta}} \rho\, d\rho/\sqrt{25 - \rho^2} = 50\pi$$

4.41. Coordenadas polares:

$$A = (4/a) \int_0^{\pi/4} d\theta \int_0^{(a/2)\sqrt{\cos 2\theta}} (a^2 + 4\rho^2)\rho\, d\rho = a^2(20 - 3\pi)/36$$

4.42. Coordenadas cilíndricas:

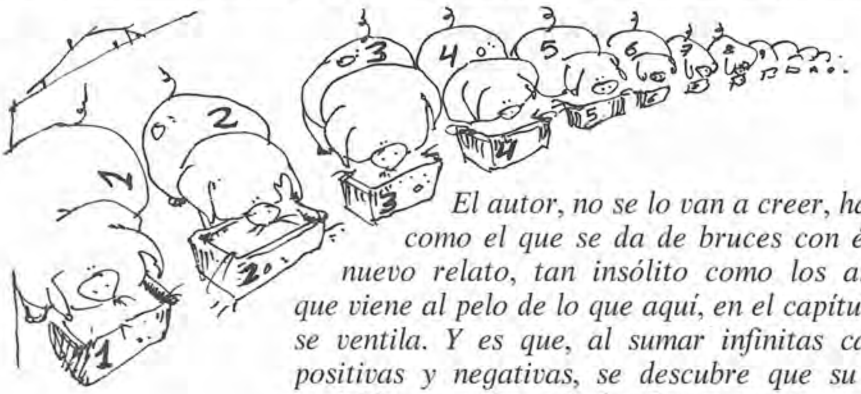
$$V = 2 \int_0^{\pi/2} d\theta \int_0^{2\cos\theta} \rho\, d\rho \int_0^{\rho^2} dz = \frac{3\pi}{2}$$

4.43. Coordenadas polares:

$$A = \int_0^1 \sqrt{1 + \rho^2} \int_0^{2\pi} d\theta = \pi\left[\sqrt{2} + \mathrm{L}\left(1 + \sqrt{2}\right)\right]$$

4.44. $A = 2\sqrt{1 + 1/a} \displaystyle\int_0^b dy \int_0^{\sqrt{by - y^2}} dx = 2\sqrt{1 + 1/a} \int_0^b \sqrt{by - y^2}\, dy = (1/4)\pi b^2 \sqrt{1 + 1/a}$

4.45. $V = 4 \displaystyle\int_0^2 dx \int_0^{\sqrt{4 - x^2}} dy \int_{x^2}^{4 - y^2} dz = (8/3) \int_0^2 (4 - x^2)\sqrt{4 - x^2}\, dx = 8\pi$ (cambio $x = 2\cos\theta$)



El autor, no se lo van a creer, ha hallado, como el que se da de bruces con él, con un nuevo relato, tan insólito como los anteriores, que viene al pelo de lo que aquí, en el capítulo quinto, se ventila. Y es que, al sumar infinitas cantidades, positivas y negativas, se descubre que su suma es variable, pues depende de cómo se vayan sumando los sumandos. Vea el amable lector que no se le engaña.

EL CAPORAL INGENUO Y SU MALICIOSO HERMANO

Cierto día, en nuestra habitual tertulia, discutían mis dos primos, On-Otín Ifni e Is-Otín Ifni, acerca de un problema que habían puesto en el colegio al hijo de uno de ellos. En él se hablaba de una escalera de mano que, apoyada por arriba en una pared vertical y por abajo en un suelo horizontal, caía de modo que su pie se movía con velocidad constante; se pedía la velocidad a la que descendía, entonces, el extremo de la escalera que se desliza por la pared. El motivo de la polémica era que, al hacer los cálculos pertinentes, de ellos se desprendía que dicho extremo de la escalera debía llegar al suelo con velocidad infinita. Mi primo On-Otín sostenía que lo infinito nunca se daba en el mundo físico, que el infinito es una abstracción matemática y que, en la solución real del problema, las cosas ocurrirían de modo que no se llegara a alcanzar velocidad infinita en ningún momento. Su hermano Is-Otín no veía, por el contrario, especial inconveniente en que alguna magnitud física, como una velocidad o una aceleración, pudiera valer infinito ocasionalmente, en un determinado instante.

Acabada la tertulia, me quedé dándole vueltas a la cabeza con aquel asunto del «infinito físico». Pensé en que, por ejemplo, muy bien pudiera pensarse en distancias infinitas, ya que nuestro espacio parecía no tener fin, y que, por ello, quizá pudiera haber cantidades infinitas de materia. Cuando ya de noche me acosté, seguía pensando en aquello de si lo infinito forma parte, o no, de nuestro mundo real. Al dormirme, soñé que vivía en Emro Oone, un país de extensión infinita y, también en sueños, estuve hablando con mis dos primos. Is-Otín era un habitante de aquel país, pero On-Otín hablaba conmigo desde fuera; él no estaba en Emro Oone aunque le veía con toda claridad; era algo así como un fantasma o como el apuntador en una obra teatral, que habla con los actores, pero no está entre ellos.

Allí, mi primo Is-Otín era el encargado o caporal de una gran casa de labranza, de extensión infinita, con infinitas cabezas de ganado, con pastos infinitos, y estaba mostrándome comó habían dispuesto infinitos pesebres, repartidos de trecho a trecho a lo largo de una línea recta; desde donde estábamos veía el primer pesebre, el segundo, el tercero, etc., pero llegaba un momento en que los pesebres se perdían en lontananza. Por allí, en las proximidades de los pesebres, había infinitas vacas, esparcidas más o menos uniformemente por los alrededores de aquella recta. Por cada vaca se disponía de un pesebre en el que, diariamente, se depositaba una paca de forraje; el peso de cada una de estas pacas era de una «átraj», cantidad que, según allí se estimaba, era suficiente para tener sobradamente alimentado a uno de aquellos animales. Is-Otín tomaba las debidas precauciones para que la granja se abasteciera de las pacas necesarias para dar de comer a todo aquel ganado; el cómputo diario de forraje era nulo: lo servido era igual a lo comido.

Se nos apareció entonces On-Otín, como si fuese una visión, con evidentes ganas de polemizar: quería demostrarnos que él tenía la razón, quería evidenciar que, si seguimos las reglas que rigen en el mundo de lo finito, al manejar lo infinito corremos el riesgo de llegar a situaciones absurdas. Su hermano aceptó el reto, se dispuso a entrar en el juego. On-Otín dijo entonces a Is-Otín que, si le parecía bien y para facilitar las cosas, mandase numerar las vacas y numerar los pesebres, con los números naturales 1, 2, 3, 4, ..., sin saltarse ninguno; como había infinitos de unas y de otros, para ello se precisarían de todos los números naturales. Is-

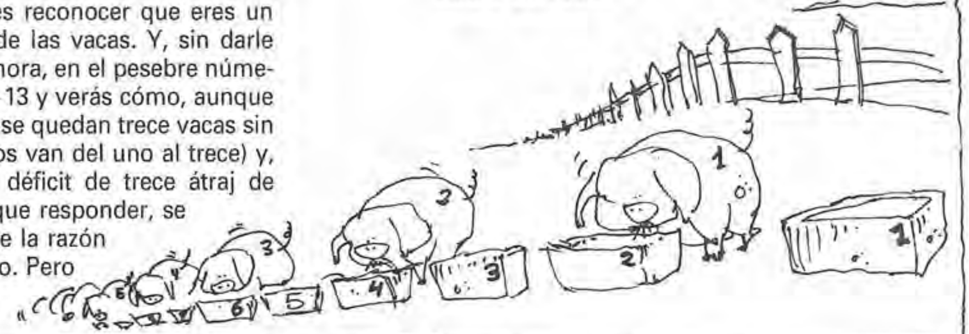
5

Otín accedió a ello y mandó que así se realizase; al rato la cosa estuvo hecha: empezando por donde nos encontrábamos, se fue poniendo un número en cada pesebre y un número sobre cada vaca y ello se hizo debidamente, pues cada uno de los números naturales se usó una vez y sólo una para numerar a un pesebre y, también, para numerar a una vaca.

Una vez que aquello estuvo hecho, dijo On-Otín que, aquel día, cada vaca comiese en el pesebre que tenía su mismo número; al hacerse de este modo, resultó que, según las previsiones de Is-Otín, no hubo ni superávit ni déficit de forraje, siendo así que comieron todas las vacas y comieron la ración que les correspondía. Pero al día siguiente, On-Otín pidió que cada vaca comiese en el pesebre cuyo número era el siguiente al que ella tenía marcado; es decir, la vaca número n comería en el pesebre número $n+1$, y ello para cualquiera que fuese el número natural n. Su hermano entrevió que aquello traería consecuencias desagradables para él, pero no pudo oponerse objetivamente a lo que se le pedía. Al hacer las cosas como había dispuesto On-Otín, resultó que todas y cada una de las vacas recibió su átraj de forraje; no obstante, había sobrado una átraj, la del pesebre número uno, en él que no comió ninguna vaca. Así pues, aquel día, en contra de lo previsto por Is-Otín, hubo superávit de forraje. Is-Otín quedó aturdido; pero, antes de que saliera de su desconcierto, On-Otín volvió a la carga diciendo: si, en lugar de enviar la vaca número n al pesebre $n+1$, la enviamos al pesebre $n+13$, pongo por caso, nos habríamos encontrado con un superávit de trece átraj de forraje, debes reconocer que eres un mal administrador del pasto de las vacas. Y, sin darle tiempo a responder, añadió: ahora, en el pesebre número n pon a comer a la vaca $n+13$ y verás cómo, aunque se ocupan todos los pesebres, se quedan trece vacas sin comer (aquellas cuyos números van del uno al trece) y, consecuentemente, habrá un déficit de trece átraj de forraje. Is-Otín, que no sabía que responder, se dio por vencido, reconoció que la razón estaba del lado de su hermano. Pero este siguió adelante con sus argumentaciones y dijo: ¿qué pasaría si a la vaca número n la envías al pesebre número $2n$ (y ello para todo número natural)? Te encontrarías entonces con que das de comer una átraj a cada una de las vacas y, no obstante, aún te quedan íntegras las infinitas átraj de los pesebres de número impar, qué puedes usar para dar de comer a las vacas el día siguiente, sin más que mandar a la vaca número n al pesebre número $2n-1$, con lo que tendrías un superávit infinito de forraje: podríamos decir que te sobró la mitad del forraje. Y, sin dar un momento de respiro a su hermano, añadió: ¿qué pasaría si a la vaca número n la envías al pesebre número $13n$, pongamos por caso? pues pasaría que, con el forraje que tenías previsto para un día, podrás dar de comer a las vacas durante trece días. Pero esto no es todo, continuó diciendo On-Otín, más su hermano le interrumpió y, visiblemente enojado, dijo: ya sabemos los dos, desde hace un rato, que tú tenías razón, no hace falta que te regodees por ello. Y dicho esto, se marchó. Entonces desperté y lo hice con el mal sabor de boca que deja haber presenciado un comportamiento tan prepotente y desafortunado como el que tuvo On-Otín con su hermano.

Nuestras razones serán tanto más convincentes cuanto menos alarde hagamos de ellas. Jactarse de estar en posesión de la verdad en nada ayuda a su aceptación, si no que, muy al contrario, mueve a rechazarla.

Gudor Ben Jusá



CAPÍTULO 5

Integrales impropias y paramétricas

5.1. Integrales múltiples impropias.—5.2. Integrales paramétricas.—5.3. Integrales paramétricas impropias.—Ejercicios y problemas.

5.1. INTEGRALES MÚLTIPLES IMPROPIAS

Hasta este momento, siempre que se ha considerado una integral $\int_C f$, el campo de integración C ha sido un conjunto acotado (es más, se ha supuesto que C era medible) y al integrando f se le ha exigido que fuera una función acotada en C. Interesa extender el concepto de integral para dar cabida a la integración en algunos supuestos en los que C o f, o ambos, no son acotados; a estas nuevas integrales se las llama «impropias» o «generalizadas». De ellas nos ocupamos ahora y lo haremos abordando simultáneamente ambas posibilidades, que se estudian de modo unificado. Cuando interesa señalar que una integral múltiple no es impropia, es decir, cuando se trata de una integral de las estudiadas en el capítulo anterior, se acostumbra a decir que la tal integral es «propia».

Una integral impropia $\int_C f$ se define como límite, para $n \to \infty$, de integrales propias $\int_{C_n} f$, cuando C_n «crece monótonamente tendiendo a C» y en el supuesto de que f esté acotada en todos los C_n. Aquí (C_n) es cualquier sucesión creciente, de subconjuntos medibles de C en los que f está acotada, que «aproxime adecuadamente» a C (esto significa que, dado cualquier conjunto medible $C' \subset C$, ha de ser $C' \subset C_n$ a partir de cierto n). Se dice que la integral impropia es convergente si dicho límite existe y es el mismo para cualquiera que sea la tal sucesión (C_n) y se toma $\int_C f = \lim \int_{C_n} f$. Interesa notar que la exigencia de que este límite sea el mismo para cualquiera que sea la susodicha sucesión (C_n) es muy restrictivo(*).

(*) En el caso de integrales simples, se es mucho menos exigente que aquí: una integral $\int_a^{\to b} f$, impropia en b, se define como límite de $\int_a^x f$ cuando $x \to b^-$; es decir, ahora que se integra en $C = [a, b]$, los conjuntos C_n son sólo los del tipo $C_n = [a, x_n]$, con $a \leqslant x_n < b$. Una condición análoga a ésta, para las integrales múltiples, sería considerar para conjuntos C_n sólo los del tipo $C_n = C \cap I_n$, donde I_n es un intervalo que crece al crecer n o sólo los del tipo $C_n = C \cap B_n$, donde (B_n) es una sucesión creciente de bolas, por ejemplo.

Esta limitación tan fuerte conduce a que, contrariamente a lo que ocurre en el caso de las integrales simples, ahora la convergencia de una integral múltiple impropia coincida con su convergencia absoluta y a que no exista la «convergencia condicional» al estilo de la de las series.

DEFINICIÓN DE INTEGRAL IMPROPIA

[73]

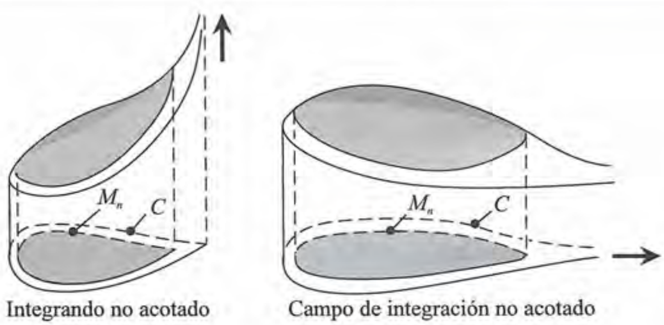
Sean dados un conjunto $C \subset \mathbb{R}^p$ y una función $f: C \to \mathbb{R}$ tales que existe alguna «sucesión básica para la integración de f en C», entendiendo por tal a una sucesión (M_n) de subconjuntos de C verificando que:

1. M_n es compacto y medible para todo $n \in \mathbb{N}$.
2. (M_n) es creciente, o sea: $M_n \subset M_{n+1}$ para todo $n \in \mathbb{N}$.
3. Si $M \subset C$ es compacto y medible, entonces $M \subset M_n$ para algún $n \in \mathbb{N}$.
4. La función f es integrable en M_n, para todo $n \in \mathbb{N}$.

Entonces, de la función f (que es integrable en todo $M \subset C$ compacto y medible) se dice que es integrable en C si, para cualquiera que sea la «sucesión básica» (M_n) que verifique a las cuatro condiciones anteriores, acontece que existe el límite (finito o infinito) de $\int_{M_n} f$, cuando $n \to \infty$, y dicho límite es el mismo para cualquiera que sea la tal sucesión (M_n); a este límite se le llama integral (convergente o divergente) de f en C y se le denota por $\underset{\to}{\int}_C f$ (o simplemente poniendo $\int_C f$):

$$\underset{\to}{\int}_C f = \lim_{n \to \infty} \int_{M_n} f$$

Nota: Si este límite no existe o es diferente para dos «sucesiones básicas» distintas, se dice que f no es integrable en C, o que no existe la integral $\underset{\to}{\int}_C f$ o, también, que dicha integral es oscilante.



Integrando no acotado — Campo de integración no acotado

[73]$_1$ Observaciones

1.° Con la anterior definición se sale al paso, a la vez, de las dos situaciones «patológicas» que aquí interesan: integrando no acotado y campo de integración no acotado. Como caso simple de integral impropia $\underset{\to}{\int}_C f$, consideremos el siguiente ejemplo, en el que C está acotado pero f no lo está: el conjunto C es compacto y medible y la función f es continua en $C - \{\boldsymbol{a}\}$ (donde $\boldsymbol{a} \in C$) y no está acotada «cerca» del punto $\boldsymbol{a}$ (esto es, f no está acotada en un entorno de $\boldsymbol{a}$). En este supuesto, acudiendo a una sucesión decreciente $B_1, B_2, ..., B_n, ...$, de entornos abiertos medibles de centro $\boldsymbol{a}$ cuyos diámetros tiendan a 0, los conjuntos $M_n = C - B_n$ forman lo que hemos llamado una «sucesión básica para la integración de f en C»; según la anterior definición, f será integrable en C y su integral es I si acontece que $\lim \int_{M_n} f = I$ (cuando $n \to \infty$) y ello es independiente de la forma en que se elijan los entornos B_n (con diámetros tendiendo a cero para $n \to \infty$).

2.º En la anterior definición [73], de integral impropia, se da acogida a la inmensa mayoría de los casos de interés, sobre todo en el campo de las aplicaciones. Aunque para nosotros sea satisfactoria, no está de más que señalemos otro modo de proceder a este respecto, que generaliza al nuestro:

- Si la función $f: C \to \mathbb{R}$, definida en $C \subset \mathbb{R}^p$, es integrable en todo subconjunto medible de C, se dice que f es integrable en C si existe un número I (que será único y se pondrá $I = \underline{\int}_C f$) tal que, para cada $\varepsilon > 0$, existe un conjunto medible $M_\varepsilon \subset C$ tal que $|\underline{\int}_C f - \int_M f| < \varepsilon$ para cualquiera que sea el conjunto medible M tal que $C \supset M \supset M_\varepsilon$.
- Sea $f: C \to \mathbb{R}$ una función, definida en $C \subset \mathbb{R}^p$, que no esté acotada en algún subconjunto medible de C. Si f es integrable en todo subconjunto medible de C en el que f esté acotada, se dice que f es integrable en C si existe un número I (que será único y se pondra $I = \underline{\int}_C f$) tal que, para cada $\varepsilon > 0$, existe un conjunto medible en el que f está acotada $M_\varepsilon \subset C$ tal que $|\underline{\int}_C f - \int_M f| < \varepsilon$ para cualquiera que sea el conjunto medible M de C en el que f esté acotada y tal que $M \supset M_\varepsilon$.

3.º Cuando existe la integral impropia $\underline{\int}_C f$ (según la anterior definición [73]) y el límite que la define es finito, entonces se dice también que dicha integral es convergente. Si el límite de $\int_{M_n} f$ (cuando $n \to \infty$), del que se habla en [73], es infinito, se dice entonces que la integral en cuestión es divergente. La definición de integral múltiple impropia no acoge a aquellos casos en los que el susodicho límite o no existe o depende de la «sucesión básica» que se considere; cuando acontezca alguna de estas cosas, diremos que la integral es oscilante.

[73]$_2$ Ejercicio

Considérese la función real $(x, y) \mapsto f(x, y)$ definida, en el primer cuadrante del plano xy, mediante la siguiente expresión y considérense también los conjuntos $C(\varepsilon, \delta)$ y C:

$$f(x, y) = \frac{y - x}{xy} \qquad \begin{aligned} C(\varepsilon, \delta) &= \{(x, y) \in \mathbb{R}^2 \;/\; 0 < \varepsilon \leqslant x \leqslant 1,\ 0 \leqslant \delta \leqslant y \leqslant 1\} \\ C &= \{(x, y) \in \mathbb{R}^2 \;/\; 0 < x \leqslant 1,\ 0 < y \leqslant 1\} \end{aligned}$$

1.º Calcular la integral (no impropia) de f en $C(\varepsilon, \delta)$.

2.º Comprobar que no existe la integral impropia $\underline{\int}_C f$ (o, si se prefiere, que dicha integral es oscilante).

Resolución

1.° $\displaystyle\int_{C(\varepsilon,\,\delta)} f = \int_\delta^1 dy \int_\varepsilon^1 \left(\frac{1}{x} - \frac{1}{y}\right) dx = \int_\delta^1 dy \left[\mathrm{L}\, x - \frac{x}{y}\right]_{x=\varepsilon}^{x=1} = \int_\delta^1 \left(\frac{\varepsilon - 1}{y} - \mathrm{L}\,\varepsilon\right) dy =$

$= [(\varepsilon - 1)\, \mathrm{L}\, y - y\, \mathrm{L}\, \varepsilon]_\delta^1 = (1 - \varepsilon)\, \mathrm{L}\, \delta - (1 - \delta)\, \mathrm{L}\, \varepsilon.$

2.° Para que exista la integral impropia $\underline{\int}_C f$ es necesario que exista el límite de $\int_{C(\varepsilon,\,\delta)} f$ cuando $\varepsilon \to 0$ y $\delta \to 0$ y que dicho límite sea el mismo para cualquiera que sea la forma en que ε y δ tienden a cero. Esto no es así, ya que, por ejemplo:

si $\delta = \varepsilon \to 0$, entonces $\int_{C(\varepsilon,\,\delta)} f = 0 \to 0$

si $\delta = 2\varepsilon \to 0$, entonces $\int_{C(\varepsilon,\,\delta)} f = (1 - \varepsilon)\, \mathrm{L}\, 2 \to \mathrm{L}\, 2$

si $\delta = \varepsilon^2 \to 0$, entonces $\int_{C(\varepsilon,\,\delta)} f = (1 - 2\varepsilon + \varepsilon^2)\, \mathrm{L}\, \varepsilon \to -\infty$

Así pues, $\underline{\int}_C f$ es oscilante (o no existe).

[73]$_3$ Ejercicio

Considérese la función $(x, y) \mapsto f(x, y)$ definida, en $\mathbb{R}^2 - \{(0, 0)\}$, mediante la expresión

$$f(x, y) = \frac{y}{(x^2 + y^2)^{3/2}}$$

Sea $U(\varepsilon)$ y $V(\varepsilon)$ las regiones del plano interiores a las siguientes curvas (dadas en coordenadas polares):

$\rho = \varepsilon$ (circunferencia de radio ε y centro el origen de coordenadas)

$\rho = \varepsilon e^{\operatorname{sen}\theta}$ (óvalo que encierra al origen; nótese que sus distancias máximas y mínimas al origen son $\rho(\pi/2) = e$ y $\rho(-\pi/2) = 1/e$)

1.° Calcular las integrales de f en los conjuntos $C_u = C - U(\varepsilon)$ y $C_v = C - V(\varepsilon)$, donde C es círculo cerrado de centro el origen y radio unidad.

2.° Comprobar que no existe la integral impropia de f en el conjunto $C_0 = C - \{(0, 0)\}$.

Resolución

1.° Cambiando a coordenadas polares, se obtiene que:

$$\int_{C_u} f = \int_0^{2\pi} \operatorname{sen}\theta\, d\theta \int_\varepsilon^1 \frac{d\rho}{\rho} = \int_0^{2\pi} (0 - \mathrm{L}\,\varepsilon) \operatorname{sen}\theta\, d\theta = (-\mathrm{L}\,\varepsilon)0 = 0$$

$$\int_{C_v} f = \int_0^{2\pi} \operatorname{sen}\theta\, d\theta \int_{\varepsilon e^{\operatorname{sen}\theta}}^1 \frac{d\rho}{\rho} = \int_0^{2\pi} (-\mathrm{L}\,\varepsilon - \operatorname{sen}\theta) \operatorname{sen}\theta\, d\theta =$$

$$= (-\mathrm{L}\,\varepsilon) \int_0^{2\pi} \operatorname{sen}\theta\, d\theta - \int_0^{2\pi} \operatorname{sen}^2\theta\, d\theta = (-\mathrm{L}\,\varepsilon)0 - \pi = -\pi$$

(obsérvese que, curiosamente, estas dos integrales son independientes de ε).

2.° Si existiese la integral impropia de f en C_0, esta integral sería el límite de las integrales $\int_{C_u} f$ y $\int_{C_v} f$, cuando $\varepsilon \to 0$. Pero estos límites son distintos (valen 0 y $-\pi$), por lo que la referida integral impropia no existe (o es oscilante).

[73]$_4$ Integrales «seudoimpropias»(*)

Sea $C \subset \mathbb{R}^p$ un conjunto medible y sea $f: C \to \mathbb{R}$ una función acotada en C. Se verifica que: si existe la integral impropia $\underset{\to C}{\int} f$, entonces también existe la integral (no impropia, ordinaria) $\int_C f$ y ambas integrales son iguales.

Demostración

Como existe la integral impropia $\underset{\to C}{\int} f$, sabemos que ella es el límite (cuando $n \to \infty$) de $\int_{M_n} f$, donde (M_n) es una cierta «sucesión básica para la integración de f en C».

Para comprobar que existe $\int_C f$, hemos de comprobar (véase [67]) que el conjunto D de los puntos de discontinuidad de f en C tiene medida nula, esto es, que dado $\varepsilon > 0$ (cualquiera) existe una sucesión numerable de intervalos compactos de $\mathbb{R}^p$, $I_1, ..., I_n, ...$, tales que

$$D \subset \bigcup_{n=1}^{\infty} I_n \qquad \text{y} \qquad \sum_{n=1}^{\infty} \mu(I_n) < \varepsilon \tag{1}$$

Según ya sabemos (véase [66]$_6$) existe un conjunto compacto $C' \subset C$ tal que $\mu(C) - \mu(C') < \varepsilon/4$ y, como (M_n) es una «sucesión básica» y C' es compacto, es $C' \subset M_h$ para cierto $h \in \mathbb{N}$, de lo que se desprende que $\mu(C - M_h) < \varepsilon/4$. Según ya se demostró anteriormente (véase [66]$_1$, 4), existen ciertos intervalos compactos $J_1, ..., J_m$ (en cantidad finita) tales que

$$C - M_h \subset J_1 \cup \cdots \cup J_m \qquad \text{y} \qquad \mu(J_1) + \cdots + \mu(J_m) < \mu(C - M_h) + \frac{\varepsilon}{4} < \frac{\varepsilon}{2} \tag{2}$$

Sean D_1 y D_2 los conjuntos $D_1 = D \cap M_h$ y $D_2 = D - D_1$. Como f es integrable en M_h, se sabe que D_1 tiene medida nula, por lo que existe una sucesión de intervalos compactos, $K_1, ..., K_n, ...$, tal que

$$D_1 \subset \bigcup_{n=1}^{\infty} K_n \qquad \text{y} \qquad \sum_{n=1}^{\infty} \mu(K_n) < \varepsilon/2 \tag{3}$$

Como $D = D_1 \cup D_2$ y $D_2 \subset C - M_h$, de [2] y [3] se desprende que [1] se verifica para la sucesión (I_n) siguiente

$$(I_n) = J_1, ..., J_m, K_1, ..., K_n, ...$$

Visto ya que existe la integral $\int_C f$, comprobemos que ella es igual a la $\underset{\to C}{\int} f$. Empecemos observando que, según se acaba de ver, dado $\varepsilon > 0$, existe $h \in \mathbb{N}$ tal que $\mu(C - M_h) < \varepsilon/4$ y como además (M_n) es creciente, de ello resulta que $\mu(C - M_n)$ tiene límite cero cuando $n \to \infty$.

(*) Dada una integral impropia $\underset{\to C}{\int} f$ (o, si se prefiere, definida al modo de las integrales impropias en [73]), diremos que ella es «seudoimpropia» si C es un conjunto medible (y, por tanto, acotado) y la función f está acotada en C.

Por otra parte, si es $H > 0$ una cota de $|f|$ en C y de acuerdo con la aditividad de la integral (véase [69]) y con el teorema de la media (véase [68], 3), se puede poner para $n \in \mathbb{N}$:

$$\left|\int_C f - \int_{M_n} f\right| = \left|\int_{C-M_n} f\right| \leqslant H\mu(C - M_n)$$

y como $\mu(C - M_n)$ tiene límite cero (para $n \to \infty$), también lo tiene el primer miembro de la anterior relación, o sea, $\int_C f$ es el límite de $\int_{M_n} f$; como este último límite es igual a la integral impropia $\int_{\to C} f$, de ello resulta que $\int_C f = \int_{\to C} f$, como había que comprobar.

CARACTERIZACIÓN DE LA INTEGRABILIDAD IMPROPIA

[74]

Sean dados un conjunto $C \subset \mathbb{R}^p$ y una función $f\colon C \to \mathbb{R}$ tales que existe alguna «sucesión básica» (M_n) para la integración de f en C (esto es, para (M_n) se satisfacen las cuatro condiciones de [73]). Se verifica entonces que:

1. Si $f(x) \geqslant 0$ para todo $x \in C$, entonces existe la integral impropia de f en C y es:

$$\int_{\to C} f = \sup\left\{\int_M f \;\middle/\; \begin{matrix} M \subset C \text{ es un conjunto} \\ \text{compacto y medible}\end{matrix}\right\} \qquad (\text{si } f \geqslant o)$$

(la integral es convergente o divergente según que este supremo sea un número finito o sea infinito).

2. Si f tiene signo variable, acudiendo a las funciones f^+ y f^- (parte positiva y parte negativa de f, de C en $\mathbb{R}$), definidas por $f^+ = \text{máx}\,\{f, 0\}$ y $f^- = \text{máx}\,\{-f, 0\}$, se verifica que la integral impropia $\int_{\to C} f$ existe y es convergente o existe y es divergente según que, respectivamente, las integrales $\int_{\to C} f^+$ e $\int_{\to C} f^-$ (que existen) sean ambas convergentes o sea una convergente y la otra divergente(*). En ambos casos es:

$$\int_{\to C} f = \int_{\to C} f^+ - \int_{\to C} f^-$$

(*) Si las integrales impropias $\int_{\to C} f^+$ e $\int_{\to C} f^-$ son ambas divergentes, entonces $\int_{\to C} f$ no existe (o es oscilante), pues lím $\int_{M_n} f$ (cuando $n \to \infty$) depende de la «sucesión básica» (M_n) que se considere. La demostración de este aserto se pude realizar siguiendo un camino similar al de la demostración de [76], 2.

Demostración

1. Veamos primero que la integral existe cuando f es positiva. Para cualquiera que sea la «sucesión básica» (M_n), como ella es creciente y f es positiva, la sucesión $\left(\int_{M_n} f\right)$ es creciente y, por ello, tiene límite, al que llamaremos s, el cual es, pues, al supremo del conjunto $I = \{\int_{M_n} f \,/\, n \in \mathbb{N}\}$. Si llamamos $I' = \{\int_M f \,/\, M \subset C$ compacto y medible$\}$

y $s' = \sup I'$, lo que hemos de comprobar es que $s = s'$. Veamos que, en efecto, ha de ser $s \leqslant s'$ y $s' \leqslant s$:

- Como $I \subset I'$, es claro que $\sup I \leqslant \sup I'$, es decir, que $s \leqslant s'$.
- Como (M_n) es una «sucesión básica», dado cualquier $M \subset C$ compacto y medible, existe $n \in \mathbb{N}$ tal que $M \subset M_n$ y, consecuentemente, $\int_M f \leqslant \int_{M_n} f$; resulta, pues, que dado cualquier elemento de I', existe algún elemento de I que es mayor que él, luego $\sup I' \leqslant \sup I$, es decir, $s' \leqslant s$.

2. Nótese que $f = f^+ - f^-$ y $|f| = f^+ + f^-$ y que, por tanto, es $f^+ = \frac{1}{2}(f + |f|)$ y $f^- = \frac{1}{2}(|f| - f)$. Ahora bien, como f es integrable en todos los M_n y, por tanto (véase [68], 1), también lo es $|f|$, de las anteriores expresiones se desprende que f^+ y f^- son integrables en M_n para todo $n \in \mathbb{N}$. Así pues, (M_n) es una «sucesión básica» para la integración de f^+ y para la de f^- en C. Dado que las funciones f^+ y f^- son funciones positivas en C y de acuerdo con lo ya obtenido en el apartado anterior, dichas funciones son integrables en C, es decir, existen $s^+ = \int_{\to C} f^+$ y $s^- = \int_{\to C} f^-$; de acuerdo con la definición de integral impropia, estas integrales son:

$$s^+ = \int_{\to C} f^+ = \lim_{n\to\infty} \int_{M_n} f^+ \qquad \text{e} \qquad s^- = \int_{\to C} f^- = \lim_{n\to\infty} \int_{M_n} f^-$$

(estos límites no dependen de la «sucesión básica» (M_n) que se considere). De ello resulta que (siempre que $s^+ - s^-$ tenga sentido, es decir, salvo si es $s = +\infty$ y $s' = +\infty$):

$$\int_{\to C} f = \lim_{n\to\infty} \int_{M_n} f = \lim_{n\to\infty} \left(\int_{M_n} f^+ - \int_{M_n} f^- \right) =$$
$$= \lim_{n\to\infty} \int_{M_n} f^+ - \lim_{n\to\infty} \int_{M_n} f^- = \begin{cases} s - s' \in \mathbb{R} \\ \infty - s' = +\infty \\ s - \infty = -\infty \end{cases}$$

Así pues, si $\int_{\to C} f^+$ y $\int_{\to C} f^-$ son ambas convergentes o una es convergente y la otra divergente, entonces $\int_{\to C} f$ converge o diverge, respectivamente.

[74]₁ Ejercicio

Considérese la siguiente integral doble impropia $\int_{\to C} f$, siendo:

$$f(x, y) = \frac{x^2 y}{(x^2+1)(y^4 - x^4)} \qquad \text{y} \qquad \begin{gathered} C = \{(x, y) \in \mathbb{R}^2 \,/\, x \geqslant 0,\ m_1 x \leqslant y \leqslant m_2 x\} \\ m_1 = -\tfrac{1}{3} \text{ y } m_2 = \tfrac{1}{2} \end{gathered}$$

Compruébese que dicha integral es convergente y hállese su valor.

Resolución

El integrando es, respectivamente, positivo y negativo en los conjuntos C^+ y C^- siguientes:

$$C^+ = \{(x, y) \in \mathbb{R}^2 \,/\, x \geqslant 0, m_1 x \leqslant y \leqslant 0\} \qquad \text{y} \qquad C^- = \{(x, y) \in \mathbb{R}^2 \,/\, x \geqslant 0, 0 \leqslant y \leqslant m_2 x\}$$

Las integrales de la función dada en C^+ y en C^- son convergentes, ya que son finitos los límites (cuando $n \to \infty$) de las integrales en

$C_n^+ = \{(x, y) \in \mathbb{R}^2 / 0 \leqslant x \leqslant n, m_1 x \leqslant y \leqslant 0\}$ y $C_n^- = \{(x, y) \in \mathbb{R}^2 / 0 \leqslant x \leqslant n, 0 \leqslant y \leqslant m_2 x\}$

Así ocurre, en efecto, pues para C^- (por ejemplo):

$$\int_{C_n^-} f = \int_0^n \frac{x^2\,dx}{x^2+1} \int_0^{m_2 x} \frac{y}{y^4 - x^4}\,dy = \int_0^n \frac{x^2}{x^2+1}\,dx \left[\frac{1}{4x^2}\,\mathrm{L}\left|\frac{x^2-y^2}{x^2+y^2}\right|\right]_{y=0}^{y=m_2 x} =$$

$$= \frac{1}{4}\,\mathrm{L}\,\frac{1-m_2^2}{1+m_2^2}\int_0^n \frac{dx}{x^2+1} = \frac{1}{4}\,\mathrm{L}\,\frac{1-m_2^2}{1+m_2^2}\,[\operatorname{arctg} n] \xrightarrow[n\to\infty]{} \int_{\to C^-} f = \frac{\pi}{8}\,\mathrm{L}\,\frac{3}{5}$$

Análogamente, se obtiene que $\int_{\to C^+} f = (\pi/8)\ \mathrm{L}\ (10/8)$, por lo cual es:

$$\int_{\to C} f = \int_{\to C^+} f + \int_{\to C^-} f = \frac{\pi}{8}\left(\mathrm{L}\,\frac{10}{8} + \mathrm{L}\,\frac{3}{5}\right) = \frac{\pi}{8}\,\mathrm{L}\,\frac{3}{4}$$

$[74]_2$ Ejercicios

Estudiar la convergencia de las siguientes integrales (dobles y triples) impropias $\int_{\to C} f$ e $\int_{\to D} f$, llegando a los resultados que se indican (donde $r > 0$ es fijo, cualquiera):

1.º $$f(x, y) = \frac{1}{(x^2+y^2)^\alpha} \begin{cases} C = \{(x, y) \in \mathbb{R}^2 \ / \ 0 < x^2 + y^2 \leqslant r^2\} \ \Rightarrow \\ D = \{(x, y) \in \mathbb{R}^2 \ / \ 0 < r^2 \leqslant x^2 + y^2\} \ \Rightarrow \end{cases}$$

$$\Rightarrow \int_{\to C} f \begin{cases} \text{converge si es } \alpha < 1 \\ \text{diverge si es } \alpha \geqslant 1 \end{cases}$$

$$\Rightarrow \int_{\to D} f \begin{cases} \text{converge si es } \alpha > 1 \\ \text{diverge si es } \alpha \leqslant 1 \end{cases}$$

2.º $$f(x, y, z) = \frac{1}{(x^2+y^2+z^2)^\alpha} \begin{cases} C = \{(x, y, z) \in \mathbb{R}^3 \ / \ 0 < x^2 + y^2 + z^2 \leqslant r^2\} \ \Rightarrow \\ D = \{(x, y, z) \in \mathbb{R}^3 \ / \ 0 < r^2 \leqslant x^2 + y^2 + z^2\} \ \Rightarrow \end{cases}$$

$$\Rightarrow \int_{\to C} f \begin{cases} \text{converge si es } \alpha < 3/2 \\ \text{diverge si es } \alpha \geqslant 3/2 \end{cases}$$

$$\Rightarrow \int_{\to D} f \begin{cases} \text{converge si es } \alpha > 3/2 \\ \text{diverge si es } \alpha \leqslant 3/2 \end{cases}$$

Resolución

Comprobemos que las cosas son como se dicen para la integral $\int_{\to D} f$ del caso 1.º y para la $\int_{\to C} f$ del caso 2.º; las otras dos integrales se estudian de modo análogo. En ambos casos, el integrando es positivo, por lo que bastará con acudir a una «sucesión básica» cualquiera, para la correspondiente integración:

1.º (Caso $\underset{\to D}{\int} f$ bidimensional). Una «sucesión básica» para esta integral es la constituida por las coronas circulares de centro el origen y radios r y $n \in \mathbb{N}$ (con $n > r$); llamando M_n a una tal corona, se tiene (operando en coordenadas polares: $x = \rho \cos \theta$, $y = \rho$ sen θ):

$$\int_{M_n} f = \int_0^{2\pi} d\theta \int_r^n \rho^{1-2\alpha}\, d\rho = \frac{\pi}{1-\alpha}\left[n^{2(1-\alpha)} - r^{2(1-\alpha)}\right]$$

$$\left(\text{si } \alpha = 1, \int_{M_n} f = 2\pi(\mathrm{L}\, n - \mathrm{L}\, r)\right)$$

$$\underset{\to D}{\int} f = \lim_{n\to\infty} \int_{M_n} f = \begin{cases} r^{2(1-\alpha)}\pi/(\alpha - 1), & \text{si } \alpha > 1 \\ \infty \text{ (divergente)}, & \text{si } \alpha \leqslant 1 \end{cases}$$

2.º (Caso $\underset{\to C}{\int} f$ tridimensional). Una «sucesión básica» para esta integral es la constituida por las coronas esféricas de centro el origen y radios $1/n$ y r (para $n \in \mathbb{N}$, con $1/n < r$); llamando M_n a una tal corona, se tiene (operando en coordenadas esféricas: $x = \rho \cos \theta$ sen φ, $y = \rho \cos \theta \cos \varphi$, $z = \rho$ sen θ):

$$\int_{M_n} f = 2\int_0^{2\pi} d\theta \int_0^{\pi/2} \operatorname{sen} \varphi\, d\varphi \int_{1/n}^{r} \rho^{2-2\alpha}\, d\rho = \frac{4\pi}{3-2\alpha}\left[r^{3-2\alpha} - \left(\frac{1}{n}\right)^{3-2\alpha}\right]$$

$$\left(\text{si } \alpha = \frac{3}{2}, \int_{M_n} f = 4\pi(\mathrm{L}\, r + \mathrm{L}\, n)\right)$$

$$\underset{\to C}{\int} f = \lim_{n\to\infty} \int_{M_n} f = \begin{cases} 4\pi r^{3-2\alpha}/(3 - 2\alpha), & \text{si } \alpha < 3/2 \\ \infty \text{ (divergente)}, & \text{si } \alpha \geqslant 3/2 \end{cases}$$

[74]$_3$ Observaciones

1.ª Para el estudio de cualquier integral múltiple impropia, $\underset{\to C}{\int} f$, se supondrá que, como hasta ahora se ha venido haciendo, se dispone de una «sucesión básica para la integración de f en C». Es interesante observar que los resultados que así se obtienen son igualmente válidos si, en lugar de suponer que existe la tal «sucesión básica», se admite que f es continua en C y que existe una sucesión creciente (M_n) de conjuntos compactos y medibles tal que $M_1 \cup M_2 \cup \cdots \cup M_n \cup \cdots = C$.

Un conjunto $C \subset \mathbb{R}^p$ que se pueda expresar en la forma $C = M_1 \cup M_2 \cup \cdots \cup M_n \cup \cdots$, donde (M_n) es una sucesión creciente de conjuntos compactos, se dice que es σ-compacto. Todo conjunto abierto es σ-compacto.

2.ª Cuando, para calcular la integral impropia $\underset{\to C}{\int} f$, haya que recurrir a una «sucesión básica» (M_n) y poner $\underset{\to C}{\int} f = \lim \int_{M_n} f$, la tal sucesión, que podía ser cualquiera (con tal de que cumpla con las cuatro condiciones pertinentes; véase [73]), se elegirá de manera que la obtención efectiva de las integrales $\int_{M_n} f$ y de su límite sea, cuanto más fácil, mejor.

CASO DE INTEGRANDO POSITIVO: CRITERIOS DE CONVERGENCIA

[75]

I. CARACTERIZACIÓN DE LA CONVERGENCIA. Sean dados un conjunto $C \subset \mathbb{R}^p$ y una función $f: C \to \mathbb{R}$ tales que existe alguna «sucesión básica (M_n) para la integración de f en C» (véase [73]). Si f es positiva en C (o sea, si $f(x) \geqslant 0$, $\forall\, x \in C$), para que la integral impropia $\int_{\to C} f$ (que existe; véase [74],1) sea convergente es necesario y suficiente que, para una determinada «sucesión básica» (M_n), la sucesión numérica $(\int_{M_n} f)$ esté acotada superiormente o, lo que es equivalente, que $(\int_{M_n} f)$ tenga límite finito, el cual será, entonces, el valor de la integral impropia $\int_{\to C} f$.

II. CRITERIO DE WEIERSTRASS O DE COMPARACIÓN. Sean dados un conjunto $C \subset \mathbb{R}$ y dos funciones $f: C \to \mathbb{R}$ y $g: C \to \mathbb{R}$ tales que existe alguna «sucesión básica (M_n) para la integración de f y de g en C». Si $o \leqslant f \leqslant g$ en C (o sea, si $0 \leqslant f(x) \leqslant g(x)$, $\forall\, x \in C$) y la integral impropia $\int_{\to C} g$ es convergente, entonces también es convergente la integral impropia $\int_{\to C} f$.

Demostración

I. 1.° Veamos primero que, si $\int_{\to C} f$ es convergente, entonces $(\int_{M_n} f)$ está acotada. Esto es así ya que, como $(\int_{M_n} f)$ es creciente (pues (M_n) es creciente y f es positiva) y tiene por límite a $\int_{\to C} f$, este límite (que es finito) es cota superior de la sucesión $(\int_{M_n} f)$.

2.° Suponiendo ahora que $(\int_{M_n} f)$ es una sucesión acotada superiormente, hemos de ver que entonces $\int_{\to C} f$ es convergente. Nótese primeramente que, según ya se dijo en [74], 1, existe la integral impropia $\int_{\to C} f$; sólo hay, pues, que comprobar que esta integral es finita. Como ella existe y de acuerdo con la definición [73], es $\int_{\to C} f = \lim \int_{M_n} f$ (cuando $n \to \infty$); ahora bien, como $(\int_{M_n} f)$ es una sucesión creciente (pues (M_n) es creciente y f es positiva) y está acotada (por hipótesis), resulta que ella tiene límite finito, por lo que $\int_{\to C} f$ es finito, como había que comprobar.

II. Como f y g son positivas, sabemos que (véase [74], 1) existen las integrales impropias $\int_{\to C} f$ e $\int_{\to C} g$, las cuales son (de acuerdo con la definición [73]):

$$\int_{\to C} f = \lim_{n\to\infty} \int_{M_n} f \qquad \text{e} \qquad \int_{\to C} g = \lim_{n\to\infty} \int_{M_n} g \tag{1}$$

Ahora bien, como $f(x) \leqslant g(x)$ para todo $x \in M_n \subset C$, según la propiedad de monotonía de las integrales (no impropias; véase [69]) se verifica que $\int_{M_n} f \leqslant \int_{M_n} g$ para todo $n \in \mathbb{N}$. Llevando este resultado a [1], se concluye que $\int_{\to C} f \leqslant \int_{\to C} g$ y, como esta última integral es un número finito (es convergente), también lo es la otra, o sea, $\int_{\to C} f$ es convergente.

[75]$_1$ Ejercicio

Considérese la integral doble impropia

$$I = \iint_{\mathbb{R}^2} e^{-x^2-y^2}\, dx\, dy$$

1.° Comprobar que es convergente hallando su valor (intégrese en círculos cerrados de centro el origen y radio $r \to \infty$).
2.° Calcular el valor de la siguiente integral (integral de Gauss):

$$J = \int_{-\infty}^{+\infty} e^{-t^2}\, dt$$

(Indicación: calcúlese ahora la anterior integral I integrando en cuadrados de centro el origen y lados tendiendo a infinito.)

Resolución

1.° El integrando es positivo, por lo que I existe (finita o infinita) y es igual al límite de la sucesión (I_n) siguiente:

$$I_n = \iint_{B_n} e^{-x^2-y^2}\, dx\, dy, \qquad \text{donde } B_n = \{(x, y) \in \mathbb{R}^2 \;/\; x^2 + y^2 \leqslant n^2\}$$

Cambiando a coordenadas polares ($x = \rho \cos \theta$, $y = \rho \operatorname{sen} \theta$) e integrando, luego, por iteración, se obtiene que

$$I_n = \int_0^{2\pi} d\theta \int_0^n e^{-\rho^2} \rho\, d\rho = \int_0^{2\pi} \frac{1}{2}(1 - ne^{-n^2})\, d\theta = \pi(1 - ne^{-n^2})$$

Por consiguiente, para I se tiene:

$$I = \lim_{n\to\infty} I_n = \pi$$

2.° Si, para $n \in \mathbb{N}$, recurrimos al cuadrado C_n:

$$C_n = \{(x, y) \in \mathbb{R}^2 \;/\; |x| \leqslant n,\ |y| \leqslant n\}$$

podemos también poner que:

$$I = \lim_{n\to\infty} I'_n, \qquad \text{donde } I'_n = \iint_{C_n} e^{-x^2-y^2}\, dx\, dy$$

Integrando por iteración, se obtiene que

$$I = \lim_{n\to\infty} \int_{-n}^{n} e^{-x^2}\,dx\left(\int_{-n}^{n} e^{-y^2}\,dy\right) = \lim_{n\to\infty}\left(\int_{-n}^{n} e^{-t^2}\,dt\right)^2 = \left(\int_{-\infty}^{+\infty} e^{-t^2}\,dt\right)^2 = J^2$$

por lo que, como $I = \pi$, resulta que $J = \sqrt{\pi}$, esto es:

$$\int_{-\infty}^{+\infty} e^{-t^2}\,dt = \sqrt{\pi}$$

$[75]_2$ Algún criterio particular de convergencia

Si f es una función positiva, definida en los conjuntos C o D que se indican, se verifica lo que sigue (donde $\alpha \in \mathbb{R}$ y $k > 0$; el radio $r > 0$ es cualquiera):

$$1.^{\text{o}}\quad \begin{cases} f(x,y)(\sqrt{x^2+y^2})^{\alpha} \begin{cases} \leqslant k \text{ para } (x,y)\in C, \text{ con } \alpha < 2 & \Rightarrow\ \int_{\to C} f \text{ converge} \\ \geqslant k \text{ para } (x,y)\in C, \text{ con } \alpha \geqslant 2 & \Rightarrow\ \int_{\to C} f \text{ diverge} \end{cases} \\ f(x,y)(\sqrt{x^2+y^2})^{\alpha} \begin{cases} \leqslant k \text{ para } (x,y)\in D, \text{ con } \alpha > 2 & \Rightarrow\ \int_{\to C} f \text{ converge} \\ \geqslant k \text{ para } (x,y)\in D, \text{ con } \alpha \leqslant 2 & \Rightarrow\ \int_{\to C} f \text{ diverge} \end{cases} \end{cases}$$

Siendo:

$$C = \{(x,y)\in\mathbb{R}^2 \,/\, 0 < x^2+y^2 \leqslant r^2\} \qquad D = \{(x,y)\in\mathbb{R}^2 \,/\, 0 < r^2 \leqslant x^2+y^2\}^{(*)}$$

$$2.^{\text{o}}\quad \begin{cases} f(x,y,z)(\sqrt{x^2+y^2+z^2})^{\alpha} \begin{cases} \leqslant k \text{ para } (x,y,z)\in C, \text{ con } \alpha < 3 & \Rightarrow\ \int_{\to C} f \text{ converge} \\ \geqslant k \text{ para } (x,y,z)\in C, \text{ con } \alpha \geqslant 3 & \Rightarrow\ \int_{\to C} f \text{ diverge} \end{cases} \\ f(x,y,z)(\sqrt{x^2+y^2+z^2})^{\alpha} \begin{cases} \leqslant k \text{ para } (x,y,z)\in D, \text{ con } \alpha > 3 & \Rightarrow\ \int_{\to D} f \text{ converge} \\ \geqslant k \text{ para } (x,y,z)\in D, \text{ con } \alpha \leqslant 3 & \Rightarrow\ \int_{\to D} f \text{ diverge} \end{cases} \end{cases}$$

Siendo:

$$C = \{(x,y,z)\in\mathbb{R}^3 \,/\, 0 < x^2+y^2+z^2 \leqslant r^2\} \qquad D = \{(x,y,z)\in\mathbb{R}^3 \,/\, 0 < r^2 \leqslant x^2+y^2+z^2\}^{(**)}$$

Comprobación

La veracidad de los dos criterios precedentes se obtiene sin más que aplicar el «criterio de comparación» [75], II a las integrales impropias $\int_{\to C} f$ e $\int_{\to C} g$ o $\int_{\to D} f$ e $\int_{\to D} g$, tomando para funciones g las de los ejercicios 1.º y 2.º de $[74]_2$.

(*) Los conjuntos C o D pueden sustituirse por una región angular (limitada entre dos rectas que parten del origen y por los mismos radios que en C o D), de iguales características que C o D, respectivamente.

(**) Los conjuntos C o D pueden sustituirse por una región de ángulo sólido (limitado por rectas que parten del origen y por los mismos radios que en C o D), de iguales características que C o D, respectivamente.

$[75]_3$ Ejercicio

Estudiar si es o no convergente la integral impropia:

$$\iint_{\to C} \frac{e^{3x}+e^{2y}-2}{x^4+x^2y^2+y^4}\sqrt[4]{xy}\,dx\,dy \qquad \text{donde } C=\{(x,y)\in\mathbb{R}^2 / 0<x^2+y^2<1, x\geqslant 0, y\geqslant 0\}$$

Resolución

El integrando es positivo y se verifica que, para $(x, y)\in C$, es (acudiendo a coordenadas polares: $x=\rho\cos\theta$, $y=\rho\,\text{sen}\,\theta$):

$$\frac{e^{3x}+e^{2y}-2}{x^4+x^2y^2+y^4}\sqrt[4]{xy}\leqslant\frac{(e^3-1)x+(e^2-1)y}{x^4+x^2y^2+y^4}\sqrt[4]{xy}<\frac{21x+8y}{x^4+x^2y^2+y^4}\sqrt[4]{xy}=$$

$$=\frac{1}{\rho^{2,5}}\frac{(21\cos\theta+8\,\text{sen}\,\theta)\sqrt[4]{\text{sen}\,\theta\cos\theta}}{\cos^4\theta+\cos^2\theta\,\text{sen}^2\theta+\text{sen}^4\theta}<\frac{1}{\rho^{2,5}}\frac{29}{\cos^4\theta+\cos^2\theta\,\text{sen}^2\theta+\text{sen}^4\theta}=$$

$$=\frac{1}{\rho^{2,5}}\frac{29}{1-(1/4)\,\text{sen}^2 2\theta}<\frac{1}{\rho^{2,5}}\frac{29\times 4}{3}$$

De acuerdo entonces con el anterior criterio $[75]_2$, 2.°, que se verifica para $\alpha=2{,}5<3$ y $k=(29\times 4)/3$, concluimos que $\iint_{\to C} f$ es convergente.

CONVERGENCIA Y CONVERGENCIA ABSOLUTA: EQUIVALENCIA ENTRE ELLAS

Según bien sabemos, una integral múltiple impropia $\int_{\to C} f$ es (véase [73]) es el límite de integrales del tipo $\int_{M_n} f$, donde (M_n) es una «sucesión básica» que tiende monótonamente a C cuando $n\to\infty$. Nótese que se está pidiendo que dicho límite exista y sea el mismo para cualquiera que sea la tal «sucesión básica» (M_n) y ello es una gran exigencia. Tanto es así que, con la referida definición, sólo resultan ser convergentes aquellas integrales múltiples impropias que son absolutamente convergentes. Más exactamente, se verifica que:

[76]

Sean dados un conjunto $C\subset\mathbb{R}^p$ (con $p\geqslant 2$) y una función $f: C\to\mathbb{R}$ tales que existe alguna «sucesión básica» (M_n) para la integración en C de f (y, por tanto, también de $|f|$; véase la definición dada en [73]). Se verifica entonces que:

1.° Si la integral impropia $\int_{\to C}|f|$ (que existe) es convergente, entonces existe y es convergente la integral impropia $\int_{\to C} f$. Se dice que $\int_{\to C} f$ es «absolutamente convergente» si $\int_{\to C}|f|$ es convergente.

2.° Si f es integrable en C y la integral impropia $\int_{\to C} f$ es convergente, entonces la integral $\int_{\to C}|f|$ (que existe) es convergente.

Así pues, para las integrales impropias múltiples, los conceptos de convergencia y de convergencia absoluta son equivalentes.

Demostración

1.º Como f es integrable en todos los conjuntos M_n, entonces lo es $|f|$ (véase [68], 1) y, por ello, también lo son las funciones f^+ y f^- (partes positiva y negativa de f; véase [74], 2), puesto que para $x \in C$ es

$$f^+(x) = \tfrac{1}{2}\left(|f(x)| + f(x)\right) \qquad \text{y} \qquad f^-(x) = \tfrac{1}{2}\left(|f(x)| - f(x)\right)$$

De acuerdo con el criterio de comparación [75], II y por ser

$$0 \leqslant f^+(x) \leqslant |f(x)| \qquad \text{y} \qquad 0 \leqslant f^-(x) \leqslant |f(x)|, \qquad \forall\, x \in C$$

de la convergencia de $\int_{\to C} |f|$ (esta integral existe por ser su integrando positivo; véase [74], 1) se infiere la convergencia de $\int_{\to C} f^+$ y de $\int_{\to C} f^-$; por ser estas convergentes y según lo dicho en [74], 2, también lo es entonces $\int_{\to C} f$, como había que comprobar.

2.º Suponiendo ahora que $\int_{\to C} f$ es convergente, veamos que también lo es $\int_{\to C} |f|$. Para ello procederemos por reducción al absurdo: suponiendo que $\int_{\to C} f$ es convergente y que $\int_{\to C} |f|$ es divergente, vamos a obtener una contradicción.

Si $\int_{\to C} |f|$ es divergente, entonces la sucesión $(\int_{M_n} |f|)$ es monótona creciente y con límite infinito. Es entonces evidente que para cada $n \in \mathbb{N}$ existe un $p > 0$ tal que

$$\int_{M_{n+p}} |f| > 3 \int_{M_n} |f| + 2n$$

lo que permite asegurar que la «sucesión básica» (M_n) se puede elegir de manera que sea:

$$\int_{M_{n+1}} |f| > 3 \int_{M_n} |f| + 2n, \qquad \forall\, n \in \mathbb{N} \tag{1}$$

(basta con quedarse con la subsucesión de (M_n) que, empezando en M_1, resulta de elegir como siguiente de cada M_n el M_{n+p}). Si llamamos $D_n = M_{n+1} - M_n$, con lo que $\int_{M_{n+1}} |f| = \int_{M_n} |f| + \int_{D_n} |f|$, de la relación [1] se desprende que

$$\int_{D_n} |f| > 2 \int_{M_n} |f| + 2n, \qquad \forall\, n \in \mathbb{N} \tag{2}$$

Echando ahora mano de las funciones f^+ y f^- (parte positiva y parte negativa de f) y como $|f| = f^+ + f^-$, se puede poner que:

$$\int_{D_n} |f| = \int_{D_n} f^+ + \int_{D_n} f^-, \qquad \forall\, n \in \mathbb{N}$$

De las dos integrales $\int_{D_n} f^+$ e $\int_{D_n} f^-$, una de ellas será no menor que la otra para infinitos valores de n; supongamos que esto ocurre con la primera, es decir, que es $\int_{D_n} f^+ \geqslant \frac{1}{2} \int_{D_n} |f|$ para infinitos valores de $n \in \mathbb{N}$ (igual se puede razonar si lo anterior se verifica para f^-, en lugar de para f^+). Quedémonos con la subsucesión de (D_n) para la que ello se verifica y, para simplificar la notación, sigamos llamando (D_n) y (M_n) a dichas subsucesiones de (D_n) y (M_n); nótese que esta última subsucesión sigue siendo una «subsucesión básica» para la integración de f en C. Por ser, pues, $\int_{D_n} f^+ \geqslant \frac{1}{2}\int_{D_n} |f|$, de [2] se desprende que:

$$\int_{D_n} f^+ > \int_{M_n} |f| + n, \qquad \forall\, n \in \mathbb{N} \tag{3}$$

A causa de la integrabilidad de f^+ en D_n, si I es un intervalo compacto que incluye a D_n y llamando $f_n^+ : I \to \mathbb{R}$ a la función definida por $f_n^+(x) = f^+(x)$ si $x \in D_n$ y $f_n^+(x) = 0$ si $x \notin D_n$, existe una partición $P_n = \{I_1, ..., I_{m_n}\}$ de I tal que la suma inferior $s(f_n^+, P_n)$ difiere de $\int_{D_n} f^+$ en menos de lo que diferen los dos miembros de la desigualdad [3], es decir, tal que

$$s(f_n^+, P_n) > \int_{M_n} |f| + n, \qquad \text{o sea } \sum_{i=1}^{m_n} \inf\,(f_n^+(I_i))\mu(I_i) > \int_{M_n} |f| + n \tag{4}$$

para todo $n \in \mathbb{N}$. Como $\inf\,(f_n^+(I_i)) \geqslant 0$ para todo I_i, la anterior suma puede extenderse solamente a los I_i en los que $\inf\,(f_n^+(I_i)) > 0$, en cuyo caso es $\inf\,(f_n^+(I_i)) = \inf\,(f^+(I_i))$; llamemos R_n al conjunto que forman estos $I_i \in P$. Nótese que $R_n \subset D_n$ (pues si I_i no está incluido en D_n, entonces $\inf\,(f_n^+(I_i)) = 0$) y que en f es positiva en todos los puntos de R_n, con lo que $f(x) = f^+(x)$ para todo $x \in R_n$. Así pues, la relación [4] se puede poner en la forma:

$$\sum_{I_i \subset R_n} \inf\,(f(I_i))\mu(I_i) > \int_{M_n} |f| + n, \qquad \forall\, n \in \mathbb{N}$$

El primer miembro de esta desigualdad es una suma inferior de la integral $\int_{R_n} f$, por lo que es menor que ésta; consecuentemente, se puede poner que:

$$\int_{R_n} f > \int_{M_n} |f| + n, \qquad \forall\, n \in \mathbb{N} \tag{5}$$

Si tenemos en cuenta que, por ser $f \geqslant -|f|$, se verifica que $\int_{M_n} f > -\int_{M_n} |f|$, al sumar esta desigualdad con la [5] se obtiene que:

$$\int_{M_n'} f > n, \qquad \text{donde } M_n' = M_n \cup R_n \quad \text{y} \quad \forall\, n \in \mathbb{N} \tag{6}$$

Obsérvese que $R_n \cap M_n = \varnothing$, pues $R_n \subset D_n = M_{n+1} - M_n$. Es evidente que la sucesión (M_n') es una «sucesión básica» para la integración de f en C, por lo que, como $\underset{\to C}{\int} f$ es convergente, la sucesión $(\int_{M_n'} f)$ ha de ser convergente, lo que entra en contradicción con la relación [6], pues según ella dicha sucesión $(\int_{M_n'} f)$ no está acotada. Con esta contradicción concluimos la demostración.

[76]$_1$ Observación

Supongamos que se desea estudiar la convergencia y, en su caso, hallar el valor de una cierta integral impropia $\underline{\int}_C f$. Como ella es convergente si, y sólo si, lo es la integral $\underline{\int}_C |f|$ (véase [76]) y para estudiar la convergencia de ésta (que será convergente o divergente, pero no oscilante) se dispone de criterios de convergencia (los de [75] y [75]$_2$, entre otros), resulta que aplicando éstos a $\underline{\int}_C |f|$ se obtiene la convergencia o no convergencia de $\underline{\int}_C f$; es decir: 1.°, si $\underline{\int}_C |f|$ resulta ser convergente, entonces también converge $\underline{\int}_C f$, y 2.°, si $\underline{\int}_C |f|$ resulta ser divergente, entonces $\underline{\int}_C f$ es divergente u oscilante, es decir, no convergente.

En el supuesto de que, procediendo como se ha dicho, se supiera que $\underline{\int}_C f$ era convergente, para calcularla podríamos entonces acudir a cualquier «sucesión básica» (M_n) para la convergencia de f en C: la integral impropia $\underline{\int}_C f$ es el límite de $\int_{M_n} f$ cuando $n \to \infty$, el cual no depende de la «sucesión básica» (M_n) que se tome.

Ejemplo

Considérese la integral doble impropia $\underline{\int}_C f$, siendo

$$f(x, y) = \frac{3x^2 - 2y^2}{\sqrt[3]{(x^2 + y^2)^5}} \qquad \text{y} \qquad C = \{(x, y) \in \mathbb{R}^2 \;/\; 0 < x^2 + y^2 \leqslant 1\}$$

Esta integral es convergente, pues es absolutamente convergente, como se comprueba aplicando el criterio [75]$_2$, 1.° (con $\alpha = 4/3 < 2$ y $k = 3$), puesto que (llamando $\rho^2 = x^2 + y^2$):

$$|f(x, y)|\rho^{4/3} = \frac{|3x^2 - 2y^2|}{\rho^{10/3}}\,\rho^{4/3} = \frac{|3x^2 - 2y^2|}{\rho^2} < \frac{3\rho^2}{\rho^2} = 3$$

Sabiendo ya que $\underline{\int}_C f$ es convergente, para calcularla, podemos acudir a una cualquiera de las respectivas sucesiones básicas. Acudiendo a la sucesión (M_n), donde M_n es la corona circular

$$M_n = \{(x, y) \in \mathbb{R}^2 \;/\; 1/n \leqslant x^2 + y^2 \leqslant 1\}$$

y cambiando a coordenadas polares ($x = \rho\cos\theta$, $y = \rho \operatorname{sen}\theta$), se obtiene que

$$\underline{\int}_C f = \lim_{n\to\infty} \int_{M_n} f = \lim_{n\to\infty} \int_0^{2\pi} (3\cos^2\theta - 2\operatorname{sen}^2\theta)\,d\theta \int_{1/n}^1 \rho^{-1/3}\,d\rho =$$

$$= \lim_{n\to\infty} \frac{3}{2}\left(1 - \frac{1}{n^{2/3}}\right) \int_0^{2\pi} (3\cos^2\theta - 2\operatorname{sen}^2\theta)\,d\theta =$$

$$= \frac{3}{2}\left[3\left(\frac{\theta}{2} + \frac{\operatorname{sen}^2\theta}{4}\right) - 2\left(\frac{\theta}{2} - \frac{\operatorname{sen} 2\theta}{4}\right)\right]_{\theta=0}^{\theta=2\pi} = \frac{3}{2}\pi$$

PROPIEDADES DE LAS INTEGRALES MÚLTIPLES IMPROPIAS

Muchas de las propiedades de las integrales propias (u ordinarias; no impropias) se siguen verificando, debidamente adaptadas, en el caso de las integrales múltiples impropias:

[77]

PROPIEDADES BÁSICAS. Sean $f: C \to \mathbb{R}$ y $g: C \to \mathbb{R}$ dos funciones, definidas en un conjunto $C \subset \mathbb{R}^p$, tales que existe alguna «sucesión básica» (M_n) para la integración en C de f y de g. Se verifica que:

1.° Si las integrales impropias $\int_{\to C} f$ e $\int_{\to C} g$ existen y son convergentes, entonces también existe y es convergente la integral impropia $\int_{\to C} (hf + kg)$, para cualesquiera $h, k \in \mathbb{R}$, la cual vale:

$$\int_{\to C} (hf(x) + kg(x))\, dx = h \int_{\to C} f(x)\, dx + k \int_{\to C} g(x)\, dx \qquad \left(\text{LINEALIDAD DE LA INTEGRAL}\right)$$

2.° Si existen las integrales impropias $\int_{\to C} f$ e $\int_{\to C} g$, entonces es:

$$(f(x) \leqslant g(x), \quad \forall\, x \in C) \quad \Rightarrow \quad \int_{\to C} f \leqslant \int_{\to C} g \qquad \left(\text{MONOTONÍA DE LA INTEGRAL}\right)$$

3.° Si existe la integral impropia $\int_{\to C} f$, entonces (nótese que $\int_{\to C} |f|$ existe; ver [74], 1):

$$\left| \int_{\to C} f \right| \leqslant \int_{\to C} |f|$$

PROPIEDAD ADITIVA. Sea f una función real definida en la unión de dos conjuntos $C \subset \mathbb{R}^p$ y $D \subset \mathbb{R}^p$ tales que existen algunas «sucesiones básicas» (M_n) y (N_n) para la integración de f en C y en D, respectivamente. Se verifica que:

4.° Si $D \subset C$ y si la integral impropia $\int_{\to C} f$ existe y es convergente, entonces también existe y es convergente la integral impropia $\int_{\to D} f$.

5.° Si existen y son convergentes las integrales impropias $\int_{\to C} f$ e $\int_{\to D} f$ y si $(M_n \cup N_n)$ es una «sucesión básica» para la integración de f en $C \cup D$, entonces también existen y son convergentes las integrales impropias $\int_{\to C \cup D} f$ e $\int_{\to C \cap D} f$ y es:

$$\int_{\to C \cup D} f = \int_{\to C} f + \int_{\to D} f - \int_{\to C \cap D} f \qquad \left(\text{ADITIVIDAD DE LA INTEGRAL}\right)$$

Demostración

1.° De acuerdo con la «linealidad» de las integrales propias (véase [68], 1), sabemos que se verifica (para todo $n \in \mathbb{N}$) que, para cualquiera que sea la «sucesión básica» (M_n), es:

$$\int_{M_n} (hf + kg) = h \int_{M_n} f + k \int_{M_n} g$$

Tomando límites para $n \to \infty$, como el segundo miembro tine límite $h \int_{\to C} f + k \int_{\to C} g$ (finito) y éste no depende de la sucesión (M_n), el primer miembro también lo tiene y de ello resulta ya obvia la igualdad a demostrar.

2.° De acuerdo con la «monotonía» de las integrales propias (véase [68], 2), sabemos que se verifica (para todo $n \in \mathbb{N}$) que, para cualquiera que sea la «sucesión básica» (M_n), es:

$$\int_{M_n} f \leqslant \int_{M_n} g$$

Tomando límites para $n \to \infty$, se concluye que $\int_{\to C} f \leqslant \int_{\to C} g$.

3.° Como $-|f| \leqslant f \leqslant |f|$ en C, de la propiedad anterior se desprende que

$$-\int_{\to C} |f| \leqslant \int_{\to C} f \leqslant \int_{\to C} |f|, \qquad \text{o sea} \quad \left|\int_{\to C} f\right| \leqslant \int_{\to C} |f|$$

4.° Para cualquiera que sea la «sucesión básica» (M_n) y para cada conjunto N_i (para $i \in \mathbb{N}$), como éste es compacto y medible y está incluido en C, existe $n_i \in \mathbb{N}$ tal que $N_i \subset M_{n_i}$. De ello se deduce que existe una subsucesión de (M_n) que es «sucesión básica» para la integración de f en C y tal que cada uno de los conjuntos que la forman incluye al conjunto de (N_n) de igual índice. Llamemos (M'_n) a la referida subsucesión. Como $N_n \subset M'_n$ (para $n \in \mathbb{N}$) y dado que $|f(\boldsymbol{x})| \geqslant 0$ para todo $\boldsymbol{x} \in C$, de las propiedades de las integrales propias (véase [68] y [69]) se desprende que:

$$\int_{N_n} |f| \leqslant \int_{M_n} |f|, \qquad \forall\, n \in \mathbb{N}$$

Tomando límites para $n \to \infty$ (nótese que los dos miembros de la anterior desigualdad tienen límite), se obtiene que $\int_{\to D} |f| \leqslant \int_{\to C} |f|$. Como $\int_{\to C} f$ es convergente, también lo es $\int_{\to C} |f|$ y, de acuerdo con la anterior desigualdad, de ello se deduce que $\int_{\to D} |f|$ converge y, por tanto, resulta que $\int_{\to D} f$ es convergente, como había que comprobar.

5.° Supongamos primero que f es positivo, esto es, que $f(\boldsymbol{x}) \geqslant 0$ para todo $\boldsymbol{x} \in C \cup D$. Según sabemos (véase [75], I), para este caso de f positiva bastará con cerciorarse de que la sucesión $(\int_{P_n} f)$ tiene límite, donde (P_n) es una «sucesión básica» para la integración de f en $C \cup D$, y comprobar que, para esa determinada sucesión básica, el límite de $\int_{P_n} f$, cuando $n \to \infty$, es igual al segundo miembro de la última igualdad (la de 5.°) del enunciado. Tomemos $(P_n) = (M_n \cup N_n)$, que sabemos que es una «sucesión básica» para la integración de f en $C \cup D$; nótese que, como se comprueba trivialmente, $(M_n \cap N_n)$ es una «sucesión básica» para la integración de f en $C \cap D$.

De acuerdo con la «aditividad» de las integrales propias (véase [69], 2), se verifica que, para $n \in \mathbb{N}$, es:

$$\int_{M_n \cup N_n} f = \int_{M_n} f + \int_{N_n} f - \int_{M_n \cap N_n} f$$

Como f es integrable en C y en D y, según la propiedad anterior (la 4.ª), también lo es en $C \cap D$ (pues éste es un subconjunto de C y de D), ocurre que los tres sumandos del anterior segundo miembro tienen límite cuando $n \to \infty$; tomando límites se obtiene, pues, que:

$$\lim_{n\to\infty} \int_{M_n \cup N_n} f = \int_{\to C} f + \int_{\to D} f - \int_{\to C\cap D} f$$

como había que comprobar.

En general, si f no es positiva en $C \cup D$, vamos a recurrir a las funciones $f^+ = \text{máx}\,\{f, 0\}$ y $f^- = \text{máx}\,\{-f, 0\}$ (parte positiva y parte negativa de f). Como existen y son convergentes las integrales de f en C, en D y en $C \cap D$, también lo son (véase [74], 2) las de f^+ y f^-; como f^+ y f^- son positivas en $C \cup D$, para ellas se verifica la propiedad a demostrar. De acuerdo con lo que se acaba de decir y echando mano de la linealidad de las integrales impropias (propiedad 1.ª), como $f = f^+ - f^-$, se puede poner:

$$\int_{\to C\cup D} f = \int_{\to C\cup D} f^+ - \int_{\to C\cup D} f^- = \left(\int_{\to C} f^+ + \int_{\to D} f^+ - \int_{\to C\cap D} f^+\right) -$$

$$-\left(\int_{\to C} f^- + \int_{\to D} f^- - \int_{\to C\cap D} f^-\right) = \left(\int_{\to C} f^+ - \int_{\to C} f^-\right) + \left(\int_{\to D} f^+ - \int_{\to D} f^-\right) -$$

$$-\left(\int_{\to C\cap D} f^+ - \int_{\to C\cap D} f^-\right) = \int_{\to C} f + \int_{\to D} f - \int_{\to C\cap D} f$$

luego $\int_{\to C\cup D} f$ es convergente y vale lo que en el enunciado se dice que vale.

[77]$_1$ Observación

De acuerdo con nuestra definición de integral impropia (véase [73]), para poder hablar de una tal integral $\int_{\to C} f$ (donde f es una función real definida en el conjunto $C \subset \mathbb{R}^p$) es necesario que exista, al menos, una «sucesión básica (M_n) para la integración de f en C» (recuérdese que hemos llamado así a una sucesión creciente (M_n) de subconjuntos compactos y medibles de C en los que f es propiamente integrable y tal que, para todo subconjunto compacto y medible M de C, se verifica que $M \subset M_n$ para algún $n \in \mathbb{N}$).

Así pues, la ausencia de una tal sucesión nos impide, al menos de momento, construir o considerar la referida integral. Es por ello que en todo cuanto venimos diciendo, se empieza por garantizar la existencia de la correspondiente «sucesión básica». En ocasiones, este requisito es excesivo, sobre todo en lo tocante a la anterior condición $M \subset M_n$

(esto es, aquella que dice: cualquier subconjunto compacto y medible de C es recubierto por alguno de los conjuntos de la «sucesión básica»); para algunas cuestiones, este último requisito puede sustituirse por este otro, menos exigente: $C = \bigcup_{n=1}^{\infty} M_n$.

Para dar cabida a ciertas integrales impropias de interés que, con nuestra anterior definición (véase [73]), no tendrían sentido, vamos a dar una generalización de dicho concepto, para lo que acudiremos a la aditividad:

AMPLIACIÓN DEL CONCEPTO DE INTEGRAL IMPROPIA. Sea dada una función $f: C \to \mathbb{R}$, definida en un conjunto $C \subset \mathbb{R}^p$. Si $\{C_1, ..., C_k\}$ es una partición de C (esto es, si: 1.º, $C = \bigcup_{i=1}^{i=k} C_i$, y 2.º, $C_i \cap C_j = \emptyset$, para $i \neq j$) y si (ateniéndose a la definición [73]) existen y son convergentes las integrales impropias de f en C_1, ..., en C_k, diremos entonces que existe y es convergente la integral impropia $\int_{\to C} f$, de f en C, y que ella es:

$$\int_{\to C} f = \sum_{i=1}^{k} \int_{\to C_i} f$$

[77]$_2$ Ejemplo

Considérese la siguiente función $f: C \to \mathbb{R}$, donde e es la función «parte entera»:

$$f(x, y) = y \, \mathrm{L}\, [x^2 + e(x + 1)], \qquad \text{siendo } C = [-1, 1] \times [0, 1]$$

Nótese que, para cualquiera que sea $a \in]0, 1[$,

$$\lim_{\substack{x \to 0^+ \\ y \to a}} f(x, y) = 0 \qquad \lim_{\substack{x \to 0^- \\ y \to a}} f(x, y) = -\infty$$

por lo que no hay ninguna «sucesión básica» para la integración de f en C. En efecto, si hubiera una tal sucesión (M_n), entonces, para todo conjunto $M \subset C$ compacto y medible, habría de ser $M \subset M_n$ para algún $n \in \mathbb{N}$, lo que no se cumple en el caso de ser, por ejemplo:

$$M = [-1/2, 1/2] \times [1/3, 2/3]$$

ya que f no está acotada en M, por lo que no lo estaría en el tal M_n y, en consecuencia, f no ería propiamente integrable en C.

No obstante, existen las integrales de f en los conjuntos

$$C_1 = [-1, 0[\times [0, 1] \qquad \text{y} \qquad C_2 = [0, 1] \times [0, 1]$$

(la primera integral es impropia; la segunda es propia), las cuales valen:

$$\int_{\to C_1} f = \lim_{\varepsilon \to 0^-} \int_{-1}^{\varepsilon} (\mathrm{L}\, x^2)\, dx \int_0^1 y\, dy = \frac{1}{2} [x \,\mathrm{L}\, x^2 - 2x]_{x=-1}^{x \to 0^-} = -1$$

$$\int_{C_2} f = \int_0^1 \mathrm{L}\, (x^2 + 1)\, dx \int_0^1 y\, dy = \frac{1}{2} [x \,\mathrm{L}\, (x^2 + 1) - 2x + 2 \operatorname{arctg} x]_0^1 = \frac{1}{2} \mathrm{L}\, 2 + \frac{\pi}{4} - 1$$

De acuerdo, pues, con la anterior ampliación del concepto de integral impropia, convendremos en que:

$$\int_{\to C} f = \int_{\to C_1} f + \int_{C_2} f = \frac{1}{2}\,\mathrm{L}\,2 + \frac{\pi}{4} - 2 = 0{,}17169\ldots$$

ALGO SOBRE LA INTEGRACIÓN (IMPROPIA) POR ITERACIÓN Y POR CAMBIO DE VARIABLE

No es tarea fácil el comprobar la veracidad de las proposiciones que enunciaremos a continuación, acerca de la integración iterada y del cambio de variable para integrales múltiples impropias. Dichas demostraciones, además de ser notablemente intrincadas, precisan de algunos resultados que no se han estudiado en nuestra anterior exposición. Por ello, y como estas dos técnicas de integración son cuestiones terminales, se ha optado por prescindir de sus demostraciones.

[78]

INTEGRACIÓN ITERADA. Sea $f: C \to \mathbb{R}$ una función, definida en un conjunto $C \subset \mathbb{R}^p$. Sea $C' \subset \mathbb{R}^r$ la proyección de C sobre el subespacio de las r primeras coordenadas de $\mathbb{R}^p$; para cada $x' \in C'$, sea $C''(x') = \{x'' \in \mathbb{R}^{p-s} \,/\, x = (x', x'') \in C\}$. Si existen las pertinentes «sucesiones básicas» para las siguientes integrales impropias(*) y estas integrales existen y son convergentes, entonces se verifica la siguiente igualdad:

$$\int_{\to C} f(x)\,dx = \int_{\to C'} \left(\int_{\to C''(x')} f(x, x'')\,dx'' \right) dx'$$

INTEGRACIÓN POR CAMBIO DE VARIABLE. Sean $A, B \subset \mathbb{R}^p$ dos conjuntos abiertos entre los que hay definida una aplicación $\varphi: A \to B$ que es biyectiva, de clase $\mathscr{C}^1$ y tal que su jacobiano no se anula (o sea, es $\det J\varphi(x) \neq 0,\ \forall\ x \in A$). Sea $f: B \to \mathbb{R}$ una función continua y tal que existe y es convergente la integral impropia $\int_{\to B} f$. Si existe una sucesión (M_n) de subconjuntos compactos y medibles de A tal que todo subconjunto compacto y medible de A es recubierto por alguno de los conjuntos de (M_n), entonces existe la otra integral que abajo se da y es:

$$\int_{\to B} f(y)\,dy = \int_{\to A} f(\varphi(x))|\det J\varphi(x)|\,dx$$

(*) Existe alguna «sucesión básica» (M_n) para la integración de f en C; para cada $x' \in C$, la $(M_n''(x'))$ es una «sucesión básica» para la integración de $x' \mapsto f(x', x'')$ en $C''(x')$; la (M_n') es una «sucesión básica» para la integración en C' de $x' \mapsto \int_{\to C''(x')} f(x', x'')\,dx''$.

[78]$_1$ Ejercicio

Calcular por iteración la siguiente integral doble impropia

$$I = \iint_{C} \frac{x^3 e^{-x} y}{(x^2 + y^2)^2}\, dx\, dy, \qquad \text{donde } C = \{(x, y) \in \mathbb{R}^2 \;/\; x > 0,\, y > x\}$$

Resolución

$$I = \int_0^{+\infty} x^3 e^{-x}\, dx \int_x^{+\infty} \frac{y\, dy}{(x^2 + y^2)^2} = \int_0^{+\infty} x^3 e^{-x}\, dx \left[\frac{-1}{2(x^2 + y^2)}\right]_{y=x}^{y \to +\infty} =$$

$$= \int_0^{+\infty} \frac{1}{4}\, x e^{-x}\, dx = \left[-\frac{1}{4}\, e^{-x}(x + 1)\right]_0^{+\infty} = \frac{1}{16}$$

[78]$_2$ Ejercicio

Calcular la siguiente integral doble impropia acudiendo a un cambio de variable:

$$I = \iint_{C} \frac{dx\, dy}{\sqrt[4]{(4x^2 + y^2)^3}}, \qquad \text{donde } C = \{(x, y) \in \mathbb{R}^2 \;/\; x^2 + (y/2)^2 \leqslant 1,\, (x, y) \neq (0, 0)\}$$

Resolución

Hagamos el cambio de variable $x = \rho \cos \theta$, $y = 2\rho \operatorname{sen} \theta$, cuyo determinante jacobiano es $2|\rho|$:

$$I = \iint_{D} \frac{d\theta\, d\rho}{\sqrt{2\rho}}, \qquad \text{donde } D = \{(\rho, \theta) \in \mathbb{R}^2 \;/\; 0 < \rho \leqslant 1\}$$

integrando por iteración, se obtiene que

$$I = \int_0^1 \frac{d\rho}{\sqrt{2\rho}} \int_0^{2\pi} d\theta = \sqrt{2}\pi \int_0^1 \frac{d\rho}{\sqrt{\rho}} = \sqrt{2}\pi \left[2\sqrt{\rho}\right]_0^1 = 2\sqrt{2}\pi$$

[78]$_3$ Observación

Respecto de la fórmula de integración iterada (véase [78]), conviene advertir que, si el integrando es positivo, entonces la existencia de la integral iterada (segundo miembro de la fórmula) garantiza la existencia de la integral múltiple (primer miembro); no obstante esto puede no ser cierto si el integrando toma valores de distinto signo. Así ocurre en el siguiente caso.

Considérese la integral $\iint_C f$, en la que f y C son:

$$f(x, y) = \frac{\operatorname{sen} x}{y} \qquad \text{y} \qquad C = [-\pi/4, \pi/4] \times]0, 1]$$

Una de las integrales iteradas de esta $\iint_C f$ es:

$$I_1 = \int_0^1 \frac{dy}{y} \int_{-\pi/4}^{\pi/4} \operatorname{sen} x \, dx = \int_0^1 \frac{0}{y} \, dy = 0$$

Pero la otra integral iterada no existe, como es obvio, y en consecuencia no existe tampoco la $\iint_C f$. Nótese que llamando

$$C_1 = [-\pi/4, 0] \times]0, 1] \qquad \text{y} \qquad C_2 = [0, \pi/4] \times]0, 1]$$

las integrales de f en C_1 y C_2 valen $-\infty$ y $+\infty$, lo que confirma la no existencia de la integral de partida.

5.2. INTEGRALES PARAMÉTRICAS

Muchas de las funciones que se manejan en el Cálculo Infinitesimal no se conocen mediante expresiones elementales, sino que vienen dadas de otros modos que, al principiante, pueden parecerle extraños, como puede ser a través de series o de integrales.

Nos ocupamos aquí de un tipo especial de funciones, que se llaman «funciones definidas por integrales» o también «integrales dependientes de parámetros». Supongamos, para fijar las ideas, que hablamos de una expresión del tipo $\int_a^b f(\lambda, x)\, dx$, donde se supone que, para cada valor de λ de un cierto conjunto Λ, la función real $x \mapsto f(\lambda, x)$ es integrable (respecto de la variable x) en un cierto intervalo $[a, b]$. La anterior integral define, pues, a una función $F: \Lambda \to \mathbb{R}$, que viene dada por $F(\lambda) = \int_a^b f(\lambda, x)\, dx$. Generalmente no es posible obtener una expresión explícita de $F(\lambda)$ (es decir, libre de las integrales), por no conocerse una primitiva (para cada $\lambda \in \Lambda$) de la función $x \mapsto f(\lambda, x)$, por lo que no se puede recurrir a la regla de Barrow. En tal supuesto, que es el que nos preocupa, no es difícil, como luego veremos, estudiar una tal función F, analizando su regularidad y deduciendo sus propiedades, y ello a partir de la regularidad y propiedades de la función, f, del integrando.

❏ INTEGRALES DEPENDIENTES DE PARÁMETROS

Se consideran los casos de integral simple, primero, y de integral múltiple, después. Obsérvese que las integrales paramétricas simples tienen por integrando a una función de, al menos, dos variables y que, por ello, su estudio no puede realizarse, debidamente, en los cursos de Cálculo Infinitesimal de una variable; es en un curso como éste, de varias variables, donde pueden ser abordadas.

[79]

INTEGRALES PARAMÉTRICAS SIMPLES

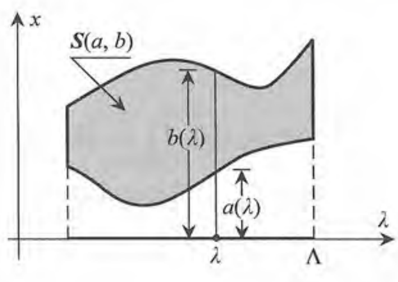
- *Caso de límites de integración fijos.* Sean dados un conjunto $\Lambda \subset \mathbb{R}^p$, un intervalo compacto real $I = [a, b]$ y una función $f: \Lambda \times I \to \mathbb{R}$, $(\lambda, x) \mapsto f(\lambda, x)$. Si, para cada $\lambda \in \Lambda$, la función $x \mapsto f(\lambda, x)$ es integrable en I, entonces se dice que $\int_a^b f(\lambda, x)\,dx$ es una integral paramétrica (con parámetro $\lambda \in \Lambda$), la cual determina a la función $\Lambda \to \mathbb{R}$, $\lambda \mapsto \int_a^b f(\lambda, x)\,dx$ (función definida por una integral).
- *Caso de límites de integración variables.* Sea dado un conjunto $\Lambda \subset \mathbb{R}^p$, sea $S(a, b)$ el «conjunto simple»(*) determinado por dos funciones acotadas $a: \Lambda \to \mathbb{R}$ y $b: \Lambda \to \mathbb{R}$ y considérese una función $f: S(a, b) \to \mathbb{R}$, $(\lambda, x) \mapsto f(\lambda, x)$. Si, para cada $\lambda \in \Lambda$, la función $x \mapsto f(\lambda, x)$ es integrable en $[a(\lambda), b(\lambda)]$, entonces se dice que $\int_{a(\lambda)}^{b(\lambda)} f(\lambda, x)\,dx$ es una integral paramétrica (con parámetro $\lambda \in \Lambda$), la cual determina a la función $\Lambda \to \mathbb{R}$, $\lambda \mapsto \int_{a(\lambda)}^{b(\lambda)} f(\lambda, x)\,dx$ (función definida por una integral).

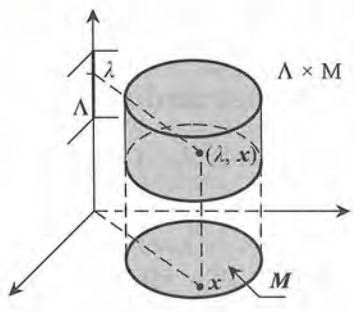
INTEGRALES PARAMÉTRICAS MÚLTIPLES

Sean dadas un conjunto $\Lambda \subset \mathbb{R}^p$, un conjunto medible $M \subset \mathbb{R}^q$ y una función $f: \Lambda \times M \to \mathbb{R}$, $(\lambda, \mathbf{x}) \mapsto f(\lambda, \mathbf{x})$. Si para cada $\lambda \in \Lambda$, la función $\mathbf{x} \mapsto f(\lambda, \mathbf{x})$ es integrable en M, entonces se dice que $\int_M f(\lambda, \mathbf{x})\,d\mathbf{x}$ es una integral paramétrica (con parámetro $\lambda \in \Lambda$), la cual determina a la función $\Lambda \to \mathbb{R}$, $\lambda \mapsto \int_M f(\lambda, \mathbf{x})\,d\mathbf{x}$ (función definida por una integral)(**).

(*) Según se dijo en [66], es $S(a, b) = \{(\lambda, x) \in \mathbb{R}^{p+1} \,/\, \lambda \in \Lambda,\ a(\lambda) \leqslant x \leqslant b(\lambda)\}$, donde se supone que $a(\lambda) \leqslant b(\lambda)$ para todo $\lambda \in \Lambda$.

(**) No se consideran las integrales del tipo $\int_{M(\lambda)} f(\lambda \mathbf{x})\,d\mathbf{x}$, en las que el campo de integración M depende del parámetro λ. Son de escasa utilidad y demasiado complejas.





[79]$_1$ Ejemplo

Considérese la integral doble dependiente de un parámetro λ, con $\lambda \in\,]-\pi/2, \pi/2[$, dada por la expresión:

$$F(\lambda) = \iint_M \operatorname{tg}(\lambda x)[1 + \operatorname{tg}^2(\lambda y)]\,dx\,dy, \qquad \text{donde } M = \{(x, y) \in \mathbb{R}^2 \,/\, 0 \leqslant x \leqslant 1,\ 0 \leqslant y \leqslant x\}$$

Es evidente que $F(0) = 0$ y que para $\lambda \neq 0$ es

$$F(\lambda) = \int_0^1 \operatorname{tg}(\lambda x)\,dx \int_0^x [1 + \operatorname{tg}^2(\lambda x)]\,dy = \int_0^1 \frac{1}{\lambda}\operatorname{tg}^2(\lambda x)\,dx = \frac{\operatorname{tg}\lambda - \lambda}{\lambda^2}$$

Obsérvese que el integrando, esto es, la función

$$]-\pi/2, \pi/2[\times M \to \mathbb{R}, \qquad (\lambda, xy) \mapsto \operatorname{tg}(\lambda x)[1 + \operatorname{tg}^2(\lambda y)]$$

es de clase $\mathscr{C}^\infty$ en su campo de definición y que la integral paramétrica $\lambda \mapsto F(\lambda)$ también es de clase $\mathscr{C}^\infty$ en su intervalo de definición $]-\pi/2, \pi/2[$ (nótese que $F(\lambda)$ es desarrollable en serie de MacLaurin en dicho intervalo, $F(\lambda) = x/3 + 2x^3/15 + 17x^5/315 + \cdots$). Como luego veremos, este resultado se verifica con carácter general; es decir, si $(\lambda, \boldsymbol{x}) \mapsto f(\lambda, \boldsymbol{x})$ es una función de clase $\mathscr{C}^r$ en $\Lambda \times M$, entonces la función $\lambda \mapsto F(\lambda) = \int_M f(\lambda, \boldsymbol{x})\,d\boldsymbol{x}$ es, al menos, de clase $\mathscr{C}^r$ en Λ.

[79]$_2$ Reducción al caso de límites de integración fijos

Considérese el conjunto simple $S(a, b)$, determinado por dos funciones acotadas $a : \Lambda \to \mathbb{R}$ y $b : \Lambda \to \mathbb{R}$ (definidas en un cierto conjunto $\Lambda \subset \mathbb{R}^p$), y sea $f : S(a, b) \to \mathbb{R}$ una función tal que, para cualquier $\lambda \in \Lambda$, la función $x \mapsto f(\lambda, x)$ es continua en el intervalo $[a(\lambda), b(\lambda)]$. Se verifica entonces que la integral paramétrica $\int_{a(\lambda)}^{b(\lambda)} f(\lambda, x)\,dx$ se puede expresar en forma de integral paramétrica con límites de integración fijos; en concreto, es:

$$\int_{a(\lambda)}^{b(\lambda)} f(\lambda, x)\,dx = \int_0^1 g(\lambda, t)\,dt, \qquad \text{donde } g(\lambda, x) = f\big(\lambda, a(\lambda) + [b(\lambda) - a(\lambda)]t\big)[b(\lambda) - a(\lambda)]$$

Nótese que si f es de clase $\mathscr{C}^r$ en $S(a, b)$ y a y b son de clase $\mathscr{C}^r$ en Λ, para cualquier $r \geqslant 0$, entonces g es de la misma clase $\mathscr{C}^r$ en $\Lambda \times [0, 1]$.

Demostración

Para cada $\lambda \in \Lambda$, considérese la función $\varphi_\lambda : [0, 1] \to \mathbb{R}$ definida mediante la expresión

$$\varphi_\lambda(t) = a(\lambda) + [b(\lambda) - a(\lambda)]t$$

que es una biyección de clase $\mathscr{C}^\infty$, de $[0, 1]$ en $[a(\lambda), b(\lambda)]$, tal que $\varphi_\lambda(0) = a(\lambda)$ y $\varphi_\lambda(1) = b(\lambda)$. Así pues, esta φ_λ puede utilizarse como función de cambio de variable para aplicar la regla de integración por sustitución(*) a la integral $\int_{a(\lambda)}^{b(\lambda)} f(\lambda, x)\,dx$, para cada $\lambda \in \Lambda$. Al aplicar esta regla se obtiene, obviamente, que dicha integral es igual a la $\int_0^1 g(\lambda, t)\,dt$, donde g es la función que se cita en el enunciado, como había que comprobar.

(*) Si $f : [a, b] \to \mathbb{R}$ es una función continua en el intervalo compacto $[a, b]$ y si $\varphi : [\alpha, \beta] \to \mathbb{R}$ es una función de clase $\mathscr{C}^1$ en $[\alpha, \beta]$ y tal que $\varphi(\alpha) = a$, $\varphi(\beta) = b$ y $\varphi([\alpha, \beta]) = [a, b]$, entonces:

$$\int_a^b f(x)\,dx = \int_\alpha^\beta f(\varphi(t))\varphi'(t)\,dt$$

CONTINUIDAD DE LAS INTEGRALES PARAMÉTRICAS

Si, en una integral paramétrica, el integrando es continuo (como función de las variables de integración y de los parámetros) y sus conjuntos de definición (campo de integración y dominio de parámetros) son compactos, entonces la función definida por la integral es continua y su continuidad es uniforme:

[80]

I. CONTINUIDAD DE LAS INTEGRALES PARAMÉTRICAS SIMPLES

a) *Caso de límites de integración fijos.* Sea $\Lambda \subset \mathbb{R}^p$ un conjunto compacto; sea $I = [a, b]$ un intervalo compacto real; sea $f: \Lambda \times I \to \mathbb{R}$, $(\lambda, x) \mapsto f(\lambda, x)$, una función continua en $\Lambda \times I$. Se verifica entonces que la función $\lambda \mapsto \int_a^b f(\lambda, x)\, dx$ (que existe para $\lambda \in \Lambda$) es uniformemente continua en Λ.

b) *Caso de límites de integración variables.* Sea $\Lambda \subset \mathbb{R}^p$ un conjunto compacto; sea $S(a, b)$ el «conjunto simple»(*) definido por dos funciones continuas $a: \Lambda \to \mathbb{R}$ y $b: \Lambda \to \mathbb{R}$; sea $f: S(a, b) \to \mathbb{R}$ una función continua en $S(a, b)$. Se verifica entonces que la función $\lambda \mapsto \int_{a(\lambda)}^{b(\lambda)} f(\lambda, x)\, dx$ (que existe para $\lambda \in \Lambda$) es uniformemente continua en Λ.

II. CONTINUIDAD DE LAS INTEGRALES PARAMÉTRICAS MÚLTIPLES

Sea $\Lambda \subset \mathbb{R}^p$ un conjunto compacto; sea $M \subset \mathbb{R}^p$ un conjunto compacto y medible; sea $f: \Lambda \times M \to \mathbb{R}$, $(\lambda, \mathbf{x}) \mapsto f(\lambda, \mathbf{x})$, una función continua en $\Lambda \times M$. Se verifica entonces que la función $\lambda \mapsto \int_M f(\lambda, \mathbf{x})\, d\mathbf{x}$ (que existe para $\lambda \in \Lambda$) es uniformenete continua en Λ.

(*) $S(a, b) = \{(\lambda, x) \in \mathbb{R}^{p+1} \,/\, \lambda \in \Lambda, a(\lambda) \leqslant x \leqslant b(\lambda)\}$, donde se supone que $a(\lambda) \leqslant b(\lambda)$ para todo $\lambda \in \Lambda$.

Demostración

La existencia de la integral paramétrica, para todo $\lambda \in \Lambda$, es obvia en los tres casos que se consideran en el enunciado. Para demostrar la continuidad, empezaremos reduciendo dichos tres casos a uno solo de ellos (el último), que será, pues, el único que probaremos.

1.° Consideremos el caso I.*b*), de integral simple con límites de integración variables. Sabemos que es $\int_{a(\lambda)}^{b(\lambda)} f(\lambda, x)\, dx = \int_0^1 g(\lambda, x)\, dx$, donde $g: \Lambda \times [0, 1]$ es la función que se da en $[79]_2$, con lo que el problema se reduce a probar la continuidad de la función $\lambda \mapsto \int_0^1 g(\lambda, x)\, dx$ para $\lambda \in \Lambda$. Como g es continua en $C \times [0, 1]$, ya que es composición de funciones continuas, este caso ha quedado, pues, reducido al I.*a*) de integral simple con límites de integración fijos.

2.° Es obvio que la demostración del último caso del enunciado (el de integral múltiple) implicaría, en particular, la veracidad del I.*a*) (integral simple con límites fijos), que es caso particular de aquel. Nos debemos ocupar, pues, solamente de la continuidad a la integral paramétrica del apartado II.

3.° (Caso II(*): continuidad uniforme en Λ de la función $F: C \to \mathbb{R}$ dada por $F(\lambda) = \int_M f(\lambda, x)\, dx$). Dado que C es compacto, sólo hay que probar la continuidad de F en cualquier $\lambda_0 \in \Lambda$. Como $\Lambda \times M$ es compacto (pues lo son Λ y M), la continuidad de f en $\Lambda \times M$ es uniforme y por ello, dado $\varepsilon > 0$ (cualquiera), existe $\delta > 0$ tal que

$$\|\lambda - \lambda_0\| < \delta \quad \Rightarrow \quad \left[|f(\lambda, x) - f(\lambda_0, x)| < \frac{\varepsilon}{\mu(M)}, \quad \forall\, x \in M \right]$$

donde $\mu(M)$ es la medida de M; por ello, del teorema de la media (véase [68], 3) se deduce que

$$\|\lambda - \lambda_0\| < \delta \quad \Rightarrow \quad |F(\lambda) - F(\lambda_0)| \leqslant \int_M |f(\lambda, x) - f(\lambda_0, x)|\, dx < \frac{\varepsilon}{\mu(M)}\, \mu(M) = \varepsilon$$

lo que prueba la continuidad de F en λ_0.

[80]$_1$ Observaciones

1. Bajo las hipótesis del anterior teorema [80], de continuidad de las integrales paramétricas, se puede asegurar que, para cualquiera que sea $\lambda \in \Lambda$, existen los siguientes límites y se verifica:

$$\lim_{\lambda \to \lambda_0} \int_a^b f(\lambda, x)\, dx = \int_a^b \lim_{\lambda \to \lambda_0} f(\lambda, x)\, dx \qquad \text{(para el caso I}a\text{)}$$

$$\lim_{\lambda \to \lambda_0} \int_{a(\lambda)}^{b(\lambda)} f(\lambda, x)\, dx = \int_{\lim_{\lambda \to \lambda_0} a(\lambda)}^{\lim_{\lambda \to \lambda_0} b(\lambda)} \lim_{\lambda \to \lambda_0} f(\lambda, x)\, dx \qquad \text{(para el caso I}b\text{)}$$

$$\lim_{\lambda \to \lambda_0} \int_M f(\lambda, x)\, dx = \int_M \lim_{\lambda \to \lambda_0} f(\lambda, x)\, dx \qquad \text{(para el caso II)}$$

2. Si, en las hipótesis de los anteriores teoremas [80], en lugar de suponer que M es compacto, se supone que M es un entorno de cierto $\lambda_0 \in \mathbb{R}^p$, entonces se concluye la continuidad de la integral paramétrica en un entorno de λ_0. Para cerciorarse de ello basta con sustituir, en la demostración, M por un entorno compacto de λ_0 incluido en M.
3. En estos teoremas [80], para garantizar la continuidad (respecto del parámetro) de la integral paramétrica, se es muy exigente; la continuidad del integrando, que allí se pide, no es condición necesaria para ello. Así, por ejemplo, si sg y d son las funciones «signo» y de Dirichlet, es decir, si

$$\begin{cases} \operatorname{sg}(t) = 1,\ 0 \ \text{o} \ -1 & \text{según que sea } t > 0,\ t = 0 \text{ o } t < 0 \\ \operatorname{d}(\lambda) = 1 \text{ o } 0 & \text{según que sea } \lambda \in \mathbb{Q} \text{ o } \lambda \in \mathbb{R} - \mathbb{Q} \end{cases}$$

es evidente que se verifica que

$$\int_{-1}^{1} \operatorname{sg}[\operatorname{d}(\lambda) x]\, dx = 0, \qquad \forall\, \lambda \in \mathbb{R}$$

(*) El caso I.a) se prueba como éste: póngase $[a, b]$ en lugar de M y, en particular, $b - a$ en lugar de $\mu(M)$.

donde el integrando (función definida en $\mathbb{R} \times [0, 1]$) es discontinuidad en todos los puntos de su campo de definición y, sin embargo, la integral (función, de λ, definida en $\mathbb{R}$) es continua en todo $\mathbb{R}$.

4. Abundando en lo dicho en la anterior observación, es interesante saber que en los teoremas [80], sobre la continuidad de las integrales paramétricas, se pueden reducir las exigencias que se imponen al integrando. Así, el teorema [80], II puede «debilitarse» pidiendo que f, en lugar de ser continua en $\Lambda \times M$, sea tal que las funciones $f_\lambda : M \to \mathbb{R}$, $x \mapsto f_\lambda(\boldsymbol{x}) = f(\lambda, \boldsymbol{x})$, sean continuas en M, para todo $\boldsymbol{\lambda} \in \Lambda$, y converjan uniformemente(*) en M (hacia una cierta función) cuando $\lambda \to \lambda_0$; en este supuesto, el teorema asegura que:

$$\lim_{\lambda \to \lambda_0} \int_M f(\lambda, \boldsymbol{x})\, d\boldsymbol{x} = \int_M \lim_{\lambda \to \lambda_0} f(\lambda, \boldsymbol{x})\, d\boldsymbol{x}$$

$[80]_2$ Ejercicio

Sea $f: [\alpha, \beta] \times [a, b]$ una función continua en el intervalo compacto $I = [\alpha, \beta] \times [a, b]$ de $\mathbb{R}^2$. Pruébese que la siguiente función F (que evidentemente existe) es continua en I:

$$F : I \to \mathbb{R}, \qquad (u, y) \mapsto F(u, y) = \int_\alpha^u f(x, y)\, dx$$

Resolución

Para cualesquiera que sean los puntos (u, y) y $(u + \Delta u, y + \Delta y)$ de I, se tiene:

$$F(u + \Delta u, y + \Delta y) - F(u, y) = \big(F(u + \Delta u, y + \Delta y) - F(u, y + \Delta y)\big) + \big(F(u, y + \Delta y) - F(u, y)\big) =$$
$$= \int_u^{u+\Delta u} f(x, y + \Delta y)\, dx + \int_\alpha^u \big(f(x, y + \Delta y) - f(x, y)\big)\, dx$$

Como f es continua e I es compacto, sabemos que f está acotada en I; sea $K > 0$ una cota suya. Como f es continua en (u, y), dado $\varepsilon > 0$ (cualquiera) existe $\delta > 0$ tal que

$$|\Delta y| < \delta \quad \Rightarrow \quad |f(x, y + \Delta y) - f(x, y)| < \frac{\varepsilon}{2(\beta - \alpha)}$$

Llamando $\rho = \text{máx}\ \{\delta, \varepsilon/2K\}$, resulta entonces evidente que

$$\left.\begin{array}{l} |\Delta u| < \rho \\ |\Delta y| < \rho \end{array}\right\} \ \Rightarrow \ |F(u + \Delta u, y + \Delta y) - F(u, y)| \leqslant \left| \int_u^{u+\Delta u} f(x, y + \Delta y)\, dx \right| +$$
$$+ \int_\alpha^u |f(x, y + \Delta y) - f(x, y)|\, dx < |\Delta u| K + (u - \alpha)\ \frac{\varepsilon}{2(\beta - \alpha)} < \frac{\varepsilon}{2K}\ K + (\beta - \alpha)\ \frac{\varepsilon}{2(\beta - \alpha)} = \varepsilon$$

lo que prueba la continuidad de F en cualquier punto $(u, y) \in I$.

(*) Se dice que una familia de funciones $\{f_\lambda : M \to \mathbb{R} \ / \ \lambda \in \Lambda\}$ converge uniformemente en M hacia una función $\varphi : M \to \mathbb{R}$ cuando $\lambda \to \lambda_0$ (donde λ_0 es un punto de acumulación de Λ) si, para cada $\varepsilon > 0$, existe un $\delta > 0$ tal que:

$$[\lambda \in C,\ \|\lambda - \lambda_0\| < \delta] \quad \Rightarrow \quad |f_\lambda(\boldsymbol{x}) - \varphi(\boldsymbol{x})| < \varepsilon, \qquad \forall\ \boldsymbol{x} \in M$$

DERIVACIÓN DE LAS INTEGRALES PARAMÉTRICAS

Si en una integral paramétrica el integrando es derivable respecto de uno de los parámetros y, además, dicho integrando y la correspondiente derivada son continuas, entonces la integral es derivable respecto de dicho parámetro y resulta que se pueden intercambiar la integración y la derivación:

$$\frac{\partial}{\partial \lambda_i} \int_M f(\lambda_1, ..., \lambda_p, \mathbf{x})\, d\mathbf{x} = \int_M \frac{\partial f(\lambda_1, ..., \lambda_p, \mathbf{x})}{\partial \lambda_i}\, d\mathbf{x}$$

Más exactamente, se verifica la siguiente propiedad, conocida como la regla de Leibniz:

[81]

I. DERIVACIÓN DE LAS INTEGRALES PARAMÉTRICAS SIMPLES

a) *Caso de límites de integración fijos.* Sea $\Lambda \subset \mathbb{R}^p$ un conjunto abierto; sea $I = [a, b]$ un intervalo compacto real; sea $f\colon \Lambda \times I \to \mathbb{R}$, $(\lambda, x) \mapsto f(\lambda, x)$, una función continua en $\Lambda \times I$ y que admite la derivada parcial $\partial f/\partial \lambda_i$ (donde λ_i es la componente i-ésima de λ, para cierto $i = 1, ..., p$), la cual es continua en $\Lambda \times I$. Se verifica entonces que la función $\lambda \mapsto \int_a^b f(\lambda, x)\, dx$ (que existe para $\lambda \in \Lambda$) es derivable respecto de λ_i en Λ y su derivada viene dada por:

$$\frac{\partial}{\partial \lambda_i}\left(\int_a^b f(\lambda, x)\, dx\right) = \int_a^b \frac{\partial f(\lambda, x)}{\partial \lambda_i}\, dx \qquad \binom{\text{REGLA DE}}{\text{LEIBNIZ}}$$

b) *Caso de límites de integración variables.* Sea $\Lambda \subset \mathbb{R}^p$ un conjunto abierto; sea $S(a, b)$ el «conjunto simple»(*) definido por dos funciones acotadas $a\colon \Lambda \to \mathbb{R}$ y $b\colon \Lambda \to \mathbb{R}$; y considérese una función $f\colon U \to \mathbb{R}$, donde U es un abierto y $S(a, b) \subset U$. Si, en sus respectivos campos de definición, las funciones a, b y f son continuas y admiten derivadas parciales $\partial a/\partial \lambda_i$, $\partial b/\partial \lambda_i$ y $\partial f/\partial \lambda_i$ (donde λ_i es la componente i-ésima de λ e $i = 1, ..., p$), las cuales son continuas, entonces la función $\lambda \mapsto \int_{a(\lambda)}^{b(\lambda)} f(\lambda, x)\, dx$ (que existe para $\lambda \in \Lambda$) es derivable respecto de λ_i en Λ y su derivada viene dada por

$$\frac{\partial}{\partial \lambda_i}\left(\int_{a(\lambda)}^{b(\lambda)} f(\lambda, x)\, dx\right) = \int_{a(\lambda)}^{b(\lambda)} \frac{\partial f(\lambda, x)}{\partial \lambda_i}\, dx + f(\lambda, b(\lambda))\, \frac{\partial b(\lambda)}{\partial \lambda_i} - f(\lambda, a(\lambda))\, \frac{\partial a(\lambda)}{\partial \lambda_i}$$

II. DERIVACIÓN DE LAS INTEGRALES PARAMÉTRICAS MÚLTIPLES

Sea $\Lambda \subset \mathbb{R}^p$ un conjunto abierto; sea $M \subset \mathbb{R}^q$ un conjunto compacto y medible; sea $f\colon \Lambda \times M$, $(\lambda, \mathbf{x}) \mapsto f(\lambda, \mathbf{x})$, una función continua en $\Lambda \times M$ y que admite derivada parcial $\partial f/\partial \lambda_i$ (donde λ_i es la componente i-ésima de λ e $i = 1, ..., p$), la cual es continua en $\Lambda \times M$. Se verifica entonces que la función $\lambda \mapsto \int_M f(\lambda, \mathbf{x})\, d\mathbf{x}$ (que existe para $\lambda \in \Lambda$) es derivable respecto de λ_i en Λ y su derivada viene dada por:

$$\frac{\partial}{\partial \lambda_i}\left(\int_M f(\lambda, \mathbf{x})\, d\mathbf{x}\right) = \int_M \frac{\partial f(\lambda, \mathbf{x})}{\partial \lambda_i}\, d\mathbf{x} \qquad \binom{\text{REGLA DE}}{\text{LEIBNIZ}}$$

(*) $S(a, b) = \{(\lambda, x) \in \mathbb{R}^{p+1} / \lambda \in \Lambda,\ a(\lambda) \leqslant x \leqslant b(\lambda)\}$, donde se supone que $a(\lambda) \leqslant b(\lambda)$ para todo $\lambda \in \Lambda$.

Demostración

1.º Consideremos primero el teorema II; el I.*a* es un caso particular del II, por lo que no se hará una demostración específica para él. Nótese en primer lugar que las integrales del enunciado existen y son funciones continuas (para $\lambda \in \Lambda$), según se desprende de [80]. Empecemos señalando que, como Λ es abierto, para cada $\lambda \in \Lambda$ existe un $r_\lambda > 0$ tal que $\bar{B}(\lambda, r_\lambda) \subset \Lambda$. Si llamamos $F : \Lambda \to \mathbb{R}$ a la función definida por $F(\lambda) = \int_M f(\lambda, x)\, dx$ y si representamos por e el vector i-ésimo de la base canónica de $\mathbb{R}^p$, lo que debemos demostrar es que

$$\lim_{\rho \to 0} \frac{F(\lambda + \rho e) - F(\lambda)}{\rho} = \int_M \frac{\partial f(\lambda, x)}{\partial \lambda_i}\, dx, \qquad \forall\ \lambda \in \Lambda \qquad [1]$$

Fijémonos en la función $\rho \mapsto f(\lambda + \rho e, x)$, la cual depende de $x \in M$ y de $\lambda \in \Lambda$ y está definida en $[-r_\lambda, r_\lambda]$. Aplicando a dicha función el teorema de los incrementos finitos, se puede poner:

$$\frac{F(\lambda + \rho e) - F(\lambda)}{\rho} = \int_M \frac{f(\lambda + \rho e, x) - f(\lambda, x)}{\rho}\, dx = \int_M \frac{\partial f(\lambda + \theta \rho e, x)}{\partial \lambda_i}\, dx$$

para cierto $\theta \in]0, 1[$ (el cual depende de λ, x y ρ) y siempre que sea $|\rho| \leqslant r_\lambda$. Por tanto, se puede poner

$$\left| \frac{F(\lambda + \rho e) - F(\lambda)}{\rho} - \int_M \frac{\partial f(\lambda, x)}{\partial \lambda_i}\, dx \right| = \left| \int_M \left(\frac{\partial f(\lambda + \theta \rho e, x)}{\partial \lambda_i} - \frac{\partial f(\lambda, x)}{\partial \lambda_i} \right) dx \right| \leqslant$$

$$\leqslant \int_M \left| \frac{\partial f(\lambda + \theta \rho e, x)}{\partial \lambda_i} - \frac{\partial f(\lambda, x)}{\partial \lambda_i} \right| dx \qquad [2]$$

Como $\bar{B}(\lambda, r_\lambda) \times M$ es un conjunto compacto y $\partial f / \partial \lambda_i$ es una función continua en dicho conjunto, resulta que esta función es uniformemente continua en el referido conjunto; por tanto, dado $\varepsilon > 0$ (cualquiera), existe un $\delta > 0$ tal que

$$(|\rho| < \delta, x \in I) \quad \Rightarrow \quad \left| \frac{\partial f(\lambda + \theta \rho e, x)}{\partial \lambda_i} - \frac{\partial f(\lambda, x)}{\partial \lambda_i} \right| < \frac{\varepsilon}{\mu(M)}$$

(donde $\mu(M)$ es la medida del conjunto M). Por tanto, llevando este último resultado a [2], se concluye que

$$|\rho| < \text{mín}\,(r_\lambda, \delta) \quad \Rightarrow \quad \left| \frac{F(\lambda + \rho e) - F(\lambda)}{\rho} - \int_M \frac{\partial f(\lambda, x)}{\partial \lambda_i}\, dx \right| < \int_M \frac{\varepsilon}{\mu(M)}\, dx = \varepsilon$$

Esta última relación confirma que se verifica [1], como había que comprobar.

2.° (Caso I.*b*). Aquí nos vamos a apoyar en el anterior caso I.*a*, ya demostrado. Para ello, nótese en primer lugar que, por la continuidad de a y de b en Λ y como para el abierto U es $S(a, b) \subset U$, se puede asegurar que, para cada $\lambda_0 \in \Lambda$, existen ciertos $\delta_0 > 0$ y $h_0 > 0$ tales que, llamando

$$C_0 = B(\lambda_0, \delta) \times [a(\lambda_0) - h_0, b(\lambda_0) + h_0], \qquad \text{es } C_0 \subset U$$

Según esto, siempre que sea $\|\lambda - \lambda_0\| < \delta_0$, la integral $\int_{a(\lambda)}^{b(\lambda)} f(\lambda, x)\, dx$ se puede poner en la forma siguiente (en la que existen las integrales del segundo miembro)

$$\int_{a(\lambda)}^{b(\lambda)} f(\lambda, x)\, dx = I(\lambda) + I_b(\lambda) - I_a(\lambda) \tag{3}$$

siendo

$$I(\lambda) = \int_{a(\lambda_0)}^{b(\lambda_0)} f(\lambda, x)\, dx, \quad I_a(\lambda) = \int_{a(\lambda_0)}^{a(\lambda)} f(\lambda, x)\, dx, \quad I_b(\lambda) = \int_{b(\lambda_0)}^{b(\lambda)} f(\lambda, x)\, dx$$

A la integral $I(\lambda)$, que tiene límites de integración fijos, le es de aplicación el resultado del caso I.*a* (sustituyendo $\Lambda \times I$ por C_0), del cual se obtiene que, para $\lambda = \lambda_0$, es:

$$\frac{\partial I(\lambda_0)}{\partial \lambda_i} = \int_{a(\lambda_0)}^{b(\lambda_0)} \frac{\partial f(\lambda_0, x)}{\partial \lambda_i}\, dx \tag{4}$$

Veamos ahora que $I_a(\lambda)$ e $I_b(\lambda)$ también son derivables en λ_0, respecto de λ_i, y obtengamos sus derivadas (sólo consideraremos $I_b(\lambda)$; para $I_a(\lambda)$ se procede de manera análoga). Dicha derivada es, si existe, el límite (se llama $\boldsymbol{e}$ al vector i-ésimo de la base canónica):

$$l = \lim_{\rho \to 0} \frac{I_b(\lambda_0 + \rho \boldsymbol{e}) - I_b(\lambda_0)}{\rho} = \lim_{\rho \to 0} \frac{1}{\rho} \int_{b(\lambda_0)}^{b(\lambda_0 + \rho \boldsymbol{e})} f(\lambda_0 + \rho \boldsymbol{e}, x)\, dx$$

Acudiendo al teorema de la media integral(*), se puede asegurar que existe cierto $\xi \in \mathbb{R}$ comprendido entre $b(\lambda_0)$ y $b(\lambda_0 + \rho \boldsymbol{e})$ tal que

$$l = \lim_{\rho \to 0} \frac{b(\lambda_0 + \rho \boldsymbol{e}) - b(\lambda_0)}{\rho} f(\lambda_0 + \rho \boldsymbol{e}, \xi) = b'(\lambda_0) f(\lambda_0, x)$$

(como b es continua, $\xi \to x$ cuando $\rho \to 0$; como f es continua, $f(\lambda_0 + \rho \boldsymbol{e}, \xi)$ tiende a $f(\lambda_0, x)$ cuando $\rho \to 0$), es decir, se ha obtenido que

$$\frac{\partial I_b(\lambda_0)}{\partial \lambda_i} = b'(\lambda_0) f(\lambda_0, x), \qquad \text{igualmente es } \frac{\partial I_a(\lambda_0)}{\partial \lambda_i} = a'(\lambda_0) f(\lambda_0, x) \tag{5}$$

Llevando ahora los resultados [4] y [5] a [3], se concluye que $\int_{a(\lambda)}^{b(\lambda)} f(\lambda, x)\, dx$ es derivable respecto de λ_i y que su derivada viene dada por la expresión que había que comprobar.

(*) Si $\varphi : [\alpha, \beta] \to \mathbb{R}$ es una función continua en el intervalo compacto $[\alpha, \beta]$, entonces existe $\xi \in [\alpha, \beta]$ tal que $\int_\alpha^\beta \varphi(x)\, dx = (\beta - \alpha)\varphi(\xi)$.

[81]$_1$ Observaciones

1. Bajo las mismas hipótesis de los teoremas [81], de derivabilidad de las integrales paramétricas $\int_a^b f(\lambda, x)\,dx$, $\int_{a(\lambda)}^{b(\lambda)} f(\lambda, x)\,dx$ e $\int_M f(\boldsymbol{\lambda}, \mathbf{x})\,d\mathbf{x}$ (funciones de $\boldsymbol{\lambda} \in \Lambda$) respecto de λ_i, dichas derivadas son funciones continuas en Λ, según se desprende obviamente de las expresiones [81] de dichas derivadas (nótese que las integrales que en ellas aparecen son funciones continuas, luego se pueden aplicar los teoremas [80]).
2. A la vista de las expresiones [81], de las derivadas de las integrales paramétricas es evidente que si la función f fuese de clase $\mathscr{C}^h$ (con $h \geqslant 1$) en su campo de definición y si, para el caso I.b, también lo fuesen las funciones a y b, entonces las referidas derivadas serían de clase $\mathscr{C}^h$ en Λ.
3. Si, en los anteriores teoremas [81], el parámetro λ fuese real (esto es, si $\Lambda \subset \mathbb{R}$), pongamos que Λ fuera un intervalo abierto, la derivación de la que se habla en [81] no sería una derivada parcial, sino derivada «total» respecto λ (que ahora es escalar); así, por ejemplo, la fórmula [81], I.b será, en este caso:

$$\frac{d}{d\lambda}\int_{a(\lambda)}^{b(\lambda)} f(\lambda, x)\,dx = \int_{a(\lambda)}^{b(\lambda)} \frac{\partial f(\lambda, x)}{\partial \lambda}\,dx + f(\lambda, b(\lambda))\,\frac{db(\lambda)}{d\lambda} - f(\lambda, a(\lambda))\,\frac{da(\lambda)}{d\lambda}$$

En este mismo supuesto de λ real, si el intervalo de variación de λ fuese cerrado, en lugar de abierto (como se ha supuesto), el resultado obtenido (la regla de Leibniz) seguiría siendo válido si se entiende que las derivadas en los extremos de dicho intervalo son derivadas laterales.

[81]$_2$ Ejercicio

A partir del resultado (elemental):

$$\int_0^a \frac{dx}{x^2 + a^2} = \frac{1}{a}\left[\operatorname{arctg}\frac{x}{a}\right]_0^a = \frac{\pi}{4a} \qquad [1]$$

calcular las siguientes integrales

$$\int_0^a \frac{dx}{(x^2 + a^2)^2} \qquad \text{e} \qquad \int_0^a \frac{dx}{(x^2 + a^2)^3}$$

Resolución

Derivando en [1] respecto del parámetro a (regla de Leibniz), se tiene:

$$\int_0^a \frac{-2a\,dx}{(x^2 + a^2)^2} + \frac{1}{2a^2} = -\frac{\pi}{4a^2}, \qquad \text{luego} \int_0^a \frac{dx}{(x^2 + a^2)^2} = \frac{2 + \pi}{8a^3}$$

Derivando nuevamente respecto de a, de la última igualdad se obtiene que

$$\int_0^a \frac{-4a\,dx}{(x^2 + a^2)^3} + \frac{1}{4a^2} = \frac{-3(2 + \pi)}{8a^4}, \qquad \text{luego} \int_0^a \frac{dx}{(x^2 + a^2)^3} = \frac{8 + 3\pi}{32a^5}$$

INTEGRACIÓN DE LAS INTEGRALES PARAMÉTRICAS

[82]

I. INTEGRACIÓN DE LAS INTEGRALES PARAMÉTRICAS SIMPLES

a) *Caso de límites de integración fijos.* Sean $[\alpha, \beta]$ y $[a, b]$ dos intervalos compactos reales; sea $f: [\alpha, \beta] \times [a, b] \to \mathbb{R}$, $(\lambda, x) \mapsto f(\lambda, x)$ una función continua en $[\alpha, \beta] \times [a, b]$. Se verifica entonces que la función $\lambda \mapsto \int_a^b f(\lambda, x)\, dx$ (que existe para $\lambda \in [\alpha, \beta]$) es integrable en $[\alpha, \beta]$ y se tiene:

$$\int_\alpha^\beta \left(\int_a^b f(\lambda, x)\, dx \right) d\lambda = \int_a^b \left(\int_\alpha^\beta f(\lambda, x)\, d\lambda \right) dx = \iint_{[\alpha, \beta] \times [a, b]} f(\lambda, x)\, d\lambda\, dx \qquad [\text{I}.a]$$

b) *Caso de límites de integración variables.* Sea $S \subset \mathbb{R}^2$ un «conjunto simple» respecto de las dos coordenadas, esto es

$$S = \{(\lambda, x) \in \mathbb{R}^2 \;/\; \alpha \leqslant \lambda \leqslant \beta,\ \varphi(\lambda) \leqslant x \leqslant \psi(\lambda)\} = \\ = \{(\lambda, x) \in \mathbb{R}^2 \;/\; a \leqslant x \leqslant b,\ h(x) \leqslant \lambda \leqslant k(x)\}$$

donde φ y ψ son funciones continuas en un intervalo compacto $[\alpha, \beta]$ y h y k son funciones continuas en un intervalo compacto $[a, b]$; sea $f: S \to \mathbb{R}$, $(\lambda, x) \mapsto f(\lambda, x)$, una función continua en S. Se verifica entonces que la función $\lambda \to \int_{\varphi(\lambda)}^{\psi(\lambda)} f(\lambda, x)\, dx$ (que existe para $\lambda \in [\alpha, \beta]$) es integrable en $[\alpha, \beta]$ y se tiene:

$$\int_\alpha^\beta \left(\int_{\varphi(\lambda)}^{\psi(\lambda)} f(\lambda, x)\, dx \right) d\lambda = \int_a^b \left(\int_{h(x)}^{k(x)} f(\lambda, x)\, d\lambda \right) dx = \iint_S f(\lambda, x)\, d\lambda\, dx \qquad [\text{I}.b]$$

II. INTEGRACIÓN DE LAS INTEGRALES PARAMÉTRICAS MÚLTIPLES

Sean $\Lambda \subset \mathbb{R}^p$ y $M \subset \mathbb{R}^q$ dos conjuntos compactos y medibles; sean $f: \Lambda \times M \to \mathbb{R}$, $(\boldsymbol{\lambda}, \boldsymbol{x}) \mapsto f(\boldsymbol{\lambda}, \boldsymbol{x})$, una función continua en $\Lambda \times M$. Se verifica entonces que la función $\boldsymbol{\lambda} \mapsto \int_M f(\boldsymbol{\lambda}, \boldsymbol{x})\, d\boldsymbol{x}$ (que existe para $\boldsymbol{\lambda} \in \Lambda$) es integrable en Λ y se tiene:

$$\int_\Lambda \left(\int_M f(\boldsymbol{\lambda}, \boldsymbol{x})\, d\boldsymbol{x} \right) d\boldsymbol{\lambda} = \int_M \left(\int_\Lambda f(\boldsymbol{\lambda}, \boldsymbol{x})\, d\boldsymbol{\lambda} \right) d\boldsymbol{x} = \int_{\Lambda \times M} f(\boldsymbol{\lambda}, \boldsymbol{x})\, d\boldsymbol{\lambda}\, d\boldsymbol{x} \qquad [\text{II}]$$

Demostración

En los tres casos que se consideran en el enunciado, la integral paramétrica, $\int_a^b f(\lambda, x)\, dx$, $\int_{\varphi(\lambda)}^{\psi(\lambda)} f(\lambda, x)\, dx$ o $\int_M f(\boldsymbol{\lambda}, \boldsymbol{x})\, d\boldsymbol{x}$, es una función continua del parámetro (según se probó en [80]), por lo que es integrable (respecto del parámetro) en su campo de definición, es decir, existe la primera de las dos integrales iteradas de I.*a*, I.*b* y II. Intercambiando los papeles de λ y x se concluye que existen las segundas integrales iteradas. También existen las integrales dobles de I.*a*, I.*b* y II, ya que los integrandos son continuos y se integra sobre conjuntos medibles. Según el teorema de Fubini(*) (véase [70]), en los tres casos se verifica que la integral doble del final es igual a cada una de las dos integrales iteradas respectivas (en el caso II, téngase en cuenta que $\Lambda \times M$ es un conjunto medible), por lo que son efectivamente ciertas las igualdades de I.a, I.b y II.

[82]₁ Ejercicio

Calcular las siguientes integrales iteradas:

$$I = \int_0^1 \left(\int_x^1 e^{-y^2}\, dy \right) dx \qquad \text{y} \qquad J = \int_0^{\sqrt{\pi}} \left(\int_{y^2}^1 \sqrt{x} \operatorname{sen} x\, dx \right) dy$$

Resolución

Como las funciones $y \mapsto e^{-y^2}$ y $x \mapsto \sqrt{x} \operatorname{sen} x$ no admiten primitiva elemental, nos va a interesar cambiar, en I y en J, el orden de integración, lo cual es correcto según el teorema [80],I.b:

$$I = \int_0^1 \left(\int_0^y e^{-y^2}\, dx \right) dy = \int_0^1 (e^{-y^2} y)\, dy = \left[-\frac{1}{2}\, e^{-y^2} \right]_0^1 = \frac{e-1}{2e}$$

$$J = \int_0^{\pi} \left(\int_0^{\sqrt{x}} \sqrt{x} \operatorname{sen} x\, dy \right) dx = \int_0^{\pi} (x \operatorname{sen} x)\, dx = [\operatorname{sen} x - x \cos x]_0^{\pi} = \pi$$

[82]₂ Interpretación geométrica

Sea $S \subset \mathbb{R}^2$ un «conjunto simple» respecto de las dos coordenadas; S es un conjunto, incluido en un cierto intervalo compacto $[a, b] \times [c, d]$, que se puede expresar de las dos maneras siguientes, para unas ciertas funciones $\varphi_1, \varphi_2, \psi_1, \psi_2$, que se suponen continuas:

$$S = \{(x, y) \in \mathbb{R}^2 \ / \ a \leqslant x \leqslant b,\ \varphi_1(x) \leqslant y \leqslant \varphi_2(x)\}, \qquad \varphi_1, \varphi_2 : [a, b] \to [c, d]$$
$$S = \{(x, y) \in \mathbb{R}^2 \ / \ c \leqslant y \leqslant d,\ \psi_1(y) \leqslant x \leqslant \psi_2(y)\}, \qquad \psi_1, \psi_2 : [c, d] \to [a, b]$$

Consideremos una cierta función $f: S \to \mathbb{R}$ continua en S y supongamos que nos interesamos por el volumen V del cuerpo sólido que (en la representación cartesiana usual, con coordenadas rectangulares xyz) ocupa la porción del cilindro vertical de base S que está delimitada entre el plano $z = 0$ y la superficie $z = f(x, y)$, es decir, el volumen de

$$C = \{(x, y, z) \in \mathbb{R}^2 \ / \ (x, y) \in S,\ 0 \leqslant z \leqslant f(x, y)\}$$

(*) En el caso de integrales paramétricas simples, se pueden hacer demostraciones directas (sin acudir a las integrales múltiples). Para el caso de límites de integración fijos, véase el ejercicio [82]₃.

Dicho volumen es, según se indicó cuando hablábamos de la integral doble

$$V = \iint_S f(x, y)\, dx\, dy \qquad [1]$$

Tomemos un valor cualquiera de $x_0 \in [a, b]$ y consideremos la sección que, en el cuerpo C, produce el plano $x = x_0$; el área de esta sección es

$$a(x_0) = \int_{\varphi_1(x_0)}^{\varphi_2(x_0)} f(x_0, y)\, dy$$

Si consideramos un «pequeño» Δx (incremento de x) en torno del valor x_0, la expresión $a(x_0)\Delta x$ es una «buena aproximación» del volumen de la «rebanada» que, en el cuerpo C, producen dos planos paralelos al $x = x_0$, próximos a él, y situados a distancia Δx. Así pues, dando una partición del intervalo $[a, b]$ y tomando en ella unos puntos intermedios, la correspondiente suma de Riemann $\Sigma\, a(x_i)\Delta x_i$, de la integral $\int_a^b a(x)\, dx$, será una «buena aproximación» de volumen V, pues es la suma de «buenas aproximaciones» de todas las «rebanadas» que con ello se produce en V. Por tanto, al acudir a las «rebanadas» de C paralelas al plano $x = 0$, las anteriores consideraciones nos inducen a pensar que debe ser:

$$V = \int_a^b a(x)\, dx = \int_a^b \left(\int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y)\, dy \right) dx \qquad [2]$$

Intercambiando ahora los papeles de x e y, un razonamiento análogo al anterior nos llevaría a suponer que:

$$V = \int_c^d \left(\int_{\psi_1(y)}^{\psi_2(y)} f(x, y)\, dx \right) dy \qquad [3]$$

A la vista de [1], [2] y [3], las anteriores consideraciones geométricas nos llevan a ver como plausible que sea

$$\int_a^b \left(\int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y)\, dy \right) dx = \int_c^d \left(\int_{\psi_1(y)}^{\psi_2(y)} f(x, y)\, dx \right) dy = \iint_S f(x, y)\, dx\, dy$$

que es la fórmula (ya demostrada; véase [80], I.*b*) de integración de una integral paramétrica.

[82]$_3$ Ejercicio

Hacer una demostración directa (sin recurrir a las integrales dobles) del teorema [82], I.*a*; es decir, pruébese que si $f: [\alpha, \beta] \times [a, b]$ es una función continua en el intervalo compacto $I = [\alpha, \beta] \times [a, b]$ de $\mathbb{R}^2$, entonces se verifica que

$$\int_\alpha^\beta \left(\int_a^b f(x, y)\, dy \right) dx = \int_a^b \left(\int_\alpha^\beta f(x, y)\, dx \right) dy$$

Resolución

Sean G y H las funciones de $[\alpha, \beta]$ en $\mathbb{R}$, definidas por las expresiones:

$$G(u) = \int_\alpha^u \left(\int_a^b f(x, y)\, dy \right) dx \qquad \text{y} \qquad H(u) = \int_a^b \left(\int_\alpha^u f(x, y)\, dx \right) dy$$

Según sabemos (véase [80], I.a), la función $x \mapsto \int_a^b f(x, y)\, dy$ (que existe) es continua en $[\alpha, \beta]$. Por tanto, a $u \mapsto G(u)$ le es de aplicación el teorema fundamental del cálculo(*), del que se deduce que G es derivable en $[\alpha, \beta]$ y que

$$G'(u) = \int_a^b f(u, y)\, dy, \qquad \forall\, u \in [\alpha, \beta]$$

La función $(u, y) \mapsto \int_\alpha^u f(x, y)\, dx$ es continua en I (véase $[80]_2$) y derivable respecto de u en I (teorema fundamental del cálculo) y su derivada es la función $(u, y) \mapsto f(u, y)$. Por ello (véase [81], I.a), se puede asegurar que H es derivable en $[\alpha, \beta]$ y que $H'(u) = \int_a^b f(u, y)\, dy$. Nos encontramos, pues, con que $G'(u) = H'(u)$ para todo $u \in [\alpha, \beta]$, luego $G(u) - H(u)$ es constante para $u \in [\alpha, \beta]$; como $G(\alpha) = 0$ y $H(\alpha) = 0$, resulta que $G(u) = H(u)$ para $u \in [\alpha, \beta]$ y, en particular, $G(\beta) = H(\beta)$, que es la igualdad que se quería comprobar.

CÁLCULO DE INTEGRALES RECURRIENDO A LAS INTEGRALES PARAMÉTRICAS

[83]

Supongamos que se quiere hallar una integral $I = \int_a^b \varphi(x)\, dx$, donde $\varphi : [a, b] \to \mathbb{R}$, es una función continua. Para ello, vamos a acudir a una cierta función real $(\lambda, x) \mapsto f(\lambda, x)$, definida en cierto conjunto de la forma $[\alpha, \beta] \times [a, b]$, en el que es continua y con derivada $\partial f / \partial \lambda$ continua. De acuerdo con los teoremas anteriores, se verifica que:

1.º Si para cierto $\lambda_0 \in [\alpha, \beta]$ y todo $x \in [a, b]$ es $\varphi(x) = \partial f(\lambda_0, x)/\partial \lambda$ y llamando $F(\lambda) = \int_a^b f(\lambda, x)\, dx$ (suponemos que esta integral se sabe calcular), es entonces $I = F'(\lambda_0)$.

2.º Si para cierto $\lambda_0 \in [\alpha, \beta]$ y todo $x \in [a, b]$ es $\varphi(x) = f(\lambda_0, x)$ y llamando $G(\lambda) = \int_a^b (\partial f(\lambda, x)/\partial \lambda)\, dx$ y $J = \int_a^b f(\alpha, x)\, dx$ (suponemos que estas integrales se deben calcular), es entonces $I = \int_\alpha^{\lambda_0} G(\lambda)\, d\lambda + J$.

Nota: También puede incorporarse el parámetro λ a los límites de integración, poniendo $a(\lambda)$ o $b(\lambda)$ variables, en lugar de los valores fijos a o b.

(*) Si $\varphi : [\alpha, \beta] \to \mathbb{R}$ es una función continua en $[\alpha, \beta]$, entonces la función $u \mapsto G(u) = \int_\alpha^u \varphi$ es derivable en $[\alpha, \beta]$ y su derivada es $G'(u) = \varphi(u)$ para todo $u \in [\alpha, \beta]$.

[83]$_1$ Ejercicio

Calcular la siguiente integral I (indicación: derivar respecto del parámetro a en la integral $J(a)$):

$$I = \int_0^1 \frac{x-1}{\operatorname{L} x}\, dx \qquad\qquad J(a) = \int_0^1 \frac{x^a - 1}{\operatorname{L} x}\, dx$$

Resolución

Derivando respecto de a, como se indica en el enunciado, de la regla de Leibniz se obtiene:

$$J'(a) = \int_0^1 \frac{x^a \operatorname{L} x}{\operatorname{L} x}\, dx = \int_0^1 x^a\, dx = \left[\frac{x^{a+1}}{a+1}\right]_{x=0}^{x=1} = \frac{1}{a+1}$$

Integrando ahora (la variable es a), se tiene que, para cierto $c \in \mathbb{R}$, es:

$$J(a) = \int \frac{da}{a+1} = \operatorname{L}\,(a+1) + c$$

Como es $J(0) = 0$, de esta última expresión se obtiene que $\operatorname{L} 1 + c = 0$, o sea, que es $c = 0$. Como $I = J(1)$, se concluye que

$$I = J(1) = \operatorname{L}\,(1+1) + 0 = \operatorname{L} 2$$

[83]$_2$ Ejercicio

Calcular la siguiente integral I (indicación: integrar entre 0 y a la integral paramétrica $J(a)$):

$$I = \int_0^{\pi/2} \frac{x}{\operatorname{tg} x}\, dx \qquad\qquad J(a) = \int_0^{\pi/2} \frac{dx}{1 + a^2 \operatorname{tg}^2 x}$$

Resolución

Empecemos calculando $J(a)$, para lo que realizamos el cambio de variable $\operatorname{tg} x = u$, del que se tiene:

$$J(a) = \int_0^{+\infty} \frac{du}{(1+u^2)(1+a^2u^2)}$$

Descomponiendo este integrando de $J(a)$ en suma de fracciones simples, será:

$$\begin{cases} \text{si } a \neq \pm 1, & J(a) = \dfrac{1}{1-a^2} \displaystyle\int_0^{+\infty} \left(\dfrac{1}{1+u^2} - \dfrac{a^2}{1+a^2u^2} \right) du = \dfrac{\pi}{2(1+|a|)} \\ \text{si } a = \pm 1, & J(a) = \displaystyle\int \dfrac{du}{(1+u^2)^2} = \left[\dfrac{u}{2(1+u^2)} + \dfrac{1}{2} \operatorname{arctg} u \right]_{u=0}^{u \to +\infty} = \dfrac{\pi}{4} \end{cases}$$

Por tanto, para $a \geqslant 0$, que son los valores que nos van a interesar, es:

$$J(a) = \frac{\pi}{2(1+a)}$$

Según se indica en el enunciado, integremos $J(a)$ entre 0 y a_0; intercambiando el orden de integración (véase [82], I.a) se tiene sucesivamente:

$$\int_0^{a_0} \left(\int_0^{\pi/2} \frac{dx}{1+a^2 \operatorname{tg}^2 x} \right) da = \int_0^{a} \frac{\pi \, da}{2(1+a)}, \qquad \int_0^{\pi/2} \left(\int_0^{a_0} \frac{da}{1+a^2 \operatorname{tg}^2 x} \right) dx = \frac{\pi}{2} \operatorname{L}(1+a_0)$$

$$\int_0^{\pi/2} \frac{\operatorname{arctg}(a_0 \operatorname{tg} x)}{\operatorname{tg} x} \, dx = \frac{\pi}{2} \operatorname{L}(1+a_0)$$

Tomando para a_0 el valor 1, se concluye que

$$\int_0^{\pi/2} \frac{x}{\operatorname{tg} x} \, dx = \frac{\pi}{2} \operatorname{L} 2, \qquad \text{o sea } I = \frac{\pi}{2} \operatorname{L} 2$$

[83]$_3$ Ejercicio

Calcular la siguiente integral I (indicación: derivar respecto del parámetro a en la integral $J(a)$):

$$I = \int_0^1 \frac{\operatorname{L}(1+x)}{1+x^2} \, dx \qquad\qquad J(a) = \int_0^a \frac{\operatorname{L}(1+ax)}{1+x^2} \, dx$$

Resolución

Derivando $J(a)$, según se dice en el enunciado, se obtiene (téngase en cuenta que el límite superior de integración también depende de a):

$$J'(a) = \int_0^a \frac{x}{(1+x^2)(1+ax)} \, dx + \frac{\operatorname{L}(1+a^2)}{1+a^2}$$

Descomponiendo este último integrando en suma de fracciones simples, se llega a:

$$J'(a) = \frac{1}{1+a^2}\left[\frac{1}{2}\,\mathrm{L}\,(1+x^2) + a \operatorname{arctg} x - \mathrm{L}\,(1+ax)\right]_{x=0}^{x=a} + \frac{\mathrm{L}\,(1+a^2)}{1+a^2} =$$

$$= \frac{\mathrm{L}\,\sqrt{1+a^2} + a \operatorname{arctg} a}{1+a^2}$$

Integrando ahora $J'(a)$ entre 0 y a_0 y como $J(0) = 0$, se tiene:

$$J(a_0) = \int_0^{a_0} J'(a)\,da = \int_0^{a_0} \frac{\mathrm{L}\,\sqrt{1+a^2}}{1+a^2}\,da + \int_0^{a_0} \frac{a \operatorname{arctg} a}{1+a^2}\,da$$

Para la primera integral del segundo miembro, integrando por partes ($u = \mathrm{L}\,\sqrt{1+a^2}$, $dv = da/(1+a^2)$), se tiene:

$$\int_0^{a_0} \frac{\sqrt{1+a^2}}{1+a^2}\,da = (\operatorname{arctg} a_0)\,\mathrm{L}\,\sqrt{1+a_0^2} - \int_0^{a_0} \frac{a \operatorname{arctg} a}{1+a^2}\,da$$

Llevando este resultado a la anterior expresión de $J(a_0)$, resulta que

$$J(a_0) = (\operatorname{arctg} a_0)\,\mathrm{L}\,\sqrt{1+a_0^2}$$

Como $I = J(1)$, de la última igualdad se desprende que:

$$I = J(1) = (\operatorname{arctg} 1)\,\mathrm{L}\,\sqrt{2} = \frac{\pi\sqrt{2}}{4}$$

5.3. INTEGRALES PARAMÉTRICAS IMPROPIAS

En el análisis que se acaba de realizar, acerca de las integrales paramétricas, sólo se han considerado integrales propias. Por ello, vamos a completar dicho estudio, considerando ahora las integrales impropias dependientes de parámetros, las cuales tienen aplicaciones muy estimables. De dichas integrales impropias, las realmente interesantes son las integrales simples con uno o dos parámetros, mientras que las integrales paramétricas impropias múltiples tienen escasa importancia; por ello, y por la dificultad que encierra su estudio, no vamos a ocuparnos de estas integrales. Por análogas razones, no consideraremos las integrales con parámetros en los límites de integración, limitándonos a integrales impropias del tipo $\int_{\to a}^{b} f(\lambda, x)\,dx$ o $\int_a^{\to b} f(\lambda, x)\,dx$; es más, como el estudio de la primera de éstas se reduce obviamente al de la segunda, sólo hablaremos de integrales (paramétricas, simples) que son impropias en el extremo de su intervalo de integración.

Las integrales impropias que hemos estudiado anteriormente (véase 5.1) eran integrales impropias múltiples; las integrales impropias simples no son objeto de este curso de Cálculo Infinitesimal, pues su estudio corresponde a un curso de Cálculo Infinitesimal de una variable. Recuérdese que una integral impropia simple del tipo $\int_a^{\to b} \varphi(x)\,dx$ (impropia en b) se define como sigue:

Sea $[a, b[$ un intervalo no compacto, donde $-\infty < a < b \leqslant +\infty$; sea $\varphi : [a, b[\to \mathbb{R}$ una función integrable en todo intervalo compacto $[a, x]$, con $x < b$. Se dice que φ es integrable en $[a, b[$ si existe el límite lím $\int_a^x \varphi$ cuando $x \to b^-$, al cual se le llama integral (impropia en b) de φ en $[a, b[$ y se pone

$$\int_a^{\to b} \varphi(t)\,dt = \lim_{x \to b^-} \int_a^x \varphi(t)\,dt$$

Si este límite es finito, la integral se dice convergente; si el límite es infinito, la integral se llama divergente. Cuando, a lo largo de lo que sigue, se recurra a algún resultado específico de estas integrales impropias (simples) que no sea una consecuencia obvia de lo que antes se estudió para las integrales impropias múltiples(*), se hará un «recordatorio» en el que se resuma el resultado en cuestión.

INTEGRALES PARAMÉTRICAS IMPROPIAS: CONVERGENCIA Y CONVERGENCIA UNIFORME

Consideremos aquí una función $\lambda \mapsto F(\lambda)$ definida por una integral impropia, esto es, del tipo $F(\lambda) = \int_a^{\to b} f(\lambda, x)\,dx$, donde la integral se supone que converge para todos los valores λ de un cierto conjunto Λ, que es el campo de definición de F. Si «las cosas funcionan bien», cabe esperar que las propiedades del integrando f (continuidad, derivabilidad, etc.) las «herede» la integral F. En general, esto no ocurre, salvo que se exija que la convergencia de la integral sea una convergencia especial, más «fuerte» que la ordinaria, de la que nos ocupamos a continuación.

Se dice que un conjunto Λ es el «campo de convergencia» de una integral $\int_a^{\to b} f(\lambda, x)\,dx$ (impropia en b y con un parámetro λ) si para cada $\lambda_0 \in \Lambda$ es convergente la integral $\int_a^{\to b} f(\lambda_0, x)\,dx$. Si la «rapidez» con la que $\int_a^{b-\varepsilon} f(\lambda, x)\,dx$ converge hacia $\int_a^{\to b} f(\lambda, x)\,dx$ cuando $\varepsilon \to 0$ es «similar» para los distintos valores que se toman de λ en Λ, se dirá que la integral impropia es uniformemente convergente en Λ. De un modo más preciso, se dan las siguientes definiciones:

(*) Conviene notar que lo dicho en 5.1, para integrales impropias múltiples, es válido, en general, para las integrales impropias simples. Lo dicho deja de ser cierto en lo que se refiere a la convergencia absoluta, que, en el caso de las integrales múltiples, equivale a la convergencia ordinaria; en el caso de integrales impropias simples, la convergencia absoluta implica la convergencia ordinaria, pero el recíproco no se verifica, en general.

[84]

Sea $I = [a, b[$ un intervalo no compacto de $\mathbb{R}$, con $-\infty < a < b \leqslant +\infty$, y considérese un conjunto $\Lambda \subset \mathbb{R}^p$. Sea $f: \Lambda \times I \to \mathbb{R}$ una función tal que, para cada $\lambda \in \Lambda$ y cada $k \in I$, existe la integral (propia) $\int_a^k f(\lambda, x)\, dx$.

I. CONVERGENCIA. Si esta integral tiene límite finito, cuando $k \to b^-$ y para cualquier $\lambda \in \Lambda$, entonces a dicho límite se le representa poniendo $\int_a^{\to b} f(\lambda, x)\, dx$ (integral paramétrica impropia) y se dice que $\int_a^k f(\lambda, x)\, dx$ converge en Λ hacia la integral paramétrica impropia $\int_a^{\to b} f(\lambda, x)\, dx$, cuando $k \to b^-$. Esta convergencia significa, pues, que: para cada $\lambda \in \Lambda$ y cada $\varepsilon > 0$, existe $k_{\lambda\varepsilon} \in I$ (que dependerá de λ y ε) tal que

$$(k_{\lambda\varepsilon} < k < b) \quad \Rightarrow \quad \left| \int_k^{\to b} f(\lambda, x)\, dx \right| < \varepsilon \qquad [1]$$

Lo dicho se puede resumir así: la integral $\int_a^{\to b} f(\lambda, x)\, dx$ es la siguiente función, definida mediante un límite:

$$\Lambda \to \mathbb{R}, \quad \lambda \mapsto \int_a^{\to b} f(\lambda, x)\, dx = \lim_{k \to b^-} \int_a^k f(\lambda, x)\, dx$$

II. CONVERGENCIA UNIFORME. Considérese la anterior integral paramétrica impropia $\int_a^{\to b} f(\lambda, x)\, dx$, que se supone convergente en Λ (al menos). Se dice que la convergencia de $\int_a^k f(\lambda, x)\, dx$ hacia $\int_a^{\to b} f(\lambda, x)\, dx$, cuando $k \to b^-$, es convergencia uniforme en Λ si, para cada $\varepsilon > 0$, existe un $k_\varepsilon \in I$ (que sólo dependerá de ε) tal que

$$(k_\varepsilon < k < b) \quad \Rightarrow \quad \left| \int_k^{\to b} f(\lambda, x)\, dx \right| < \varepsilon, \qquad \forall\, \lambda \in \Lambda \qquad [2]$$

(esto es si la condición [*I*], de convergencia (ordinaria), se verifica para algún $k_{\lambda\varepsilon}$ que no depende de λ, al cual se le ha llamado k_ε).

Nota: Las integrales paramétricas del tipo $\int_{\to a}^{b} f(\lambda, x)\, dx$, impropias en a, se definen, mutatis mutandis, como las del tipo $\int_a^{\to b} f(\lambda, x)\, dx$ que se han considerado aquí.

[84]$_1$ Observaciones

1. Las precedentes definiciones de convergencia y de convergencia uniforme son válidas, tanto si las integración impropia (en b) se realiza en un intervalo no acotado, o sea, si $[a, b[= [a, +\infty[$, cuanto si se realiza en un intervalo acotado en el que el integrando no está acotado (para $x \to b^-$).
2. Es obvio que la convergencia uniforme, de $\int_a^{\to b} f(\lambda, x)\, dx$ en Λ (para $\lambda \in \Lambda$), implica su convergencia ordinaria, ya que, de verificarse la condición [2], entonces se cumple [1] para $k_{\lambda\varepsilon} = k_\varepsilon$. El recíproco es falso en general, como se comprueba fácilmente con algún contraejemplo (véase [84]$_3$); es decir, la condición de convergencia uniforme es una exigencia mayor que la de convergencia (ordinaria).

3. Igual que se han definido las integrales paramétricas impropias del tipo $\int_a^{\to b} f(\lambda, x)\,dx$ (impropias en b), se definen las del tipo $\int_{\to a}^{b} f(\lambda, x)\,dx$ (impropias en a), así como su convergencia uniforme. Para ello, basta con sustituir, en lo dicho en [84], $\int_a^{\to b}$, $\int_k^{\to b}$ e $\int_a^k$ por $\int_{\to a}^{b}$, $\int_{\to a}^{k}$ e $\int_k^b$, respectivamente, $[a, b[$ con $-\infty < a < b \leqslant +\infty$ por $]a, b]$ con $-\infty \leqslant a < b < +\infty$, «impropia en b» por «impropia en a», $k \to b^-$ por $k \to a^+$ y $k_{\lambda\varepsilon} < k < b$ y $k_\varepsilon < k < b$ por $a < k < k_{\lambda\varepsilon}$ y $a < k < k_\varepsilon$, respectivamente.

[84]$_2$ Caracterización, mediante el supremo, de la convergencia uniforme

Sea $I = [a, b[$ un intervalo no compacto de $\mathbb{R}$, con $-\infty < a < b \leqslant +\infty$, y considérese un conjunto $\Lambda \subset \mathbb{R}^p$. Sea $f: \Lambda \times I \to \mathbb{R}$ una función tal que, para cada $\lambda \in \Lambda$ y cada $k \subset I$, existe la integral (propia) $\int_a^k f(\lambda, x)\,dx$. Para que exista la integral paramétrica $\int_a^{\to b} f(\lambda, x)\,dx$ (impropia en b) y ella sea uniformemente convergente en Λ (para $\lambda \in \Lambda$) es necesario y suficiente que se verifique la siguiente condición:

$$\lim_{k \to b^-} \left[\sup \left\{ \left| \int_k^{\to b} f(\lambda, x)\,dx \right| \Big/ \lambda \in \Lambda \right\} \right] = 0 \qquad [1]$$

Demostración

1.° Supongamos primero que la integral impropia es uniformemente convergente en Λ. Entonces, dado $\varepsilon > 0$ (cualquiera), existe un $k_\varepsilon \in I$ para el que se cumple la relación [2] de [84] y, por consiguiente, también se verifica la relación

$$(k_\varepsilon < k < b) \;\Rightarrow\; \sup \left\{ \left| \int_k^{\to b} f(\lambda, x)\,dx \right| \Big/ \lambda \in \Lambda \right\} \leqslant \varepsilon$$

de la que se desprende que se satisface la anterior condición [1].

2.° Recíprocamente, supongamos ahora que se cumple la condición [1]. De ser así, es obvio que, dado $\varepsilon > 0$ (cualquiera), existe un $k_\varepsilon \in I$ tal que

$$(k_\varepsilon < k < b) \;\Rightarrow\; \sup \left\{ \left| \int_k^{\to b} f(\lambda, x)\,dx \right| \Big/ \lambda \in \Lambda \right\} < \varepsilon$$

De esta relación se desprende obviamente la condición [2] de [84], luego nuestra integral impropia es uniformemente convergente en Λ.

[84]$_3$ Ejemplo

Consideremos la integral impropia $F(\lambda) = \int_0^{+\infty} \lambda e^{-\lambda x}\,dx$, donde $\lambda \in \mathbb{R}$. Como $x \mapsto \lambda e^{-x\lambda}$ tiene como primitiva $x \mapsto e^{-\lambda x}$, resulta que

$$F(\lambda) = [e^{-\lambda x}]_{x=0}^{x \to +\infty} = \begin{cases} +\infty, & \text{si } \lambda < 0 \\ 0, & \text{si } \lambda = 0 \\ -1, & \text{si } \lambda > 0 \end{cases}$$

Por tanto, la inegral diverge para $\lambda < 0$ y es convergente en $\mathbb{R}^+ = \{\lambda \in \mathbb{R} \mid \lambda \geqslant 0\}$.

Para cualquiera que sea el número $a > 0$, la integral es uniformemente convergente en $[a, +\infty[$, ya que

$$\lim_{k\to+\infty} \sup \left\{ \left| \int_k^{+\infty} \lambda e^{-\lambda x}\, dx \right| \Big/ \lambda \geqslant a \right\} = \lim_{k\to+\infty} \sup \{e^{-\lambda k} / \lambda \geqslant a\} = \lim_{k\to+\infty} e^{-ak} = 0$$

luego se cumple la condición $[84]_2$ de convergencia uniforme para $\lambda \in [a, +\infty[$. Sin embargo, la integral no converge uniformemente en $]0, +\infty[$, ya que

$$\lim_{k\to+\infty} \sup \left\{ \left| \int_k^{+\infty} \lambda e^{-\lambda x}\, dx \right| \Big/ \lambda > 0 \right\} = \lim_{k\to+\infty} \sup \{e^{-\lambda k} / \lambda > 0\} = \lim_{k\to+\infty} 1 = 1 \neq 0$$

luego la condición $[84]_2$ de convergencia uniforme no se cumple en $]0, +\infty[$.

$[84]_4$ Ejercicio

Considérese la integral paramétrica $\int_a^{\to b} f(\lambda, x)\, dx$ (impropia en b) para $\lambda \in \Lambda \subset \mathbb{R}^p$, y sea (k_n) una sucesión de puntos de $[a, b[$ tal que $\lim k_n = b$ y siendo $k_1 = a$. Pruébese que si la citada integral impropia es, respectivamente, convergente o es uniformemente convergente en Λ, entonces también es convergente o es uniformemente convergente en Λ la serie de funciones $\sum_{i=1}^{\infty} F_i(\lambda)$, donde $F_i(\lambda) = \int_{k_i}^{k_{i+1}} f(\lambda, x)\, dx$.

Comprobación

La suma n-ésima de la serie es $S_n(\lambda) = \int_a^{k_n} f(\lambda, x)\, dx$, es obvio que las condiciones de convergencia y de convergencia uniforme de la integral (esto es, las condiciones [1] y [2] de [84]) implican que, dado $\varepsilon > 0$ (cualquiera), se verifica que:

1.º (Si la convergencia de la integral es simple, no uniforme) Para cada $\lambda \in \Lambda$ existe un $\nu_\lambda \in \mathbb{N}$ tal que

$$n \geqslant \nu_\lambda \;\Rightarrow\; k_{\lambda\varepsilon} < k_n < b \;\Rightarrow\; |F(\lambda) - S_n(\lambda)| = \left| \int_{k_n}^{\to b} f(\lambda, x)\, dx \right| < \varepsilon$$

lo que prueba la convergencia en Λ de la serie de funciones.

2.º (Si la convergencia de la integral es uniforme) Existe un $\nu \in \mathbb{N}$ tal que

$$n \geqslant \nu \;\Rightarrow\; k_\varepsilon < k_n < b \;\Rightarrow\; |F(\lambda) - S_n(\lambda)| = \left| \int_{k_n}^{\to b} f(\lambda, x)\, dx \right| < \varepsilon,\ \forall\ \lambda \in \Lambda$$

lo que prueba la convergencia uniforme en Λ de la serie de funciones.

CRITERIOS DE CONVERGENCIA UNIFORME

[85]

Considérense un conjunto $\Lambda \subset \mathbb{R}^p$, un intervalo no compacto $I = [a, b[\subset \mathbb{R}$, con $-\infty < a < b \leqslant +\infty$, y una función $f: \Lambda \times I \to \mathbb{R}$, $(\lambda, x) \mapsto f(\lambda, x)$, tales que para cada $\lambda \in \Lambda$ y cada $k \in I$ existe la integral $\int_a^k f(\lambda, x)\, dx$. Se verifica que:

1. CRITERIO GENERAL DE CAUCHY. La integral paramétrica impropia (en b) $\int_a^{\to b} f(\lambda, x)\, dx$ existe y es uniformemente convergente en Λ si, y sólo si, se cumple la siguiente condición (de Cauchy): para cada $\varepsilon > 0$ existe un $k_\varepsilon \in I$ tal que

$$(k_\varepsilon < k < k' < b) \quad \Rightarrow \quad \left| \int_k^{k'} f(\lambda, x)\, dx \right| < \varepsilon, \qquad \forall\, \lambda \in \Lambda$$

2. CRITERIO DE WEIERSTRASS. Si $g: I \to \mathbb{R}$ es una función positiva tal que la integral $\int_a^{\to b} g(x)\, dx$ (impropia en b) es convergente y si es $|f(\lambda, x)| \leqslant g(x)$ para cualesquiera $\lambda \in \Lambda$ y $x \in I$, entonces la integral paramétrica impropia $\int_a^{\to b} f(\lambda, x)\, dx$ existe y es uniformemente convergente en Λ.

Demostración

1.º Si $\int_a^{\to b} f(\lambda, x)\, dx$ existe y es uniformemente convergente para $\lambda \in \Lambda$, sabemos que (véase [84], II), dado $\varepsilon > 0$ (cualquiera), existe $k_\varepsilon \in I$ tal que

$$(k_\varepsilon < k < b) \quad \Rightarrow \quad \left| \int_k^{\to b} f(\lambda, x)\, dx \right| < \frac{\varepsilon}{2}, \qquad \forall\, \lambda \in \Lambda$$

De ello se desprende obviamente que:

$$(k_\varepsilon < k < k' < b) \quad \Rightarrow \quad \left| \int_k^{k'} f(\lambda, x)\, dx \right| \leqslant \left| \int_k^{\to b} f(\lambda, x)\, dx \right| + \\ + \left| \int_{k'}^{\to b} f(\lambda, x)\, dx \right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \qquad \forall\, \lambda \in \Lambda$$

que es la condición de Cauchy. Recíprocamente, suponiendo ahora que se verifica esta condición, se tiene: según el criterio general de Cauchy de las integrales impropias (no paramétricas), conocido de cursos anteriores(*), la integral impropia $\int_a^{\to b} f(\lambda, x)\, dx$ existe y es convergente para todo $\lambda \in \Lambda$. De acuerdo con la condición

(*) Criterio de Cauchy para integrales impropias (no paramétricas). Una integral $\int_a^{\to b} \varphi(x)\, dx$ (impropia en b; donde $-\infty < a < b \leqslant +\infty$) existe y es convergente si, y sólo si, para cada $\varepsilon > 0$, existe un $k_\varepsilon \in [a, b[$ tal que:

$$(k_\varepsilon < k < k' < b) \quad \Rightarrow \quad \left| \int_k^{k'} \varphi(x)\, dx \right| < \varepsilon$$

de Cauchy, dado $\varepsilon > 0$ (cualquiera), existe $k_\varepsilon \in I$ tal que $|\int_k^{k'} f(\lambda, x)\,dx| < \varepsilon/2$ para todo $\lambda \in \Lambda$ y siempre que sea $k_\varepsilon < k < k' < b$; por tanto:

$$(k_\varepsilon < k < k' < b) \quad \Rightarrow \quad \left|\int_k^{\to b} f(\lambda, x)\,dx\right| = \left|\lim_{k' \to b^-} \int_k^{k'} f(\lambda, x)\,dx\right| \leqslant \frac{\varepsilon}{2} < \varepsilon, \qquad \forall\ \lambda \in \Lambda$$

lo que prueba la convergencia uniforme en Λ de la integral impropia $\int_a^{\to b} f(\lambda, x)\,dx$.

2.° De acuerdo con el criterio de Weierstrass de las integrales simples impropias, conocido ya de cursos anteriores(*) (y generalizando anteriormente, en [75], II), se sabe que la integral impropia $\int_a^{\to b} f(\lambda, x)\,dx$ existe y es convergente para todo $\lambda \in \Lambda$. Para cualesquiera que sean $k \in I$ y $\lambda \in \Lambda$, es:

$$\left|\int_k^{\to b} f(\lambda, x)\,dx\right| \leqslant \int_k^{\to b} |f(\lambda, x)|\,dx \leqslant \int_k^{\to b} g(x)\,dx \qquad [1]$$

Para la convergencia de $\int_a^{\to b} g(x)\,dx$, para cada $\varepsilon > 0$ existe $k_\varepsilon \in I$ tal que $\int_k^{\to b} g(x)\,dx < \varepsilon$ siempre que sea $k_\varepsilon < k < b$. Por ello y de [1] se desprende que

$$(k_\varepsilon < k < b) \quad \Rightarrow \quad \left|\int_k^{\to b} f(\lambda, x)\,dx\right| \leqslant \int_k^{\to b} g(x)\,dx < \varepsilon, \qquad \forall\ \lambda \in \Lambda$$

lo que prueba la convergencia uniforme de $\int_a^{\to b} f(\lambda, x)\,dx$ en Λ (véase [84],[2]).

[85]$_1$ Ejemplo

La integral impropia, dependiente del parámetro $y \in \mathbb{R}$:

$$I(y) = \int_{-\infty}^{+\infty} \frac{xy - 3}{(2 + 3x^2 + 4y^2)\sqrt{x^2y^2 + 3}}\,dx$$

es uniformemente convergente en $\mathbb{R}$ (para $y \in \mathbb{R}$). En efecto: es fácil comprobar que, para cualesquiera $x \in \mathbb{R}$ e $y \in \mathbb{R}$, es $-2 \leqslant (xy - 3)/\sqrt{x^2y^2 + 3} < 1$, luego

$$\left|\frac{xy - 3}{(2 + 3x^2 + 4y^2)\sqrt{x^2y^2 + 3}}\right| \leqslant \frac{2}{2 + 3x^2} < \frac{1}{1 + x^2}, \qquad \forall\ x \in \mathbb{R} \text{ y } \forall\ y \in \mathbb{R}$$

Como, por otra parte, es

$$\int_{-\infty}^{+\infty} \frac{1}{1 + x^2}\,dx = [\operatorname{arctg} x]_{-\infty}^{+\infty} = \pi \qquad \text{(esta integral es convergente)}$$

del criterio de Weierstrass se infiere la convergencia uniforme en $\mathbb{R}$ de $I(y)$.

(*) Criterio de Weierstrass o de comparación para integrales impropias (no paramétricas). Sean $\int_a^{\to b} \varphi(x)\,dx$ e $\int_a^{\to b} \psi(x)\,dx$ dos integrales (impropias en b; donde $-\infty \leqslant a < b \leqslant +\infty$). Si $\psi(x) \geqslant 0$ y $|\varphi(x)| \leqslant \psi(x)$ para todo $x \in [a, b[$ y si $\int_a^{\to b} \psi(x)\,dx$ existe y es convergente, entonces $\int_a^{\to b} \varphi(x)\,dx$ existe y es absolutamente convergente (y, por tanto, convergente).

$[85]_2$ Un criterio consecuencia del de Weierstrass

Sean dados un conjunto $\Lambda \subset \mathbb{R}^p$, un intervalo no compacto $I = [a, b[\subset \mathbb{R}$, con $-\infty < a < b \leqslant +\infty$, una función acotada $\varphi : \Lambda \times I \to \mathbb{R}$, $(\lambda, x) \mapsto \varphi(\lambda, x)$, tal que existe la integral $\int_a^k \varphi(\lambda, x)\, dx$ para cualquiera $\lambda \in \Lambda$ y $k \in I$, y una función $\psi : I \to \mathbb{R}$, $x \mapsto \psi(x)$, tal que la integral $\int_a^{\vec{b}} |\psi(x)|\, dx$ (impropia en b) es convergente. Se verifica entonces que la integral $\int_a^{\vec{b}} \varphi(\lambda, x)\psi(x)\, dx$ (impropia en b) existe y es uniformemente convergente en Λ (para $\lambda \in \Lambda$).

Comprobación

Si es K el supremo de $|\varphi(\lambda, x)|$ para $(\lambda, x) \in \Lambda \times I$, el anterior criterio se obtiene de aplicar el criterio de Weierstrass a las funciones (con la notación de [85], 2):

$$f(\lambda, x) = \varphi(\lambda, x)\psi(x) \qquad \text{y} \qquad g(x) = K|\psi(x)|$$

$[85]_3$ Observación (un criterio al estilo de los de Abel y Dirichlet)

Enunciamos, sin demostrar, un criterio para la convergencia de las integrales paramétricas impropias que es una adaptación, a estos casos, de los criterios de Abel y Dirichlet para series de funciones:

Sean dados un conjunto $\Lambda \subset \mathbb{R}^p$, un intervalo real no compacto $I = [a, b[$, con $-\infty < a < b \leqslant +\infty$, una función $\varphi : \Lambda \times I \to \mathbb{R}$, $(\lambda, x) \mapsto \varphi(\lambda, x)$, que es integrable en $[a, k]$ y existe una constante $H > 0$ tal que $|\int_a^k \varphi(\lambda, x)\, dx| < H$ para cualesquiera $k \in I$ y $\lambda \in \Lambda$, y una función $\psi : I \to \mathbb{R}$, $x \mapsto \psi(x)$, que decrece monótonamente hacia 0 cuando $x \to b^-$. Se verifica entonces que la integral $\int_a^{\vec{b}} \varphi(\lambda, x)\psi(x)\, dx$ es uniformemente convergente en Λ (para $\lambda \in \Lambda$).

Ejemplo

La siguiente integral $I(y)$ es uniformemente convergente para y variando en $[\alpha, +\infty[$, para cualquiera que sea $\alpha > 0$:

$$I(y) = \int_1^{+\infty} \frac{x \operatorname{sen}(xy)}{1 + x^2}\, dx$$

Para comprobarlo, basta con acudir al criterio anterior, en el que se toma $\varphi(x, y) = \operatorname{sen}(xy)$ y $\psi(x) = x/(1 + x^2)$, ya que, para $k \in [1, +\infty[$ y para $y \in [\alpha, +\infty[$, es:

$$\begin{cases} \left|\displaystyle\int_1^k \operatorname{sen}(xy)\, dx\right| = \left|\left[-\dfrac{1}{y} \cos(xy)\right]_{x=1}^{x=k}\right| = \left|\dfrac{1 - \cos(ky)}{y}\right| < \dfrac{2}{\alpha} \qquad (\text{es } H = 2/\alpha) \\ \psi(x) = \dfrac{x}{1 + x^2} \qquad \text{tiende monótonamente a 0 cuando } x \to +\infty \end{cases}$$

LÍMITES DE LAS INTEGRALES PARAMÉTRICAS IMPROPIAS

[86]

Considérense un conjunto $\Lambda \subset \mathbb{R}^p$, un punto $\lambda_0 \in \mathbb{R}^p$ de acumulación de Λ, un intervalo real no compacto $I = [a, b[$, con $-\infty < a < b \leqslant +\infty$, y una función $f: \Lambda \times I \to \mathbb{R}$, $(\lambda, x) \mapsto f(\lambda, x)$. Si, cuando $\lambda \to \lambda_0$, la función $\lambda \mapsto f(\lambda, x)$ tiende uniformemente(*) en I hacia una función $g: I \to \mathbb{R}$, $x \mapsto g(x)$, si existe y es convergente la integral $\int_a^{\to b} g(x)\,dx$ (impropia en b) y si la integral paramétrica $\int_a^{\to b} f(\lambda, x)\,dx$ (impropia en b) existe y es uniformemente convergente en Λ (para $\lambda \in \Lambda$), entonces $\int_a^{\to b} f(\lambda, x)\,dx$ tiende a $\int_a^{\to b} g(x)\,dx$ cuando $\lambda \to \lambda_0$, esto es:

$$\lim_{\lambda \to \lambda_0} \left(\int_a^{\to b} f(\lambda, x)\,dx \right) = \int_a^{\to b} \left(\lim_{\lambda \to \lambda_0} f(\lambda, x) \right) dx$$

(*) Este límite se dice «uniforme» en I (para $x \in I$) si, para cada $\varepsilon > 0$, existe un $\delta > 0$ (que depende sólo de ε; no depende de $x \in I$) tal que $|f(\lambda, x) - g(x)| < \varepsilon$ siempre que sea $\|\lambda - \lambda_0\| < \delta$ y para cualquiera que sea $x \in I$.

Demostración

Para cualesquiera que sean $k \in I$ y $\lambda \in \Lambda$, se puede poner:

$$\left| \int_a^{\to b} g(x)\,dx - \int_a^{\to b} f(\lambda, x)\,dx \right| \leqslant \left| \int_k^{\to b} g(x)\,dx \right| + \left| \int_a^{k} |g(x) - f(\lambda, x)|\,dx \right| + \left| \int_k^{\to b} f(\lambda, x)\,dx \right| \qquad [1]$$

Dado $\varepsilon > 0$ (cualquiera), como la integral $\int_a^{\to b} g(x)\,dx$ es convergente, existe $k'_\varepsilon \in I$ tal que

$$(k'_\varepsilon \leqslant k < b) \quad \Rightarrow \quad \left| \int_k^{\to b} g(x)\,dx \right| < \frac{\varepsilon}{3} \qquad [2]$$

Debido a la convergencia uniforme en Λ de la integral $\int_a^{\to b} f(\lambda, x)\,dx$, existe $k''_\varepsilon \in I$ tal que

$$(k''_\varepsilon \leqslant k < \varepsilon) \quad \Rightarrow \quad \left| \int_k^{\to b} f(\lambda, x)\,dx \right| < \frac{\varepsilon}{3}, \qquad \forall\ \lambda \in \Lambda \qquad [3]$$

Llamemos $k_\varepsilon = \text{máx}\,\{k'_\varepsilon, k''_\varepsilon\}$ y sea $d_\varepsilon = k_\varepsilon - a$. Por la convergencia uniforme en I de $f(\lambda, x)$ hacia $g(x)$ cuando $\lambda \to \lambda_0$, existe $\delta > 0$ tal que

$$\|\lambda - \lambda_0\| < \delta \quad \Rightarrow \quad |f(\lambda, x) - g(x)| < \frac{\varepsilon}{3d_\varepsilon}, \qquad \forall\ x \in I \qquad [4]$$

Resulta entonces que, por [1], [2], [3] y [4], se puede asegurar que:

$$\|\lambda - \lambda_0\| < \delta \quad \Rightarrow \quad \left| \int_a^{\to b} g(x)\,dx - \int_a^{\to b} f(\lambda, x)\,dx \right| \leqslant \left| \int_{k_\varepsilon}^{\to b} g(x)\,dx \right| + \left| \int_a^{k_\varepsilon} |g(x) - f(\lambda, x)|\,dx \right| +$$

$$+ \left| \int_{k_\varepsilon}^{\to b} f(\lambda, x)\,dx \right| < \frac{\varepsilon}{3} + d_\varepsilon \frac{\varepsilon}{3d_\varepsilon} + \frac{\varepsilon}{3} = \varepsilon$$

lo que prueba el teorema.

[86]$_1$ Ejercicio

Hállese el siguiente límite:

$$l = \lim_{\lambda \to 0^+} \lambda \int_0^{+\infty} e^{-\lambda x} \varphi(x)\,dx$$

sabiendo que $\varphi: [0, +\infty[\to \mathbb{R}$ es una función integrable en todo intervalo compacto de $[0, +\infty[$ y tal que $\varphi(x) \to 0$ cuando $x \to +\infty$.

Resolución

Comprobemos que es $l = 0$, para lo que acudimos al teorema anterior: con su notación, si es $f(\lambda, x) = \lambda e^{-\lambda x}\varphi(x)$, hemos de comprobar que: 1.º, la integral $\int_0^{+\infty} f(\lambda, x)\,dx$ es uniformemente convergente en un cierto intervalo $]0, a[$, con $a > 0$ (por ejemplo, para $0 < \lambda < 1$); 2.º, que el límite de $\lambda \mapsto f(\lambda, x)$ cuando $\lambda \to 0^+$, que evidentemente es $g(x) = 0$, es un límite uniforme para $x \in [0, +\infty[$, y 3.º, que $\int_0^{+\infty} g(x)\,dx$ converge y vale 0, lo que es evidente, pues $g(x) = 0$. Veamos que así ocurre:

1.º Dado $\varepsilon > 0$, como $\varphi(x) \to 0$ cuando $x \to +\infty$, existe $k_\varepsilon \in [0, +\infty[$ tal que $|x| > k_\varepsilon \Rightarrow |\varphi(x)| < \varepsilon$. Por tanto,

$$(k_\varepsilon < k) \quad \Rightarrow \quad \left| \int_k^{+\infty} f(\lambda, x)\,dx \right| = \left| \lambda \int_k^{+\infty} e^{-\lambda x}\varphi(x)\,dx \right| \leqslant \varepsilon \left| \lambda \int_k^{+\infty} e^{-\lambda x}\,dx \right| =$$

$$= \varepsilon \left| \lambda \left[\frac{e^{-\lambda x}}{\lambda} \right]_{x=k}^{x \to +\infty} \right| = \varepsilon |e^{-\lambda k}| < \varepsilon, \qquad \forall\, \lambda \in]0, 1[$$

2.º La función $\varphi: [0, +\infty[\to \mathbb{R}$ está acotada; sea H una cota suya. Dado $\varepsilon > 0$, si es $0 < \lambda < \delta$ con $\delta = \varepsilon/H$, se tiene:

$$0 < \lambda < \delta \quad \Rightarrow \quad |f(\lambda, x) - g(x)| = |\lambda e^{-\lambda x}\varphi(x)| < \delta \frac{1}{e^{\lambda x}} H \leqslant \varepsilon, \qquad \forall\, x \in [0, +\infty[$$

CONTINUIDAD DE LAS INTEGRALES PARAMÉTRICAS IMPROPIAS

[87]

Considérense un conjunto compacto $\Lambda \subset \mathbb{R}^p$, un intervalo real no compacto $I = [a, b[$, con $-\infty < a < b \leqslant +\infty$, y una función $f: \Lambda \times I$, $(\lambda, x) \mapsto f(\lambda, x)$. Si la integral paramétrica $\int_a^{\to b} f(\lambda, x)\, dx$ (impropia en b) existe y es uniformemente convergente en Λ (para $\lambda \in \Lambda$) y si la función f es continua en $\Lambda \times I$, entonces la función

$$\Lambda \to \mathbb{R}, \quad \lambda \mapsto \int_a^{\to b} f(\lambda, x)\, dx, \qquad \text{es continua en } \Lambda$$

(como Λ es compacto, la continuidad de esta función es continuidad uniforme en Λ).

Demostración

Para simplificar la notación, llamemos $\varphi(\lambda) = \int_a^{\to b} f(\lambda, x)\, dx$. Hemos de probar la continuidad de $\lambda \mapsto \varphi(\lambda)$ en cualquier punto $\lambda_0 \in \Lambda$. Para ello, tomando $k \in I$ y $\lambda \in \Lambda$ cualesquiera, pongamos:

$$|\varphi(\lambda) - \varphi(\lambda_0)| \leqslant \left| \int_a^k f(\lambda, x)\, dx - \int_a^k f(\lambda_0, x)\, dx \right| + \left| \int_k^{\to b} f(\lambda, x)\, dx \right| + \left| \int_k^{\to b} f(\lambda_0, x)\, dx \right| \quad [1]$$

Dado $\varepsilon > 0$ (cualquiera), como la integral $\int_a^{\to b} f(\lambda, x)\, dx$ es uniformemente convergente para $\lambda \in \Lambda$, existe un $k_\varepsilon \in I$ tal que, si en los dos últimos sumandos de la anterior desigualdad [1] se toma $k = k_\varepsilon$, cada uno de ellos es menor que $\varepsilon/3$, pues λ y λ_0 son de Λ. Tomando ahora $k = k_\varepsilon$ (fijo), sabemos (véase [80], I.*a*) que la función $\lambda \mapsto \int_a^{k_\varepsilon} f(\lambda, x)\, dx$ es continua en $\lambda_0 \in \Lambda$, por lo que existe $\delta > 0$ tal que

$$\|\lambda - \lambda_0\| < \delta \quad \Rightarrow \quad \left| \int_a^{k_\varepsilon} f(\lambda, x)\, dx - \int_a^{k_\varepsilon} f(\lambda_0, x)\, dx \right| < \frac{\varepsilon}{3}$$

Llevando estos resultados a [1], en donde se toma $k = k_\varepsilon$, nos encontramos con que, si se toma $\lambda \in \Lambda$ tal que $\|\lambda - \lambda_0\| < \delta$, entonces es $|\varphi(\lambda) - \varphi(\lambda_0)| < \varepsilon$, lo que prueba la continuidad de φ en λ_0.

[87]$_1$ Continuidad de la función Γ de Euler

Considérese la integral paramétrica (impropia en $+\infty$; impropia en 0 si es $p < 1$):

$$\Gamma(p) = \int_{\to 0}^{+\infty} e^{-x} x^{p-1}\, dx$$

De este modo se define una función continua $]0, +\infty[\to \mathbb{R}$, $p \mapsto \Gamma(p)$, que se llama «función gamma de Euler», o función euleriana de segunda especie. Más exactamente, se verifica que:

1.º $\Gamma(p)$ converge en $]0, +\infty[$ (o sea, si es $p > 0$) y diverge si es $p \leqslant 0$.
2.º Si es $0 < h < k < +\infty$, entonces $\Gamma(p)$ es uniformemente convergente en $[h, k]$.
3.º $\Gamma :]0, +\infty[\to \mathbb{R}$, $p \mapsto \Gamma(p)$ es una función continua en $]0, +\infty[$.

Comprobación

1.º Descompongamos $\Gamma(p)$ en suma de dos integrales impropias, cada una de ellas en un extremo, poniendo

$$\Gamma(p) = \int_{\to 0}^{1} e^{-x}x^{p-1}\,dx + \int_{1}^{+\infty} e^{-x}x^{p-1}\,dx = I_1(p) + I_2(p)$$

$I_2(p)$ converge para todo p, como se desprende al compararla con $\int_1^{+\infty} x^{-2}\,dx$ (que es convergente), ya que $e^{-x}x^{p-1}/x^2 \to 0$ para cualquiera que sea $p \in \mathbb{R}$. La integral $I_1(p)$ converge para $p > 0$ (nótese que esta integral es propia para $p \geqslant 1$) y diverge para $p \leqslant 0$, pues $I_1(p)$ tiene el mismo carácter que la integral $\int_{\to 0}^{1} x^{p-1}\,dx$, ya que $e^{-x} \to 1$ para $t \to 0^+$.

2. Para todo $p \in [h, k]$, y sin olvidar que es $x > 0$, se verifica que

$$e^{x}x^{p-1} \leqslant e^{-x}(x^{h-1} + x^{k-1}) \qquad [1]$$

Como h y k pertenecen al campo de convergencia de Γ, se sabe que la integral

$$\int_{\to 0}^{+\infty} e^{-x}(x^{h-1} + x^{k-1})\,dx$$

es convergente. Por ello y por [1], del criterio de Weierstrass (véase [85], 2) se desprende la convergencia uniforme de $\int_{\to 0}^{+\infty} e^{-x}x^{p-1}\,dx$ en $[h, k]$.

3. La función $p \mapsto \Gamma(p)$ es continua en su campo de definición $]0, +\infty[$. En efecto: para cualquiera que sea $p \in]0, +\infty[$, como la integral impropia $\Gamma(p)$ es uniformemente convergente en $[h, k]$, con $0 < h < p < k < +\infty$ (según se vio en el apartado anterior) y como su integrando $(p, x) \mapsto e^{-x}x^{p-1}$ es una función continua para $x \in]0, +\infty[$ y $p \in [h, k]$, del teorema [87] (de continuidad de las integrales paramétricas impropias) se desprende la continuidad de Γ en $[h, k]$ y, por tanto, en cualquier $p \in]0, +\infty[$.

[87]$_2$ Continuidad de la función B de Euler

Considérese la integral paramétrica (impropia en 0 si es $p < 1$ e impropia en 1 si es $q < 1$):

$$\mathrm{B}(p, q) = \int_{\to 0}^{\to 1} x^{p-1}(1 - x)^{q-1}\,dx$$

De este modo se define una función continua $]0, +\infty[\times]0, +\infty[, (p, q) \mapsto \mathrm{B}(p, q)$, que se llama «función beta de Euler», o función euleriana de primera especie. Más exactamente, se verifica que:

1.º $\mathrm{B}(p, q)$ converge en $]0, +\infty[\times]0, +\infty[$ (o sea, cuando es $p > 0$ y $q > 0$) y diverge si es $p \leqslant 0$ o $q \leqslant 0$.
2.º Si es $0 < h$ y $0 < k$, entonces la integral $\mathrm{B}(p, q)$ es uniformemente convergente en $[h, +\infty[\times [k, +\infty[$.
3.º $\mathrm{B}:]0, +\infty[\times]0, +\infty[\to \mathbb{R}, (p, q) \mapsto \mathrm{B}(p, q)$, es una función continua en su campo de definición $]0, +\infty[\times]0, +\infty[$.

Comprobación

1.º Si es $p \geqslant 1$ y $q \geqslant 1$, el integrando de $\mathrm{B}(p, q)$ es continuo y, por ello, la integral es propia y, obviamente, convergente. Supondremos, pues, que es $0 < p < 1$ o $0 < q < 1$, en cuyo caso la integral es impropia en 0 o en 1, respectivamente; para estudiar entonces por separado lo que acontece en uno y otro extremo de integración, pondremos:

$$\mathrm{B}(p, q) = \int_{\to 0}^{1/2} x^{p-1}(1-x)^{q-1}\, dx + \int_{1/2}^{\to 1} x^{p-1}(1-x)^{q-1}\, dx = \\ = I_1(p, q) + I_2(p, q)$$

$I_1(p, q)$ tiene el mismo carácter que la $\int_{\to 0}^{1/2} x^{p-1}\, dx$, como se deduce de que $(1-x)^{q-1} \to 1$ cuando $x \to 0^+$; como $\int_{\to 0}^{1/2} x^{p-1}\, dx$ converge para $p > 0$ y diverge para $p \leqslant 0$, lo mismo le acontece a $I_1(p, q)$. Para $I_2(p, q)$, comparándola con $\int_{1/2}^{\to 1}(1-x)^{q-1}\, dx$, se concluye igualmente que converge para $q > 0$ y diverge para $q \leqslant 0$. De todo ello se infiere la convergencia de $\mathrm{B}(p, q)$ siempre que sea $p > 0$ y $q > 0$.

2.º Para $p \geqslant h > 0$ y $q \geqslant k > 0$, como entonces es $h - 1 \leqslant p - 1$ y $k - 1 \leqslant q - 1$, se verifica que

$$x^{p-1}(1-x)^{q-1} < x^{h-1}(1-x)^{k-1}, \qquad \forall\, x \in]0, 1[$$

Por ello y por ser $\mathrm{B}(h, k)$ convergente, del criterio de Weierstrass (véase [85], 2) se desprende la convergencia de $\mathrm{B}(p, q)$ para $p \geqslant h$ y $q \geqslant k$.

3.º La función $(p, q) \mapsto \mathrm{B}(p, q)$ es continua en su campo de definición. En efecto: para cualesquiera que sean $p > 0$ y $q > 0$, como la integral impropia $\mathrm{B}(p, q)$ es uniformemente convergente en $[h, +\infty[\times [k, +\infty[$, con $0 < h < p$ y $0 < k < q$ (según se acaba de ver en el apartado anterior) y como su integrando $(p, q, x) \mapsto x^{p-1}(1-x)^{q-1}$ es una función continua para $x \in]0, 1[$, $p \in [h, +\infty[$ y $q \in [k, +\infty[$, del teorema [87] (de continuidad de las integrales paramétricas impropias) se desprende la continuidad de B en $[h, +\infty[\times [k, +\infty[$ y, por tanto, en cualquier $(p, q) \in]0, +\infty[\times]0, +\infty[$.

INTEGRACIÓN DE LAS INTEGRALES IMPROPIAS CON UN PARÁMETRO

[88]

1. INTEGRACIÓN PROPIA (de una integral impropia). Considérense un intervalo compacto $\Lambda = [\alpha, \beta] \subset \mathbb{R}$, un intervalo no compacto $I = [a, b[\subset \mathbb{R}$, con $-\infty < a < b \leqslant +\infty$, y una función $f: \Lambda \times I \to \mathbb{R}$, $(\lambda, x) \mapsto f(\lambda, x)$. Si la integral $\int_a^{\to b} f(\lambda, x)\, dx$ (impropia en b) existe y es uniformemente convergente en Λ (para $\lambda \in \Lambda$) y si la función f es continua en $\Lambda \times I$, entonces la función $\lambda \mapsto \int_a^{\to b} f(\lambda, x)\, dx$ (de Λ en $\mathbb{R}$) es integrable para $\lambda \in \Lambda$ y se verifica que

$$\int_\alpha^\beta \left(\int_a^{\to b} f(\lambda, x)\, dx \right) d\lambda = \int_a^{\to b} \left(\int_\alpha^\beta f(\lambda, x)\, d\lambda \right) dx \qquad [1]$$

(esto es: se puede permutar el orden en el que se realizan las integraciones).

2. INTEGRACIÓN IMPROPIA (de una integral impropia). Considérense dos intervalos no compactos $\Lambda = [\alpha, \beta[$ e $I = [a, b[$, con $-\infty < \alpha < \beta \leqslant +\infty$ y $-\infty < a < b \leqslant +\infty$, y una función $f: \Lambda \times I \to \mathbb{R}$, $(\lambda, x) \mapsto f(\lambda, x)$. Si la función f es continua y no negativa en $\Lambda \times I$, si la integral $\int_a^{\to b} f(\lambda, x)\, dx$ (impropia en b) es uniformemente convergente en Λ (para $\lambda \in \Lambda$), si la integral $\int_\alpha^{\to \beta} f(\lambda, x)\, d\lambda$ (impropia en β) es uniformemente convergente en I (para $x \in I$) y si una de las integrales iteradas

$$\int_\alpha^{\to \beta} \left(\int_a^{\to b} f(\lambda, x)\, dx \right) d\lambda \qquad \text{o} \qquad \int_a^{\to b} \left(\int_\alpha^{\to \beta} f(\lambda, x)\, d\lambda \right) dx \qquad [2]$$

es convergente, entonces la otra también converge y ambas son iguales.

Demostración

1. De acuerdo con el teorema anterior (véase [87]), la función $\lambda \mapsto \int_a^{\to b} f(\lambda, x)\, dx$ es continua en $\Lambda = [\alpha, \beta]$, por lo que es entonces integrable en $[\alpha, \beta]$; existe, pues, el primer miembro de [1], que es la igualdad que tenemos que verificar. El teorema quedará, entonces, demostrado si se comprueba que

$$\exists\ l = \lim_{k \to b^-} \int_a^k \left(\int_\alpha^\beta f(\lambda, x)\, d\lambda \right) dx, \qquad \text{y es} \quad l = \int_\alpha^\beta \left(\int_a^{\to b} f(\lambda, x)\, dx \right) d\lambda \qquad [3]$$

Ahora bien, según sabemos (por el teorema de integración de las integrales paramétricas; véase [82], I.a), para cualquiera que sea $k \in [a, b[$ se puede poner:

$$\int_\alpha^\beta \left(\int_a^{\to b} f(\lambda, x)\, dx \right) d\lambda = \int_\alpha^\beta \left(\int_a^k f(\lambda, x)\, dx \right) d\lambda + \int_\alpha^\beta \left(\int_k^{\to b} f(\lambda, x)\, dx \right) d\lambda =$$

$$= \int_a^k \left(\int_\alpha^\beta f(\lambda, x)\, d\lambda \right) dx + \int_\alpha^\beta \left(\int_k^{\to b} f(\lambda, x)\, dx \right) d\lambda$$

A la vista de esta igualdad y de lo dicho en [3], resulta obvio que todo se reduce a comprobar que

$$\lim_{k \to b^-} \int_\alpha^\beta \left(\int_k^{\to b} f(\lambda, x)\, dx \right) d\lambda = 0 \qquad [4]$$

Como la integral $\int_a^{\to b} f(\lambda, x)\, dx$ es uniformemente convergente para $\lambda \in [\alpha, \beta]$, dado $\varepsilon > 0$ (cualquiera), existe $k_\varepsilon \in [a, b[$ tal que

$$(k_\varepsilon < k < b) \quad \Rightarrow \quad \left| \int_k^{\to b} f(\lambda, x)\, dx \right| < \frac{\varepsilon}{\beta - \alpha}, \qquad \forall\ \lambda \in [\alpha, \beta]$$

por tanto, se verifica también que:

$$(k_\varepsilon < k < b) \quad \Rightarrow \quad \left| \int_\alpha^\beta \left(\int_k^{\to b} f(\lambda, x)\, dx \right) d\lambda \right| \leqslant \int_\alpha^\beta \left| \int_k^{\to b} f(\lambda, x)\, dx \right| d\lambda < \int_\alpha^\beta \frac{\varepsilon}{\beta - \alpha}\, d\lambda = \varepsilon$$

lo cual prueba que, en efecto, se verifica la igualdad [4].

2. Supongamos que es la segunda, de las dos integrales iteradas [2], la que se sabe convergente. Para cualquiera que sea $\beta' \in [\alpha, \beta[$, de acuerdo con el teorema anterior (aplicado en $[\alpha, \beta']$), se puede asegurar que es

$$\int_\alpha^{\beta'} \left(\int_a^{\to b} f(\lambda, x)\, dx \right) d\lambda = \int_a^{\to b} \left(\int_\alpha^{\beta'} f(\lambda, x)\, d\lambda \right) dx \leqslant \int_a^{\to b} \left(\int_\alpha^{\to \beta} f(\lambda, x)\, d\lambda \right) dx$$

(la desigualdad es cierta por ser f no negativa). El primer miembro de esta relación crece al crecer β' (pues f es no negativa) y está acotado por el último miembro (que es finito por hipótesis); por tanto, existirá su límite para $\beta' \to \beta^-$ y será:

$$\int_\alpha^\beta \left(\int_a^{\to b} f(\lambda, x)\, dx \right) d\lambda \leqslant \int_a^{\to b} \left(\int_\alpha^{\to \beta} f(\lambda, x)\, d\lambda \right) dx \qquad [5]$$

Esta relación [5] permite asegurar que su primer miembro es una integral convergente, pues lo es el segundo. Así pues, la primera de las integrales [2] es entonces convergente. Puede, por tanto, repetirse el anterior razonamiento intercambiando los papeles de las integrales iteradas de [2], de lo que se desprende que también se ha de verificar la desigualdad contraria de la [5]; de estas dos desigualdades se desprende, pues, que las dos integrales iteradas [2] son iguales, como había que comprobar.

[88]$_1$ Observación

En los teoremas anteriores ([87] y [88], sobre continuidad e integrabilidad de las integrales paramétricas impropias), si deja de cumplirse la condición de convergencia uniforme de las correspondientes integrales impropias, puede ocurrir que dejen de cumplirse las tesis de dichos teoremas. Así ocurre en el siguiente contraejemplo:

Para la función f, continua en su campo de definición:

$$f: [0, 1] \times [0, +\infty[\to \mathbb{R}, (\lambda, x) \mapsto f(\lambda, x) = e^{-\lambda x}\lambda x(\lambda x - 2)$$

es evidente que al integrar iteradamente en uno y otro orden, se tiene:

$$\int_0^{+\infty} \left(\int_0^1 f(\lambda, x)\, dx \right) d\lambda = \int_0^{+\infty} [-e^{-\lambda x}\lambda x^2]_{x=0}^{x=1}\, d\lambda =$$
$$= \int_0^{+\infty} -e^{-\lambda}\lambda\, d\lambda = [e^{-\lambda}(\lambda - 1)]_0^{+\infty} = -1$$

$$\int_0^1 \left(\int_0^{+\infty} f(\lambda, x)\, d\lambda \right) dx = \int_0^1 [-e^{-\lambda x}\lambda^2 x]_{\lambda=0}^{\lambda \to +\infty}\, dx = \int_0^1 0\, dx = 0$$

El que estas dos integrales iteradas sean distintas no contradice (aunque lo parezca) al teorema [88], 1 de integración, ya que la integral impropia $\int_0^{+\infty} f(\lambda, x)\, dx$ no es uniformemente convergente para $\lambda \in [0, 1]$, aun cuando sí que es convergente en este intervalo.

[88]$_2$ Relación entre las funciones Γ y B de Euler

Para cualesquiera que sean $p > 0$ y $q > 0$, se verifica que

$$\mathrm{B}(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p + q)}$$

Comprobación

Realizando el cambio de variable $x \mapsto t = yx$ (donde $y > 0$ es cualquiera, fijo) en la integral $\int_0^{+\infty} e^{-yx}x^{p-1}\, dx$, se obtiene que

$$\int_0^{+\infty} e^{-yx}x^{p-1}\, dx = \int_0^{+\infty} e^{-t}y^{-p}t^{p-1}\, dt = y^{-p}\,\Gamma(p) \qquad [1]$$

De ello se desprende que, para cualquiera que sea $y > 0$, es:

$$\Gamma(p) = \int_0^{+\infty} e^{-xy}x^{p-1}y^p\, dx$$

Multiplicando miembro a miembro la anterior igualdad por la $\Gamma(q) = \int_0^{+\infty} e^{-y}y^{q-1}\, dy$, es evidente que se obtiene:

$$\Gamma(p)\Gamma(q) = \left(\int_0^{+\infty} e^{-xy}x^{p-1}y^p\, dx \right)\left(\int_0^{+\infty} e^{-y}y^{q-1}\, dy \right)$$

Ahora bien, llamando $f(x, y) = e^{-y(x+1)}x^{p-1}y^{p+q-1}$, la integral $\int_0^{+\infty} \left(\int_0^{+\infty} f(x, y)\, dx\right) dy$ es convergente y, por ello, es aquí de aplicación el teorema [88], 2 (de integración de integrales impropias paramétricas), el cual permite poner que:

$$\Gamma(p)\Gamma(q) = \int_0^{+\infty} \left(\int_0^{+\infty} e^{-y(x+1)}x^{p-1}y^{p+q-1}\, dy\right) dx = \int_0^{+\infty} \left(\int_0^{+\infty} e^{-y(x+1)}y^{p+q-1}\, dy\right) x^{p-1}\, dx =$$

$$\overset{(*)}{=} \int_0^{+\infty} (1 + x)^{-(p+q)}\Gamma(p + q)x^{p+1}\, dx = \Gamma(p + q) \int_0^{+\infty} x^{p-1}(1 + x)^{-(p+q)}\, dx \overset{(**)}{=}$$

$$= \Gamma(p + q)\mathrm{B}(p, q)$$

(*) Acudir a la relación [1], cambiando los papeles de x por y y de y por $x + 1$.
(**) En $\mathrm{B}(p, q) = \int_0^1 t^{p-1}(1 - t)^{q-1}\, dt$, hágase el cambio $t = 1/(1 + x)$.

DERIVACIÓN DE LAS INTEGRALES PARAMÉTRICAS IMPROPIAS

[89]

Considérense un conjunto abierto y conexo $\Lambda \subset \mathbb{R}^p$, un intervalo real no compacto $I = [a, b[$, con $-\infty < a < b \leqslant +\infty$, y una función $f: \Lambda \times I \to \mathbb{R}$, $(\lambda, x) \mapsto f(\lambda, x)$, continua en $\Lambda \times I$ y que admite la derivada parcial $\partial f/\partial \lambda_i$ (donde λ_i es la componente i-ésima de λ, para cierto $i = 1, ..., p$), la cual es continua en $\Lambda \times I$. Si, además, la integral $\int_a^{\vec{b}} [\partial f(\lambda, x)/\partial \lambda_i]\, dx$ (impropia en b) es uniformemente convergente en Λ y, para algún $\lambda_0 \in \Lambda$, la integral $\int_a^{\vec{b}} f(\lambda_0, x)\, dx$ (impropia en b) es constante, entonces la función $\lambda \mapsto \int_a^{\vec{b}} f(\lambda, x)\, dx$ es convergente en Λ, es uniformemente convergente en todo compacto incluido en Λ y es derivable respecto de λ_i en Λ, siendo

$$\frac{\partial}{\partial \lambda_i} \int_a^{\vec{b}} f(\lambda, x)\, dx = \int_a^{\vec{b}} \frac{\partial f(\lambda, x)}{\partial \lambda_i}\, dx \qquad \left(\begin{smallmatrix}\text{REGLA DE}\\ \text{LEIBNIZ}\end{smallmatrix}\right)$$

Demostración

La demostración se realiza para el caso de ser $\lambda \in \mathbb{R}$ (es decir, $\Lambda \subset \mathbb{R}$, $p = 1$ o $\lambda_i = \lambda$) y ello no sólo porque sea más fácil (que lo es), sino debido a que, en lo referente a la derivación parcial respecto de λ_i, las restantes componentes de λ permanecen fijas, lo que hace que no sea difícil reducir el caso general al caso $\lambda \in \mathbb{R}$. En lo que sigue, supondremos, pues, que Λ es un intervalo abierto. Llamemos φ a la siguiente función, definida por una integral uniformemente convergente:

$$\varphi: \Lambda \to \mathbb{R},\ \lambda \mapsto \varphi(\lambda) = \int_a^{\vec{b}} \frac{\partial f(\lambda, x)}{\partial \lambda}\, dx$$

Para cualquiera que sea $\lambda \in \Lambda$, llamando Λ_λ al intervalo compacto de extremos λ y λ_0, es claro que a la función $\partial f/\partial\lambda$ le es de aplicación el anterior teorema [88], 1 en $\Lambda_\lambda \times I$, el cual permite poner que:

$$\begin{aligned}\int_{\lambda_0}^{\lambda} \varphi(\mu)\, d\mu &= \int_{\lambda_0}^{\lambda} \left(\int_a^{\to b} \frac{\partial f(\mu,\, x)}{\partial \mu}\, dx \right) d\mu = \int_a^{\to b} \left(\int_{\lambda_0}^{\lambda} \frac{\partial f(\mu,\, x)}{\partial \mu}\, d\mu \right) dx = \\ &= \int_a^{\to b} [f(\mu,\, x)]_{\mu=\lambda_0}^{\mu=\lambda}\, dx = \int_a^{\to b} f(\lambda,\, x)\, dx - \int_a^{\to b} f(\lambda_0,\, x)\, dx\end{aligned} \qquad [1]$$

Nótese que el primer miembro de la anterior cadena de igualdades existe para $\lambda \in \Lambda$ (pues φ es continua, por serlo $\partial f/\partial\lambda$ y según el teorema [87]), luego existe el último; como sabemos que $\int_a^{\to b} f(\lambda_0, x)\, dx$ es convergente, resulta que también lo es el otro sumando de este último miembro, que llamaremos $\psi(\lambda)$, es decir, que la integral

$$\psi(\lambda) = \int_a^{\to b} f(\lambda,\, x)\, dx$$

converge para todo $\lambda \in \Delta$. Como el anterior resultado [1] se puede expresar en la forma

$$\int_{\lambda_0}^{\lambda} \varphi(\mu)\, d\mu = \psi(\lambda) - \psi(\lambda_0)$$

y de esta expresión se infiere (por el teorema fundamental del cálculo; no se olvide que φ es continua en Λ_λ) que $\psi : \Lambda_\lambda \to \mathbb{R}$ es derivable en Λ_λ y que su derivada es φ, esto es, que

$$\frac{d}{d\lambda} \int_a^{\to b} f(\lambda,\, x)\, dx = \int_a^{\to b} \frac{\partial f(\lambda,\, x)}{\partial \lambda}\, dx, \qquad \forall\ \lambda \in \Lambda$$

como había que probar. Sólo nos queda, pues, por comprobar que la convergencia de la integral $\int_a^{\to b} f(\lambda, x)\, dx$ es uniforme en todo intervalo compacto $K \subset \Lambda$. Sea l la amplitud del intervalo K y tómese arbitrariamente un valor $\lambda_1 \in K$. Obsérvese en primer lugar que la relación [1] es igualmente válida si, en lugar de λ_0 y de $[a, b[$, se consideran λ_1 y $[k_1, k_2] \subset [a, b[$ cualquiera; al hacerlo de ese modo, [1] queda en la forma

$$\int_{\lambda_1}^{\lambda} \left(\int_{k_1}^{k_2} \frac{\partial f(\mu,\, x)}{\partial \mu}\, dx \right) d\mu = \int_{k_1}^{k_2} f(\lambda,\, x)\, dx - \int_{k_1}^{k_2} f(\lambda_1,\, x)\, dx, \qquad \forall\ \lambda \in K \qquad [2]$$

Por otra parte, como $\int_a^{\to b} [\partial f(\lambda, x)/\partial\lambda]\, dx$ es uniformemente convergente en K, por el criterio de Cauchy (véase [85], 2) sabemos que, dado $\varepsilon > 0$ (cualquiera), existe $k'_\varepsilon \in I$ tal que

$$(k'_\varepsilon < k_1 < k_2 < b) \quad \Rightarrow \quad \left| \int_{k_1}^{k_2} \frac{\partial f(\lambda,\, x)}{\partial \lambda}\, dx \right| < \frac{\varepsilon}{2l}, \qquad \forall\ \lambda \in K \qquad [3]$$

Nótese que como la integral impropia (no paramétrica) $\int_a^{\to b} f(\lambda_1, x)\,dx$ es convergente, se sabe que(*) (para el $\varepsilon > 0$ dado) existe un $k''_\varepsilon \in I$ tal que

$$(k''_\varepsilon < k_1 < k_2 < b) \quad \Rightarrow \quad \left| \int_{k_1}^{k_2} f(\lambda_1, x)\,dx \right| < \frac{\varepsilon}{2} \qquad [4]$$

Por tanto, de [2], [3] y [4] y llamando $k_\varepsilon = \text{máx}\,\{k'_\varepsilon, k''_\varepsilon\}$, se desprende que:

$$(k_\varepsilon < k_1 < k_2 < b) \quad \Rightarrow \quad \left| \int_{k_1}^{k_2} f(\lambda, x)\,dx \right| \leqslant \left| \int_{\lambda_1}^{\lambda} \left(\int_{k_1}^{k_2} \frac{\partial f(\mu, x)}{\partial \mu}\,dx \right) d\mu \right| +$$

$$+ \left| \int_{k_1}^{k_2} f(\lambda_1, x)\,dx \right| < |\lambda - \lambda_1| \frac{\varepsilon}{2l} + \frac{\varepsilon}{2} \leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \qquad \forall\, \lambda \in K$$

lo que, según el criterio de Cauchy [85], 2, asegura la convergencia uniforme de la integral $\int_a^{\to b} f(\lambda, x)\,dx$ en todo intervalo compacto $K \subset \Lambda$.

[89]₁ Cálculo de la integral de Dirichlet

Se llama integral de Dirichlet a la integral impropia $\int_0^{+\infty} [(\text{sen}\, x)/x]\,dx$ (nótese que en $x = 0$ es «seudoimpropia»). Se verifica que

$$\int_0^{+\infty} \frac{\text{sen}\, x}{x}\,dx = \frac{\pi}{2}$$

Demostración

Recurramos a la siguiente integral paramétrica:

$$\int_0^{+\infty} e^{-y} \cos(xy)\,dy = \left[\frac{e^{-y}}{x^2+1} (x\,\text{sen}\,(xy) - \cos(xy)) \right]_{y=0}^{y \to +\infty} = \frac{1}{x^2+1}, \qquad \forall\, x \in \mathbb{R} \quad [1]$$

Esta integral es uniformemente convergente para $x \in \mathbb{R}$, ya que $|e^{-y} \cos(xy)| \leqslant e^{-y}$ e $\int_0^{+\infty} e^{-y}\,dy$ converge. Por ello, al integrar en [1] para $x \in [0, k]$, se tiene:

$$\int_0^k \left(\int_0^{+\infty} e^{-y} \cos(xy)\,dy \right) dx = \int_0^k \frac{dx}{x^2+1} = \text{arctg}\, k$$

(*) Recuérdese que el criterio de Cauchy para una integral impropia $\int_a^{\to b} g(x)\,dx$ dice que ésta es convergente, si y sólo si, para cada $\varepsilon > 0$, existe un $h_\varepsilon \in [a, b[$ tal que

$$(h_\varepsilon < k < k' < b) \quad \Rightarrow \quad \left| \int_k^{k'} g(x)\,dx \right| < \varepsilon$$

Aquí se puede invertir el orden de integración y poner que:

$$\operatorname{arctg} k = \int_0^{+\infty} \left(\int_0^k \cos(xy)\, dx \right) e^{-y}\, dy = \int_0^{+\infty} e^{-y} \frac{\operatorname{sen}(ky)}{y}\, dy \qquad [2]$$

Esta última integral es uniformemente convergente, ya que:

$$\left| e^{-y} \frac{\operatorname{sen}(ky)}{y} \right| \leqslant \left| e^{-y} \frac{ky}{y} \right| = ke^{-y}$$

y la integral $\int_0^{+\infty} ke^{-y}\, dy$ es convergente. Por tanto, en [2] se puede integrar, para k variando en $[0, x]$, de lo cual se obtiene:

$$\int_0^x \left(\int_0^{+\infty} e^{-y} \frac{\operatorname{sen}(ky)}{y}\, dy \right) dk = \int_0^x \operatorname{arctg} k\, dk = x \operatorname{arctg} x - \frac{1}{2} \operatorname{L}(1 + x^2)$$

Como, por la convergencia uniforme, se puede invertir el orden de integración, al hacerlo, se obtiene obviamente que:

$$\int_0^{+\infty} e^{-y} \frac{1 - \cos(xy)}{y^2}\, dy = x \operatorname{arctg} x - \frac{1}{2} \operatorname{L}(1 + x^2)$$

Como esta igualdad se verifica para todo $x > 0$, seguirá siendo cierta si se sustituye x por $1/x$, de lo que se obtiene:

$$\int_0^{+\infty} e^{-y} \frac{1 - \cos(y/x)}{y^2}\, dy = \frac{1}{x} \operatorname{arctg} \frac{1}{x} - \frac{1}{2} \operatorname{L}\left(1 + \frac{1}{x^2}\right)$$

En esta última integral, hagamos el cambio de variable $y \mapsto t = y/x$; con ello, se llega a:

$$\int_0^{+\infty} e^{-xt} \frac{1 - \cos t}{t^2}\, dt = \operatorname{arctg} \frac{1}{x} - \frac{x}{2} \operatorname{L}\left(1 + \frac{1}{x^2}\right)$$

Si ahora tomamos límites para $x \to 0^+$ (véase [86]; la integral es uniformemente convergente), obtenemos que:

$$\int_0^{+\infty} \frac{1 - \cos t}{t^2}\, dt = \frac{\pi}{2}$$

Al integrar por partes en la integral de Dirichlet, se concluye:

$$\int_0^{+\infty} \frac{\operatorname{sen} x}{x}\, dx = \left[\frac{1 - \cos x}{x} \right]_0^{+\infty} + \int_0^{+\infty} \frac{1 - \cos x}{x^2}\, dx = 0 + \frac{\pi}{2}$$

LA FUNCIÓN GAMMA (Γ) DE EULER

Esta función, de la que ya algo se adelantó en $[87]_1$ y que tiene no poco interés, se define mediante una integral $\left(\Gamma(p) = \int_0^{+\infty} e^{-x}x^{p-1}\,dx\right)$ y, al no poderse expresar mediante funciones elementales, hay que estudiarla acudiendo a las propiedades de las integrales paramétricas impropias, que permiten obtener información, acerca de la integral, a partir del análisis de su integrando.

[90]

Se llama «gamma de Euler» (o integral euleriana de segunda especie) a la función $p \mapsto \Gamma(p)$ que viene dada por la expresión:

$$\Gamma(p) = \int_{\to 0}^{+\infty} e^{-x}x^{p-1}\,dx, \qquad \forall\, p \in\,]0, +\infty[\tag{1}$$

Esta función es continua e infinitas veces derivable en $]0, +\infty[$, siendo

$$\Gamma^{(n)}(p) = \int_{\to 0}^{+\infty} (\mathrm{L}\,x)^n e^{-x}x^{p-1}\,dx, \qquad \forall\, n \in \mathbb{N} \tag{2}$$

Otras expresiones útiles de $\Gamma(p)$ se obtienen al realizar en [1] los cambios de variable $x \mapsto y$ dados por $x = ky$, por $x = e^y$, por $x = -\mathrm{L}\,y$ y por $x = y^2$.

Se verifica la relación de recurrencia $\Gamma(p+1) = p\Gamma(p)$, para $p > 0$; con ella, si se conociera $\Gamma(p)$ para $p \in\,]0, 1]$, se conociera $\Gamma(p)$ para todo $p > 0$. Como $\Gamma(1) = 1$, resulta que:

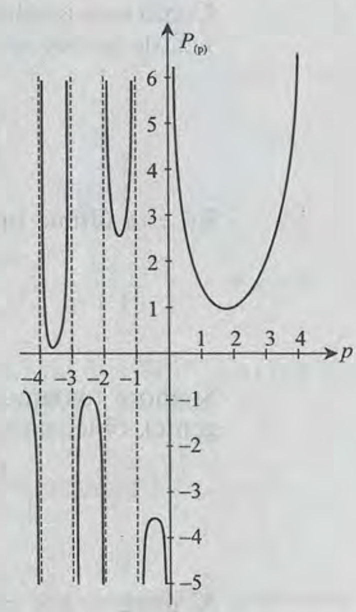
$$\Gamma(n+1) = n!, \qquad \forall\, n \in \mathbb{N} \tag{3}$$

Para $p \to 0^+$, $\Gamma(p)$ es un infinito equivalente al $1/p$.

La anterior relación de recurrencia, aplicada reiteradamente, conduce a:

$$\Gamma(p) = \frac{\Gamma(p+n)}{(p+n-1)(p+n-2)\cdots p}, \quad \forall\, p > 0 \text{ y } \forall\, n \in \mathbb{N} \tag{4}$$

Se utiliza esta expresión para extender Γ a $\mathbb{R}-\mathbb{Z}^-$, lo cual se hace así: dado $p \in \mathbb{R}-\mathbb{Z}^-$, existe $n_p \in \mathbb{N}$ tal que $-(n_p+1) < p < -n_p$; se define $\Gamma(p)$ por la expresión [4] tomando en ella $n = n_p$.



Siempre que sea $0 < p < 1$, se verifica que

$$\Gamma(p)\Gamma(1-p) = \frac{\pi}{\operatorname{sen}(p\pi)} \quad \left(\text{en particular } \Gamma(1/2) = \sqrt{\pi}\right)$$

Esta fórmula, junto a la $\Gamma(p+1) = p\Gamma(p)$ reduce el cálculo de $\Gamma(p)$ para todo $p > 0$ al cálculo de los $\Gamma(p)$ correspondientes a $1/2 \leqslant p < 1$.

Comprobación

1. La existencia y continuidad de Γ para $p \in]0, +\infty[$ ya se estudió en $[87]_1$; ocupémonos ahora de su derivabilidad. Para ello, acudiremos al teorema de derivación de las integrales paramétricas impropias (véase [89]); de él se deduce que Γ admite derivada de cualquier orden y que, según la regla de Leibniz, es:

$$\Gamma'(p) = \int_{\to 0}^{+\infty} (\mathrm{L}\,x) e^{-x} x^{p-1}\,dx, \quad \ldots, \quad \Gamma^{(n)}(p) = \int_{\to 0}^{+\infty} (\mathrm{L}\,x)^n e^{-x} x^{p-1}\,dx$$

ya que la integral impropia que define a cada una de las funciones que se derivan es uniformemente convergente en todo intervalo compacto $[p_1, p_2]$, como pasamos a comprobar. Nótese que existe una constante k tal que:

$$(0 < x \leqslant 1) \;\Rightarrow\; |(\mathrm{L}\,x)^n x^{(p_1/2)}| \leqslant k \qquad \text{y} \qquad (1 \leqslant x < +\infty) \;\Rightarrow\; |(\mathrm{L}\,x)^n x^{-p_2}| \leqslant k$$

lo que permite poner que para todo $p \in [p_1, p_2]$ es:

$$|(\mathrm{L}\,x)^n e^{-x} x^{p-1}| \leqslant \begin{cases} k e^{-x} x^{p-(p_1/2)-1} \leqslant k e^{-x} x^{(p_1/2)-1}, & \text{si } 0 < x \leqslant 1 \\ k e^{-x} x^{p+p_2-1} \leqslant k e^{-x} x^{2p_2-1}, & \text{si } 1 \leqslant x < \infty \end{cases}$$

es decir, el integrando está mayorado, en módulo y para $p \in [p_1, p_2]$, por una función cuya integral es convergente (ya que $p_1/2$ y $2p_2$ son positivos). Del criterio de Weierstrass (véase [85], 2) se desprende, pues, la convergencia uniforme de la integral $\int_{\to 0}^{+\infty} (\mathrm{L}\,x)^n e^{-x} x^{p-1}\,dx$.

2. Con el cambio de variable $x = y^2$, se obtiene que:

$$\Gamma(p) = 2 \int_{\to 0}^{+\infty} e^{-y^2} y^{2p-1}\,dy$$

Si aquí integramos por partes (para $p > 0$), se obtiene que

$$\Gamma(p+1) = \int_{0}^{+\infty} e^{-x} x^p\,dx = [-e^{-x} x^p]_{x=0}^{x\to+\infty} + p \int_{\to 0}^{+\infty} e^{-x} x^{p-1}\,dx = 0 + p\Gamma(p)$$

El valor de $\Gamma(1)$ se obtiene integrando directamente:

$$\Gamma(1) = \int_{0}^{+\infty} e^{-x}\,dx = [-e^{-x}]_0^{+\infty} = 1$$

Si $n \in \mathbb{N}$, aplicando reiteradamente la anterior fórmula recurrente $\Gamma(p+1) = p\Gamma(p)$, se obtiene que:

$$\Gamma(n+1) = n\Gamma(n) = n(n-1)\Gamma(n-1) = \cdots = n(n-1)\cdots 1\Gamma(1) = n!$$

Como, para $p > 0$, es $\Gamma(p) = \Gamma(p + 1)/p$ y como $\Gamma(p + 1) \to \Gamma(1) = 1$ cuando $p \to 0^+$, resulta que, cuando $p \to 0^+$, $\Gamma(p)$ es un infinito equivalente al $1/p$.

Al aplicar reiteradamente la relación $\Gamma(p) = \Gamma(p + 1)/p$, resulta que para $p > 0$ y $n \in \mathbb{N}$ es:

$$\Gamma(p) = \frac{\Gamma(p+1)}{p} = \frac{\Gamma(p+2)}{(p+1)p} = \cdots = \frac{\Gamma(p+n)}{(p+n-1)(p+n-2)\cdots p}$$

3. La fórmula $\Gamma(p)\Gamma(1-p) = \pi/[\text{sen}\,(p\pi)]$ no es fácil de probar; no obstante, algo se dirá sobre ella en el próximo apartado, como consecuencia de las relaciones que ligan a las funciones Γ y B (véase el apartado 5 de la demostración de [91]).

[90]$_1$ Ejercicio

Calcular el valor de las siguientes integrales impropias:

$$I = \int_0^{+\infty} e^{-ax^2}\,dx \qquad (a > 0 \text{ dado})$$

$$J = \int_0^{+\infty} x^4 e^{-5x^2}\,dx$$

Resolución

1. Hagamos el cambio de variable $ax^2 = t$:

$$I = \frac{1}{2\sqrt{a}}\int_0^{+\infty} e^{-t}t^{-1/2}\,dt = \frac{1}{2\sqrt{a}}\,\Gamma\left(\frac{1}{2}\right) = \frac{1}{2}\sqrt{\frac{\pi}{a}}$$

2. Hagamos el cambio de variable $5x^2 = t$:

$$J = \frac{1}{50\sqrt{5}}\int_0^{+\infty} e^{-t}t^{3/2}\,dt = \frac{1}{50\sqrt{5}}\,\Gamma\left(\frac{5}{2}\right) = \frac{1}{50\sqrt{5}}\cdot\frac{3}{2}\cdot\frac{1}{2}\,\Gamma\left(\frac{1}{2}\right) = \frac{3}{200}\sqrt{\frac{\pi}{5}}$$

❑ LA FUNCIÓN BETA (B) DE EULER

Al igual que con la «función gamma», también vamos a ocuparnos de esta otra «función euleriana», llamada función «beta de Euler», íntimamente relacionada con aquella y que se estudia con las mismas técnicas que se utilizaron en el apartado anterior.

[91]

Se llama «beta de Euler» (o integral euleriana de primera especie) a la función $(p, q) \mapsto \mathrm{B}(p, q)$, continua en su campo de definición, que viene dada por la expresión:

$$\mathrm{B}(p, q) = \int_0^1 x^{p-1}(1-x)^{q-1}\,dx, \qquad \forall\ p, q \in]0, +\infty[\qquad [1]$$

Mediante el cambio de variable $x \mapsto y = 1 - x$, se obtiene que $\mathrm{B}(p, q) = \mathrm{B}(q, p)$. Otras expresiones interesantes de $\mathrm{B}(p, q)$ son:

$$\mathrm{B}(p, q) = \int_0^{+\infty} \frac{t^{p-1}}{(1+t)^{p+q}}\,dt = \int_0^1 \frac{u^{p-1}+u^{q-1}}{(1+u)^{p+q}}\,du = 2\int_0^{\pi/2} (\operatorname{sen}^{2p-1}\theta)(\cos^{2q-1}\theta)\,d\theta$$

De esta última se obtiene en particular que $\mathrm{B}(1/2, 1/2) = \pi$.

Se verifican las siguientes relaciones de recurrencia:

$$\mathrm{B}(p, q+1) = \frac{q}{p+q}\,\mathrm{B}(p, q) \qquad \text{y} \qquad \mathrm{B}(p+1, q) = \frac{p}{p+q}\,\mathrm{B}(p, q)$$

con ellas, si se conociera $\mathrm{B}(p, q)$ para $p, q \in]0, 1]$, se conocería $\mathrm{B}(p, q)$ para cualesquiera $p, q > 0$. Como $\mathrm{B}(1, 1) = 1$, resulta que

$$\mathrm{B}(m, n) = \frac{(m-1)!(n-1)!}{(m+n-1)!}, \qquad \forall\ m, n \in \mathbb{N}$$

Las funciones Γ y B están ligadas mediante la relación

$$\mathrm{B}(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$$

De esta fórmula se desprende que $\Gamma(1/2) = \sqrt{\pi}$ y que $\int_0^{+\infty} e^{-x^2}\,dx = 1/2\sqrt{\pi}$.

Comprobación

1. La existencia y continuidad de B para $p, q \in]0, +\infty[$ ya se estudió en $[87]_2$.
2. Las otras tres expresiones de $\mathrm{B}(p, q)$ se obtienen, a partir de la expresión que define a $\mathrm{B}(p, q)$, del siguiente modo: la primera con el cambio $x = 1/(1+t)$; la segunda dividiendo el intervalo de integración mediante el punto $t = 1$ y, en la integral extendida a $[1, +\infty[$, realizando el cambio $u = 1/t$; la tercera con el cambio de variable $x = \cos^2\theta$ (realizado en la integral con la que se definió a B).

3. La fórmula de recurrencia se obtiene utilizando la última expresión de $B(p, q)$ (como una integral trigonométrica) y recurriendo a las fórmulas de reducción(*) para integrales del tipo $\int \operatorname{sen}^m \theta \cos^n \theta \, d\theta$.
4. La relación que liga a Γ y B ya se comprobó en $[88]_2$.
5. Comprobemos ahora la fórmula $\Gamma(p)\Gamma(p-1) = \pi/[\operatorname{sen}(p\pi)]$, que se dio en [90], pero que quedó allí sin demostrar. Para ello, nótese que, según la anterior relación de recurrencia, es

$$B(p, 1-p) = \frac{\Gamma(p)\Gamma(1-p)}{\Gamma(1)} = \Gamma(p)\Gamma(1-p) \qquad [1]$$

y que, por otra parte, se verifica que

$$B(p, 1-p) = \int_0^{+\infty} \frac{t^{p-1}}{1+t}\, dt = \frac{\pi}{\operatorname{sen}(p\pi)} \qquad [2]$$

(lo que se puede comprobar recurriendo al método de integración de los residuos, pasando al campo complejo). De [1] y [2] se concluye obviamente la fórmula a demostrar.

$[91]_1$ Ejercicio

Calcular la siguiente integral, reduciéndola a una integral euleriana de segunda especie:

$$I = \int_0^{\pi} \frac{\sqrt{\operatorname{sen} x}}{(5 + 3\cos x)^{3/2}}\, dx$$

Resolución

Con el cambio de variable $\operatorname{tg}(x/2) = t$ (con lo que se tiene: $\operatorname{sen} x = (2t)/(1+t^2)$, $\cos x = (1-t^2)/(1+t^2)$ y $dx = 2dt/(1+t^2)$), se obtiene que

$$I = \int_0^{+\infty} \frac{\sqrt{t}}{(4+t^2)^{3/2}}\, dt$$

Haciendo ahora el nuevo cambio $t^2 = 4z$, se llega a que

$$I = \frac{\sqrt{2}}{8}\int_0^{+\infty} z^{-1/4}(1+z)^{-3/2}\, dz = \frac{\sqrt{2}}{8} B\left(\frac{3}{4}, \frac{3}{4}\right) = \frac{\sqrt{2}}{8}\,\frac{\Gamma(3/4)\Gamma(3/4)}{\Gamma(3/2)} = \frac{[\Gamma(3/4)]^2}{2\sqrt{2\pi}}$$

$$(*) \int \operatorname{sen}^m \theta \cos^n \theta \, d\theta = \begin{cases} \dfrac{\operatorname{sen}^{m+1}\theta \cos^{n-1}\theta}{m+n} + \dfrac{n-1}{m+n}\displaystyle\int \operatorname{sen}^m \theta \cos^{n-2}\theta \, d\theta \\ -\dfrac{\operatorname{sen}^{m-1}\theta \cos^{n+1}\theta}{m+n} + \dfrac{m-1}{m+n}\displaystyle\int \operatorname{sen}^{m-2} \theta \cos^{n}\theta \, d\theta \end{cases}$$

Ejercicios y problemas

ENUNCIADOS

5.1. Sea $f: \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}$ la función que vale cero en todos los puntos excepto en los siguientes cuadrados abiertos de lado unidad y lados paralelos a los ejes:

$$\begin{cases} \text{los que tienen su centro en } (n/2,\ n/2), \\ \quad \text{para } n \in \mathbb{N}; \text{ en ellos } f \text{ vale } 1 \\ \text{los que tienen su centro en } (n/2,\ 1+n/2), \\ \quad \text{para } n \in \mathbb{N}; \text{ en ellos, } f \text{ vale } -1 \end{cases}$$

Estudiar si f es integrable en $C = \mathbb{R}^+ \times \mathbb{R}^+$.

5.2. Estúdiese si existe la integral doble impropia

$$\iint_C \operatorname{sen}(x^2 + y^2)\, dx\, dy, \quad \text{donde } C = \mathbb{R}^+ \times \mathbb{R}^+$$

5.3. Analizar la convergencia de la integral impropia

$$\iint_C \frac{dx\, dy}{x - y}, \qquad \text{donde}$$

$$C = \{(x, y) \in \mathbb{R}^2 \;/\; 0 \leqslant x \leqslant 1,\ 0 \leqslant y \leqslant x\}$$

5.4. Calcular las siguientes integrales, en las que $C = [0, 1] \times [0, 1]$:

$$I_1 = \iint_C \frac{dx\, dy}{\sqrt{xy}} \qquad \text{e} \qquad I_2 = \iint_C \frac{dx\, dy}{\sqrt{|x - y|}}$$

5.5. Acudiendo a la siguiente integral doble I, calcular la integral J, siendo $C = [0, +\infty[\times [1, 2]$:

$$J = \int_0^{+\infty} \frac{e^{-x} - e^{-2x}}{x}\, dx, \quad I = \iint_C e^{-xy}\, dx\, dy$$

5.6. Calcular el valor de la integral I y estudiar si es convergente la integral J, siendo

$$I = \iint_C \frac{dx\, dy}{\sqrt{1 - x^2 - y^2}}, \; J = \iint_C \frac{\operatorname{sen}^3(x + y)}{\sqrt{1 - x^2 - y^2}}\, dx\, dy$$

$$C = \{(x, y) \in \mathbb{R}^2 \;/\; x^2 + y^2 < 1\}$$

5.7. Hallar las relaciones a las que deben satisfacer las reales α, β y γ para que la siguiente integral I sea convergente, hallando entonces su valor:

$$I = \iint_{\mathbb{R}^2} e^{-(\alpha x^2 + 2\beta xy + \gamma y^2)}\, dx\, dy$$

(recúrrase a la integral $J = \iint_{\mathbb{R}^2} e^{-(u^2 + v^2)}\, du\, dv$, que vale $J = \pi$).

5.8. Calcular la siguiente integral doble impropia

$$I = \iint_C x^3 e^{-(x^2 + y^2)^{5/2}}\, dx\, dy, \qquad \text{donde}$$

$$C = \{(x, y) \in \mathbb{R}^2 \;/\; x \geqslant 0\}$$

5.9. Hallar la integral impropia I. Estudiar la existenia de la integral J y, en su caso, hallarla

$$I = \iint_{\mathbb{R}^2} |xy|\, e^{-|x+y|}\, dx\, dy,$$

$$J = \iint_{\mathbb{R}^2} xy\, e^{-|x+y|}\, dx\, dy$$

5.10. Estudiar la convergencia de la integral impropia $I = \iint_C f$, siendo:

$$f(x, y) = \left(\frac{x}{x^2 + y^2 + 1}\right)^n \quad (n > 2) \quad C = \mathbb{R}^+ \times \mathbb{R}^+$$

5.11. Calcular las siguientes integrales iteradas:

$$I = \int_0^{+\infty} dx \int_x^{+\infty} e^{-y^2}\, dy$$

$$J = \int_0^{+\infty} dx \int_x^{+\infty} (x/y^2) e^{-y} \operatorname{sen} y\, dy$$

5.12. Calcular el volumen V del cuerpo salido limitado por el plano $z = 0$ y por la superficie $z = f(x, y)$ en los dos casos siguientes:

1.º $f(x, y) = (x^2 + y^2) e^{-(x^2 + y^2)}$
2.º $f(x, y) = x^2 y^2 e^{-(x^2 + y^2)}$

5.13. Calcular las siguientes integrales dobles impropias:

$$I = \iint_C e^{x/y}\,dx\,dy,$$

donde $C = \{(x, y) \in \mathbb{R}^2 \,/\, 0 \leqslant x \leqslant y^2,\ 0 < y < 1\}$

$$J = \iint_C (x^4 + y^2)^{-1}\,dx\,dy,$$

donde $C = \{(x, y) \in \mathbb{R}^2 \,/\, x \geqslant 1,\ y \geqslant x^2\}$

5.14. Estudiar si existe y si es convergente la integral doble impropia $I = \iint_C f$, donde

$$f(x, y) = \frac{\operatorname{sen} x \operatorname{sen} y}{(\sqrt{x^2 + y^2})^\alpha} \quad (\alpha > 0)$$

$$C = \{(x, y) \in \mathbb{R}^2 \,/\, 0 < x^2 + y^2 \leqslant 1\}$$

5.15. Estudiar la convergencia de las siguientes integrales triples impropias, en las que $\rho = \sqrt{x^2 + y^2 + z^2}$ y $C = \{(x, y, z) \in \mathbb{R}^3 \,/\, 0 < \rho \leqslant 1\}$:

1.ª $\displaystyle\iiint_C \frac{dx\,dy\,dz}{\rho^3\,\mathrm{L}\,\rho^{2/3}}$ 2.ª $\displaystyle\iiint_C \frac{(\mathrm{L}\,\rho)\,dx\,dy\,dz}{\rho^2}$

3.ª $\displaystyle\iiint_C \frac{xyz}{\rho^{3/2}}\,dx\,dy\,dz$

5.16. Estudiar la convergencia de la siguiente integral doble impropia:

$$I = \iint_C x^{-3/2} e^{y-x}\,dx\,dy, \qquad \text{donde}$$

$$C = \{(x, y) \in \mathbb{R}^2 \,/\, 0 < x,\ 0 \leqslant y \leqslant x\}$$

5.17. Estudiar la convergencia de la integral en $\mathbb{R}^3$ de $\rho^{-\rho}$, donde $\rho = \sqrt{x^2 + y^2 + z^2}$.

5.18. Calcular la integral doble impropia $I = \iint_C f$, siendo

$$f(x, y) = \frac{dx\,dy}{(1+y)(1+x^2y)} \qquad \text{y}$$

$$C = \{(x, y) \in \mathbb{R}^2 \,/\, x \geqslant 0,\ y \geqslant 0\}$$

A partir del valor obtenido, calcular:

$$J = \int_0^{+\infty} \frac{\mathrm{L}\,(x^2)}{1 - x^2}\,dx$$

5.19. Calcular la derivada $f'(y)$, donde $f(y)$ es la integral paramétrica:

$$f(y) = \int_{y^2}^{y} \frac{\operatorname{sen}(xy)}{x}\,dx$$

5.20. Sea $f: \mathbb{R} \to \mathbb{R}$ una función de clase $\mathscr{C}^1$. Si f es una función impar, pruébese que la siguiente función $F: \mathbb{R} \to \mathbb{R}$ es constante

$$F(y) = \int_{-y}^{y} \big(f(x + y) + f(x - y)\big)\,dx$$

5.21. Calcular la integral $I(a) = \int_0^\pi \mathrm{L}\,(1 + a \cos x)\,dx$, donde $a > -1$.

5.22. Calcular la integral $I_n(a) = \int_0^1 x^a (\mathrm{L}\,x)^n\,dx$, para $n \in \mathbb{N}$ y siendo $a \neq -1$ (indicación: recurrir a la integral $J(a) = \int_0^1 x^a\,dx$.

5.23. Calcular la siguiente integral $I(a)$ (indicación: acudir a la integral $J(a)$):

$$I(a) = \int_0^{\pi/2} \frac{dx}{(\operatorname{tg} x + a)^2} \qquad J(a) = \int_0^{\pi/2} \frac{dx}{\operatorname{tg} x + a}$$

5.24. Pruébese que para $a > 0$ y $n \in \mathbb{N}$, se verifica que

$$\int_0^x \frac{dt}{(t^2 + a)^n} = \frac{(-1)^{n+1}}{(n-1)!}\,\frac{\partial^{n-1}\!\left(\dfrac{1}{\sqrt{a}} \operatorname{arctg} \dfrac{x}{\sqrt{a}}\right)}{\partial a^{n-1}}$$

5.25. Calcular las integrales

$$I_1 = \int_0^\pi \frac{\cos^2 x\,dx}{(a^2 \cos^2 x + b^2 \operatorname{sen}^2 x)} \qquad \text{e}$$

$$I_2 = \int_0^\pi \frac{\operatorname{sen}^2 x\,dx}{(a^2 \cos^2 x + b^2 \operatorname{sen}^2 x)}$$

(indicación: obtener precisamente que $I = \int_0^\pi f(x)\,dx = \pi/(ab)$, siendo $f(x) = (a^2 \cos^2 x + b^2 \operatorname{sen}^2 x)^{-1}$).

5.26. Calcular la siguiente integral $I(\lambda)$, donde $|\lambda| < 1$:

$$I(\lambda) = \int_0^{\pi/2} \frac{\mathrm{L}\ (1 - \lambda^2 \operatorname{sen}^2 x)}{\operatorname{sen} x}\,dx$$

5.27. Calcular la siguiente integral $I(\lambda)$, donde $\lambda > -1$:

$$I(\lambda) = \int_0^{\pi/2} \frac{\mathrm{L}\,(1 + \lambda \operatorname{sen}^2 x)}{\operatorname{sen}^2 x}\, dx$$

5.28. Calcular la integral:

$$I(\lambda) = \int_0^{\lambda} \mathrm{L}\,[1 + (\operatorname{tg} \lambda)(\operatorname{tg} x)]\, dx$$

5.29. Calcular la siguiente integral $I(\lambda)$, para $0 \leqslant \lambda \leqslant \pi/2$:

$$I(\lambda) = \int_{\pi/2-\lambda}^{\pi/2} \operatorname{sen} x \arccos\left(\frac{\cos \lambda}{\operatorname{sen} x}\right) dx$$

5.30. Calcular la siguiente integral I (indicación: calcular primero la inmtegral $J(\lambda)$):

$$I = \int_0^{\pi} \frac{dx}{(2 - \cos x)^2} \qquad J(\lambda) = \int_0^{\pi} \frac{dx}{\lambda - \cos x}$$

5.31. Calcular la siguiente integral I (indicación: recurrir a la integral $J(a)$):

$$I = \int_0^1 \frac{\mathrm{L}\, x}{\sqrt{x}}\, dx \qquad J(a) = \int_0^1 x^a\, dx$$

5.32. Calcular la siguiente integral I (indicación: recurrir a la integral $J(t)$):

$$I = \int_0^1 \frac{\operatorname{arctg} x}{x\sqrt{1-x^2}}\, dx \qquad J(t) = \int_0^1 \frac{\operatorname{arctg}(tx)}{x\sqrt{1-x^2}}\, dx$$

5.33. Calcular la integral

$$I(\lambda) = \int_0^{\pi/2} \frac{\operatorname{arctg}(\lambda \operatorname{tg} x)}{\operatorname{tg} x}\, dx$$

5.34. Calcular por derivación paramétrica:

$$I = \int_0^{+\infty} \frac{\operatorname{arctg}(ax)}{x(1 + x^2)}\, dx$$

5.35. Calcular por derivación paramétrica

$$I = \int_0^{+\infty} \frac{1 - e^{-x^2}}{x^2}\, dx$$

(recuérdese que $\int_0^{+\infty} e^{-h^2x^2}\, dx = \sqrt{\pi}/2h$).

5.36. Calcular la integral:

$$I = \int_0^1 \frac{x^{a-1} - 1}{\mathrm{L}\, x}\, dx$$

5.37. Calcular la siguiente integral:

$$I = \int_0^{+\infty} \frac{\operatorname{sen}^2 x}{x^2}\, dx$$

(recuérdese que $\int_0^{+\infty} \frac{\operatorname{sen} x}{x}\, dx = \frac{\pi}{2}$).

5.38. Calcular la integral:

$$I = \int_0^{+\infty} \frac{1}{x}\left\{\left(\frac{1}{2} + \frac{1}{x}\right)(e^{-ax} - e^{-x/2}) + \frac{2a-1}{2}\, e^{-x}\right\} dx$$

5.39. Reducir a integrales eulerianas de segunda especie la integral:

$$I = \int_0^1 \sqrt{1 - x^5}\, dx$$

5.40. Calcular la siguiente integral (recúrrase a las funciones eulerianas):

$$I = \int_0^{\pi/2} \left(\frac{1}{\sqrt{\operatorname{tg} x}} + \frac{1}{\sqrt{\operatorname{ctg} x}}\right) dx$$

5.41. Calcular la siguiente integral (echar mano de las integrales eulerianas):

$$I = \int_0^{+\infty} \operatorname{arctg} x^{\alpha}\, dx \qquad (\text{donde } \alpha < -1)$$

5.42. Pruébese que la siguiente integral es uniformemente convergente para $x \geqslant 0$:

$$\int_0^{+\infty} e^{-xt} \frac{\operatorname{sen} t}{t}\, dt$$

5.43. Probar que las siguientes integrales son uniformemente convergentes:

$$I = \int_0^{+\infty} \frac{\operatorname{sen}(\lambda x)}{1 + x^2}\, dx \quad \text{para } \lambda \in \mathbb{R}$$

$$J = \int_0^1 \frac{e^{x\lambda}}{\sqrt{1 - x}}\, dx \quad \text{para } \lambda \geqslant c \text{ (cualquiera)}$$

SOLUCIONES

5.1. No existe la integral: $A_n = [0, n] \times [0, n]$, $B_n = [0, n] \times [0, n+1]$; $\iint_{A_n} f = 1 \to 1$ e $\iint_{B_n} f = -1 \to -1$ (para $n \to \infty$).

5.2. $C_n = \{(x, y) \in C \,/\, \rho < n\}$; $\iint_{C_n} f = \pi(1 - \cos n^2)/4$ (polares) carece de límite para $n \to \infty$; $\iint_C f$ oscilante (no existe).

5.3. $C_\delta = \{(x, y) \in C \,/\, y < x - \delta\}$, $\iint_{C_\delta} f \to \infty$ cuando $\delta \to 0$; la integral diverge.

5.4. $I_1 = \int_0^1 (dx/\sqrt{x}) \int_0^1 (dy/\sqrt{y}) = 4$; $I_2 = 2 \iint_D dx\, dy/\sqrt{x - y}$, con $D = \{(x, y) \in C \,/\, y < x\}$, $I_2 = 8/3$.

5.5. Integrando por iteración (de dos maneras): $I = \mathrm{L}\, 2$ e $I = J$, luego $J = \mathrm{L}\, 2$.

5.6. Pasando a polares, $I = 2\pi$. Comparando J con I, J es convergente.

5.7. El integrando infinitésimo cuando $(x, y) \to \infty$; $\alpha x^2 + 2\beta xy + \gamma y^2$ definida positiva: $\alpha > 0$ y $\alpha\gamma - \beta^2 > 0$. Cambiando a la forma canónica, $I = J/\sqrt{\alpha\gamma - \beta^2} = \pi/\sqrt{\alpha\gamma - \beta^2}$.

5.8. Cambio a polares $I = \int_{-\pi}^{\pi} \cos^3 \theta\, d\theta \int_0^{+\infty} e^{-\rho^5} \rho^4\, d\rho = 4/15$.

5.9. $C_n = \{(x, y) \in \mathbb{R}^2 \,/\, 0 \leqslant x \leqslant n,\ 0 \leqslant y \leqslant n\}$, $I = 4 \lim \iint_{C_n} xye^{-x-y}\, dx\, dy = 2$. J es absolutamente convergente (mayorada por I); $D_n = [-n, n] \times [-n, n]$, $\iint_{D_n} xye^{-|x+y|}\, dx\, dy = 0$, $J = 0$.

5.10. $f \geqslant 0$; $C_n = \{(x, y) \in \mathbb{R}^2 \,/\, 0 \leqslant \theta \leqslant \pi/2,\ \rho < n\}$ (polares), $I = \lim \iint_{C_n} f = I_1 \cdot I_2$, $I_1 = \int_0^{\pi/2} \operatorname{sen}^n \theta\, d\theta < \pi/2$, $I_2 = \int_0^n \rho^{n+1}\, d\rho/(1 + \rho^2)^n < \int_0^n \rho^{-n+1}\, d\rho$ converge ($n > 2$); I converge.

5.11. $I = \int_0^{+\infty} dy e^{-y^2} \int_0^y dx = 1/2$; $J = \int_0^{+\infty} (e^{-y}/y^2) \operatorname{sen} y\, dy \int_0^y dx = 1/4$

5.12. 1.º Cambio a polares, $V = \iint_{\mathbb{R}^2} f = \pi$.
2.º Integración iterada (integral de Gauss $\int_0^{+\infty} e^{-x^2}\, dx = \sqrt{\pi}/2$), $V = \pi/16$.

5.13. $I = \int_0^1 dy \int_0^{y^2} e^{x/y}\, dy = 1/2$; $J = \int dx \int (y^2 + x^4)^{-1}\, dy = \pi/4$

5.14. $\iint_C f^+ = \iint_C f^-$; I converge si converge $J = \iint_D f$ con $D = \{(x, y) \in C \,/\, x > 0, y > 0\}$; en D, $xy/4 < \operatorname{sen} x \operatorname{sen} y < 4xy$, I converge sii converge $K = \iint_D (xy/\rho^\alpha)\, dx\, dy = \int_0^1 \rho^{3-\alpha}\, d\rho \int_0^{\pi/2} \operatorname{sen} \theta \cos \theta\, d\theta = (1/2) \int_0^1 \rho^{3-\alpha}\, d\rho$; si $\alpha < 4$, I converge; si $\alpha \geqslant 4$, I no existe.

5.15. 1.ª Divergente. 2.ª Convergente. 3.ª Divergente.

5.16. Converge: $\int_r^R x^{-3/2} e^{-x} \int_0^x e^y\, dy$ tiene límite finito para $r \to 0$ y $R \to +\infty$.

5.17. Converge: en $(0, 0, 0)$ seudoimpropia, pues $\rho^{-\rho} \to 1$; en ∞ converge, pues $\rho^{-\rho} < \rho^{-4}$.

5.18. Integración iterada $I = \int_0^{+\infty} dy \int_0^{+\infty} f(x, y)\, dx = \mathrm{B}(1/2, 1/2)(\pi/2) = \pi^2/2$ (B = beta de Euler). Con la otra iteración $I = -J$.

5.19. $f'(y) = (2 \operatorname{sen} y^2 - 3 \operatorname{sen} y^3)/y$.

5.20. $F'(y) = \int_{-t}^{t} (f'(x + y) - f'(x - y))\, dx + 0 = [f(x + y) - f(x - y)]_{-y}^{y} = 0$

5.21. $I'(a) = \dfrac{\pi}{a}(1 - 1/\sqrt{1 - a^2})$, $I(a) = \pi\, \mathrm{L}\, (1 + \sqrt{1 - a^2}) + c$, $c = -\pi\, \mathrm{L}\, 2$

5.22. $J(a) = (1 + a)^{-1}$, derivando sucesivas veces: $I_n(a) = (-1)^n n!/(1 + a)^{n+1}$.

5.23. $I(a) = \left. \dfrac{ax + \mathrm{L}\, (\operatorname{sen} x + a \cos x)}{a^2 + 1} \right]_{x=0}^{x=\pi/2} = \dfrac{a\pi - 2\, \mathrm{L}\, a}{2(a^2 + 1)}$. Derivando respecto de a:
$$J(a) = \frac{\pi a(1 - a^2) - 2(a^2 + 1) + 4a^2\, \mathrm{L}\, a}{2a(a^2 + 1)^2}$$

5.24. Indicación (derivando respecto de a).

5.25. $I = 2 \int_0^{\pi/2} f(x)\, dx$; cambio $t = \operatorname{tg} x$. Derivando respecto de a y respecto de b en $I = \pi/(a, b)$, $I_1 = \pi/(2a^3 b)$, $I_2 = \pi/(2ab^3)$.

5.26. Derivando respecto de λ, $I'(\lambda) = -\dfrac{2}{\sqrt{1 - \lambda^2}} \operatorname{arctg} \dfrac{\lambda}{\sqrt{1 - \lambda^2}}$; integrando, $I(\lambda) = -\left(\operatorname{arctg} \dfrac{\lambda}{\sqrt{1 - \lambda^2}}\right)^2 + c$, $c = 0$.

5.27. Derivando respecto de λ, $I'(\lambda) = \dfrac{-\pi}{2\sqrt{1+\lambda}}$; integrando $I(\lambda) = \pi\sqrt{1+} \ + c$, $c = -\pi$.

5.28. Derivando respecto de λ, $I'(\lambda) = \lambda$ tg $\lambda -$ $-$ L (cos λ); integrando, $I(\lambda) = -\lambda$ L (cos λ) $+ c$, $c = 0$.

5.29. Derivando respecto de λ, $I'(\lambda) = (\pi/2)$ sen λ; integrando, $I(\lambda) = -(\pi/2)$ cos $\lambda + c$, $c = \pi/2$.

5.30. Con el cambio tg $(x/2) = t$, $J(\lambda) = \pi/\sqrt{\lambda^2 - 1}$ (para $|\lambda| > 1$). Derivamos, $J'(\lambda) =$ $= -\int_0^\pi (\lambda - \cos x)^{-2}\, dx = -\pi\lambda/(\lambda^2 - 1)^{3/2}$, $I = -J(2) = 2\pi/(3\sqrt{3})$.

5.31. $J(a) = 1/(a + 1)$, para $a > -1$; $J'(a) = -1/(a + 1)^2$; $I = J'(-1/2) = -4$.

5.32. $J'(t) = (\pi/2)(1 + t^2)^{-1/2}$ (cambios $x = \operatorname{sen} u$, tg $u = z$); $J(t) = (\pi/2)\,\mathrm{L}\,(t + \sqrt{t^2 + 1}) + c$, $c = 0$; $I = J(1) = (\pi/2)$ L $(1 + \sqrt{2})$.

5.33. $I'(\lambda) = (\pi/2)/(1 + \lambda)$; $I(\lambda) = (\pi/2)$ L $(1 + \lambda) + c$, $c = 0$.

5.34. $I'(a)$ es racional, $I'(a) = \dfrac{-a^2}{1-a^2}\displaystyle\int_0^{+\infty} \frac{dx}{1 + a^2x^2} +$ $+ \dfrac{1}{1-a^2}\displaystyle\int_0^{+\infty} \frac{dx}{1+x^2} = \frac{\pi}{2(1+a)}$;

$I(a) = \dfrac{\pi \text{ L } 2}{2}$

5.35. $I(a) = \displaystyle\int_0^{+\infty} (1 - e^{-ax^2})\, dx/x^2$; $I'(a) = \sqrt{\dfrac{\pi}{2a}}$; $I(a) = \sqrt{\pi a} + c$, $c = 0$; $I = I(1) = \sqrt{\pi}$.

5.36. $I'(a) = 1/a$; $I(a) = \mathrm{L}\, a$.

5.37. $I(a) = \displaystyle\int_0^{+\infty} (\operatorname{sen}^2 ax)\, dx/x^2$; $I'(a) = \pi/2$; $I(a) = \pi/2 +$ $+ c$, $c = 0$; $I = I(1) = \pi/2$.

5.38. $I''(a)$ (por partes, $u = 1 + x/2) = 1/a + 1/(2a^2)$; $I'(a) = \mathrm{L}\, a - 1/(2a) + c$, $c = 0$; $I(a) =$ $= (a - 1/2)\,\mathrm{L}\, a - a + c'$, $c' = 1/2$.

5.39. Cambio $x^5 = t$, $I = \frac{1}{5}\,\mathrm{B}(\frac{1}{5}, \frac{3}{2}) = \frac{1}{5}\,\dfrac{\Gamma(\frac{1}{5})\Gamma(\frac{3}{2})}{\Gamma(\frac{17}{10})}$

5.40. $I = \int_0^{\pi/2} (\cos x)^{1/2} (\operatorname{sen} x)^{-1/2}\, dx +$ $+ \int_0^{\pi/2} (\operatorname{sen} x)^{1/2} (\cos x)^{-1/2}\, dx = \frac{1}{2}\,\mathrm{B}(\frac{1}{4}, \frac{3}{4}) +$ $+ \frac{1}{2}\,\mathrm{B}(\frac{3}{4}, \frac{1}{4}) = \Gamma(\frac{3}{4})\Gamma(\frac{1}{4}) = \pi/\operatorname{sen}(\pi/4) = \pi\sqrt{2}$

5.41. Partes ($u = $ arctg x^α) y cambio $x^{2\alpha} = t$,

$$I = \frac{1}{2}\int_0^{+\infty} \frac{t^{1/(2\alpha) - 1/2}}{1+t}\, dt = \frac{1}{2}\,\mathrm{B}\left(\frac{1}{2\alpha} + \frac{1}{2}, \frac{1}{2} - \frac{1}{2\alpha}\right)$$

$$= \frac{1}{2}\,\frac{\Gamma\left(\dfrac{1}{2\alpha} + \dfrac{1}{2}\right)\Gamma\left(\dfrac{1}{2} - \dfrac{1}{2\alpha}\right)}{\Gamma(1)} = \frac{1}{2}\,\frac{\pi}{\cos(\pi/2\alpha)}$$

5.42. Integrar por partes ($u = e^{-xt}/t$) y acudir a que $|(1 + xt)e^{-xt} \cos t| \leqslant 1$.

5.43. 1.° $|\operatorname{sen}(\lambda x)/(1 + x^2)| < 1/(1 + x^2)$
2.° $|e^{x\lambda}/\sqrt{1-x}| \leqslant e^{|c|}/\sqrt{1-x}$

Referencias bibliográficas

APÓSTOL, T. M., *Análisis matemático*. Reverté.
APÓSTOL, T. M., *Calculus*, vol. 2. Reverté.
BARTLE, R. G., *Introducción al análisis matemático*. Limusa.
BENTABOL, M. N., MARGALEF, J., y OUTERELO, E., *Análisis matemático (Cálculo diferencial)*. Pirámide.
BENTABOL, M. N., MARGALEF, J., y OUTERELO, E., *Análisis matemático (Cálculo integral)*, Pirámide.
BUCK, C., *Cálculo superior*. Ediciones del Castillo.
BURGOS, J., DE, *Cálculo infinitesimal de una variable*. McGraw-Hill.
BURGOS, J., DE, *Cálculo infinitesimal (teoría y problemas)*. Alhambra.
CASTILLO, F., *Análisis matemático II*. Alhambra.
COURANT, R., y JOHN, F., *Introducción al cálculo y análisis matemático*, vol. 2. Limusa.
ERWE, F., *Cálculo diferencial e integral*. Selecciones científicas.
FERNÁNDEZ VIÑA, J. A., *Análisis matemático II*. Tecnos.
FERNÁNDEZ VIÑA, J. A., *Análisis matemático III*. Tecnos.
FLEMING, W., *Funciones de varias variables*. C.E.C.S.A.
ILIN, V., y POZNIAK, E., *Fundamentos de análisis matemático*, vols. 2 y 3. MIR.
KUDRIAVSEV, L. D., *Curso de análisis matemático*, vols. 1 y 2. MIR.
LANG, S., *Cálculo*. Addison-Wesley Iberoamericana.
LANG, S., *Introducción al análisis matemático*. Addison-Wesley Iberoamericana.
LINES, E., *Principios de análisis matemático*. Reverté.
PISKUNOV, N., *Cálculo diferencial e integral*. MIR.
PUIG ADAM, R., *Cálculo integral*. Biblioteca Matemática.
REY PASTOR, J., *Elementos de teoría de funciones*. Madrid.
ROGOSINSKI, W. W., *Volumen e integral*. Dossat.
RUDIN, W., *Principios de análisis matemático*. McGraw-Hill.

Índice de materias

La presente obra va dirigida a aquellos estudiantes que, después de haber seguido un primer curso de cálculo infinitesimal, de una variable, deben continuar su formación en esta disciplina, ya sean alumnos de ciencias matemáticas o físicas, de ingeniería o arquitectura, de informática, de ciencias económicas o empresariales.

El texto se articula en torno a las cuestiones fundamentales, que ocupan en él lugares destacados y se presentan de forma compendiada y precisa. A ellas se les anexan los complementos pertinentes: ejemplos, comentarios, generalizaciones y, por supuesto, abundantes ejercicios y problemas con solución.

Se ha pretendido (¡ojalá se consiga!) que, al tiempo que se estudie aquí el cálculo infinitesimal, se adquiera soltura en el correcto uso del razonamiento deductivo y se aprenda el arte de usar lo conocido para resolver lo que se desconoce.

9 788448 116217

Mc Graw Hill

ISBN: 84-481-1621-6