Tema 14

Teorema de cambio de variable

Como último resultado fundamental en el estudio de la integral de Lebesgue, vamos a probar un *teorema de cambio de variable* para integrales múltiples, análogo al obtenido en su momento para funciones de una variable. Sin embargo, la situación es diferente, porque no disponemos de un teorema fundamental del cálculo, y porque las funciones diferenciables que ahora nos ocupan son más complicadas. La demostración que hemos de hacer, requiere en primer lugar estudiar el comportamiento de la medida de Lebesgue frente a transformaciones lineales, que ya no tienen por qué ser isometrías. Esto permite aprovechar la idea clave del cálculo diferencial, el hecho de que, salvo traslaciones, una función diferenciable puede aproximarse localmente por aplicaciones lineales. De esta forma se consigue analizar el comportamiento de la medida de Lebesgue, y por tanto de la integral de Lebesgue, frente a transformaciones bastante generales, que es en el fondo la información que nos da el teorema de cambio de variable. Junto con el teorema de Fubini, este resultado es otra herramienta clave para el cálculo de integrales múltiples. Ilustraremos su utilidad práctica, considerando los cambios de variable que se usan con más frecuencia para el cálculo de integrales dobles o triples.

14.1. Medida de Lebesgue y aplicaciones lineales

Usaremos la matriz asociada a una aplicación lineal de $\mathbb{R}^N$ en sí mismo, que la describe como un producto de matrices. Para ello, vemos cada vector $x \in \mathbb{R}^N$ como una matriz fila $1 \times N$ y, para cualquier matriz M, denotamos por M^t a la matriz transpuesta de M. Dado $x \in \mathbb{R}^N$, vemos por ejemplo que x^t, es una matriz columna $N \times 1$. Usaremos un hecho conocido: si el número de columnas de M coincide con el número de filas de otra matriz P, con lo que tiene sentido el producto de matrices $M \cdot P$, entonces se tiene que $\big(M \cdot P\big)^t = P^t \cdot M^t$. Si M es una matriz cuadrada, denotamos por $\det M$ a su determinante. Es sabido que $\det M^t = \det M$, y que si P es otra matriz cuadrada del mismo orden que M, se tiene: $\det(M \cdot P) = (\det M)(\det P)$.

Pues bien, si denotamos como siempre por $\{e_1, e_2 \ldots, e_N\}$ a la base usual de $\mathbb{R}^N$, a cada aplicación lineal $T : \mathbb{R}^N \to \mathbb{R}^N$ asociamos una matriz M_T, que la representa en dicha base.

Concretamente $M_T = \big(a_{jk}\big)$ es la matriz cuadrada de orden N dada por

$$a_{jk} = T\big(e_k\big)(j) \qquad \forall\, j,k \in \Delta_N$$

A partir de la matriz M_T recuperamos la aplicación lineal T, pues para todo $x \in \mathbb{R}^N$ se tiene

$$T(x)(j) = \sum_{k=1}^{N} x(k)\, T(e_k)(j) = \sum_{k=1}^{N} a_{jk}\, x(k) \qquad \forall\, j \in \Delta_N$$

igualdades que se resumen con un producto de matrices: $T(x)^t = M_T \cdot x^{\,t}$, para todo $x \in \mathbb{R}^N$. En particular, si $S : \mathbb{R}^N \to \mathbb{R}^N$ es otra aplicación lineal, de $M_S = M_T$ se deduce que $S = T$. En sentido opuesto, si M es una matriz cuadrada de orden N, definiendo $T(x) = x \cdot M^t$, o lo que es lo mismo, $T(x)^t = M \cdot x^{\,t}$, para todo $x \in \mathbb{R}^N$, se obtiene una aplicación lineal $T : \mathbb{R}^N \to \mathbb{R}^N$, tal que $M_T = M$.

En resumen, hay una correspondencia biunívoca $T \leftrightarrow M_T$ entre aplicaciones lineales de $\mathbb{R}^N$ en sí mismo, y matrices cuadradas de orden N con coeficientes reales. A la identidad en $\mathbb{R}^N$ corresponde la matriz unidad de orden N que denotaremos por Id. Además, a la composición de aplicaciones lineales $S \circ T$, corresponde el producto de matrices, pues para todo $x \in \mathbb{R}^N$ se tiene que $S\big(T(x)\big)^t = M_S \cdot T(x)^t = M_S \cdot M_T \cdot x^t$, luego la matriz asociada a $S \circ T$ es $M_S \cdot M_T$. Como consecuencia, vemos que una aplicación lineal T es biyectiva si, y sólo la matriz A_T es inversible, lo que equivale a que $\det A_T \neq 0$, en cuyo caso, la matriz asociada a T^{-1} es M_T^{-1}, la matriz inversa de M_T.

Lo dicho hasta ahora, sobre la representación matricial de las aplicaciones lineales, involucra solamente la estructura de espacio vectorial de $\mathbb{R}^N$, pero nos interesa también la interacción con el producto escalar y la norma euclídea, que es la que de momento vamos a usar. Observamos que el producto escalar de $\mathbb{R}^N$, que denotamos por $(x,y) \mapsto (x\,|\,y)$, también se expresa como un producto de matrices. Concretamente, se tiene

$$\big(x\,|\,y\big) = \sum_{k=1}^{N} x(k)\, y(k) = x \cdot y^{\,t} \qquad \forall x,y \in \mathbb{R}^N$$

Esto permite relacionar las aplicaciones lineales definidas por una matriz y su transpuesta. Concretamente, si M es una matriz cuadrada de orden N, se tiene:

$$\big(x \cdot M^t\,|\,y\big) = \big(x\,|\,y \cdot M\big) \qquad \forall x,y \in \mathbb{R}^N \tag{1}$$

La comprobación es inmediata: $\big(x \cdot M^t\,|\,y\big) = x \cdot M^t \cdot y^{\,t} = x \cdot (y \cdot M)^t = \big(x\,|\,y \cdot M\big)$

Comentemos que el cuadrado de la norma euclídea también se expresa matricialmente,

$$\big\|x\big\|^2 = \sum_{k=l}^{N} x(k)^2 = \big(x\,|\,x\big) = x \cdot x^{\,t} \qquad \forall x \in \mathbb{R}^N$$

y que el producto escalar se recupera a partir de la norma mediante la *identidad de polarización*:

$$4\,\big(x\,|\,y\big) = \big\|x+y\big\|^2 - \big\|x-y\big\|^2 \qquad \forall x,y \in \mathbb{R}^N \tag{2}$$

Sabemos ya que la medida de Lebesgue es invariante por isometrías, pero conviene conocer mejor las isometrías lineales y las matrices asociadas. Nos será útil la siguiente caracterización.

- *Sea $T : \mathbb{R}^N \to \mathbb{R}^N$ una aplicación lineal, M la matriz asociada a T y $u_k = T(e_k)$ para todo $k \in \Delta_N$. Entonces, las siguientes afirmaciones son equivalentes:*
  (i) *T es una isometría*
  (ii) *$\{u_1, u_2, \ldots u_N\}$ es una base ortonormal de $\mathbb{R}^N$*
  (iii) *$M^t \cdot M = \mathrm{Id}$, es decir, M es inversible con $M^{-1} = M^t$.*

$(i) \Rightarrow (ii)$ Si T es lineal y preserva la norma, de (2) deducimos que también preserva el producto escalar, pues para cualesquiera $x, y \in \mathbb{R}^N$, se tiene:

$$\begin{aligned} 4\big(T(x)\,|\,T(y)\big) &= \big\|T(x)+T(y)\big\|^2 - \big\|T(x)-T(y)\big\|^2 \\ &= \big\|T(x+y)\big\|^2 - \big\|T(x-y)\big\|^2 = \big\|x+y\big\|^2 - \big\|x-y\big\|^2 = 4\big(x\,|\,y\big) \end{aligned}$$

Como $\{e_1, e_2, \ldots e_N\}$ es una base ortonormal, deducimos que $\{u_1, u_2, \ldots u_N\}$ también lo es.

$(ii) \Rightarrow (iii)$ Sean $M = \big(a_{jk}\big)$, $M^t = \big(b_{jk}\big)$ y $M^t \cdot M = \big(c_{jk}\big)$. Como M es la matriz asociada a T, para $j, k \in \Delta_N$, se tiene $a_{jk} = T(e_k)(j) = u_k(j)$, luego $b_{jk} = a_{kj} = u_j(k)$. Por definición del producto de matrices, se tiene entonces que

$$c_{jk} = \sum_{h=1}^{N} b_{jh}\, a_{hk} = \sum_{h=1}^{N} u_j(h)\, u_k(h) = \big(u_j\,|\,u_k\big) \qquad \forall\, j, k \in \Delta_N$$

Como $\{u_1, u_2, \ldots u_N\}$ es una base ortonormal, tenemos $M^t \cdot M = \big((u_j\,|\,u_k)\big) = \mathrm{Id}$. Sabemos que T biyectiva, es decir, M es inversible, luego $M^{-1} = M^t$.

$(iii) \Rightarrow (i)$ Para todo $x \in \mathbb{R}^N$, se tiene que $x = x \cdot M^t \cdot M$, y usando (1) obtenemos:

$$\big\|x\big\|^2 = \big(x \cdot M^t \cdot M\,|\,x\big) = \big(x \cdot M^t\,|\,x \cdot M^t\big) = \big(T(x)\,|\,T(x)\big) = \big\|T(x)\big\|^2$$

luego T es una isometría. ■

Pasamos ya a obtener la factorización de aplicaciones lineales que permitirá analizar el comportamiento de la medida de Lebesgue frente a tales aplicaciones. Es una consecuencia sencilla de la diagonalización de una matriz simétrica, que suponemos conocida.

- *Cualquier biyección lineal $T : \mathbb{R}^N \to \mathbb{R}^N$ se puede expresar en la forma $T = U \circ D \circ V$, donde $U, V : \mathbb{R}^N \to \mathbb{R}^N$ son dos isometrías lineales, mientras que $D : \mathbb{R}^N \to \mathbb{R}^N$ es la aplicación lineal definida por una matriz diagonal con autovalores positivos, es decir, existen $\rho_1, \rho_2, \ldots, \rho_N \in \mathbb{R}^+$, tales que $D(e_k) = \rho_k e_k$ para todo $k \in \Delta_N$.*

Sea M la matriz asociada a T y sea S la aplicación lineal definida por la matriz $P = M^t \cdot M$, es decir, $S(x) = x \cdot P^{\,t} = x \cdot P$ para todo $x \in \mathbb{R}^N$, ya que P es simétrica: $P^{\,t} = P$. Además, usando (1), observamos que, para todo $x \in \mathbb{R}^N \setminus \{0\}$, se tiene

$$\big(S(x)\,|\,x\big) = \big(x \cdot M^t \cdot M\,|\,x\big) = \big(x \cdot M^t\,|\,x \cdot M^t\big) = \big(T(x)\,|\,Tx\big)\big) = \big\|T(x)\big\|^2 > 0$$

ya que T es inyectiva. Si α es un autovalor de P, existe $x \in \mathbb{R}^N$ con $\|x\| = 1$ y $S(x) = \alpha x$, con lo que $\alpha = \big(\alpha x\,|\,x\big) = \big(S(x)\,|\,x\big) \in \mathbb{R}^+$.

Así pues, todos los autovalores de P son números reales positivos. Repitiendo cada uno de ellos tantas veces como indique su multiplicidad, los escribimos en la forma $\rho_1^2, \rho_2^2, \ldots, \rho_N^2$, donde $\rho_1, \rho_2, \ldots, \rho_N \in \mathbb{R}^+$. La diagonalización de la matriz P nos proporciona entonces una base ortonormal $\{y_1, y_2, \ldots, y_N\}$, que verifica $S(y_k) = \rho_k^2\, y_k$, para todo $k \in \Delta_N$.

Para definir las aplicaciones lineales que buscamos, basta indicar su valores en cualquier base de $\mathbb{R}^N$, lo que las determina de manera única. Consideramos entonces las aplicaciones lineales $U, D, V : \mathbb{R}^N \to \mathbb{R}^N$ que, para todo $k \in \Delta_N$, verifican:

$$V(y_k) = e_k, \quad D(e_k) = \rho_k\, e_k, \quad U(e_k) = \frac{1}{\rho_k}\, T(y_k)$$

Es claro que V es biyectiva, y V^{-1} aplica la base usual en la base ortonormal $\{y_1, y_2, \ldots, y_N\}$, así que V^{-1} es una isometría, y lo mismo le ocurre a V. También es evidente que D verifica la condición requerida en el enunciado. Además, se tiene

$$\big(U \circ D \circ V\big)(y_k) = \big(U \circ D\big)(e_k) = U\big(\rho_k\, e_k\big) = T(y_k) \qquad \forall\, k \in \Delta_N$$

de donde deducimos que $U \circ D \circ V = T$, y sólo queda probar que U es una isometría. Para ello, tomando $u_k = U(e_k) = (1/\rho_k)T(y_k)$, debemos comprobar que $\{u_1, u_2, \ldots, u_N\}$ también es una base ortonormal. Para $j, k \in \Delta_N$, usando de nuevo (1), se tiene:

$$\begin{aligned}\rho_j\, \rho_k\, \big(u_j \,|\, u_k\big) &= \big(T(y_j) \,|\, T(y_k)\big) = \big(y_j \cdot M^t \,|\, y_k \cdot M^t\big) \\ &= \big(y_j \,|\, y_k \cdot M^t \cdot M\big) = \big(y_j \,|\, S(y_k)\big) = \rho_k^2\, \big(y_j \,|\, y_k\big)\end{aligned}$$

Por tanto, si $j \neq k$ tenemos $\big(u_j \,|\, u_k\big) = (\rho_k/\rho_j)\big(y_j \,|\, y_k\big) = 0$, mientras que si $j = k$ obtenemos que $\big(u_k \,|\, u_k\big) = \big(y_k \,|\, y_k\big) = 1$. ■

El siguiente resultado describe el comportamiento de la medida de Lebesgue frente a las aplicaciones lineales, generalizando la invariancia por isometrías lineales. Será nuestro punto de partida en el estudio del teorema de cambio de variable. Como la dimensión N estará fija, volvemos a evitar subíndices, denotando por $\lambda : \mathcal{M} \to [0, \infty]$ a la medida de Lebesgue en $\mathbb{R}^N$.

- *Sea $T : \mathbb{R}^N \to \mathbb{R}^N$ una aplicación lineal y M su matriz asociada. Entonces T preserva los conjuntos medibles y verifica:*

$$\lambda\big(T(E)\big) = \big|\det M\big|\, \lambda(E) \qquad \forall\, E \in \mathcal{M} \tag{3}$$

El caso más sencillo se presenta cuando T no es biyectiva, pues entonces su imagen es un subespacio vectorial de $\mathbb{R}^N$, distinto de $\mathbb{R}^N$, luego es un conjunto de medida nula. Deducimos que $T(E)$ es medible, con $\lambda\big(T(E)\big) = 0$ para todo conjunto $E \subset \mathbb{R}^N$, aún cuando E no sea medible. Además, en este caso tenemos $\det M = 0$, luego se cumple (3).

Tratemos el caso de una biyección lineal $D : \mathbb{R}^N \to \mathbb{R}^N$ definida por una matriz diagonal R con autovalores positivos, es decir, existen $\rho_1, \rho_2, \ldots, \rho_N \in \mathbb{R}^+$, tales que $D(e_k) = \rho_k\, e_k$ para todo $k \in \Delta_N$, con lo que $\det R = \prod_{k=1}^{N} \rho_k$. Como D es un homeomorfismo de $\mathbb{R}^N$, preserva los conjuntos de Borel, lo que nos permite definir una función $\mu : \mathcal{B} \to [0, \infty]$ escribiendo:

$$\mu(B) = \frac{\lambda\big(D(F)\big)}{\det R} \qquad \forall\, B \in \mathcal{B}$$

Es rutinario comprobar que μ es una medida, definida en la σ-álgebra de Borel $\mathcal{B}$, y vamos a demostrar que μ coincide con λ en la familia $\mathcal{J}$ de los intervalos acotados. Sea $I = \prod_{k=1}^{N} I_k \in \mathcal{J}$, escrito como producto cartesiano de intervalos acotados en $\mathbb{R}$. Para cada $x \in \mathbb{R}^N$ y $k \in \Delta_N$ tenemos $D(x)(k) = \rho_k x(k)$, de donde deducimos que $J = T(I)$ viene dado por $J = \prod_{k=1}^{N} J_k$ donde, para cada $k \in \Delta_N$ hemos escrito $J_k = \{\rho_k t : t \in I_k\}$, que también es un intervalo acotado en $\mathbb{R}$, con $\inf J_k = \rho_k \inf I_k$ y $\sup J_k = \rho_k \sup I_k$. Por tanto $J \in \mathcal{J}$ con

$$\lambda(J) = \prod_{k=1}^{N} \big(\sup J_k - \inf J_k \big) = \prod_{k=1}^{N} \rho_k \big(\sup J_k - \inf J_k \big) = \big(\det R\big) \lambda(I)$$

Deducimos que $\mu(I) = \lambda(I)$, como se quería. El primer teorema de unicidad de la medida de Lebesgue nos dice ahora que $\mu(B) = \lambda(B)$ para todo $B \in \mathcal{B}$, es decir

$$\lambda\big(D(B)\big) = \big(\det R\big)\lambda(B) \qquad \forall B \in \mathcal{B}$$

Dado ahora $E \in \mathcal{M}$, podemos escribir $A \subset E \subset B$, donde $A, B \in \mathcal{B}$ con $\lambda(B \setminus A) = 0$. Entonces, $D(A) \subset D(E) \subset D(B)$ y, usando (4) para el conjunto $B \setminus A \in \mathcal{B}$ tenemos

$$\lambda\big(D(B) \setminus D(A)\big) = \lambda\big(D(B \setminus A)\big) = \big(\det R\big)\lambda(B \setminus A) = 0$$

Como $D(E) \setminus D(A) \subset D(B) \setminus D(A)$, deducimos que $D(E) \setminus D(A)$ también tiene medida nula. Por tanto $D(E) = D(A) \uplus \big(D(E) \setminus D(A)\big)$ es medible, como unión de dos conjuntos medibles. Además, usando de nuevo (4), concluimos que

$$\lambda\big(D(E)\big) = \lambda\big(D(A)\big) = \big(\det R\big)\lambda(A) = \big(\det R\big)\lambda(E)$$

Para completar la demostración, bastará ahora usar la factorización obtenida previamente. Si $T : \mathbb{R}^N \to \mathbb{R}^N$ es ya una biyección lineal arbitraria, escribimos $T = U \circ D \circ V$ donde U, V son isometrías y D una biyección lineal del tipo recién estudiado. Sabemos que las isometrías preservan los conjuntos medibles, y acabamos de ver que D tiene la misma propiedad, luego T también la tiene: $T(E) \in \mathcal{M}$ para todo $E \in \mathcal{M}$.

La matriz asociada a U verifica $M_U^t \cdot M_U = \mathrm{Id}$, luego $\big(\det M_U\big)^2 = \big(\det M_U^t\big)\big(\det M_U\big) = 1$, es decir, $\big|\det M_U\big| = 1$. Análogamente se tiene $\big|\det M_V\big| = 1$ para la matriz asociada a V. En cuanto a la matriz R asociada a D, sabemos que $\det R > 0$, de donde deducimos que

$$\big|\det M\big| = \big|\det\,(M_U \cdot R \cdot M_V)\big| = \big|\det M_U\big|\,\big|\det R\big|\,\big|\det M_V\big| = \det R$$

Finalmente, para probar (3) basta observar que, para todo $E \in \mathcal{M}$ se tiene

$$\begin{aligned} \lambda\big(T(E)\big) &= \lambda\big((U \circ D \circ V)(E)\big) = \lambda\big((D \circ V)(E)\big) \\ &= \big(\det R\big)\lambda\big(V(E)\big) = \big(\det R\big)\lambda(E) = \big|\det M\big|\,\lambda(E) \end{aligned}$$

donde hemos usado la invarianza de λ por isometrías y lo ya probado para D. ∎

14.2. Medida de Lebesgue y funciones diferenciables

Como segunda etapa en nuestro camino hacia el teorema de cambio de variable, vamos a estudiar la imagen de un conjunto medible por una función de clase C^1 en un abierto que lo contenga. La diferencial de tal función está localmente acotada, lo que permite usar el teorema del valor medio para concluir que la función es localmente lipschitziana. Por ello podremos razonar con funciones lipschitzianas para obtener conclusiones locales que luego habrá que globalizar. Los razonamientos necesarios para ambas cosas son laboriosos, pero no difíciles.

Sabemos que todo abierto de $\mathbb{R}^N$ se expresa como unión numerable de intervalos diádicos dos a dos disjuntos, directamente relacionados con las bolas para la norma del máximo. De ahí que a partir de ahora usemos con frecuencia dicha norma:

$$\|x\|_\infty = \text{máx}\,\big\{\,|x(k)|\,:\,k\in\Delta_N\,\big\} \qquad \forall\,x\in\mathbb{R}^N$$

Como cuestión previa, estudiamos el comportamiento de la medida exterior de Lebesgue frente a funciones lipschitzianas. Como todas las normas en $\mathbb{R}^N$ son equivalentes, las funciones lipschitzianas para una norma, lo son para cualquier otra, aunque la constante de Lipschitz sí depende de la norma que usemos.

- *Sea Ω un abierto de $\mathbb{R}^N$ y $\phi:\Omega\to\mathbb{R}^N$ una función lipschitziana. Si $K\in\mathbb{R}^+$ verifica*

$$\big\|\,\phi(y)-\phi(x)\,\big\|_\infty \leqslant K\,\big\|\,y-x\,\big\|_\infty \qquad \forall\,x,y\in\Omega \tag{4}$$

entonces se tiene:

$$\lambda^*\big(\phi(E)\big)\leqslant K^N\lambda^*(E) \qquad \forall\,E\in\mathcal{P}(\Omega) \tag{5}$$

Si B es una bola cerrada de radio $\rho\in\mathbb{R}^+$ para la norma del máximo, es claro que B es un intervalo acotado con $\lambda(B)=\big(2\rho\big)^N$. Si ahora J es un intervalo diádico, con $J\subset\Omega$, se tiene que $\overline{J}$ es una bola cerrada que tendrá un centro $x\in\Omega$ y un radio $r\in\mathbb{R}^+$, con lo que obtenemos $\lambda(J)=\lambda(\overline{J})=\big(2r\big)^N$. De la hipótesis (4) deducimos que $\phi(J)$ está contenido en la bola cerrada de centro $\phi(x)$ y radio Kr, luego $\lambda^*\big(\phi(J)\big)\leqslant\big(2Kr\big)^N=K^N\lambda(J)$.

Sea ya $E\in\mathcal{P}(\Omega)$ para probar (5), desigualdad que es obvia si $\lambda^*(E)=\infty$. En otro caso, fijado $\varepsilon>0$ existe un abierto $H\subset\mathbb{R}^N$, con $E\subset H$, tal que $\lambda(H)<\lambda^*(E)+\varepsilon$. Entonces el abierto $G=H\cap\Omega$ verifica que $E\subset G\subset\Omega$ y $\lambda(G)\leqslant\lambda(H)<\lambda^*(E)+\varepsilon$. Escribimos ahora $G=\biguplus_{n=1}^{\infty}J_n$ donde J_n es un intervalo diádico para todo $n\in\mathbb{N}$. Usando lo ya probado para intervalos diádicos tenemos

$$\begin{aligned}\lambda^*\big(\phi(E)\big)&\leqslant\lambda^*\big(\phi(G)\big)=\lambda^*\left(\bigcup_{n=1}^{\infty}\phi(J_n)\right)\leqslant\sum_{n=1}^{\infty}\lambda^*\big(\phi(J_n)\big)\\&\leqslant K^N\sum_{n=1}^{\infty}\lambda(J_n)=K^N\lambda(G)<K^N\lambda^*(E)+K^N\varepsilon\end{aligned}$$

En vista de la arbitrariedad de ε, hemos probado (5). ■

En el próximo paso, necesitamos expresar cada abierto de $\mathbb{R}^N$ como unión numerable de bolas abiertas, aunque no sean dos a dos disjuntas. De hecho probaremos un resultado más preciso, que será útil más adelante.

- *Si Γ es un conjunto no vacío y $\{G(\gamma) : \gamma \in \Gamma\}$ una familia arbitraria de abiertos de $\mathbb{R}^N$, existe un conjunto numerable $\Gamma_0 \subset \Gamma$ tal que:*

$$\bigcup_{\gamma \in \Gamma_0} G(\gamma) = \bigcup_{\gamma \in \Gamma} G(\gamma)$$

Además, si d es una distancia que genere la topología usual de $\mathbb{R}^N$, todo abierto de $\mathbb{R}^N$ se expresa como unión numerable de bolas abiertas para la distancia d.

Sea $\Omega = \bigcup_{\gamma \in \Gamma} G(\gamma)$ y observemos que todo abierto de $\mathbb{R}^N$ puede obviamente recubrirse de esta forma. Fijada la distancia d, expresaremos Ω como unión numerable de bolas abiertas para d, obteniendo al mismo tiempo el conjunto numerable Γ_0 requerido.

Para cada $x \in \Omega$, tomamos $\gamma(x) \in \Gamma$ con $x \in G(\gamma(x))$, y $r(x) \in \mathbb{Q}^+$ tal que la bola abierta de centro x y radio $2r(x)$ esté contenida en $G(\gamma(x))$. Como $\mathbb{Q}^N$ es denso en $\mathbb{R}^N$, podemos también tomar $q(x) \in \mathbb{Q}^N$ de forma que $d(q(x), x) < r(x)$. Llamando $B(x)$ a la bola abierta de centro $q(x)$ y radio $r(x)$, es obvio que $x \in B(x)$, y comprobaremos que $B(x) \subset G(\gamma(x)) \subset \Omega$. En efecto, para todo $y \in B(x)$ se tiene

$$d\big(y, x\big) \leqslant d\big(y, q(x)\big) + d\big(q(x), x\big) < 2r(x), \quad \text{luego} \quad y \in G\big(\gamma(x)\big) \subset \Omega$$

El conjunto $\big\{\big(q(x), r(x)\big) : x \in \Omega\big\} \subset \mathbb{Q}^N \times \mathbb{Q}^+$ es numerable, luego $\big\{B(x) : x \in \Omega\big\}$ también lo es, así que podemos escribir $\big\{B(x) : x \in \Omega\big\} = \big\{B_n : n \in \mathbb{N}\big\}$. Se tiene entonces claramente

$$\Omega = \bigcup_{x \in \Omega} B(x) = \bigcup_{n=1}^{\infty} B_n$$

luego Ω es unión numerable de bolas abiertas para la distancia d. Además, para cada $n \in \mathbb{N}$ existe $x_n \in \Omega$ tal que $B_n = B(x_n)$, y tomando $\Gamma_0 = \big\{\gamma(x_n) : n \in \mathbb{N}\big\}$, tenemos un conjunto numerable $\Gamma_0 \subset \Gamma$, que verifica:

$$\Omega = \bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} B(x_n) \subset \bigcup_{n=1}^{\infty} G\big(\gamma(x_n)\big) = \bigcup_{\gamma \in \Gamma_0} G(\gamma) \qquad \blacksquare$$

Podemos ya probar que las funciones de clase C^1 preservan los conjuntos medibles, y lo que es importante, también preservan los conjuntos de medida nula.

- *Sea Ω un abierto de $\mathbb{R}^N$ y $\phi : \Omega \to \mathbb{R}^N$ una función de clase C^1. Entonces:*
  - (i) *Para todo conjunto de medida nula $Z \subset \Omega$, el conjunto $\phi(Z)$ tiene medida nula.*
  - (ii) *Para todo conjunto medible $E \subset \Omega$, el conjunto $\phi(E)$ es medible*

(i) Para cada $x \in \Omega$, sea $\| D\phi(x) \|$ la norma de la diferencial de ϕ en x, como aplicación lineal, de $\mathbb{R}^N$ con la norma del máximo, en el mismo espacio normado:

$$\| D\phi(x) \| = \max \left\{ \| D\phi(x)(y) \|_\infty : y \in \mathbb{R}^N, \ \| y \|_\infty = 1 \right\}$$

Se tiene entonces que $\Omega = \bigcup_{n=1}^{\infty} G_n$ donde $G_n = \left\{ x \in \Omega : \| D\phi(x) \| < n \right\}$ para todo $n \in \mathbb{N}$.

Por hipótesis la función $x \mapsto D\phi(x)$ es continua, y en todo espacio normado, la norma es una función continua, luego la función $x \mapsto \| D\phi(x) \|$ también es continua. Fijado $n \in \mathbb{N}$, vemos por tanto que el conjunto G_n es abierto. Sea ahora B una bola abierta para la norma del máximo, con $B \subset G_n$. Como B es un abierto convexo y $\| D\phi(x) \| < n$ para todo $x \in B$, el teorema del valor medio nos dice que $\phi\big|_B$ es lipschitziana con constante n, es decir:

$$\big\| \phi(y) - \phi(x) \big\|_\infty \leqslant n \big\| y - x \big\|_\infty \qquad \forall x, y \in B$$

Por el resultado probado previamente para funciones lipschitzianas, tenemos

$$\lambda^* \big(\phi(Z \cap B) \big) \leqslant n^N \lambda^*(Z \cap B) \leqslant n^N \lambda^*(Z) = 0$$

Todavía con $n \in \mathbb{N}$ fijo, el abierto G_n se expresa en la forma $G_n = \bigcup_{k=1}^{\infty} B_k$, donde B_k es una bola abierta, siempre con la norma del máximo, para todo $k \in \mathbb{N}$. Por lo anteriormente demostrado, tenemos $\lambda^* \big(\phi(Z \cap B_k) \big) = 0$ para todo $k \in \mathbb{N}$, de donde

$$\lambda^* \big(\phi(Z \cap G_n) \big) = \lambda^* \left(\bigcup_{k=1}^{\infty} \phi(Z \cap B_k) \right) \leqslant \sum_{k=1}^{\infty} \lambda^*(Z \cap B_k) = 0$$

Como lo anterior era cierto para todo $n \in \mathbb{N}$, concluimos que

$$\lambda^* \big(\phi(Z) \big) = \lambda^* \left(\bigcup_{n=1}^{\infty} \phi(Z \cap G_n) \right) \leqslant \sum_{n=1}^{\infty} \lambda^* \big(\phi(Z \cap G_n) \big) = 0$$

luego $\phi(Z)$ es un conjunto de medida nula, como se quería.

(ii). Si K es un subconjunto compacto de Ω, como f es continua, sabemos que $\phi(K)$ es compacto, luego medible. Si ahora C es un subconjunto cerrado de $\mathbb{R}^N$ con $C \subset \Omega$ podemos escribir $C = \bigcup_{n=1}^{\infty} K_n$ donde $K_n = \{ x \in C : \| x \| \leqslant n \}$ es un subconjunto compacto de Ω, para todo $n \in \mathbb{N}$. Por tanto, $\phi(C) = \bigcup_{n=1}^{\infty} f(K_n)$ es medible, como unión numerable de conjuntos medibles. Si ahora A es un conjunto de tipo F_σ en $\mathbb{R}^N$, con $A \subset \Omega$, podemos escribir $A = \bigcup_{k=1}^{\infty} C_k$ donde C_k es un subconjunto cerrado de $\mathbb{R}^N$ que está contenido en Ω, para todo $k \in \mathbb{N}$. De nuevo vemos que $\phi(A) = \bigcup_{k=1}^{\infty} \phi(C_k)$ es medible, como unión numerable de conjuntos medibles.

Finalmente, para un conjunto medible $E \subset \Omega$, podemos escribir $E = A \cup Z$, donde A es un conjunto de tipo F_σ en $\mathbb{R}^N$, que está contenido en Ω, y $\lambda(Z) = 0$. Sabemos ya que $\phi(A)$ es medible, y (i) nos dice que $\lambda\big(\phi(Z)\big) = 0$. Por tanto, $\phi(E) = \phi(A) \cup \phi(Z)$ es medible. ■

Para el tipo de función que usaremos como cambio de variable, vamos ahora a calcular la medida de la imagen de un conjunto medible.

Recordemos que, si Ω y G son abiertos de $\mathbb{R}^N$, se dice que una aplicación $\phi : \Omega \to G$ es un **difeomorfismo** de clase C^1, cuando ϕ es biyectiva y de clase C^1 en Ω, mientras que ϕ^{-1} es de clase C^1 en G. Para cada $t \in \Omega$, denotamos por $J\phi(t)$ a la matriz jacobiana de ϕ en el punto t, que no es más que la matriz asociada a la aplicación lineal $D\phi(t)$. Sabemos que la función $t \mapsto |\det J\phi(t)|$ es continua, luego medible. Empezamos probando un resultado local, basado en la idea de que una función diferenciable se comporta localmente de manera similar a como lo hace una aplicación lineal.

- *Sea $\phi : \Omega \to G$ un difeomorfismo de clase C^1 entre dos abiertos de $\mathbb{R}^N$, y sea $\rho \in \mathbb{R}^+$ con $\rho > 1$. Entonces, cada punto $a \in \Omega$ tiene un entorno abierto $U \subset \Omega$ tal que, para todo conjunto medible $E \subset U$ se tiene:*

$$\frac{1}{\rho}\,\lambda\big(\phi(E)\big) \leqslant \int_E \big|\det J\phi(t)\big|\,dt \leqslant \rho\,\lambda\big(\phi(E)\big) \tag{6}$$

Seguimos usando la norma del máximo en $\mathbb{R}^N$ que denotaremos simplemente por $\|\cdot\|$, igual que la norma de cada aplicación lineal $S : \mathbb{R}^N \to \mathbb{R}^N$, con lo que $\|S(x)\| \leqslant \|S\|\,\|x\|$ para todo $x \in \mathbb{R}^N$. Fijado $a \in \Omega$, abreviamos escribiendo $T = D\phi(a)$, y sabemos que T es biyectiva. De hecho se tiene que $T^{-1} = D\phi^{-1}\big(\phi(a)\big)$. Es claro que

$$\lim_{\varepsilon \searrow 0} \frac{\big(1-\varepsilon\|T^{-1}\|\big)^N}{1+\varepsilon} = \lim_{\varepsilon \searrow 0} \frac{\big(1+\varepsilon\|T^{-1}\|\big)^N}{1-\varepsilon} = 1$$

luego existe $\varepsilon \in]0,1[$ tal que,

$$\frac{1}{\rho} \leqslant \frac{\big(1-\varepsilon\|T^{-1}\|\big)^N}{1+\varepsilon} \leqslant \frac{\big(1+\varepsilon\|T^{-1}\|\big)^N}{1-\varepsilon} \leqslant \rho$$

Tomando $\alpha = 1-\varepsilon\|T^{-1}\|$ y $\beta = 1+\varepsilon\|T^{-1}\|$, tenemos

$$0 < \alpha < \beta \qquad \text{y} \qquad \frac{\beta^N}{\rho} \leqslant 1-\varepsilon < 1+\varepsilon \leqslant \rho\,\alpha^N \tag{7}$$

Usando ahora que las funciones $t \mapsto D\phi(t)$ y $t \mapsto \det J\phi(t)$ son continuas, obtenemos una bola abierta $U \subset \Omega$, centrada en el punto a, tal que para todo $t \in U$ se tiene

$$\big\|D\phi(t) - T\big\| \leqslant \varepsilon \qquad \text{y} \qquad 1-\varepsilon \leqslant \frac{|\det J\phi(t)|}{|\det J\phi(a)|} \leqslant 1+\varepsilon \tag{8}$$

Recordemos que, como toda aplicación lineal, T es diferenciable en $\mathbb{R}^N$ y su diferencial en cualquier punto coincide con la propia T. Por tanto, la función $\phi - T$ es diferenciable en U, y se tiene que $D(\phi - T)(t) = D\phi(t) - T$ para todo $t \in U$.

Pues bien, como U es un abierto convexo, podemos usar el teorema del valor medio, y en vista de (8), obtenemos que $\phi - T$ es lipschitziana en U con constante ε, es decir,

$$\|\phi(x) - \phi(y) - T(x-y)\| = \|(\phi - T)(x) - (\phi - T)(y)\| \leqslant \varepsilon \|x-y\| \quad \forall x,y \in U \qquad (9)$$

Esta desigualdad indica que ϕ se comporta en U de manera similar a T, pero necesitamos controlar mejor esa similitud. Para cualesquiera $x,y \in U$, usando (9) tenemos

$$\begin{aligned}\|\phi(x) - \phi(y)\| &\leqslant \|T(x-y)\| + \varepsilon\|x-y\| = \|T(x-y)\| + \varepsilon\|T^{-1}(T(x-y))\| \\ &\leqslant \|T(x-y)\| + \varepsilon\|T^{-1}\|\,\|T(x-y)\| = \beta\,\|T(x)-T(y)\|\end{aligned}$$

De forma bastante análoga, tenemos también

$$\begin{aligned}\|T(x) - T(y)\| &\leqslant \|\phi(x)-\phi(y)\| + \varepsilon\|x-y\| = \|\phi(x)-\phi(y)\| + \varepsilon\|T^{-1}(T(x-y))\| \\ &\leqslant \|\phi(x)-\phi(y)\| + \varepsilon\|T^{-1}\|\,\|T(x-y)\|\end{aligned}$$

lo que equivale a $\|\phi(x)-\phi(y)\| \geqslant (1 - \varepsilon\|T^{-1}\|)\,\|T(x)-T(y)\| = \alpha\,\|T(x)-T(y)\|$. En resumen, hemos probado que

$$\alpha\,\|T(x)-T(y)\| \leqslant \|\phi(x)-\phi(y)\| \leqslant \beta\,\|T(x)-T(y)\| \qquad \forall x,y \in U \qquad (10)$$

Como ϕ es un homeomorfismo de Ω sobre G, el conjunto $\phi(U)$ es un abierto de G y por tanto de $\mathbb{R}^N$. Para $z,w \in \phi(U)$, podemos tomar $x = \phi^{-1}(z) \in U$, así como $y = \phi^{-1}(w) \in U$, y la primera desigualdad de (10) nos dice que

$$\|(T\circ\phi^{-1})(z) - (T\circ\phi^{-1})(w)\| = \|T(x)-T(y)\| \leqslant \frac{1}{\alpha}\,\|\phi(x)-\phi(y)\| = \frac{1}{\alpha}\,\|z-w\|$$

Esto prueba que la función $T\circ\phi^{-1}$ es lipschitziana en $\phi(U)$ con constante $1/\alpha$.

Repetimos el razonamiento anterior, intercambiando los papeles de T y ϕ. Como T es un homeomorfismo de $\mathbb{R}^N$, el conjunto $T(U)$ es abierto y, para cualesquiera $z,w \in T(U)$, usamos la segunda desigualdad de (10), con $x = T^{-1}(z) \in U$ e $y = T^{-1}(w) \in U$. Obtenemos:

$$\|(\phi\circ T^{-1})(z) - (\phi\circ T^{-1})(w)\| = \|\phi(x)-\phi(y)\| \leqslant \beta\,\|T(x)-T(y)\| = \beta\,\|z-w\|$$

así que la función $\phi\circ T^{-1}$ es lipschitziana en $T(U)$ con constante β.

Ha llegado el momento de usar lo que sabemos sobre el comportamiento de la medida de Lebesgue frente a varios tipos de funciones. Para ello fijamos un conjunto medible $E \subset U$. Tenemos entonces que $\phi(E)$ es un subconjunto medible de $\phi(U)$, el abierto en el que $T\circ\phi^{-1}$ es lipschitziana con constante $1/\alpha$, para la norma del máximo. Obtenemos que

$$\lambda(T(E)) = \lambda((T\circ\phi^{-1})(\phi(E))) \leqslant \frac{1}{\alpha^N}\,\lambda(\phi(E))$$

Nótese que hemos usado λ en lugar de λ^* porque sabemos que $\phi(E)$ y $T(E)$ son conjuntos medibles. Por otra parte, como T es lineal y su matriz asociada es $J\phi(a)$, conocemos la relación entre las medidas de E y $T(E)$, con lo que obtenemos

$$|\det J\phi(a)|\,\lambda(E) = \lambda(T(E)) \leqslant \frac{1}{\alpha^N}\,\lambda(\phi(E))$$

Para tener una desigualdad en sentido opuesto, intercambiamos de nuevo los papeles de T y ϕ. Como $T(E)$ es un subconjunto de $T(U)$, abierto en el que $\phi \circ T^{-1}$ es lipschitziana con constante β, tenemos

$$\lambda\big(\phi(E)\big) = \lambda\big(\big(\phi \circ T^{-1}\big)\big(T(E)\big)\big) \leqslant \beta^N \lambda\big(T(E)\big) = \beta^N \,\big|\det J\phi(a)\big|\, \lambda(E)$$

Enlazando las dos últimas desigualdades, obtenemos que

$$\alpha^N \,\big|\det J\phi(a)\big|\, \lambda(E) \leqslant \lambda\big(\phi(E)\big) \leqslant \beta^N \,\big|\det J\phi(a)\big|\, \lambda(E) \tag{11}$$

La última desigualdad indica que, en un entorno de a, la función ϕ transforma la medida de Lebesgue de forma muy similar a como lo hace T. Para llegar por fin al resultado que buscamos, bastará aprovechar la segunda desigualdad de (8), que hasta ahora no hemos usado. La escribimos en la forma

$$\big(1-\varepsilon\big)\,\big|\det J\phi(a)\big| \leqslant \big|\det J\phi(t)\big| \leqslant \big(1+\varepsilon\big)\,\big|\det J\phi(a)\big| \qquad \forall t \in U$$

de donde deducimos claramente que

$$\big(1-\varepsilon\big)\,\big|\det J\phi(a)\big|\,\lambda(E) \leqslant \int_E \big|\det J\phi(t)\big|\,dt \leqslant \big(1+\varepsilon\big)\,\big|\det J\phi(a)\big|\,\lambda(E) \tag{12}$$

Finalmente, enlazando (11) y (12), y usando también (7), obtenemos

$$\begin{aligned}\frac{1}{\rho}\,\lambda\big(\phi(E)\big) &\leqslant \frac{\beta^N}{\rho}\,\big|\det J\phi(t)\big|\,\lambda(E) \leqslant \big(1-\varepsilon\big)\,\big|\det J\phi(t)\big|\,\lambda(E)\\ &\leqslant \int_E \big|\det J\phi(t)\big|\,dt \leqslant \big(1+\varepsilon\big)\,\big|\det J\phi(t)\big|\,\lambda(E)\\ &\leqslant \rho\,\alpha^N \big|\det J\phi(t)\big|\,\lambda(E) \leqslant \rho\;\lambda\big(\phi(E)\big)\end{aligned}$$

y hemos probado la doble desigualdad (6), concluyendo la demostración. ■

Pasamos ahora a globalizar y hacer exacta, la información local y aproximada que nos da el resultado anterior. Obtenemos así lo que después quedará como caso particular del teorema de cambio de variable.

- *Si $\phi : \Omega \to G$ es un difeomorfismo de clase C^1 entre dos abiertos de $\mathbb{R}^N$, para todo conjunto medible $A \subset \Omega$, se tiene:*

$$\lambda\big(\phi(A)\big) = \int_A \big|\det J\phi(t)\big|\,dt \tag{13}$$

Fijamos un conjunto medible $A \subset \Omega$, y $\rho \in \mathbb{R}$ con $\rho > 1$. Para cada $x \in \Omega$, el resultado anterior nos da un entorno abierto $U(x)$ del punto x, con $U(x) \subset \Omega$, verificando (6) para todo conjunto medible $E \subset U(x)$. Entonces la familia de conjuntos $\big\{U(x) : x \in \Omega\big\}$ es un recubrimiento por abiertos de Ω del que podemos extraer un subrecubrimiento numerable. Así pues, tenemos $\Omega = \bigcup_{n=1}^{\infty} U_n$ donde, para cada $n \in \mathbb{N}$ el conjunto U_n es abierto y verifica (6) para todo conjunto medible $E \subset U_n$.

Observamos ahora que $A = \biguplus_{n=1}^{\infty} E_n$ donde $E_1 = A \cap U_1$ y $E_{n+1} = \big(A \cap U_{n+1}\big) \setminus \bigcup_{k=1}^{n} U_k$, para todo $n \in \mathbb{N}$. Como cada E_n es un subconjunto medible de U_n, usando (6) tenemos

$$\frac{1}{\rho}\,\lambda\big(\phi(E_n)\big) \leqslant \int_{E_n} \big|\det J\phi(t)\big|\,dt \leqslant \rho\,\lambda\big(\phi(E_n)\big) \qquad \forall n \in \mathbb{N}$$

Como ϕ es inyectiva, tenemos $\phi(A) = \biguplus_{n=1}^{\infty} \phi(E_n)$, lo que nos permite usar la σ-aditividad, tanto de λ como de la integral de una función medible positiva, para obtener por una parte que

$$\frac{1}{\rho}\,\lambda\big(\phi(A)\big) = \frac{1}{\rho}\sum_{n=1}^{\infty}\lambda\big(\phi(E_n)\big) \leqslant \sum_{n=1}^{\infty}\int_{E_n}\big|\det J\phi(t)\big|\,dt = \int_A \big|\det J\phi(t)\big|\,dt$$

y por otra que

$$\int_A \big|\det J\phi(t)\big|\,dt = \sum_{n=1}^{\infty}\int_{E_n}\big|\det J\phi(t)\big|\,dt \leqslant \rho\sum_{n=1}^{\infty}\lambda\big(\phi(E_n)\big) = \rho\,\lambda\big(\phi(A)\big)$$

Enlazando las dos últimas desigualdades obtenemos

$$\frac{1}{\rho}\,\lambda\big(\phi(A)\big) \leqslant \int_A \big|\det J\phi(t)\big|\,dt \leqslant \rho\,\lambda\big(\phi(A)\big)$$

Esto es válido para todo $\rho \in \mathbb{R}$ con $\rho > 1$, y haciendo $\rho \searrow 1$ obtenemos (13). ■

14.3. El teorema principal

Como se ha dicho, el último resultado es caso particular del teorema que buscamos, pero de él se deduce fácilmente el caso general mediante rutinas que ya hemos usado anteriormente.

Teorema de cambio de variable para funciones medibles positivas. *Sea $\phi : \Omega \to G$ un difeomorfismo de clase C^1 entre dos abiertos de $\mathbb{R}^N$. Dado un conjunto medible $E \subset \Omega$ y una función medible positiva $f : \phi(E) \to [0,\infty]$, consideremos la función $g : E \to [0,\infty]$ definida por $g(t) = f\big(\phi(t)\big)\big|\det J\phi(t)\big|$ para todo $t \in E$. Entonces g es medible y su integral sobre E coincide con la de f sobre $\phi(E)$, es decir,*

$$\int_{\phi(E)} f(x)\,dx = \int_E f\big(\phi(t)\big)\,\big|\det J\phi(t)\big|\,dt \tag{14}$$

Demostración. Sabemos que ϕ^{-1} preserva los conjuntos medibles, por ser una función de clase C^1 en G. Escribiendo $h(t) = f\big(\phi(t)\big)$ para todo $t \in E$, empezamos viendo que h es medible. Dado un abierto $U \subset [0,\infty]$, como f es medible, $f^{-1}(U)$ es un subconjunto medible de $\phi(E)$, luego el conjunto $h^{-1}(U) = \phi^{-1}\big(f^{-1}(U)\big)$ es medible, porque ϕ^{-1} preserva los conjuntos medibles. Como la función $t \mapsto |\det J\phi(t)|$ es continua en E, luego medible, vemos que g es medible, como producto de dos funciones medibles.

La igualdad (13) que ya conocemos, nos da directamente (14), pero sustituyendo f por una función característica. Concretamente, para $C \in \mathcal{M}$ tomamos $A = \phi^{-1}\big(C \cap \phi(E)\big)$ que es un subconjunto medible de E. Entonces tenemos $\phi(A) = C \cap \phi(E)$, y en particular $\chi_C\big(\phi(t)\big) = 1$ para todo $t \in A$, mientras que $\chi_C\big(\phi(t)\big) = 0$ para todo $t \in E \setminus A$. Usando (13) obtenemos:

$$\begin{aligned}\int_{\phi(E)} \chi_C(x)\,dx &= \lambda\big(C \cap \phi(E)\big) = \lambda\big(\phi(A)\big) = \int_A \big|\det J\phi(t)\big|\,dt \\ &= \int_A \chi_C\big(\phi(t)\big)\,\big|\det J\phi(t)\big|\,dt = \int_E \chi_C\big(\phi(t)\big)\,\big|\det J\phi(t)\big|\,dt\end{aligned}$$

que es (14) con χ_C en lugar de f.

De forma completamente rutinaria, extendemos el resultado al caso de una función simple positiva $s = \sum_{k=1}^{m} \rho_k \chi_{C_k}$ con $m \in \mathbb{N}$, $\rho_1, \rho_2, \ldots, \rho_m \in \mathbb{R}_0^+$ y $C_1, C_2, \ldots C_m \in \mathcal{M}$:

$$\begin{aligned}\int_{\phi(E)} s(x)\,dx &= \sum_{k=1}^{m} \rho_k \int_{\phi(E)} \chi_{C_k}(x)\,dx = \sum_{k=1}^{m} \rho_k \int_E \chi_{C_k}\big(\phi(t)\big)\,\big|\det J\phi(t)\big|\,dt \\ &= \int_E \left(\sum_{k=1}^{m} \rho_k \chi_{C_k}\big(\phi(t)\big)\right)\big|\det J\phi(t)\big|\,dt = \int_E s\big(\phi(t)\big)\,\big|\det J\phi(t)\big|\,dt\end{aligned}$$

El último paso es ya fácil de adivinar. El teorema de aproximación de Lebesgue nos da una sucesión $\{s_n\}$ de funciones simples positivas, tal que $\{s_n(x)\} \nearrow f(x)$ para todo $x \in \phi(E)$. Definimos entonces

$$g_n(t) = s_n\big(\phi(t)\big)\,|\det J\phi(t)| \qquad \forall\, t \in E,\ \forall\, n \in \mathbb{N}$$

Sabemos que $\{g_n\}$ es una sucesión de funciones medibles positivas, definidas en E. Además, para $t \in E$, podemos tomar $x = \phi(t)$ para obtener $\big\{s_n(\phi(t))\big\} = \big\{s_n(x)\big\} \nearrow f(x) = f(\phi(t))$, de donde $\{g_n(t)\} \nearrow g(t)$ para todo $t \in E$. Así pues, podemos usar el teorema de la convergencia monótona, tanto con la sucesión $\{s_n\}$ como con $\{g_n\}$. Usando también lo ya probado para funciones simples, obtenemos que

$$\begin{aligned}\int_{\phi(E)} f(x)\,dx &= \lim_{n\to\infty} \int_{\phi(E)} s_n(x)\,dx = \lim_{n\to\infty} \int_E s_n\big(\phi(t)\big)\,\big|\det J\phi(t)\big|\,dt \\ &= \lim_{n\to\infty} \int_E g_n(t)\,dt = \int_E g(t)\,dt = \int_E f\big(\phi(t)\big)\,\big|\det J\phi(t)\big|\,dt\end{aligned}$$

y hemos probado (14) en el caso general, lo que concluye la demostración. ∎

Deducimos inmediatamente la versión definitiva del teorema que nos ocupa.

Teorema de cambio de variable. *Sea $\phi : \Omega \to G$ un difeomorfismo de clase C^1 entre dos abiertos de $\mathbb{R}^N$. Dado un conjunto medible $E \subset \Omega$ y una función medible $f : \phi(E) \to \mathbb{R}$, se considera la función $g : E \to \mathbb{R}$ definida por $g(t) = f\big(\phi(t)\big)\,|\det J\phi(t)|$ para todo $t \in E$. Entonces f es integrable en $\phi(E)$ si, y sólo si, g es integrable en E, en cuyo caso se tiene:*

$$\int_{\phi(E)} f(x)\,dx = \int_E f\big(\phi(t)\big)\,\big|\det J\phi(t)\big|\,dt \tag{15}$$

Demostración. Para probar que g es medible, se razona igual que en el teorema anterior. Escribimos $h(t) = f\big(\phi(t)\big)$ para todo $t \in E$, y dado un abierto $U \subset \mathbb{R}$, se tiene que $f^{-1}(U)$ es medible, luego $h^{-1}(U) = \phi^{-1}\big(f^{-1}(U)\big)$ es medible, porque ϕ^{-1} preserva los conjuntos medibles. Como la función $t \mapsto |\det J\phi(t)|$ es continua en E, luego medible, vemos que g es medible, como producto de dos funciones reales medibles

Como $|f|$ es una función medible positiva, la versión anterior del teorema nos dice que:

$$\int_{\phi(E)} |f(x)|\,dx = \int_E |f|\big(\phi(t)\big)\,\big|\det J\phi(t)\big|\,dt = \int_E |g(t)|\,dt$$

La primera de estas integrales es finita si, y sólo si, lo es la tercera, es decir, $f \in \mathcal{L}_1\big(\phi(E)\big)$ si, y sólo si, $g \in \mathcal{L}_1(E)$.

Finalmente, cuando ambas funciones son integrables, usamos el teorema anterior para la parte positiva y la negativa de f, que también son funciones medibles positivas.

$$\begin{aligned}\int_{\phi(E)} f(x)\,dx &= \int_{\phi(E)} f^+(x)\,dx - \int_{\phi(E)} f^-(x)\,dx \\ &= \int_E f^+\big(\phi(t)\big)\,\big|\det J\phi(t)\big|\,dt - \int_E f^-\big(\phi(t)\big)\,\big|\det J\phi(t)\big|\,dt \\ &= \int_E \big[f^+\big(\phi(t)\big) - f^-\big(\phi(t)\big)\big]\,\big|\det J\phi(t)\big|\,dt = \int_E f\big(\phi(t)\big)\,\big|\det J\phi(t)\big|\,dt\end{aligned}$$

Tenemos así la igualdad (15), como queríamos. ■

A la hora de usar en la práctica el teorema anterior, el problema suele ser estudiar la integral de una función medible $f : A \to \mathbb{R}$, para cierto conjunto medible $A \subset \mathbb{R}^N$. Si queremos usar para ello un difeomorfismo $\phi : \Omega \to G$, necesitamos $A \subset G$, para poder tomar $E = \phi^{-1}(A)$, y así estudiar la integral de f sobre $\phi(E) = A$. Sin embargo, el teorema puede usarse en una situación ligeramente más general, que en la práctica se presenta con frecuencia. De hecho, basta que se tenga $\lambda(A \setminus G) = 0$, pues entonces se puede obviamente sustituir A por $A \cap G$. El teorema toma entonces la siguiente forma, que es la más adecuada para usarlo en la práctica:

- *Sea $\phi : \Omega \to G$ un difeomorfismo de clase C^1 entre dos abiertos de $\mathbb{R}^N$. Dado un conjunto medible $A \subset \mathbb{R}^N$ con $\lambda(A \setminus G) = 0$, y una función medible $f : A \to \mathbb{R}$, consideramos el conjunto medible $E = \phi^{-1}(A \cap G) = \{t \in \Omega : \phi(t) \in A\}$, y la función $g : E \to \mathbb{R}$ definida por $g(t) = f\big(\phi(t)\big)\,|\det J\phi(t)|$ para todo $t \in E$. Entonces f es integrable en A si, y sólo si, g es integrable en E, en cuyo caso se tiene:*

$$\int_A f(x)\,dx = \int_E f\big(\phi(t)\big)\,\big|\det J\phi(t)\big|\,dt \tag{16}$$

La hipótesis $\lambda(A \setminus G) = 0$ hace que f sea integrable en A si, y sólo si, es integrable en el conjunto $A \cap G = \phi(E)$, y por el teorema anterior, esto equivale a que g sea integrable en E. Pero además, el teorema anterior también nos da la igualdad (16), ya que

$$\int_A f(x)\,dx = \int_{A\cap G} f(x)\,dx = \int_{\phi(E)} f(x)\,dx = \int_E f\big(\phi(t)\big)\,\big|\det J\phi(t)\big|\,dt \qquad ■$$

En los ejemplos de cambio de variable que vamos a presentar, ocurre que $\lambda(\mathbb{R}^N \setminus G) = 0$, con lo que el resultado anterior puede usarse para cualquier conjunto medible $A \subset \mathbb{R}^N$.

14.4. Coordenadas polares

En el cálculo de integrales dobles, es frecuente que la descripción del conjunto sobre el que se integra, sea mucho más sencilla usando las coordenadas polares de sus elementos que si usamos coordenadas cartesianas. Alternativamente, la función integrando puede simplificarse al expresarla en términos de las coordenadas polares, e incluso se pueden dar simultáneamente las dos circunstancias. En todos esos casos, puede ser muy útil usar el cambio de variable a coordenadas polares, que pasamos a explicar.

Consideramos el abierto $\Omega = \mathbb{R}^+ \times]-\pi, \pi[$ y la función $\phi : \Omega \to \mathbb{R}^2$ dada por

$$\phi(\rho, \theta) = \big(\rho \cos\theta, \rho \operatorname{sen}\theta\big) \qquad \forall (\rho, \theta) \in \Omega$$

La elección de Ω hace que ϕ sea inyectiva, como vamos a comprobar. Si $(\rho_1, \theta_1), (\rho_2, \theta_2) \in \Omega$ verifican que $\phi(\rho_1, \theta_1) = \phi(\rho_2, \theta_2) = (x, y)$, tenemos $\rho_1^2 = \rho_2^2 = x^2 + y^2$, de donde $\rho_1 = \rho_2$, ya que $\rho_1, \rho_2 \in \mathbb{R}^+$. Entonces $\cos\theta_1 = \cos\theta_2$ y $\operatorname{sen}\theta_1 = \operatorname{sen}\theta_2$, de donde $\theta_1 - \theta_2 = 2k\pi$ con $k \in \mathbb{Z}$, pero siendo $|\theta_1| < \pi$ y $|\theta_2| < \pi$, se tiene $2|k|\pi = |\theta_1 - \theta_2| < 2\pi$, de donde $k = 0$, y tenemos $\theta_1 = \theta_2$.

Para calcular la imagen de ϕ, observamos que si $(\rho, \theta) \in \Omega$ verifica que $\phi(\rho, \theta) = (x, 0)$ con $x \in \mathbb{R}$, al ser $|\theta| < \pi$, se ha de tener $\theta = 0$ luego $x = \rho > 0$. Por tanto, para $x \in \mathbb{R}_0^-$ se tiene que $(x, 0) \notin \phi(\Omega)$, es decir, tomando $G = \mathbb{R}^2 \setminus \{(x, 0) : x \in \mathbb{R}_0^-\}$, tenemos claramente un abierto $G \subset \mathbb{R}^2$ tal que $\phi(\Omega) \subset G$, y vamos a probar que esta inclusión es una igualdad.

Fijado $(x, y) \in G$, empezamos tomando $\rho = \sqrt{x^2 + y^2} > 0$ y observamos que $\rho + x \neq 0$, pues en otro caso, $x = -\rho \in \mathbb{R}^-$, y de $x^2 = \rho^2$ se deduce $y = 0$, con lo que $(x, y) \notin G$. Ahora tomamos $\theta = 2 \operatorname{arc\,tg} u \in]-\pi, \pi[$, donde $u = y/(\rho + x)$. Recordando el cambio de variable que usábamos para las integrales simples de funciones trigonométricas, tenemos

$$\cos\theta = \frac{1 - u^2}{1 + u^2} = \frac{x}{\rho} \qquad \text{y} \qquad \operatorname{sen}\theta = \frac{2u}{1 + u^2} = \frac{y}{\rho}$$

Por tanto, $(x, y) = \phi(\rho, \theta) \in \phi(\Omega)$ y hemos probado que $\phi(\Omega) = G$.

En resumen, tenemos una aplicación biyectiva $\phi : \Omega \to G$, cuya inversa viene dada por

$$\phi^{-1}(x, y) = \left(\sqrt{x^2 + y^2}\,,\ 2 \operatorname{arc\,tg} \frac{y}{\sqrt{x^2 + y^2} + x} \right) \qquad \forall (x, y) \in G$$

Es claro que ϕ es de clase C^∞ en Ω, e igual le ocurre a ϕ^{-1} en G. En particular, ϕ es un difeomorfismo de clase C^1 de Ω sobre G. Para todo $(\rho, \theta) \in \mathbb{R}$ se tiene claramente:

$$\frac{\partial\phi}{\partial\rho}(\rho, \theta) = \big(\cos\theta, \operatorname{sen}\theta\big) \qquad \text{y} \qquad \frac{\partial\phi}{\partial\theta}(\rho, \theta) = \big(-\rho \operatorname{sen}\theta, \rho\cos\theta\big)$$

de donde deducimos que $\det J\phi(\rho, \theta) = \rho\big(\cos^2\theta + \operatorname{sen}^2\theta\big) = \rho > 0$, para todo $(\rho, \theta) \in \Omega$. Por último, observamos que $\lambda\big(\mathbb{R}^N \setminus G\big) = \lambda\big(\{(x, 0) : x \in \mathbb{R}_0^-\}\big) = 0$, lo que permite usar el teorema de cambio de variable en cualquier conjunto medible $A \subset \mathbb{R}^2$. Obtenemos la siguiente regla práctica para el cálculo de integrales dobles.

- **Cambio de variable a coordenadas polares.** *Dado un conjunto medible $A \subset \mathbb{R}^2$ y una función medible $f : A \to \mathbb{R}$, consideremos el conjunto medible*

$$E = \big\{ (\rho, \theta) \in \mathbb{R}^+ \times]-\pi, \pi[\, : \, (\rho \cos\theta, \rho \operatorname{sen}\theta) \in A \big\}$$

y la función $g : E \to \mathbb{R}$ dada por $g(\rho,\theta) = \rho \, f(\rho\cos\theta, \rho \operatorname{sen}\theta)$, para todo $(\rho,\theta) \in E$. Entonces $f \in \mathcal{L}_1(A)$ si, y sólo si, $g \in \mathcal{L}_1(E)$, en cuyo caso se tiene:

$$\int_A f(x,y) \, d(x,y) = \int_E \rho \, f\big(\rho\cos\theta, \rho \operatorname{sen}\theta\big) \, d(\rho,\theta)$$

Como ejemplo sencillo, veamos el área de un círculo de radio $r \in \mathbb{R}^+$ que, salvo traslación, podemos suponer centrado en el origen, es decir, del conjunto $A = \big\{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leqslant r^2\big\}$. Volvemos a denotar por λ_2 a la medida de Lebesgue en $\mathbb{R}^2$. Para $(\rho,\theta) \in \mathbb{R}^+ \times]-\pi, \pi[$, es claro que $\big(\rho\cos\theta, \rho \operatorname{sen}\theta\big) \in A$, si, y sólo si, $\rho \leqslant r$. Por tanto, en el enunciado anterior, tenemos $E =]0, r] \times]-\pi, \pi[$. Tras el cambio de variable a coordenadas polares, usamos el teorema de Fubini, obteniendo

$$\lambda_2(A) = \int_A d(x,y) = \int_E \rho \, d(\rho,\theta) = \int_{-\pi}^{\pi} \left(\int_0^r \rho \, d\rho \right) d\theta = \int_{-\pi}^{\pi} \frac{r^2}{2} \, d\theta = \pi r^2$$

Como aplicación menos evidente del resultado anterior, calculemos el área encerrada bajo la llamada *campana de Gauss*, esto es, el área de la subgráfica de la función $x \mapsto e^{-x^2}$, que sabemos viene dada por

$$\lambda_2\big(\big\{(x,y) \in \mathbb{R}^2 : 0 < y < e^{-x^2}\big\}\big) = \int_{-\infty}^{+\infty} e^{-x^2} \, dx$$

El integrando no admite una primitiva expresable en términos de funciones conocidas, por lo que la regla de Barrow no nos ayuda a evaluar la integral anterior. Aunque parezca que complicamos el problema, la relacionamos con una integral doble, mediante el teorema de Fubini:

$$\left(\int_{-\infty}^{+\infty} e^{-x^2} \, dx \right)^2 = \left(\int_{-\infty}^{+\infty} e^{-x^2} \, dx \right) \left(\int_{-\infty}^{+\infty} e^{-y^2} \, dy \right) = \int_{\mathbb{R}^2} e^{-(x^2+y^2)} \, d(x,y)$$

Sin embargo, esta integral doble se calcula fácilmente, mediante un cambio de variable a coordenadas polares. Usamos el resultado anterior con $A = \mathbb{R}^2$ y por tanto $E = \mathbb{R}^+ \times]-\pi, \pi[$. El teorema de Fubini nos lleva entonces a una integral simple, que se calcula inmediatamente con la regla de Barrow:

$$\begin{aligned} \int_{\mathbb{R}^2} e^{-(x^2+y^2)} \, d(x,y) &= \int_E \rho \, e^{-\rho^2} \, d(\rho,\theta) = \int_{-\pi}^{\pi} \left(\int_0^{+\infty} \rho \, e^{-\rho^2} \, d\rho \right) d\theta \\ &= 2\pi \int_0^{+\infty} \rho \, e^{-\rho^2} \, d\rho = 2\pi \left[-\frac{e^{-\rho^2}}{2} \right]_0^{+\infty} = \pi \end{aligned}$$

Así pues, el área bajo la campana de Gauss es: $\displaystyle\int_{-\infty}^{+\infty} e^{-x^2} \, dx = \sqrt{\pi}$.

14.5. Coordenadas cilíndricas

Hay dos sistemas de coordenadas curvilíneas en $\mathbb{R}^3$ que hacen un papel análogo al de las coordenadas polares en $\mathbb{R}^2$. En primer lugar, consideramos las coordenadas cilíndricas, que guardan una relación muy directa con las polares.

Las coordenadas cilíndricas de cada punto $(x,y,z) \in \mathbb{R}^3$, son (ρ,θ,z), donde (ρ,θ) son las coordenadas polares del punto $(x,y) \in \mathbb{R}^2$. Con esta observación, se adivina claramente todo el estudio del cambio de variable a coordenadas cilíndricas que hacemos en lo que sigue.

Consideramos el abierto $\Omega = \mathbb{R}^+ \times]-\pi,\pi[\times \mathbb{R}$ y la función $\phi : \Omega \to \mathbb{R}^3$ definida por

$$\phi(\rho,\theta,z) = \big(\rho\cos\theta,\, \rho\,\mathrm{sen}\,\theta,\, z\big) \qquad \forall\,(\rho,\theta,z) \in \Omega \tag{17}$$

Del estudio hecho anteriormente para las coordenadas polares, se deduce directamente que ϕ es inyectiva y que su imagen es el conjunto abierto $G = \mathbb{R}^3 \setminus \{(x,0,z) : x \in \mathbb{R}_0^-,\, z \in \mathbb{R}\}$. Por tanto, $\phi : \Omega \to G$ es biyectiva, y de hecho sabemos que su inversa viene dada por

$$\phi^{-1}(x,y,z) = \left(\sqrt{x^2+y^2}\,,\ 2\,\mathrm{arc\,tg}\,\frac{y}{\sqrt{x^2+y^2}+x}\,,\ z \right) \qquad \forall\,(x,y,z) \in G$$

De nuevo ϕ y ϕ^{-1} son funciones de clase C^∞ en Ω y G respectivamente, y en particular, ϕ es un difeomorfismo de clase C^1 de Ω sobre G. Para todo $(\rho,\theta,z) \in \Omega$, se tiene

$$\frac{\partial\phi}{\partial\rho}(\rho,\theta,z) = \big(\cos\theta, \mathrm{sen}\,\theta, 0\big)\,, \ \ \frac{\partial\phi}{\partial\theta}(\rho,\theta,z) = \big(-\rho\,\mathrm{sen}\,\theta, \rho\cos\theta, 0\big)\,, \ \ \frac{\partial\phi}{\partial z}(\rho,\theta,z) = \big(0,0,1\big)$$

de donde se deduce que $\det J\phi(\rho,\theta,z) = \rho > 0$. Como $\mathbb{R}^3 \setminus G = \{(x,0,z) : x \in \mathbb{R}_0^-,\ z \in \mathbb{R}\}$ está contenido en el plano de ecuación $y = 0$, tenemos $\lambda_3(\mathbb{R}^3 \setminus G) = 0$, siendo λ_3 la medida de Lebesgue en $\mathbb{R}^3$. Esto permite usar el teorema de cambio de variable en cualquier conjunto medible $A \subset \mathbb{R}^3$. Tenemos así una regla práctica para el cambio de variable a coordenadas cilíndricas, enteramente análoga a la obtenida para las coordenadas polares en el plano.

- **Cambio de variable a coordenadas cilíndricas.** *Dado un conjunto medible $A \subset \mathbb{R}^3$ y una función medible $f : A \to \mathbb{R}$, consideramos el conjunto medible*

  $$E = \big\{ (\rho,\theta,z) \in \mathbb{R}^+ \times]-\pi,\pi[\times \mathbb{R} \ : \ (\rho\cos\theta,\, \rho\,\mathrm{sen}\,\theta,\, z) \in A \big\}$$

  y definimos $g : E \to \mathbb{R}$ por $g(\rho,\theta,z) = \rho\, f(\rho\cos\theta,\, \rho\,\mathrm{sen}\,\theta,\, z)$, para todo $(\rho,\theta,z) \in E$. Entonces $f \in \mathcal{L}_1(A)$ si, y sólo si, $g \in \mathcal{L}_1(E)$, en cuyo caso se tiene:

  $$\int_A f(x,y,z)\, d(x,y,z) = \int_E \rho\, f\big(\rho\cos\theta,\, \rho\,\mathrm{sen}\,\theta,\, z\big)\, d(\rho,\theta,z)$$

Esta regla suele ser muy útil para calcular volúmenes de conjuntos que tengan una simetría axial, o integrales triples sobre conjuntos del mismo tipo. Como ejemplo más significativo, vamos a calcular el volumen de un sólido de revolución.

Consideremos un conjunto medible $S \subset \mathbb{R}_0^+ \times \mathbb{R}$, y sea $S_0 = \{(u,0,z) : (u,z) \in S\}$, que es un subconjunto medible de $\mathbb{R}^3$, contenido en el plano de ecuación $y = 0$, y tal que todos sus puntos tienen la primera coordenada no negativa. Sea entonces A el sólido de revolución engendrado al girar el conjunto S_0 alrededor del eje vertical, esto es, de la recta $\mathbb{R}(0,0,1)$. Cada punto de S_0 tiene la forma $(u,0,z)$ con $(u,z) \in S$ y al girar alrededor del eje indicado, recorre una circunferencia contenida en un plano horizontal, centrada en el punto $(0,0,z)$ y de radio u, es decir, el conjunto $\{(x,y,z) \in \mathbb{R}^3 : x^2 + y^2 = u^2\}$. Como A es la unión de todas las circunferencias de este tipo, tenemos:

$$A = \left\{ \big(x,y,z\big) \in \mathbb{R}^3 : \big(\sqrt{x^2+y^2}, z\big) \in S \right\} \tag{18}$$

Esta vez no es evidente que A sea medible, pero lo comprobaremos enseguida. Denotando por $\phi : \Omega \to G$ al difeomorfismo definido en (17), el conjunto E que aparece en el último enunciado viene dado por

$$E = \phi^{-1}(A \cap G) = \{(\rho,\theta,z) \in \Omega : (\rho,z) \in S\}$$

El intercambio de la segunda coordenada con la tercera es una isometría de $\mathbb{R}^3$ en sí mismo, que transforma E en el conjunto $\big(S \cap (\mathbb{R}^+ \times \mathbb{R})\big) \times]-\pi, \pi[$, que es medible, luego E también lo es. Como ϕ preserva los conjuntos medibles, deducimos que $\phi(E) = A \cap G$ es medible, pero sabemos que $\lambda_3(A \setminus G) = 0$, luego A es medible.

Podemos ya usar el último resultado, junto con el teorema de Fubini.

$$\lambda_3(A) = \int_A d(x,y,z) = \int_E \rho\, d(\rho,\theta,z) = \int_{-\pi}^{\pi} \left(\int_S \rho\, d(\rho,z) \right) d\theta = 2\pi \int_S \rho\, d(\rho,z)$$

En la última integral, podemos cambiar el nombre de la variable de integración, ρ por x, que representa mejor la abscisa de los puntos de S y de S_0. Mantenemos z como segunda variable, por ser la tercera coordenada de los puntos de S_0. Por tanto, hemos probado que el volumen del sólido de revolución A definido en (18) viene dado por

$$\lambda_3(A) = 2\pi \int_S x\, d(x,z) \tag{19}$$

Como ejemplo más sencillo, si $S = [0, r] \times [0, h]$ con $r,h \in \mathbb{R}^+$, entonces A es un cilindro de altura h cuya base es un círculo de radio r. El volumen de dicho cilindro es:

$$\lambda_3(A) = 2\pi \int_0^h \left(\int_0^r x\,dx \right) dz = 2\pi \frac{r^2}{2} h = \pi r^2 h$$

Otro ejemplo más interesante aparece tomando como S un círculo de radio $r \in \mathbb{R}^+$ centrado en el punto $(R,0) \in \mathbb{R}^2$ con $R > r$. Entonces $S_0 = \left\{ (u,0,z) \in \mathbb{R}^3 : (u-R)^2 + z^2 \leqslant r^2 \right\}$, es también un círculo de radio r que gira alrededor del eje vertical, de forma que su centro describe una circunferencia de radio R. Por tanto, el sólido de revolución A es un *rosco*, al que se suele llamar **toro sólido**, de radio menor r y radio mayor R.

Para calcular en este caso la integral doble que aparece en (19) usamos una traslación y un cambio de variable a coordenadas polares. Escribiendo $S = (R,0) + C$ y $F =]0\,,r]\times]-\pi\,,\pi[$, obtenemos:

$$\frac{\lambda_3(A)}{2\pi} = \int_S x\,d(x,z) = \int_C (R+x)\,d(x,z) = \int_F \rho\,(R+\rho\cos\theta)\,d(\rho,\theta)$$
$$= \int_0^r \rho\left(\int_{-\pi}^{\pi}(R+\rho\cos\theta)\,d\theta\right)d\rho = 2\pi R\int_0^r \rho\,d\rho = \pi R r^2$$

Por tanto, el volumen del toro sólido viene dado por $\lambda_3(A) = (2\pi R)(\pi r^2)$. Como se podía intuir, este volumen coincide con el de un cilindro de altura $2\pi R$, cuya base es un círculo de radio r, justo el cilindro que se obtendría al "desenroscar" el toro sólido.

Comentemos por último un hecho que puede parecer sorprendente. Aunque el conjunto S tenga un área finita, el sólido de revolución A puede tener volumen infinito. Es lo que le ocurre por ejemplo al conjunto $S = \{\,(x,z)\in\mathbb{R}^2 : x\geqslant 1\,,\ 0\leqslant z\leqslant 1/x^2\,\}$, cuya área es

$$\lambda_2(S) = \int_1^{+\infty}\left(\int_0^{1/x^2} dz\right)dx = \int_1^{+\infty}\frac{dx}{x^2} = 1$$

Sin embargo, el volumen del correspondiente sólido de revolución A viene dado por

$$\lambda_3(A) = \int_S x\,d(x,z) = \int_1^{+\infty} x\left(\int_0^{1/x^2} dz\right)dx = \int_1^{+\infty}\frac{dx}{x} = \infty$$

14.6. Coordenadas esféricas

Las coordenadas cilíndricas son útiles para problemas que presentan una simetría axial, como ocurre con los sólidos de revolución que hemos estudiado. Para problemas con simetría respecto al origen, suelen ser más útiles las coordenadas esféricas que ahora vamos a estudiar.

Las coordenadas esféricas de un punto $P = (x,y,z)\in\mathbb{R}^3$ se denotan por (r,θ,φ) y se interpretan geométricamente como sigue. En primer lugar, r es la distancia euclídea del origen al punto P, así que P está situado en la esfera de radio r centrada en el origen. Entonces φ es la *latitud* de P, entendiendo que el *ecuador* está contenido en el plano $z=0$, siendo la latitud positiva en el *hemisferio norte* y negativa en el *sur*. Vemos entonces que $\varphi\in[-\pi/2,\pi/2]$ verifica que $z = r\operatorname{sen}\varphi$. El conjunto de puntos de latitud φ forman una circunferencia, el *paralelo* de latitud φ, cuyo radio es $r\cos\varphi$. Finalmente, θ es la *longitud* del punto P en la misma esfera, entendiendo que el origen de *meridianos* se sitúa en el plano $y=0$, siendo la longitud positiva hacia el *este* y negativa hacia el *oeste*. Si P tiene longitud $\theta\in[-\pi,\pi]$, se ha de tener $x = r\cos\varphi\cos\theta$, mientras que $y = r\cos\varphi\operatorname{sen}\theta$. Esta interpretación geométrica ayuda a entender los razonamientos que siguen.

Consideramos el conjunto abierto $\Omega = \mathbb{R}^+\times]-\pi\,,\pi[\times]-\pi/2\,,\pi/2[\subset\mathbb{R}^3$, en el que se define la función $\phi:\Omega\to\mathbb{R}^3$ dada por

$$\phi\big(r\,,\theta\,,\varphi\big) = \big(r\cos\varphi\cos\theta\,,\ r\cos\varphi\operatorname{sen}\theta\,,\ r\operatorname{sen}\varphi\big)\qquad\forall\,(r,\theta,\varphi)\in\Omega$$

Para comprobar que ϕ es inyectiva, supongamos que $\phi(r_1,\theta_1,\varphi_1) = \phi(r_2,\theta_2,\varphi_2) = (x,y,z)$, donde $(r_1,\theta_1,\varphi_1)\,,(r_2,\theta_2,\varphi_2) \in \Omega$. Entonces se tiene que

$$x^2+y^2+z^2 \,=\, r_1^2\big(\cos^2\theta_1+\operatorname{sen}^2\theta_1\big)\cos^2\varphi_1 \,+\, r_1^2\operatorname{sen}^2\varphi_1 \,=\, r_1^2$$

y análogamente, $x^2+y^2+z^2 = r_2^2$, luego $r_1^2 = r_2^2$, de donde $r_1 = r_2$, ya que $r_1, r_2 \in \mathbb{R}^+$. Deducimos que $\operatorname{sen}\varphi_1 = z/r_1 = z/r_2 = \operatorname{sen}\varphi_2$, y siendo $\varphi_1, \varphi_2 \in]-\pi/2\,,\pi/2\,[$, esto implica que $\varphi_1 = \varphi_2$. Como además, $\cos\varphi_1 = \cos\varphi_2 > 0$, tenemos ahora que

$$\big(\cos\theta_1,\operatorname{sen}\theta_1\big) = \left(\frac{x}{r_1\,\cos\varphi_1},\frac{y}{r_1\,\cos\varphi_1}\right) = \left(\frac{x}{r_2\,\cos\varphi_2},\frac{y}{r_2\,\cos\varphi_2}\right) = \big(\cos\theta_2,\operatorname{sen}\theta_2\big)$$

y siendo $\theta_1,\theta_2 \in]-\pi\,,\pi\,[$, concluimos que $\theta_1 = \theta_2$. En resumen, $(r_1,\theta_1,\varphi_1) = (r_2,\theta_2,\varphi_2)$, y esto prueba que ϕ es inyectiva.

Para buscar la imagen de ϕ, sea $(r,\theta,\varphi) \in \Omega$ verificando $\phi(r,\theta,\varphi) = (x,0,z)$ con $x,z \in \mathbb{R}$. De $r\operatorname{sen}\theta\cos\varphi = 0$, siendo $r\cos\varphi > 0$, obtenemos que $\operatorname{sen}\theta = 0$. Como $\theta \in]-\pi\,,\pi\,[$, esto implica que $\theta = 0$, luego $\cos\theta = 1$ y deducimos que $x = r\cos\varphi > 0$. Así pues, para $x \in \mathbb{R}_0^-$ y $z \in \mathbb{R}$, se tiene que $(x,0,z) \notin \phi(\Omega)$.

Equivalentemente, considerando el abierto $G = \mathbb{R}^3 \setminus \big(\mathbb{R}_0^- \times \{0\} \times \mathbb{R}\big)$, tenemos $\phi(\Omega) \subset G$. Para probar la otra inclusión, fijamos $(x,y,z) \in G$, y observamos que $(x,y) \neq (0,0)$ con lo que tomamos $r = \sqrt{x^2+y^2+z^2} \in \mathbb{R}^+$. De hecho tenemos $z^2 < r^2$, es decir, $|z| < r$, lo que permite tomar $\varphi = \operatorname{arc\,sen}(z/r)$ para tener $\varphi \in]-\pi/2\,,\pi/2\,[$ que verifica $z = r\operatorname{sen}\varphi$.

Aprovechamos ahora el estudio previo de las coordenadas polares. Como $(x,y) \notin \mathbb{R}_0^- \times \mathbb{R}$, tomando $\rho = \sqrt{x^2+y^2} > 0$, existe $\theta \in]-\pi\,,\pi\,[$ tal que $\rho\cos\theta = x$ y $\rho\operatorname{sen}\theta = y$. Como por otra parte $x^2+y^2 = r^2-z^2 = r^2\cos^2\varphi$, siendo $r\cos\varphi > 0$, obtenemos que $\rho = r\cos\varphi$, de donde deducimos que $r\cos\varphi\cos\theta = \rho\cos\theta = x$ y $r\cos\varphi\operatorname{sen}\theta = \rho\operatorname{sen}\theta = y$. En resumen, tenemos $\phi(r,\theta,\varphi) = (x,y,z)$ y hemos probado que $\phi(\Omega) = G$. Así pues $\phi : \Omega \to G$ es biyectiva y de hecho hemos visto que su inversa viene dada, para todo $(x,y,z) \in G$ por:

$$\phi^{-1}(x,y,z) \,=\, \Big(\sqrt{x^2+y^2+z^2}\,,\; 2\operatorname{arc\,tg}\frac{y}{\sqrt{x^2+y^2}+x}\,,\; \operatorname{arc\,sen}\frac{z}{\sqrt{x^2+y^2+z^2}}\Big)$$

Vemos que ϕ es de clase C^∞ en Ω, mientras que ϕ^{-1} lo es en G. En particular, ϕ es un difeomorfismo de clase C^1 de Ω sobre G. En cada punto $(r,\theta,\varphi) \in \Omega$ las derivadas parciales de ϕ vienen dadas por

$$\begin{aligned}
\frac{\partial\phi}{\partial r}\big(r,\theta,\varphi\big) \;&=\; \big(\cos\theta\cos\varphi\,,\;\operatorname{sen}\theta\cos\varphi\,,\operatorname{sen}\varphi\big)\\
\frac{\partial\phi}{\partial\theta}\big(r,\theta,\varphi\big) \;&=\; \big(-r\operatorname{sen}\theta\cos\varphi\,,\; r\cos\theta\cos\varphi\,,\; 0\big)\\
\frac{\partial\phi}{\partial\varphi}\big(r,\theta,\varphi\big) \;&=\; \big(-r\cos\theta\operatorname{sen}\varphi\,,\; -r\operatorname{sen}\theta\operatorname{sen}\varphi\,,\; r\cos\varphi\big)
\end{aligned}$$

Esto permite escribir la matriz jacobiana y comprobar fácilmente que

$$\det J\phi(r,\theta,\varphi) \,=\, r^2\,\cos\varphi \,>\, 0 \qquad \forall\,(r,\theta,\varphi) \in \Omega$$

Teniendo en cuenta que una vez más se tiene $\lambda_3(\mathbb{R}^3 \setminus G) = 0$, podemos ya enunciar la regla práctica que se usa en este caso.

- **Cambio de variable a coordenadas esféricas.** *Dado un conjunto medible* $A \subset \mathbb{R}^3$ *y una función medible* $f : A \to \mathbb{R}$*, consideramos el conjunto medible*

$$E = \big\{(r,\theta,\varphi) \in \mathbb{R}^+ \times]-\pi,\pi[\times]-\pi/2,\pi/2[: (\rho\cos\varphi\cos\theta, r\cos\varphi\operatorname{sen}\theta, r\operatorname{sen}\varphi) \in A\big\}$$

y definimos $g : E \to \mathbb{R}$ *por* $g(r,\theta,\varphi) = r^2\cos\varphi\, f(r\cos\varphi\cos\theta,\, r\cos\varphi\operatorname{sen}\theta,\, r\operatorname{sen}\varphi)$*, para todo* $(r,\theta,\varphi) \in E$*. Entonces* $f \in \mathcal{L}_1(A)$ *si, y sólo si,* $g \in \mathcal{L}_1(E)$*, en cuyo caso:*

$$\int_A f(x,y,z)\,d(x,y,z) = \int_E r^2\cos\varphi\, f(r\cos\varphi\cos\theta, r\cos\varphi\operatorname{sen}\theta, r\operatorname{sen}\varphi)\,d(r,\theta,\varphi)$$

Como ejemplo muy sencillo, pero suficiente para mostrar la utilidad de la regla anterior, calculemos el volumen de una bola euclídea en $\mathbb{R}^3$, que salvo traslación podemos suponer centrada en el origen. Tomamos por tanto

$$A = \big\{\,(x,y,z) \in \mathbb{R}^3 \,:\, x^2+y^2+z^2 \leqslant R^2\,\big\}$$

con $R \in \mathbb{R}^+$. Está claro que el conjunto E, definido en el enunciado anterior, viene dado por

$$E =]\,0\,,R\,]\times]-\pi\,,\pi\,[\,\times\,]-\pi/2\,,\pi/2\,[$$

con lo que la regla anterior y el teorema de Fubini nos dan:

$$\begin{aligned}\lambda_3(A) &= \int_A d(x,y,z) = \int_E r^2\,\cos\varphi\,d(r,\theta,\varphi) = \int_0^R\left[\int_{\pi}^{\pi}\left(\int_{-\pi/2}^{\pi/2} r^2\,\cos\varphi\,d\varphi\right)d\theta\right]dr\\ &= \int_0^R\left(\int_{-\pi}^{\pi} 2\,r^2\,d\theta\right)dr = \int_0^R 4\pi r^2\,dr = \frac{4}{3}\,\pi\,R^3\end{aligned}$$

Como se puede adivinar, para calcular la integral de cualquier función, sobre la bola A siempre será conveniente usar las coordenadas esféricas, y ese cambio estará especialmente indicado cuando se trate de una función *radial*, que en coordenadas esféricas sólo dependerá de la variable r. Como ejemplo ilustrativo, fijado $\alpha \in \mathbb{R}$, consideremos la función $f : \mathbb{R}^3 \to \mathbb{R}$ dada por

$$f(x,y,z) = \big(x^2+y^2+z^2\big)^{\alpha} \qquad \forall\,(x,y,z) \in \mathbb{R}^3 \setminus \{(0,0,0)\}$$

sin que importe el valor de f en el origen. Estudiemos la integrabilidad de f en la bola unidad euclídea $B = \big\{\,(x,y,z) \in \mathbb{R}^3 \,:\, x^2+y^2+^2 \leqslant 1\big\}$, así como en el conjunto $\mathbb{R}^3 \setminus B$. Razonando como antes, pero ahora con $R = 1$, tomamos

$$E =]\,0\,,1\,]\times]-\pi,\pi\,[\,\times\,]-\pi/2\,,\pi/2\,[$$

y el cambio de variable a coordenadas esféricas nos dice que

$$\begin{aligned}\int_B f(x,y,z)\,d(x,y,z) &= \int_E r^{2\alpha+2}\cos\varphi\,d(r,\theta,\varphi) = \int_0^1\left[\int_{-\pi}^{\pi}\left(\int_{-\pi/2}^{\pi/2} r^{2\alpha+2}\cos\varphi\,d\varphi\right)d\theta\right]dr\\ &= \int_0^1\left(\int_{-\pi}^{\pi} 2\,r^{2\alpha+2}\,d\theta\right)dr = 4\pi\int_0^1 r^{2\alpha+2}\,dr\end{aligned}$$

La función potencia de exponente $2\alpha+2$ es integrable en $]0,1[$ si, y sólo si, $2\alpha+2>-1$, es decir, $\alpha>-3/2$. Deducimos que f es integrable en B si, y sólo si, $\alpha>-3/2$, en cuyo caso:

$$\int_B f(x,y,z)\,d(x,y,z) = 4\pi\left[\frac{r^{2\alpha+3}}{2\alpha+3}\right]_0^1 = \frac{4\pi}{2\alpha+3} \qquad \forall\alpha\in]-3/2,+\infty[$$

Para el conjunto $C=\mathbb{R}^3\setminus B$ el razonamiento es análogo, pero tomando

$$E=]1,+\infty[\times]-\pi,\pi[\times]-\pi/2,\pi/2[$$

Obtenemos entonces que

$$\begin{aligned}\int_B f(x,y,z)\,d(x,y,z) &= \int_E r^{2\alpha+2}\cos\varphi\, d(r,\theta,\varphi) = \int_1^{+\infty}\left[\int_{-\pi}^{\pi}\left(\int_{-\pi/2}^{\pi/2} r^{2\alpha+2}.\cos\varphi\, d\varphi\right)d\theta\right]dr\\ &= \int_1^{+\infty}\left(\int_{-\pi}^{\pi} 2\,r^{2\alpha+2}d\theta\right)dr = 4\pi\int_1^{+\infty} r^{2\alpha+2}\,dr\end{aligned}$$

La función potencia de exponente $2\alpha+2$ es integrable en $]1,+\infty[$ si, y sólo si, $2\alpha+2<-1$, es decir $\alpha<-3/2$. Estos son los valores de α para los que f es integrable en C y se tiene:

$$\int_C f(x,y,z)\,d(x,y,z) = 2\pi\left[\frac{r^{2\alpha+3}}{2\alpha+3}\right]_1^{+\infty} = -\frac{4\pi}{2\alpha+3} \qquad \forall\alpha\in]-\infty,-3/2[$$

Deducimos que, cualquiera que sea $\alpha\in\mathbb{R}$, la función f nunca es integrable en $\mathbb{R}^3$.