1. Teorema de Cambio de Variable para la Integral de Riemann.

En el caso de la integral de Riemann para funciones reales de una variable real, se puede demostrar un teorema de cambio de variable de forma muy sencilla utilizando los teoremas fundamentales del cálculo, en condiciones “buenas” sobre la función que se quiere integrar y sobre la función de cambio de variable:

Supongamos que $g : [a, b] \longrightarrow \mathbb{R}$ es una función continua en $[a, b]$, derivable en (a, b) y con derivada continua, y que $f : \mathbb{R} \longrightarrow \mathbb{R}$ es una función continua. Entonces el teorema de cambio de variable asegura que

$$\int_{g(a)}^{g(b)} f = \int_a^b (f \circ g) \cdot g'$$

En efecto, si F es una primitiva de f en $[g(a), g(b)]$, se tiene que

$$\int_{g(a)}^{g(b)} f = F(g(b)) - F(g(a))$$

Por otro lado, la regla de la cadena asegura que $(F \circ g)' = (F' \circ g) \cdot g' = (f \circ g) \cdot g'$, es decir, $F \circ g$ es una primitiva de $(f \circ g) \cdot g'$, y se tiene también

$$\int_a^b (f \circ g) \cdot g' = F \circ g(b) - F \circ g(a) = F(g(b)) - F(g(a))$$

Se pueden dar teoremas más generales de cambio de variable, con condiciones menos fuertes sobre la función f (sólo integrable) y más fuertes en la función g (difeomorfismo de clase C^1 en (a, b)), a los que llegaremos como caso particular del teorema de cambio de variable para funciones de varias variables que vamos a demostrar.

La situación en el caso de funciones en $\mathbb{R}^n$ será en términos generales la siguiente: tendremos un conjunto medible-Jordan N, una función biyectiva y diferenciable $g : N \longrightarrow \mathbb{R}^n$, y una función integrable f definida en $g(N) = M$. Y se tratará de demostrar que, en ciertas condiciones, se verifica la igualdad

$$\int_M f = \int_N (f \circ g) \cdot |Jg|$$

donde $|Jg|$ es el valor absoluto del Jacobiano de g.

Para ello habrá que asegurar primero que $g(N)$ es un conjunto medible-Jordan, para que tenga sentido la integral de f sobre $M = g(N)$; en segundo lugar habrá que demostrar que la función $(f \circ g) \cdot |Jg|$ es integrable; y por último habrá que comprobar la igualdad de las integrales.

Iremos resolviendo cada uno de estos problemas en varios pasos, empezando por casos sencillos sobre las funciones f y g. El primer teorema va encaminado a establecer condiciones suficientes sobre la función g para asegurar que transforma conjuntos medibles en conjuntos medibles.

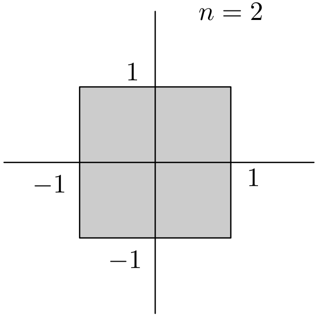
Por la mayor comodidad que supone en la utilización de las bolas como rectángulos, utilizaremos en $\mathbb{R}^n$ la norma infinito: si $x = (x_1, \dots, x_n)$, la norma infinito de x es

$$\|x\|_\infty = \max\{|x_1|, \dots, |x_n|\}$$

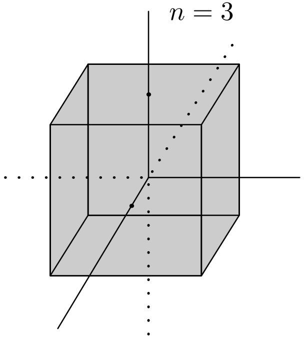
Con esta norma, la bola de centro x y radio $r > 0$ es

$$\overline{B}(x,r) = \{y \in \mathbb{R}^n : \max\{|x_1 - y_1|, \ldots, |x_n - y_n|\} \leq r\}$$

que es el rectángulo $[x_1 - r, x_1 + r] \times \cdots \times [x_n - r, x_n + r]$



$\overline{B}(0,1)$

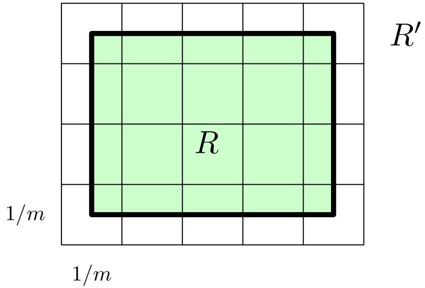


$\overline{B}(0,1)$

Antes de nada, conviene tener en cuenta la siguiente observación:

Observación 1. *En las definiciones de conjuntos de contenido cero y de medida cero, se pueden sustituir los rectángulos por cubos (rectángulos con todos lados de la misma longitud).*

En efecto, basta tener en cuenta que para todo $\epsilon > 0$, un rectángulo R de $\mathbb{R}^n$ se puede incluir en otro rectángulo R' de modo que $v(R') \leq v(R) + \epsilon$, y tal que las longitudes de los lados de R' sean números racionales. De esta forma, si $R' = [a_1, b_1] \times \cdots \times [a_n, b_n]$, con $b_i - a_i = r_i \in \mathbb{Q}$, podemos escribir $r_i = k_i/m$ poniendo común denominador, lo que quiere decir que cada segmento $[a_i, b_i]$ se puede subdividir en k_i intervalos de longitud m^{-1}; por tanto R' se puede dividir en $k_1 \cdot \cdots \cdot k_n$ cubos de lado m^{-1}, y la suma de los volúmenes de estos cubos es igual al volumen de R'.



En consecuencia para todo $\epsilon > 0$, un rectángulo R en $\mathbb{R}^n$ está contenido en una familia finita de cubos $Q_1, \ldots, Q_k$ de modo que $\sum_{i=1}^{k} v(Q_i) \leq v(R) + \epsilon$

$\square$

Utilizaremos también la siguiente definición:

Definición (Función Lipschitziana)**.** *Sea U un conjunto en $\mathbb{R}^n$, y $G : U \longrightarrow \mathbb{R}^m$ una función.*

Se dice que G es lipschitziana si existe una constante $K > 0$ tal que para todo par de puntos x e y en $\mathbb{R}^n$ se verifica

$$\|G(x) - G(y)\| \leq K \|x - y\|$$

Por ejemplo, las aplicaciones lineales son funciones lipschitizianas en todo $\mathbb{R}^n$, y si F es una función diferenciable en un punto x_0 de un abierto U, entonces es localmente lipschitziana, es decir, existe una bola centrada en x_0 y contenida en U donde F es lipschitziana.

Teorema 1.

1. *Sea $H \subset \mathbb{R}^n$ un conjunto de medida cero (resp. de contenido cero) y sea $g : H \longrightarrow \mathbb{R}^m$ $(n \leq m)$ una aplicación lipschitziana. Entonces $g(H)$ tiene medida cero (resp. contenido cero).*

2. *Sea $A \subset \mathbb{R}^n$ abierto, y $g : A \longrightarrow \mathbb{R}^m$ $(n \leq m)$ una función de clase C^1 en A. Sea $H \subset A$ un conjunto de medida cero (resp. contenido cero) tal que $\overline{H}$ esté contenido en A. Entonces $g(H)$ tiene medida cero(resp. contenido cero).*

3. *Sea $A \subset \mathbb{R}^n$ abierto y $g : A \longrightarrow \mathbb{R}^m$ $(n \leq m)$ una función de clase C^1 en A. Sea $H \subset A$ un conjunto de medida cero. Entonces $g(H)$ tiene medida cero.*

Observación 2. *El último apartado del teorema no es cierto sustituyendo medida cero por contenido cero.*

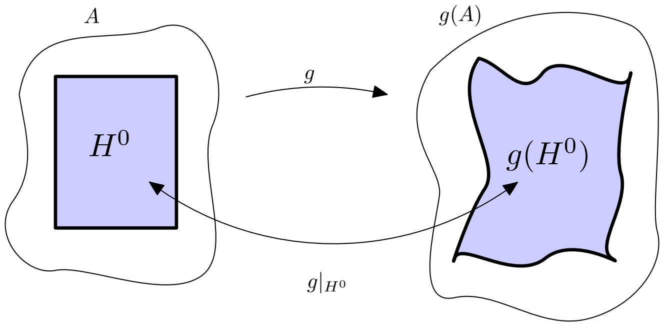
Para poner un ejemplo, considérese la función $g : (\frac{-\pi}{2}, \frac{\pi}{2}) \longrightarrow \mathbb{R}$, definida por $g(x) = \tan(x)$ y el conjunto $H = \{\frac{\pi}{2} - \frac{1}{n}, n \in I\!N\}$.

g es una función de clase C^1 en el abierto $A = (\frac{-\pi}{2}, \frac{\pi}{2})$, H es un conjunto de contenido nulo, pues es una sucesión convergente en $\mathbb{R}$, y verifica $H \subset A$, y sin embargo $g(H)$ no puede tener contenido nulo ya que no es un conjunto acotado ($g(\frac{\pi}{2} - \frac{1}{n}) = \tan(\frac{\pi}{2} - \frac{1}{n})$ tiende a infinito cuando n tiende a infinito).

$\square$

Como consecuencia del teorema anterior veamos ahora que un difeomorfismo de clase C^1 en un conjunto abierto de $\mathbb{R}^n$ transforma conjuntos medibles en conjuntos medibles. De hecho lo demostramos en condiciones un poco más generales, que incluyen la mayoría de los casos prácticos.

Proposición 1. *Sea A un abierto de $\mathbb{R}^n$ y sea $g : A \longrightarrow \mathbb{R}^n$ una función de clase C^1 en A. Sea H un conjunto medible-Jordan, tal que $\overline{H} \subset A$ y tal que la restricción de g a H^0, interior de H, sea un difeomorfismo de clase C^1. Entonces $g(H)$ es medible-Jordan.*



Demostración:
Si H es medible-Jordan, en particular es acotado y por tanto su adherencia $\overline{H}$ es compacto. Entonces como $g(H) \subset g(\overline{H})$, y $g(\overline{H})$ es compacto por ser la imagen por una función continua de un conjunto compacto, se tiene que $g(H)$ es acotado.

Además, $\overline{g(H)}$ es el menor cerrado que contiene a $g(H)$, y por tanto $\overline{g(H)} \subset g(\overline{H})$.

Por otro lado, la hipótesis de que $g|_{H^0} : H^0 \longrightarrow g(H^0)$ es un difeomorfismo de clase C^1 implica, como consecuencia del teorema de la función inversa, que $g(H^0)$ es abierto, y por tanto que $g(H^0) \subset g(H)^0$.

Por último,

$$Fr(g(H)) = \overline{g(H)} \setminus g(H)^0 \subset g(\overline{H}) \setminus g(H^0) \subset g(\overline{H} \setminus H^0) = g(Fr(H))$$

Aplicando el teorema anterior al conjunto $Fr(H)$, que es un subconjunto cerrado de A de contenido cero, se tiene que $g(Fr(H))$ tiene contenido cero, y por tanto también $Fr(g(H))$ tiene contenido cero.

En consecuencia, $g(H)$ es medible-Jordan.

$\square$

Otra aplicación sencilla del teorema anterior es la siguiente demostración de que todo subespacio vectorial propio de $\mathbb{R}^n$ (un subespacio vectorial se llama "propio" si no es el vacío ni el total) es un conjunto de medida nula, que utilizaremos más adelante.

Proposición 2. *Todo subespacio vectorial propio de $\mathbb{R}^n$ tiene medida cero.*

Demostración:
Sea H un subespacio de $\mathbb{R}^n$ de dimensión $k < n$, y sea $\{v_1, \ldots, v_k\}$ una base de H. Cada vector v de H será de la forma $v = \sum_{i=1}^{k} \lambda_i v_i$ con $(\lambda_1, \ldots, \lambda_k) \in \mathbb{R}^k$. Definamos entonces la función

$$\begin{aligned} g : \mathbb{R}^{k+1} &\longrightarrow \mathbb{R}^n \\ (\lambda_1, \ldots, \lambda_k, \lambda) &\longrightarrow \sum_{i=1}^{k} \lambda_i v_i \end{aligned}$$

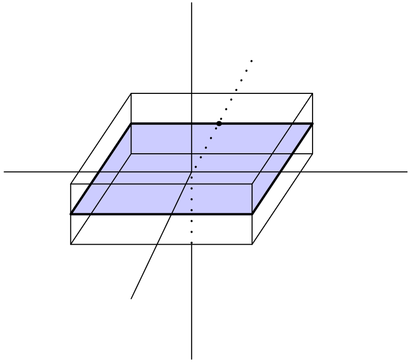
Es claro que podemos poner $H = g(\mathbb{R}^k \times \{0\})$.
Ahora bien, $\mathbb{R}^k \times \{0\}$ tiene medida cero en $\mathbb{R}^{k+1}$:

En efecto, $(\mathbb{R}^k \times \{0\}) \subset \bigcup_{i=1}^{\infty} Q_i$ donde

$$Q_i = \overbrace{[-i,i] \times \cdots \times [-i,i]}^{k} \times [\frac{-\epsilon}{(2i)^k \cdot 2^{i+1}}, \frac{\epsilon}{(2i)^k \cdot 2^{i+1}}]$$

y

$$\sum_{i=1}^{\infty} v(Q_i) = \sum_{i=1}^{\infty} (2i)^k \frac{\epsilon}{(2i)^k \cdot 2^{i+1}} = \epsilon$$



Y, por último, $g : I\!R^{k+1} \longrightarrow \mathbb{R}^n$ es lipschitziana:

$$\|g(\lambda_1, \ldots, \lambda_k, \lambda) - g(\mu_1, \ldots, \mu_k, \mu)\|_\infty =$$

$$= \| \sum_{i=1}^{k} (\lambda_i - \mu_i) v_i \|_\infty \leq \sum_{i=1}^{k} |\lambda_i - \mu_i| \|v_i\| \leq$$

$$\leq \text{máx}\{|\lambda_i - \mu_i|,\ 1 \leq i \leq k\} \cdot \sum_{i=1}^{k} \|v_i\|$$

$$\leq \|(\lambda_1, \ldots, \lambda_k, \lambda) - (\mu_1, \ldots, \mu_k, \mu)\|_\infty \cdot L$$

siendo $L = \sum_{i=1}^{k} \|v_i\|$.

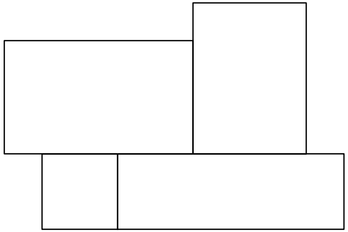
Aplicado el teorema se deduce que $H = g(I\!R^k \times \{0\})$ tiene medida nula.

□

Antes de seguir adelante con la demostración de una primera versión de cambio de variable, para aplicaciones lineales, vamos a destacar algunas observaciones técnicas que se repiten en las demostraciones siguientes.

Observación 3. *Si $\mathcal{Q}$ es una familia finita de conjuntos medibles en $\mathbb{R}^n$ que no se solapan (es decir, $M_i^0 \cap M_j^0 = \emptyset$, para todo $i \neq j$), entonces*

$$v(\bigcup_{M \in \mathcal{Q}} M) = \sum_{M \in \mathcal{Q}} v(M)$$



En efecto, podemos poner

$$v(\bigcup_{M\in\mathcal{Q}} M) = v\left((\bigcup_{M\in\mathcal{Q}} M^0) \cup (\bigcup_{M\in\mathcal{Q}} Fr(M)) \right)$$

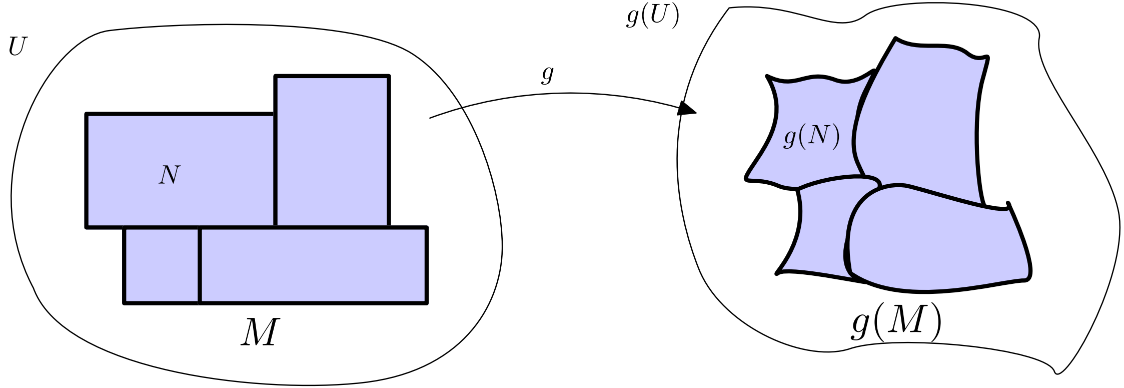
Aquí, $(\bigcup_{M\in\mathcal{Q}} M^0)$ y $(\bigcup_{M\in\mathcal{Q}} Fr(M))$ son conjuntos disjuntos. Además, para cada $M \in \mathcal{Q}$, la frontera $Fr(M)$ de M tiene contenido cero, por lo que el conjunto $(\bigcup_{M\in\mathcal{Q}} Fr(M))$ tiene contenido cero. Entonces

$$v(\bigcup_{M\in\mathcal{Q}} M) = v(\bigcup_{M\in\mathcal{Q}} M) = \sum_{M\in\mathcal{Q}} v(M^0) = \sum_{M\in\mathcal{Q}} v(M)$$

$\square$

Observación 4. *Sea $\mathcal{Q}$ una familia de conjuntos medibles en $\mathbb{R}^n$ que no se solapan, y sea U un abierto en $\mathbb{R}^n$ tal que el conjunto $M = (\bigcup_{N\in\mathcal{Q}} N)$ verifique $\overline{M} \subset U$. Sea g una función de clase C^1 de U en $\mathbb{R}^n$, tal que la restricción a M^0 sea un difeomorfismo de clase C^1. Entonces*

$$v(g(M)) = \sum_{N\in\mathcal{Q}} v(g(N))$$



En efecto, por un lado

$$v(g(M)) = v\left(g(\bigcup_{N\in\mathcal{Q}} N)\right) = v\left(g(\bigcup_{N\in\mathcal{Q}} N^0) \cup g(\bigcup_{N\in\mathcal{Q}} Fr(N))\right) =$$

$$= v\left((\bigcup_{N\in\mathcal{Q}} g(N^0)) \cup (\bigcup_{N\in\mathcal{Q}} g(Fr(N)))\right) \leq v(\bigcup_{N\in\mathcal{Q}} g(N^0)) + 0$$

$$= \sum_{N\in\mathcal{Q}} v(g(N^0))$$

puesto que, por el teorema 1, $\bigcup_{N\in\mathcal{Q}} g(Fr(N))$ tiene contenido cero.

Por otro lado, como $\bigcup_{N\in\mathcal{Q}} g(N^0)$ es un subconjunto de $g(M)$, es claro que

$$v(\bigcup_{N\in\mathcal{Q}} g(N^0)) \leq v(g(M))$$

Por tanto,

$$\begin{aligned} v(g(M)) &= \sum_{N\in\mathcal{Q}} v(g(N^0)) = \sum_{N\in\mathcal{Q}} \big(v(g(N^0)) + v(g(Fr(N)))\big) \\ &= \sum_{N\in\mathcal{Q}} v(g(N)) \end{aligned}$$

como queríamos demostrar.

□

Por último nos interesa destacar una propiedad de descomposición de los isomorfismos lineales de $\mathbb{R}^n$, (aplicaciones lineales biyectivas de $\mathbb{R}^n$ en $°R^n$):

Observación 5. *Todo isomorfismo lineal $L : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ se puede descomponer como composición de aplicaciones lineales "elementales" de los tres tipos siguientes:*
Si $x = (x_1, \cdots, x_n)$ y $Lx = ((Lx)_1, \cdots, (Lx)_n)$

- *Tipo A: Existe $\lambda \in \mathbb{R}$ y existe $i,\ 1 \leq i \leq n$ tal que $(Lx)_i = \lambda x_i$, y $(Lx)_j = x_j$ para todo $j \neq i$.*
- *Tipo B: Existen $i, k,\ 1 \leq i, k \leq n$, tales que $(Lx)_i = x_k$, $(Lx)_k = x_i$, y $(Lx)_j = x_j$, para todo $j \neq i, k$.*
- *Tipo C: Existen $i, k,\ 1 \leq i, k \leq n$ tales que $(Lx)_k = x_i + x_k$ y $(Lx)_j = x_j$ para todo $j \neq k$.*

Este resultado, que no vamos a demostrar, es el fundamento del método de Gauss para la inversión de matrices.

Con estas observaciones, podemos demostrar el siguiente teorema:

Teorema 2. *Toda aplicación lineal* $L : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ *transforma conjuntos medibles-Jordan en conjuntos medibles-Jordan. Además, para cada conjunto medible-Jordan* M *en* $\mathbb{R}^n$ *se tiene*

$$v(L(M)) = |\det L| \cdot v(M)$$

Demostración:
Toda aplicación lineal en $\mathbb{R}^n$ es una función de clase C^1 en todo el espacio $\mathbb{R}^n$. Distinguiremos dos casos: cuando L es un isomorfismo y cuando no lo es.

Caso primero: Supongamos que L no es un isomorfismo. Entonces la imagen $L(\mathbb{R}^n)$ es un subespacio vectorial propio de $\mathbb{R}^n$, y por tanto, por la proposición 2, $L(\mathbb{R}^n)$ tiene medida cero. Además todo subespacio vectorial es un cerrado, y por otro lado, si M es medible-Jordan en particular es acotado y $L(M)$ es también acotado.

Así que tenemos $L(M) \subset \overline{L(M)} \subset L(\mathbb{R}^n)$, donde $\overline{L(M)}$ es un subconjunto compacto de un conjunto de medida nula. En consecuencia $\overline{L(M)}$ tiene contenido nulo, y también $L(M)$ tiene contenido nulo.

En particular $L(M)$ es medible, y $v(L(M)) = 0$. Como por otro lado el determinante $\det L$ es cero, por no ser L un isomorfismo, también se verifica que $|\det L| \cdot v(M) = 0$, y se tiene el resultado.

Caso segundo: Supongamos ahora que L es un isomorfismo en $\mathbb{R}^n$.

En primer lugar, todo isomorfismo lineal es un difeomorfismo de clase C^1 en todo $\mathbb{R}^n$, por lo que, aplicando la proposición 1, L transforma conjuntos medibles-Jordan en conjuntos medibles-Jordan.

En segundo lugar, demostraremos que basta probar el teorema en el caso en que L es una aplicación lineal elemental como las definidas en la observación anterior. En efecto, si L es un isomorfismo, existe una descomposición de L de la forma $L = L_1 \circ L_2 \circ \cdots \circ L_k$, donde L_i son aplicaciones elementales, $1 \leq i \leq k$. Si suponemos que el resultado es cierto para estas aplicaciones elementales, y M es un conjunto medible en $\mathbb{R}^n$, se tiene

$$
\begin{aligned}
&v(L(M)) = v(L_1(L_2 \circ \cdots \circ L_k(M))) = \\
&\quad = |\det L_1| v(L_2 \circ \cdots \circ L_k(M)) = \cdots = |\det L_1| \ldots |\det L_k| \cdot v(M) = \\
&\quad = |\det L| v(M)
\end{aligned}
$$

y por tanto el resultado sería cierto también para L.

Y en tercer lugar vamos a ver que dada una aplicación elemental L, basta demostrar el resultado cuando el conjunto medible M es un rectángulo.

En efecto, supongamos que el resultado es cierto para rectángulos, y sea M un conjunto medible-Jordan cualquiera. Sea A un abierto en $\mathbb{R}^n$ tal que $\overline{M} \subset A$.

Dado $\epsilon > 0$, sea P una partición de A tal que

$$
\overline{S}(\chi_M, P) - \underline{S}(\chi_M, P) < \epsilon
$$

Si llamamos $E_1 = \bigcup\{R \in \mathcal{P},\ R \cap M \neq \emptyset\}$ y $E_2 = \bigcup\{R^0, R \in \mathcal{P}, R \subset M\}$, se tiene:

a) E_1 y E_2 son medibles-Jordan, por ser unión finita de conjuntos medibles. Además $E_2 \subset M \subset E_1$, y por tanto $v(E_2) \leq v(M) \leq v(E_1)$.

b)

$$\overline{S}(\chi_M, P) = \sum_{R \in \mathcal{P}} M_R(\chi_M) v(R) = \sum_{R \in \mathcal{P}, R \cap M \neq \emptyset} v(R) = v(E_1)$$

puesto que si $R \cap M = \emptyset$ la función característica de M vale cero en cada punto de R, y $M_R(\chi_M) = 0$, y por otro lado, si $R \cap M \neq \emptyset$ entonces hay al menos un punto de R en el que χ_M vale uno, y $M_R(\chi_M) = 1$.

Y análogamente

$$\begin{aligned} \underline{S}(\chi_M, P) &= \sum_{R \in \mathcal{P}} m_R(\chi_M) v(R) = \sum_{R \in \mathcal{P}, R \subset M} v(R) = \\ &= \sum_{R \in \mathcal{P}, R \subset M} v(R^0) = v(E_2) \end{aligned}$$

c) $L(E_1)) = L(\bigcup\{R \in \mathcal{P},\ R \cap M \neq \emptyset\}) = \bigcup\{L(R), R \in \mathcal{P},\ R \cap M \neq \emptyset\}$ y por tanto, si el resultado es cierto para rectángulos,

$$v(L(E_1)) = \sum_{R \in \mathcal{P},\ R \cap M \neq \emptyset} v(L(R)) = |\det L| v(E_1)$$

Y análogamente, $L(E_2) = \bigcup\{L(R^0), R \in \mathcal{P},\ R \subset M\}$, y

$$v(L(E_2)) = \sum_{R\in\mathcal{P},\ R\subset M} v(L(R^0)) = |\det L| v(E_2)$$

d) Por otro lado, de la desigualdad $E_2 \subset M \subset E_1$ se deduce que también $L(E_2) \subset L(M) \subset L(E_1)$, y por tanto

$$v(L(E_2)) \leq v(L(M)) \leq v(L(E_1))$$

Sustituyendo $v(L(E_2))$ y $v(L(E_1))$ por los valores obtenidos en (c), se tiene

$$|\det L| v(E_2) \leq v(L(M)) \leq |\det L| v(E_1) \qquad \text{I}$$

Por último, si en la desigualdad obtenida en (a) multiplicamos por $|\det L|$, se tiene

$$|\det L| v(E_2) \leq |\det L| v(M) \leq |\det L| v(E_1) \qquad \text{II}$$

Restando I y II, se obtiene

$$\begin{aligned} |\det L|(v(E_2) - v(E_1)) &\leq v(L(M)) - |\det L| v(M) \leq \\ &\leq |\det L|(v(E_1) - v(E_2)) \end{aligned}$$

de donde

$$|v(L(M)) - |\det L|v(M)| \leq |\det L|(v(E_1) - v(E_2)) =$$
$$= |\det L|(\overline{S}(\chi_M, P) - \underline{S}(\chi_M, P)) \leq |\det L|\epsilon$$

Como esto es cierto para todo $\epsilon > 0$, tiene que ser

$$v(L(M)) = |\det L|v(M)$$

Luego, efectivamente, si el resultado se demuestra para rectángulos, entonces es cierto para cualquier conjunto medible-Jordan.

Sea entonces R un rectángulo, $R = [a_1, b_1] \times \ldots, \times[a_n, b_n]$, y L una aplicación lineal elemental.

Primer caso: Si L es de tipo A, es decir, existe i, $1 \leq i \leq n$ tal que $(Lx)_i = \lambda x_i$ para algún número real λ, y $(Lx)_j = x_j$ para todo $j \neq i$, la matriz de la aplicación lineal es de la forma

$$
\begin{array}{cc}
 & \begin{array}{ccccccc} & & & i & & & \end{array} \\
i & \begin{pmatrix}
1 & & & & & & \\
 & \ddots & & & & \mathbf{0} & \\
 & & 1 & & & & \\
 & & & \lambda & & & \\
 & & & & 1 & & \\
 & \mathbf{0} & & & & \ddots & \\
 & & & & & & 1
\end{pmatrix}
\end{array}
$$

donde son cero todos los términos fuera de la diagonal, y unos todos los términos de la diagonal excepto el de lugar ii que vale λ. En particular, $|\det L| = |\lambda|$.

Por otro lado, la imagen del rectángulo R es un rectángulo

$$
\begin{aligned}
L(R) &= \{L(x_1, \ldots, x_i, \ldots, x_n),\ (x_1, \ldots, x_i, \ldots, x_n) \in R\} = \\
&= \{(x_1, \ldots, \lambda x_i, \ldots, x_n),\ (x_1, \ldots, x_i, \ldots, x_n) \in R\} = \\
&= [a_1, b_1] \times \cdots \times [\lambda a_i, \lambda b_i] \times \cdots \times [a_n, b_n]
\end{aligned}
$$

que tiene volumen $v(L(R)) = |\lambda| v(R) = |\det L| v(R)$, y se tiene el resultado.

Segundo caso: Si L es de tipo B, es decir, existen i, k, $1 \leq i, k \leq n$ tales que $(Lx)_i = x_k$, $(Lx)_k = x_i$ y para todo $j \neq i, k$ $(Lx)_j = x_j$, la matriz de L es de la forma

$$\begin{array}{cc}
 & \begin{array}{ccccccc} & & i & & k & & \end{array} \\
\begin{array}{c} \\ \\ i \\ \\ k \\ \\ \\ \end{array} & \left(\begin{array}{ccccccc}
1 & & & & & & \\
 & \ddots & & & & \mathbf{0} & \\
 & & 0 & & 1 & & \\
 & & & \ddots & & & \\
 & & 1 & & 0 & & \\
 & \mathbf{0} & & & & \ddots & \\
 & & & & & & 1
\end{array}\right)
\end{array}$$

donde son unos todos los elementos de la diagonal excepto los de los lugares ii y kk, que pasan a las posiciones ik y ki, y son cero todos lor términos que no están indicados explícitamente.

Es claro entonces que $|\det L| = 1$. Por otro lado, la imagen del rectángulo R es un rectángulo

$$L(R) = \{L(x_1, \dots, x_i, \dots, x_k, \dots, x_n),\ (x_1, \dots, x_i, \dots, x_k, \dots, x_n) \in R\} =$$

$$= \{(x_1, .., x_{i-1}, x_k, x_{i+1}, .., x_{k-1}, x_i, x_{k+1}, .., x_n),\ (x_1, .., x_i, .., x_k, .., x_n) \in R\} =$$

$$= [a_1, b_1] \times \cdots \times [a_{i-1}, b_{i-1}] \times [a_k, b_k] \times [a_{i+1}, b_{i+1}] \times \dots$$
$$\cdots \times [a_{k-1}, b_{k-1}] \times [a_i, b_i] \times [a_{k+1}, b_{k+1}] \times \cdots \times [a_n, b_n]$$

que tiene el mismo volumen que R. Por tanto también en este caso

$$v(L(R)) = v(R) = |\det L| v(R)$$

Tercer caso: Si L es de tipo C, es decir, existen i, k, $1 \leq i, k \leq n$ tales que $(Lx)_k = x_i + x_k$, y para todo $j \neq k$ $(Lx)_j = x_j$, la matriz de L es de la forma

$$\begin{array}{c c}
 & \begin{array}{ccccccc} & & i & & & & \end{array} \\
\begin{array}{c} \\ \\ i \\ \\ k \\ \\ \\ \end{array} & \begin{pmatrix}
1 & & & & & & \\
 & \ddots & & & & \mathbf{0} & \\
 & & 1 & & & & \\
 & & & \ddots & & & \\
 & & 1 & & 1 & & \\
 & \mathbf{0} & & & & \ddots & \\
 & & & & & & 1
\end{pmatrix}
\end{array}$$

donde son unos todos los términos de la diagonal y el de posición ki, y son cero todos los términos que no aparecen indicados explícitamente. En particular el determinante de L verifica $|\det L| = 1$.

Por otro lado, la imagen de R está contenida en el rectángulo

$$L(R) = \{L(x_1, \ldots, x_k, \ldots, x_n),\ (x_1, \ldots, x_k, \ldots, x_n) \in R\} =$$

$$= \{(x_1, ..., x_{k-1}, x_i + x_k, x_{k+1}, ..., x_n),\ (x_1, ..., x_{k-1}, x_k, x_{k+1}, ..., x_n) \in R\} \subset$$
$$\subset [a_1, b_1] \times ... \times [a_{k-1}, b_{k-1}] \times [a_i + a_k, b_i + b_k] \times [a_{k+1}, b_{k+1}] \times ... \times [a_n, b_n]$$
$$= R'$$

Llamemos R'' al rectángulo

$$R'' = [a_1, b_1] \times \cdots \times [a_{k-1}, b_{k-1}] \times [a_{k+1}, b_{k+1}] \times \cdots \times [a_n, b_n]$$

Aplicando el Teorema de Fubini para calcular $v(L(R))$, se tiene

$$(*) \qquad v(L(R)) = \int_{R'} \chi_{L(R)} =$$
$$= \int_{R''} \left(\int_{[a_i+a_k, b_i+b_k]} \chi_{L(R)}(y_1, .., y_k, .., y_n) dy_k \right) d(y_1, .., y_{k-1}, y_{k+1}, .., y_n)$$

Ahora bien, fijo $(y_1, \ldots, y_{k-1}, y_{k+1}, \ldots, y_n)$ en R'', la función $\chi_{R)}$ vale uno en $(y_1, \ldots, y_{k-1}, y_k, y_{k+}$
si y sólo si este punto está en $L(R)$, es decir, si y sólo si existe un punto $(x_1, \ldots, x_i, \ldots, x_k, \ldots, x_n) \in$
R tal que

$$y_1 = x_1;\ \ldots;\ y_i = x_i;\ \ldots$$
$$y_k = x_i + x_k = y_i + x_k$$
$$\ldots;\ y_n = x_n$$

lo que equivale a que $y_k \in [y_i + a_k, y_i + b_k]$. Sustituyendo estos límites en la integral (*), se tiene

$$\begin{aligned} v(L(R)) &= \int_{R''} \left(\int_{y_i+a_k}^{y_i+b_k} 1 dy_k \right) d(y_1, \ldots, y_{k-1}, y_{k+1}, \ldots, y_n) = \\ &= \int_{R''} (b_k - a_k) d(y_1, \ldots, y_{k-1}, y_{k+1}, \ldots, y_n) = \\ &= (b_k - a_k) v(R'') = v(R) \end{aligned}$$

lo que prueba que también en este último caso $v(L(R)) = |\det L| v(R)$, y termina la demostración del teorema.

$\square$

Este teorema que acabamos de demostrar es el primer teorema de cambio de variable. Si pensamos en la aplicación lineal L como una función de cambio de variable, L es una función de clase C^1 en todo $\mathbb{R}^n$, tal que para cada $x \in \mathbb{R}^n$ la diferencial de L en x es la propia función L ($dL(x) = L$). El teorema asegura que L transforma conjuntos medibles-Jordan en conjuntos medibles-Jordan, y, expresando los volúmenes mediante la integral de la función característica, que

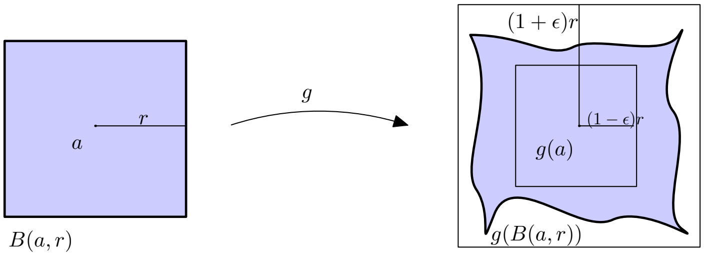
$$\begin{aligned} \int_{L(M)} 1 &= v(L(M)) = |\det L| v(M) = |\det L| \int_M 1 = \\ &= \int_M 1 \cdot |\det L| = \int_M 1 \circ L \cdot |JL| \end{aligned}$$

donde JL es el Jacobiano de L, que es la fórmula de cambio de variable para la función integrable $f \equiv 1$ y la función de cambio $g = L$.

El siguiente paso en la demostración del teorema de cambio de variable general es el caso en que f es la función constantemente uno, y g es una función de clase C^1. Para este resultado utilizaremos un lema de tipo técnico sobre el comportamiento de una función en relación con su diferencial, que suele utilizarse también en las demostraciones de los teoremas de la función inversa y de la función implícita del cálculo diferencial.

Lema 1. *Sean U un abierto de $\mathbb{R}^n$, y $g : U \longrightarrow \mathbb{R}^n$ una función de clase C^1 en U. Sea $a \in U$ tal que $dg(a) = I$, la identidad en $\mathbb{R}^n$, y sea $r > 0$ tal que la bola cerrada de centro a y radio r, $\overline{B}(a, r)$, esté contenida en U. Supongamos que existe ϵ, $0 < \epsilon < 1$, tal que $\|dg(x) - dg(z)\| \leq \epsilon$ para todos $x, z \in \overline{B}(a, r)$. Entonces*

$$\overline{B}(g(a), (1 - \epsilon)r) \subset g(\overline{B}(a, r)) \subset \overline{B}(g(a), (1 + \epsilon)r)$$



Demostración:
Empecemos con el primer contenido, y veamos en primer lugar que basta demostrarlo en el caso $a = g(a) = 0$.

En efecto, si definimos $h(x) = g(x+a)-b$, donde $b = g(a)$, definida en el abierto $V = U-a$, h es una función de clase C^1 en V, con $h(0) = 0$ y $dh(0) = dg(a) = I$. Supongamos que h cumple el teorema; entonces, como $g(y) = h(y-a)+b$ para cada $y \in U$, si

$$z \in \overline{B}(g(a),(1-\epsilon)r) = \overline{B}(b,(1-\epsilon)r) = b + \overline{B}(0,(1-\epsilon)r)$$

se tiene que $z = b+w$ con $w \in \overline{B}(0,(1-\epsilon)r)$. Por hipótesis, $\overline{B}(0,(1-\epsilon)r) \subset h(\overline{B}(0,r))$, y por tanto, existe $x \in \overline{B}(0,r)$ tal que $w = h(x)$; y tomando $y = x+a$ se tiene $y \in \overline{B}(a,r)$ y además

$$z = b + h(y-a) = g(y)$$

Es decir, $\overline{B}(g(a),(1-\epsilon)r) \subset g(\overline{B}(a,r))$

Así pues, supongamos que g es una función de clase C^1 en un abierto U de $\mathbb{R}^n$ que contiene a 0, con $g(0) = 0$ y $dg(0) = I$, y sean $r > 0,\ 0 < \epsilon < 1$ tales que $\overline{B}(0,r) \subset U$, y $\|dg(x)-dg(z)\| \le \epsilon$ para todos $x, z \in \overline{B}(0,r)$. Hay que probar que para todo $y \in \overline{B}(0,(1-\epsilon)r)$ existe $x \in \overline{B}(0,r)$ tal que $y = g(x)$.

Dado $y \in \overline{B}(0,(1-\epsilon)r)$, definimos la función

$$g_y(x) = x - g(x) + y$$

en $\overline{B}(0,r)$. La función g_y transforma $\overline{B}(0,r)$ en sí misma, pues para todo $x \in \overline{B}(0,r)$

$$\begin{aligned}
\|g_y(x)\| &= \|x - g(x) + y\| \leq \|x - g(x)\| + \|y\| = \\
&= \|g(x) - g(0) + dg(0)(x)\| + \|y\| \leq \\
&\leq \|x\| \cdot \sup_{z \in \overline{B}(0,r)} \|dg(z) - dg(0)\| + \|y\| \leq \\
&\leq r \cdot \epsilon + (1-\epsilon) \cdot r = r
\end{aligned}$$

aplicando el teorema del valor medio a la función $h(x) = g(x) - dg(0)(x)$.

El conjunto $\overline{B}(0,r)$ es un espacio métrico completo, al ser un subconjunto cerrado de $\mathbb{R}^n$, que es completo. Y además g_y es contractiva:

$$\begin{aligned}
\|g_y(x) - g_y(z)\| &= \|x - z + g(x) - g(z)\| = \\
&= \|dg(0)(x-z) - g(x) - g(z)\| \leq \\
&\leq \|x - z\| \cdot \sup_{u \in \overline{B}(0,r)} \|dg(0) - dg(u)\| \leq \\
&\leq \epsilon \|x - z\|
\end{aligned}$$

aplicando el teorema el valor medio a la función $h(x) = dg(0)(x) - g(x)$. Como $0 < \epsilon < 1$ por hipótesis, efectivamente la función g_y es contractiva, y podemos aplicar el teorema del punto fijo para asegurar que existe un único punto $x \in \overline{B}(0,r)$ tal que

$$x = g_y(x) = x - g(x) + y$$

Para este punto se tiene entonces $g(x) = y$, como queríamos demostrar.

Para el segundo contenido, dado $y \in g(\overline{B}(a,r))$, existe $x \in \overline{B}(a,r)$ tal que $y = g(x)$, y entonces

$$\|g(x) - g(a)\| \leq \sup_{z \in \overline{B}(a,r)} \|dg(z)\| \cdot \|x - a\|$$

aplicando el teorema del valor medio a la función g; por hipótesis, para cada $z \in \overline{B}(a,r)$

$$\|dg(z)\| \leq \|dg(z) - dg(a)\| + \|dg(a)\| \leq 1 + \epsilon$$

y por tanto $y \in \overline{B}(g(a),(1+\epsilon)r)$, lo que termina la demostración del teorema.

$\square$

Proposición 3. *Sea R un rectángulo en $\mathbb{R}^n$, y U un abierto tal que $R \subset U$. Sea g una función de U en $\mathbb{R}^n$, que sea un difeomorfismo de clase C^1 en U. Entonces*

$$v(g(R)) = \int_R |Jg|$$

Demostración:
Si g es una aplicación lineal, el resultado es consecuencia inmediata del teorema 2, como ya observamos después de su demostración.

El caso general se demuestra aproximando g por su diferencial. Supondremos primero que R es un cubo en $\mathbb{R}^n$, es decir, que tiene todos sus lados de la misma longitud. Entonces para cada $N \in I\!N$, se puede dividir R en una partición de N^n cubos de lado N^{-1}, y considerando en $\mathbb{R}^n$ la norma infinito, cada cubo se puede poner a la vez como una bola para la norma. Si llamamos S a estos cubos, se tiene que $v(g(R)) = \sum_S v(g(S))$, así que basta demostrar el resultado para cada cubo S, y se puede suponer que el radio de S es todo lo pequeño que sea necesario.

Como g es un difeomorfismo de clase C^1, la función $x \longrightarrow \|dg(x)^{-1}\|$ es continua en U, y alcanzará en R su supremo por ser R compacto. Existirá entonces una constante $C > 0$ tal que $\|dg(x)^{-1}\| \leq C$ para todo $x \in R$.

Por otro lado, dado $\epsilon,\ 0 < \epsilon < 1$, como la función $x \longrightarrow \|dg(x)\|$ es uniformemente continua en R, podemos tomar N suficientemente grande para que se verifique $\|g(x) - g(z)\| \leq \epsilon/C$ para todo $x, z \in S$.

Sea a el centro de S; se tiene

$$\|dg(a)^{-1} \circ dg(x) - dg(a)^{-1} \circ dg(z)\| \leq \|dg(a)^{-1}\| \cdot \frac{\epsilon}{C} \leq \epsilon$$

Por el lema anterior, $dg(a)^{-1} \circ g(S)$ contiene una bola de radio $(1 - \epsilon)$ por el radio de S, y está contenido en una bola de radio $(1 + \epsilon)$ por el radio de S; como estas bolas son a su vez cubos, las llamaremos Q y Q' respectivamente, de modo que

$$Q \subset dg(a)^{-1} \circ g(S) \subset Q'$$

y aplicando a cada conjunto la función $dg(a)$

$$dg(a)(Q) \subset g(S) \subset dg(a)(Q')$$

Como $dg(a)$ es una aplicación lineal y $dg(a)(Q)$ es un conjunto medible-Jordan, aplicando el lema 1, si s es el radio de S, el volumen de $dg(a)(Q)$ verificará

$$v(dg(a)(Q)) = |Jg(a)|v(Q) = |Jg(a)\ (1-\epsilon)^n s^n =$$
$$= |Jg(a)| \cdot v(S) - \epsilon \cdot C_1 \cdot v(S)$$

para una cierta constante C_1

$$((1-\epsilon)^n = 1 + \sum_{k=1}^{n} \binom{n}{k} (-1)^{n-k} \epsilon^{n-k} =$$

$$= \ 1 - \epsilon(\sum_{k=1}^{n} \binom{n}{k} (-1)^{n-k-1} \epsilon^{n-k-1}) = 1 - \epsilon \cdot C_1$$

Análogamente $dg(a)(Q')$ es un conjunto medible de volumen

$$v(dg(a)(Q')) = |Jg(a)|v(Q') = |Jg(a)| \cdot v(S) + \epsilon \cdot C_2 \cdot v(S)$$

para una cierta constante C_2.

En consecuencia

$$|Jg(a)|v(S) - \epsilon C_1 \le v(g(S)) \le |Jg(a)|v(S) + \epsilon C_2 v(S)$$

Tomando ínfimos y supremos cuando a recorre S,

$$m_S(|Jg|)\cdot v(S) - \epsilon C_1 v(S) \le v(g(S)) \le M_S(|Jg|)\cdot v(S) + \epsilon C_2 v(S)$$

Y sumando en S,

$$\underline{S}(|Jg|, P) - \epsilon C_1 v(R) \le v(g(R)) \le \overline{S}(|Jg|, P) + \epsilon C_2 v(R)$$

Como esto vale para todo $\epsilon,\ 0 < \epsilon < 1$, y la función $|Jg|$ es integrable al ser una función continua, se tiene que cumplir

$$\int_R |Jg| \le v(g(R)) \le \int_R |Jg|$$

lo que prueba el resultado.

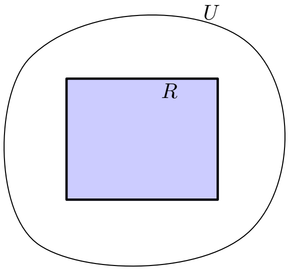
$\square$

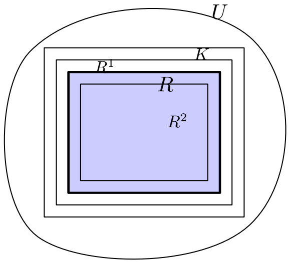
En la demostración de la proposición habíamos supuesto que R era un cubo. En el caso general en que R es un rectángulo, definimos $\delta = d(R, U^c) = \inf\{\|x - y\|,\ x \in R,\ y \in U^c\}$,

![AAVVR]

Cambio de Variable en la integral Riemann.

la distancia de R al complementario de U, que es un número estrictamente positivo al ser R un compacto contenido en el abierto U (la demostración se deja como ejercicio), y consideramos el conjunto $K = R + \overline{B}(0, \delta/2)$, que es un compacto contenido en U y que contiene a R. Dado $\epsilon > 0$ podemos escoger un rectángulo R^1 tal que $R \subset R^1 \subset K$, $v(R^1) \leq v(R) + \epsilon$, y de forma que R^1 se puede descomponer como unión finita de cubos, $R^1 = \bigcup_{i=1}^{k} Q_i$ (como en la observación 1).





Entonces

$$v(g(R)) \leq v(g(R^1)) = \sum_{i=1}^{k} v(g(Q_i)) = \sum_{i=1}^{n} \int_{Q_i} |Jg| = \int_{R^1} |Jg| =$$

$$= \int_R |Jg| + \int_{R^1 \setminus R} |Jg| \leq$$

$$\leq \int_R |Jg| + C \cdot v(R^1 \setminus R) \leq \int_R |Jg| + C \cdot \epsilon$$

siendo C una cota de $|Jg|$ en K, que existe por ser, por hipótesis, $|Jg|$ una función continua en U, y $K \subset U$ compacto.

Y análogamente, podemos escoger un rectángulo R^2 contenido en R, tal que $v(R) \leq v(R^2) + \epsilon$, y de modo que R^2 se pueda descomponer como unión finita de cubos $R^2 = \bigcup_{j=1}^{m} Q_j$ (con un proceso análogo al definido en la observación 1). De este modo

$$\int_R |Jg| \leq \int_{R^2} |Jg| + \int_{R \setminus R^2} |Jg| \leq v(g(R^2)) + C \cdot v(R \setminus R^2) \leq$$

$$\leq v(g(R)) \leq v(R) + C \cdot \epsilon$$

Se deduce entonces que para todo $\epsilon > 0$

$$-C \cdot \epsilon \leq v(R) - \int_R |Jg| \leq C \cdot \epsilon$$

y por tanto que $v(R) = \int_R |Jg|$

$\square$

Como consecuencia se obtiene con relativa facilidad el siguiente resultado, que es la tercera versión del teorema de cambio de variable:

Proposición 4. *Sea R un rectángulo en $\mathbb{R}^n$, y U un abierto tal que $\overline{R} \subset U$. Sea $g : U \longrightarrow \mathbb{R}^n$ un difeomorfismo de clase C^1 de U en $g(U)$. Y sea f una función integrable en $g(R)$. Entonces $(f \circ g)$ es integrable en R y*

$$\int_{g(R)} f = \int_R (f \circ g)|Jg|$$

Demostración:
En primer lugar, veamos que la función $(f \circ g)$ es integrable: si notamos por $D(h)$ el conjunto de puntos de discontinuidad de una función h, en general, tenemos

$$\begin{aligned} D(f \circ g) &= \{x \in U,\ f \circ g \text{ no es continua en } x\} = \\ &= \{x \in U,\ f \text{ no es continua en } g(x)\} = \\ &= g^{-1}(\{y \in g(U), f \text{ no es continua en } y\}) = \\ &= g^{-1}(D(f)) \end{aligned}$$

Como $g^{-1} : g(U) \longrightarrow U$ es una función de clase C^1, por la hipótesis sobre g, transforma conjuntos de medida cero en conjuntos de medida cero, y en particular el conjunto $D(f)$, que tiene medida cero por ser f integrable, en el conjunto de puntos de discontinuidad de la composición $f \circ g$ que tendrá medida cero.

En consecuencia, también es integrable la función $(f \circ g) \cdot |Jg|$ por ser producto de funciones integrables.

En segundo lugar, dada una partición cualquiera P de R, para cada rectángulo S definido por P y cada $y \in g(S)$ se tiene

$$m_S(f \circ g) = m_{g(S)}(f) \leq f(y) \leq M_{g(S)}(f) = M_S(f \circ g)$$

(con la interpretación habitual de $m_A(f) = \inf\{f(t), t \in A\}$, y $M_A(f) = \sup\{f(t), t \in A\}$)

$$\begin{aligned}
\int_S m_S(f \circ g) \cdot |Jg| &= \int_S m_{g(S)}(f) \cdot |Jg| = \\
&= m_{g(S)}(f) \int_S |Jg| = m_{g(S)}(f) v(g(S)) = \int_{g(S)} m_{g(S)}(f) \leq \\
&\leq \int_{g(S)} f \leq \\
&\leq \int_{g(S)} M_{g(S)}(f) = M_{g(S)}(f) v(g(S)) = M_{g(S)}(f) \int_S |Jg| = \\
&= \int_S M_{g(S)}(f) \cdot |Jg| = \int_S M_S(f \circ g) \cdot |Jg|
\end{aligned}$$

Y también se tiene, trivialmente

$$\int_S m_S(f\circ g)\cdot|Jg| \le \int_S (f\circ g)|Jg| \le \int_S M_S(f\circ g)\cdot|Jg|$$

Restando las dos desigualdades, y sumando en S,

$$\begin{aligned}\left|\int_{g(R)} f - \int_R (f\circ g)|Jg|\right| &\le \sum_S \left(M_S(f\circ g) - m_S(f\circ g)\right)\int_S |Jg| \le \\ &\le C\cdot\sum_S \left(M_S(f\circ g) - m_S(f\circ g)\right) v(S) = \\ &= C\cdot\left(\overline{S}((f\circ g),P) - \underline{S}((f\circ g),P)\right)\end{aligned}$$

siendo C una cota de $|Jg|$ en R. Como hemos probado ya que $(f\circ g)$ es integrable, esta diferencia puede hacerse tan pequeña como se quiera, lo que implica que

$$\int_{g(R)} f = \int_R (f\circ g)|Jg|$$

como queríamos demostrar.

$\square$

Por último, vamos a demostrar la versión definitiva del teorema de cambio de variable para la integral de Riemann de funciones de varias variables:

Teorema 3 (Cambio de Variable en $\mathbb{R}^n$).

Sea U un abierto de $\mathbb{R}^n$, y sea $g : U \longrightarrow \mathbb{R}^n$ una función de clase C^1 en U. Sea M un conjunto medible-Jordan tal que $\overline{M} \subset U$, y supongamos que la restricción de g al interior de M, $g|_{M^0}$ es un difeomorfismo de clase C^1. Sea f una función integrable en $g(M)$. Entonces $(f \circ g)$ es integrable en M, y

$$\int_{g(M)} f = \int_M (f \circ g)|Jg|$$

Demostración: ▶ (Saltar al final de la demostración)

Podemos suponer para la demostración que M es cerrado: en efecto, como M es acotado, entonces $\overline{M}$ es compacto y por tanto $g(\overline{M}) = \overline{g(M)}$; se tiene entonces

$$\int_{g(M)} f = \int_{\overline{g(M)}} f = \int_{g(\overline{M})} f = \qquad (1)$$

y si el resultado es cierto para conjuntos medibles cerrados

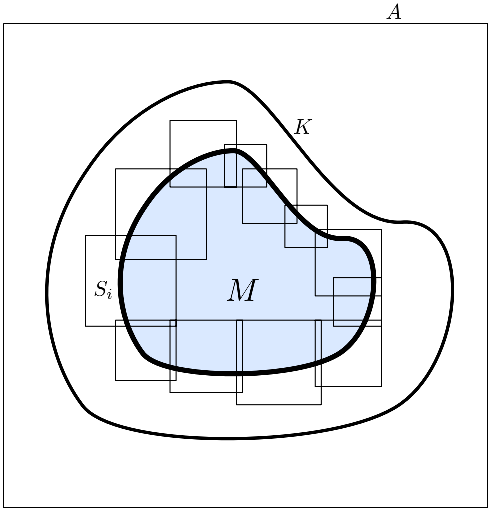
$$(1) \qquad = \int_{\overline{M}} (f \circ g)|Jg| = \int_M (f \circ g)|Jg|$$

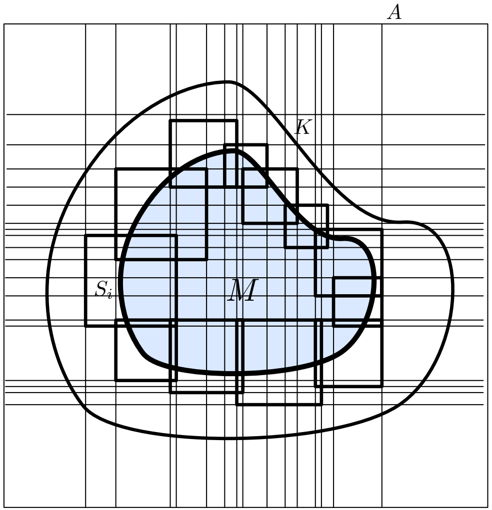
Sea entonces M un conjunto medible-Jordan, cerrado (y por tanto compacto), tal que $M \subset U$. Veamos en primer lugar que la función $(f \circ g)$ es integrable en M, estudiando el conjunto de puntos de discontinuidad.

Sea $\delta = d(M, U^c)$, la distancia de M al complementario de U, que es un número estrictamente positivo al ser M compacto, U^c cerrado y $M \subset U$; consideramos el conjunto $K = M + \overline{B}(0, \delta/2)$, que es un compacto que contiene a M en su interior, y que está contenido en U.

Como por hipótesis $Fr(M)$ tiene contenido cero, dado $0 < \epsilon < (\delta/2)^n$, existe una familia finita de cubos $S_1, \ldots, S_k$ tales que $Fr(M) \subset \bigcup_{i=1}^k S_i^0$, y $\sum_{i=1}^k v(S_i) \leq \epsilon$; podemos suponer además que para todo i, $S_i \cap Fr(M) \neq \emptyset$ (si no eliminaríamos ese rectángulo), con lo que necesariamente $S_i \cap M \neq \emptyset$. En particular el volumen de cada S_i es menor que ϵ, y por tanto, si l_i es la longitud del lado de S_i, $l_i \leq \epsilon^{\frac{1}{n}} \leq \delta/2$, y por tanto cada S_i está contenido en K.

Consideramos A un cubo en $\mathbb{R}^n$ que contenga a K, y definimos una partición P de A prolongando los lados de los cubos S_i, $1 \leq i \leq k$.



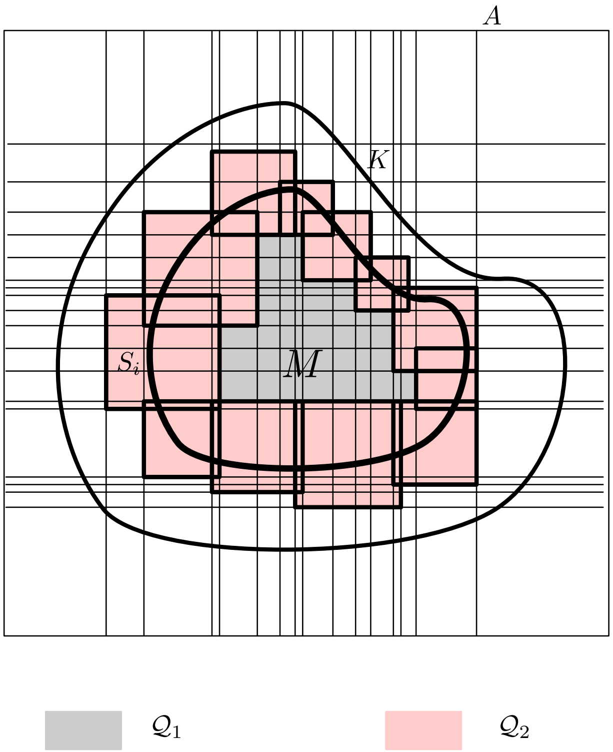


Sea $\mathcal{Q}_1$ la familia de los rectángulos definidos por P que no cortan al interior de ningún S_i, pero están contenidos en el interior de M, y sea $\mathcal{Q}_2$ la familia de los rectángulos definidos por P que están contenidos en algún S_i, $1 \leq i \leq k$.

AAVVR

Cambio de Variable en la integral Riemann.

Teorema de Cambio . . .



Obsérvese que si R es un rectángulo de los definidos por P que corta a M, necesariamente está en algunas de las dos familias:

En efecto, si $R \in \mathcal{P}$, y $R \notin \mathcal{Q}_2$, entonces $R \cap S_i^0 = \emptyset$ para todo $,i,\ 1 \leq i \leq k$, y por tanto $R \cap Fr(M) = \emptyset$. Se tiene entonces que

$$\begin{aligned} R \cap \overline{M} &= R \cap (M^0 \cup Fr(M)) = \\ &= (R \cap M^0) \cup (R \cap Fr(M)) = (R \cap M^0) \end{aligned}$$

de modo que si $R \cap \overline{M}$ es no vacío, sería a la vez abierto y cerrado en R como subespacio métrico de $\mathbb{R}^n$, y como R es conexo, tiene que ser el vacío o el total R; en el primer caso R no corta a M, y en el segundo R estaría contenido en M^0, y por tanto estaría en $\mathcal{Q}_1$.

Llamemos

$$M_1 = \bigcup \{R \in \mathcal{Q}_1\} = \bigcup \{R, R \in \mathcal{P}, R \subset M^0\}$$

$$M_2 = M \setminus M_1 \subset \bigcup \{R \in \mathcal{Q}_2\} = \bigcup_{i=1}^{k} S_i$$

y sean $D(f \circ g)$ el conjunto de puntos de discontinuidad de $f \circ g$, y $D(f)$ el conjunto de puntos de discontinuidad de f. Se tiene

$$D(f \circ g) = (D(f \circ g) \cap M_1) \cup (D(f \circ g) \cap M_2)$$

$$\begin{aligned} D(f\circ g)\cap M_1 &= \{x\in M_1,\ f\circ g \text{ no es continua en } x\} = \\ &= \{x\in M_1,\ f \text{ no es continua en } g(x)\} = \\ &= (g|_{M_1})^{-1}(\{y\in g(M_1),\ f \text{ no es continua en } y\}) = \\ &= (g|_{M_1})^{-1}(D(f)\cap g(M_1)) \end{aligned}$$

$D(f)$ tiene medida cero, ya que por hipótesis f es integrable en $g(M)$, y $(g|_{M_1})^{-1}$ es una función de clase C^1 en $g(M^0)$, que es un abierto que contiene a $g(M_1)$, por la hipótesis de que g es un difeomorfismo de clase C^1 en M^0. Entonces $(g|_{M_1})^{-1}(D(f)\cap g(M_1))$ tiene medida cero, y existirá una familia numerable de rectángulos $\{R_i\}$ tal que

$$(g|_{M_1})^{-1}(D(f)\cap g(M_1)) \subset \bigcup_{i=1}^{\infty} R_i$$

y

$$\sum_{i=1}^{\infty} v(R_i) \leq \epsilon$$

Por otro lado,

$$D(f\circ g)\cap M_2 = \{x\in M\setminus M_1,\ f \text{ no es continua en } g(x)\}$$

es un subconjunto de M_2, que está contenido en $\bigcup_{i=1}^{k} S_i$, unión finita de cubos, cuya suma de volúmenes es menor que ϵ.

En consecuencia, $D(f \circ g)$ está contenido en una unión numerable de rectángulos, con suma de volúmenes menor o igual que 2ϵ.

Como esto se puede hacer para cualquier $\epsilon > 0$, se tiene que $D(f \circ g)$ tiene medida cero, y que $f \circ g$ es integrable.

Como $|Jg|$ es una función continua en U, en particular es continua y acotada en el compacto M, y por tanto es integrable en M; luego la función $(f \circ g) \cdot |Jg|$ es también integrable en M. Para demostrar el teorema queda por probar la igualdad de las integrales.

Sea $\epsilon > 0$, y consideremos A, P, $\mathcal{Q}_1$ y $\mathcal{Q}_2$, M_1 y M_2 como antes.

Como M_1, M_2, $g(M_1)$ y $g(M_2)$ son medibles-Jordan, $M = M_1 \cup M_2$, $g(M) = g(M_1) \cup g(M_2)$, y $g(M_1) = \bigcup_{R \in \mathcal{Q}_1} g(R)$, se tiene

$$\begin{aligned}
\int_{g(M)} f &= \int_{g(M_1)} f + \int_{g(M)\backslash g(M_1)} f = \\
&= \sum_{R\in\mathcal{Q}_1} \int_{g(R)} f + \int_{g(M)\backslash g(M_1)} f = \\
&= \sum_{R\in\mathcal{Q}_1} \int_{R} (f\circ g)|Jg| + \int_{g(M)\backslash g(M_1)} f = \\
&= \int_{M_1} (f\circ g)|Jg| + \int_{g(M)\backslash g(M_1)} f
\end{aligned}$$

de donde

$$\int_{g(M)} f - \int_{M_1} (f\circ g)|Jg| = \int_{g(M)\backslash g(M_1)} f$$

y

$$\left|\int_{g(M)} f - \int_{M_1} (f\circ g)|Jg|\right| = \left|\int_{g(M)\setminus g(M_1)} f\right| \leq$$

$$\leq \int_{g(M)\setminus g(M_1)} |f| \leq \int_{g(M_2)} |f| \leq \qquad (*)$$

$$\leq C_1 \cdot v(g(M_2)) = C_1 \cdot v(g(\bigcup\{S_i \cap M,\ 1 \leq i \leq k\})) =$$

$$= C_1 \cdot \sum_{i=1}^{k} v(g(S_i \cap M)) \leq C_1 \cdot \sum_{i=1}^{k} v(g(S_i))$$

teniendo en cuenta que $g(M)\setminus g(M_1) \subset g(M_2)$ y $|f|$ es una función positiva, y utilizando después la observación 4; la constante C_1 es una cota de $|f|$ en el compacto K que contiene a M.

Aplicando el teorema del valor medio a la función g en cada rectángulo S_i, si a es el centro de R, y C_2 es una cota de $\|dg\|$ en K, se tiene que para cada $x \in S_i$ existe $z \in S_i$ tal que

$$\|g(x) - g(a)\| \leq \|dg(z)\| \cdot \|x - a\| \leq C_2 \cdot \|x - a\|$$

es decir, $g(S_i)$ está contenido en un cubo de centro $g(a)$ y lado C_2 veces el lado de S_i; de este modo, $v(g(S_i)) \leq C_2^n \cdot v(S_i)$, para cada i, $1 \leq i \leq k$. Por tanto

$$\int_{g(M_2)} |f| \leq C \sum_{i=1}^{k} v(S_i) \leq C\epsilon \qquad (**)$$

para una cierta constante $C = C_1 \cdot C_2^n$.

Por otro lado,

$$\left|\int_M (f\circ g)|Jg| - \int_{M_1}(f\circ g)|Jg|\right| = \left|\int_{M\setminus M_1}(f\circ g)|Jg|\right| =$$

$$= \left|\int_{M_2}(f\circ g)|Jg|\right| \le \int_{M_2} |f\circ g|\cdot|Jg| \le$$

$$\le C_1\cdot C_3 v(M_2) \le C'\cdot \epsilon \qquad \text{(***)}$$

donde C_3 es una cota de $|Jg|$ en K, y $C' = C_1 \cdot C_3$.

Como consecuencia de las desigualdades (*), (**), y (***),

$$-\epsilon\cdot C \le \int_{g(M)} f - \int_{M_1}(f\circ g)|Jg| \le \epsilon\cdot C$$

y

$$-\epsilon\cdot C' \le \int_{M_1}(f\circ g)|Jg| - \int_M (f\circ g)|Jg| \le \epsilon\cdot C'$$

y sumando las dos desigualdades

$$-\epsilon(C+C') \le \int_{g(M)} f - \int_M (f\circ g)|Jg| \le \epsilon(C+C')$$

Como esto es cierto para cualquier $\epsilon > 0$, se tiene la igualdad de las integrales y el fin de la demostración del teorema.

◀(Volver al enunciado) $\square$