UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO

FACULTAD DE CIENCIAS

Prueba de algunos teoremas de Cálculo de varias variables

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

M A T E M Á T I C O

P R E S E N T A:

ALEJANDRO DARÍO ROJAS SÁNCHEZ

DIRECTOR DE TESIS:
DR. JAVIER PÁEZ CÁRDENAS
2010

Datos del Alumno
Rojas
Sánchez
Alejandro Darío
5585447387
Universidad Nacional Autónoma de México
Facultad de Ciencias
Matemáticas
405017377

Datos del tutor
Dr
Javier
Páez
Cárdenas

Datos del sinodal 1
Dr
Héctor
Méndez
Lango

Datos del sinodal 2
Dr
Jefferson Edwin
King
Dávalos

Datos del sinodal 3
Dr
Francisco
Marmolejo
Rivas

Datos del sinodal 4
M en C
Francisco Manuel
Barrios Paniagua

Datos del trabajo escrito
Prueba de algunos de teoremas de Cálculo de
varias variables
98 p
2010

Índice General

Introducción

El conocimento y dominio del cálculo diferencial e integral por el estudiante de ciencias es, sin lugar a dudas, fundamental para la formación del mismo, pues a través del rigor, el necesario desarrollo de la intuición y la dedicación que el estudio del cálculo exige, se dan las bases para una buena preparación hacia el infinito mundo de las matemáticas. Citando a Michael Spivak, el Cálculo no sólamente es el preludio de las matemáticas, sino el primer encuentro real con las mismas. Aunque en algunos textos se nos presenta al cálculo como una colección de materias por su inevitable contacto con la geometría analítica y el álgebra, es posible ver al cálculo como una bella evolución de una idea: la de función.

El objetivo de este trabajo, es el de presentar la prueba rigurosa y formal de algunos teoremas de cálculo de varias variables, tanto diferencial como integral, los cuales, ya sea por su complejidad, y en algunos casos por su extensión, resulta poco práctico realizar dentro del curso de cálculo correspondiente. Las ideas para la demostración de algunos de estos teoremas están tomadas del libro Advanced Calculus de Hans Sagan. Concretamente se generaliza y corrige la prueba del teorema de Lebesgue para funciones Riemann integrables realizada por dicho texto, llevando el teorema del caso en $\mathbb{R}$ al caso en $\mathbb{R}^n$. También los lineamientos para la prueba del teorema de cambio de variable son tomados de este libro, siendo en esta parte donde más correcciones se realizaron. Las correcciones a estas pruebas, fueron desarrolladas por el autor del presente trabajo.

Al ser éste un trabajo que pretende ser un complemento a los cursos de cálculo de varias variables, se espera que el lector tenga conocimientos de los temas que se suelen ver en estos cursos; se espera que se domine el concepto de función continua y función diferenciable así como los conceptos de ínfimo y supremo de un conjunto acotado; también es indispensable un manejo básico de la topología euclideana de $\mathbb{R}^n$ ya que conceptos como los de conjuntos abiertos, cerrados, conexos y compactos son utilizados con mucha frecuencia. Asimismo, se requiere que el lector tenga conocimientos de lo que usualmente se estudia en un primer curso de álgebra lineal, específiamente, que conozca las nociones de espacio vectorial y transformaciones lineales.

Es importante destacar que las pruebas aquí presentadas tratan, en la medida de lo posible, de mantenerse dentro del campo del cálculo o, en otras palabras, tratan de no escaparse del nivel de conocimientos que tiene un estudiante de cálculo en varias variables.

Esta tesis consta de cuatro capítulos y cada uno está dividido en secciones. Conforme se avance en la lectura, posiblemente sea necesario regresar de un capítulo a otro, debido a que en cada capítulo subsecuente se va haciendo referencia a alguna herramienta desarrollada en uno previo.

En el capítulo uno, se demuestran el teorema de la función inversa así como el teorema de la función implícita; para ello se trabaja con transformaciones lineales y se da una caracterización de las funciones de clase C^1.

En el capítulo dos, se demuestra el teorema de Lebesgue, que da condiciones necesarias y suficientes para que una función acotada sea Riemann integrable, previamente se repasa el concepto de integral de Riemann y todo lo relacionado con ella. Se habla un poco de rectángulos y se analizan algunas de sus propiedades con la intención de estudiar el concepto de conjunto con medida de Lebesgue cero.

En el capítulo tres, no se realiza la prueba de algún teorema importante en específico. Aquí se da una rápida revisión de la medida de Jordan y la integración de funciones sobre conjuntos Jordan medibles. El objetivo de este capítulo es brindar algunas herramientas necesarias en el capítulo siguiente, pero también el de analizar la relación que existe entre los conjuntos con medida de Lebesgue cero y los conjuntos con medida de Jordan cero; más aún, en la mayoría de los resultados demostrados en este capítulo, se procura utilizar el teorema de Lebesgue tratando de exhibir lo mucho que simplifica el trabajo el uso del concepto de medida de Lebesgue cero.

Finalmente, en el capítulo cuatro se da una prueba, con mucho cuidado, del teorema de cambio de variable. Su prueba es la más extensa de todos los teoremas expuestos, y de alguna manera unifica todo el trabajo previo, pues en su demostración están aplicadas todas las herramientas desarrolladas a lo largo de la tesis.

Como es natural en los textos de matemáticas, cada lema, proposición, teorema, y demás resultados, están numerados para lograr una rápida referencia cuando sea necesario, y únicamente en el último capítulo se ha tenido la necesidad de numerar también algunas ecuaciones en aquellos resultados cuya prueba resulta bastante técnica, para así facilitar su lectura.

Capítulo 1

Teorema de la función inversa

En este capítulo se dará una prueba con mucho detalle del teorema de la función inversa en varias variables. Para ello iniciaremos con algunos resultados de álgebra lineal, mismos que nos serán de gran utilidad a lo largo de este trabajo. Específicamente hablaremos de transformaciones lineales en espacios vectoriales normados, y en particular daremos una norma al espacio vectorial de las transformaciones lineales de $\mathbb{R}^n$ en $\mathbb{R}^m$.

1.1 El espacio $\pounds(n,m)$

En esta sección, nuestro objetivo principal es proveer de una norma al espacio de las transformaciones lineales de $\mathbb{R}^n$ en $\mathbb{R}^m$, que denotaremos por $\pounds(n,m)$; antes demostraremos algunas propiedades de las tranformaciones lineales.

Comencemos con un poco de notación para las normas. Recordemos que la norma euclideana en $\mathbb{R}^n$ está definida como $\|\hat{x}\| = \sqrt{x_1^2 + \cdots + x_n^2}$ si $\hat{x} = (x_1, \ldots, x_n)$ y ésta será la notación, es decir, la doble barra. Cuando estemos trabajando con una norma en algún espacio vectorial, no necesariamente $\mathbb{R}^n$, la notación que se usará será tambien la doble barra pero agregaremos un subíndice particular a la doble barra; por ejemplo la norma caja $\|\hat{x}\|_\infty = \max\{|x_1|, ..., |x_n|\}$ en donde $|x_i|$ denota simplemente al valor absoluto para números reales. De momento con estas notaciones será suficiente para poder iniciar nuestra labor.

Ya que hemos tocado a la norma euclideana, mencionemos (sin demostrar) algunas propiedades que cumple esta norma:

$$
\begin{array}{lll}
i) & |x_j| \leq \|\hat{x}\| & \text{para toda } j = 1, ..., n \\
ii) & \|\hat{x}\| \leq |x_1| + \cdots + |x_n| & \\
iii) & \|\hat{x}\| \leq \sqrt{n}\max\{|x_1|, ..., |x_n|\} & \\
iv) & |\hat{x} \cdot \hat{y}| \leq \|\hat{x}\| \, \|\hat{y}\| &
\end{array}
$$

esta última propiedad, como todos sabemos, es conocida como la desigualdad de Cauchy-Schwarz.

Una vez establecido esto demos paso a nuestra primera proposición que nos obsequiará una útil relación entre el valor de un vector en $\mathbb{R}^n$ bajo una transformación lineal y el vector mismo concluyendo de paso la continuidad uniforme de cualquier transformación lineal $\mathbb{R}^n \xrightarrow{L} \mathbb{R}^m$.

Proposición 1.1 *Si $L : \mathbb{R}^n \to \mathbb{R}^m$ es una transformación lineal entonces existe k positivo tal que $\|L(\hat{x})\| \leq k\|\hat{x}\|$ para toda $\hat{x} \in \mathbb{R}^n$.*

Dem. Consideremos la base canónica para $\mathbb{R}^n$, $\hat{e}_j$ para $j = 1, ..., n$, es decir, si $\hat{x} = (x_1, \ldots, x_n)$, entonces $\hat{x} = x_1\hat{e}_1 + \cdots + x_n\hat{e}_n$. De la linealidad de L se sigue que

$$\begin{aligned}\|L(\hat{x})\| &= \|x_1L(\hat{e}_1) + \cdots + x_nL(\hat{e}_n)\| \leq |x_1|\,\|L(\hat{e}_1)\| + \cdots + |x_n|\,\|L(\hat{e}_n)\| \\ &\leq (\|L(\hat{e}_1)\| + \cdots + \|L(\hat{e}_n)\|)\,\|\hat{x}\|\end{aligned}$$

ya que $|x_j| \leq \|\hat{x}\|$; salvo que L sea la transformación cero (en cuyo caso podemos tomar cualquier $k \in \mathbb{R}$, $k > 0$), $k = (\|L(\hat{e}_1)\| + \cdots + \|L(\hat{e}_n)\|) > 0$ satisface la propiedad deseada. ■

Corolario 1.1 *Si $L : \mathbb{R}^n \to \mathbb{R}^m$ es una transformación lineal entonces L es uniformemente continua en $\mathbb{R}^n$.*

Dem. Por la proposición anterior existe $k > 0$ tal que $\|L(\hat{x})\| \leq k\|\hat{x}\|$ para toda $\hat{x} \in \mathbb{R}^n$ de manera que

$$\|L(\hat{x}) - L(\hat{y})\| = \|L(\hat{x} - \hat{y})\| \leq k\|\hat{x} - \hat{y}\|$$

y como este último término puede hacerse tan pequeño como se quiera, concluimos la continuidad uniforme de L. ■

Para el caso en el que la transformación lineal $L : \mathbb{R}^n \to \mathbb{R}^n$ es inyectiva, se tiene una relación inversa a la de la proposición 1.1

Lema 1.1 *Sea $L : \mathbb{R}^n \to \mathbb{R}^n$ una transformación lineal. L es uno a uno si y sólo si existe un número real positivo m tal que para toda $\hat{x}$ en $\mathbb{R}^n$ $\|L(\hat{x})\| \geq m\|\hat{x}\|$.*

Dem. Supongamos primero que L no es uno a uno, es decir, existe $\hat{x}_0 \neq \hat{0}$ tal que $L(\hat{x}_0) = \hat{0}$, pero entonces para cualquier $m > 0$ se tiene que $0 = \|L(\hat{x}_0)\| < m\|\hat{x}_0\|$.

Supongamos ahora que $L : \mathbb{R}^n \to \mathbb{R}^n$ es uno a uno, entonces la función inversa de L es una transformación lineal, $L^{-1} : \mathbb{R}^n \to \mathbb{R}^n$, de manera que por la proposición 1.1 sabemos que existe $k > 0$ tal que $\|L^{-1}(\hat{y})\| \leq k\|\hat{y}\|$ para toda $\hat{y} \in \mathbb{R}^m$; tomemos $m = \frac{1}{k} > 0$ y entonces para toda $\hat{x} \in \mathbb{R}^n$

$$m\|\hat{x}\| = m\left\|L^{-1}(L(\hat{x}))\right\| \leq mk\|L(\hat{x})\| = \|L(\hat{x})\|$$

■

El lema anterior también es válido cuando $L : \mathbb{R}^n \to \mathbb{R}^m$, con $n \leq m$; y aunque pudimos plantear el lema anterior de esta manera, no lo hicimos porque no nos hace falta para los objetivos que se tienen planteados. Ahora bien, es necesario que $n \leq m$ para poder hablar de inyectividad; pues como ya sabemos, si la dimensión del domino de una trasformación lineal es mayor que la dimensión del codominio entonces la transformación lineal no puede ser inyectiva.

Definición 1.1 *Sea*

$$\pounds(n,m) = \left\{ L : \mathbb{R}^n \to \mathbb{R}^m \mid \ L \quad es \quad lineal \ \right\}$$

es decir, $\pounds(n,m)$ es el conjunto de las transformaciones lineales de $\mathbb{R}^n$ en $\mathbb{R}^m$, que junto con la suma y la multiplicación por escalares usuales en las funciones, podemos decir que $\pounds(n,m)$ es un espacio vectorial.

Para ahorrarnos un poco de notación, cuando se trate de transformaciones lineales de $\mathbb{R}^n$ en $\mathbb{R}^n$ denotaremos simplemente $\pounds = \pounds(n,n)$.

Gracias a la proposición 1.1 sabemos que si $L : \mathbb{R}^n \to \mathbb{R}^m$ es una transformación lineal entonces el conjunto

$$E_L = \left\{ k > 0 \mid \ \|L(\hat{x})\| \leq k \|\hat{x}\| \quad \forall \hat{x} \in \mathbb{R}^n \ \right\}$$

es no vacío, y claramente acotado inferiormente por cero, de modo que existe el ínfimo de estas constantes, con lo que nuestra siguiente definición queda bien justificada.

Definición 1.2 *Sea $\|\cdot\|_{\pounds(n,m)} : \pounds(n,m) \to \mathbb{R}$ definida como*

$$\|L\|_{\pounds(n,m)} = \inf E_L$$

Una vez más, para simplificar las cosas denotaremos $\|L\|_{\pounds}$ cuando estemos trabajando con transformaciones de $\mathbb{R}^n$ en $\mathbb{R}^n$.

El número definido arriba es muy importante, pues dota de una norma al espacio $\pounds(n,m)$; además tiene sentido que siendo el ínfimo de las constantes $k > 0$ para las cuales se cumple que $\|L(\hat{x})\| \leq k\|\hat{x}\|$ para toda $\hat{x} \in \mathbb{R}^n$ entonces $\|L\|_{\pounds(n,m)}$ sea una constante que conserva esta propiedad, como lo demostraremos a continuación.

Proposición 1.2 *Si $L \in \pounds(n,m)$ entonces para toda $\hat{x} \in \mathbb{R}^n$*

$$\|L(\hat{x})\| \leq \|L\|_{\pounds(n,m)} \|\hat{x}\|$$

Dem. Sea $L \in \pounds(n,m)$, supongamos que existe $\hat{x}_0 \in \mathbb{R}^n$ tal que

$$\|L\|_{\pounds(n,m)} \|\hat{x}_0\| < \|L(\hat{x}_0)\|$$

entonces $\hat{x}_0$ no puede ser el vector cero, de modo que

$$\|L\|_{\pounds(n,m)} < \frac{1}{\|\hat{x}_0\|} \|L(\hat{x}_0)\|$$

De la definición de $\|L\|_{\pounds(n,m)}$ se sigue que existe $k \in E_L$ tal que

$$k < \frac{1}{\|\hat{x}_0\|} \|L(\hat{x}_0)\|$$

es decir, por una parte se debe tener que $k\|\hat{x}_0\| < \|L(\hat{x}_0)\|$ y por otra que $\|L(\hat{x})\| \leq k\|\hat{x}\|$ para toda $\hat{x} \in \mathbb{R}^n$ lo cual no puede ser. Por lo tanto

$$\|L(\hat{x})\| \leq \|L\|_{\pounds(n,m)} \|\hat{x}\|$$

para toda $\hat{x} \in \mathbb{R}^n$. ■

Una consecuencia importante de este resultado es el hecho de que el número $\|L\|_{\pounds(n,m)}$ es, en efecto, una norma para el espacio $\pounds(n,m)$.

Teorema 1.1 $\|\cdot\|_{\pounds(n,m)} : \pounds(n,m) \to \mathbb{R}$ *es una norma para* $\pounds(n,m)$.

Dem. Para que $\|\cdot\|_{\pounds(n,m)} : \pounds(n,m) \to \mathbb{R}$ sea una norma para $\pounds(n,m)$, debe cumplir tres condiciones:

1) $\forall L \in \pounds(n,m)$, $\|L\|_{\pounds(n,m)} \geq 0$ y $\|L\|_{\pounds(n,m)} = 0$ si y sólo si $L(\hat{x}) = \hat{0}$ $\forall \hat{x} \in \mathbb{R}^n$.

2) $\forall \lambda \in \mathbb{R}$

$$\|\lambda L\|_{\pounds(n,m)} = |\lambda| \|L\|_{\pounds(n,m)}$$

3)$\forall L_1, L_2 \in \pounds(n,m)$

$$\|L_1 + L_2\|_{\pounds(n,m)} \leq \|L_1\|_{\pounds(n,m)} + \|L_2\|_{\pounds(n,m)}$$

Por definición queda claro que para toda $L \in \pounds(n,m)$, $\|L\|_{\pounds(n,m)} \geq 0$ así como también es claro que si $L(\hat{x}) = \hat{0}$ para toda $\hat{x} \in \mathbb{R}^n$ entonces $\|L\|_{\pounds(n,m)} = 0$; supongamos ahora que $L \in \pounds(n,m)$ no es la transformación cero, es decir, existe $\hat{x}_0 \in \mathbb{R}^n$ tal que $L(\hat{x}_0) \neq \hat{0}$, y supongamos que $\|L\|_{\pounds(n,m)} = 0$. Entonces para cualquier $\varepsilon > 0$, existe $k \in E_L$ con $k < \frac{\varepsilon}{\|\hat{x}_0\|}$, de donde $\|L(\hat{x}_0)\| \leq k\|\hat{x}_0\| < \varepsilon$, pero eso implica que $\|L(\hat{x}_0)\| = 0$ y por tanto $L(\hat{x}_0) = \hat{0}$ contradiciendo la propiedad de $\hat{x}_0$, de modo que $\|L\|_{\pounds(n,m)} > 0$ finalizando así la prueba de 1).

Para probar 2) sea $\lambda \in \mathbb{R}$, si $\lambda = 0$ la igualdad es trivial, así que supongamos $\lambda \neq 0$; si $k \in E_{\lambda L}$ entonces $\|\lambda L(\hat{x})\| \leq k\|\hat{x}\|$ para toda $\hat{x} \in \mathbb{R}^n$ de donde

$$\|L(\hat{x})\| \leq \frac{k}{|\lambda|}\|\hat{x}\|$$

para toda $\hat{x} \in \mathbb{R}^n$, es decir, que $\frac{k}{|\lambda|} \in E_L$ y por lo tanto

$$\|L\|_{\pounds(n,m)} \leq \frac{k}{|\lambda|}$$

o lo que es lo mismo $|\lambda| \|L\|_{\pounds(n,m)} \leq k$, y como esto ocurre para toda $k \in E_{\lambda L}$ se tiene que

$$|\lambda| \|L\|_{\pounds(n,m)} \leq \|\lambda L\|_{\pounds(n,m)}$$

Recíprocamente si $k \in E_L$ entonces para toda $\hat{x} \in \mathbb{R}^n$

$$|\lambda| \|L(\hat{x})\| \leq |\lambda| k \|\hat{x}\|$$

es decir,

$$\|\lambda L(\hat{x})\| \leq |\lambda| k \|\hat{x}\|$$

de manera que $|\lambda| k \in E_{\lambda L}$ y por lo tanto $\|\lambda L\|_{\pounds(n,m)} \leq |\lambda| k$ o equivalentemente

$$\frac{1}{|\lambda|}\|\lambda L\|_{\pounds(n,m)} \leq k$$

y esto $\forall k \in E_L$, de forma que

$$\frac{1}{|\lambda|}\|\lambda L\|_{\pounds(n,m)} \leq \|L\|_{\pounds(n,m)}$$

ó $\|\lambda L\|_{\pounds(n,m)} \leq |\lambda| \|L\|_{\pounds(n,m)}$ y por lo tanto $\|\lambda L\|_{\pounds(n,m)} = |\lambda| \|L\|_{\pounds(n,m)}$.

Finalmente para probar 3) sean $L_1, L_2 \in \pounds(n,m)$; dado $\hat{x} \in \mathbb{R}^n$, por la proposición 1.1

$$\|(L_1 + L_2)(\hat{x})\| \leq \|L_1(\hat{x})\| + \|L_2(\hat{x})\| \leq \left(\|L_1\|_{\pounds(n,m)} + \|L_2\|_{\pounds(n,m)}\right)\|\hat{x}\|$$

es decir,

$$\|L_1\|_{\pounds(n,m)} + \|L_2\|_{\pounds(n,m)} \in E_{L_1+L_2}$$

y por lo tanto

$$\|L_1 + L_2\|_{\pounds(n,m)} \leq \|L_1\|_{\pounds(n,m)} + \|L_2\|_{\pounds(n,m)}$$

■

Ya que tenemos que $\|\cdot\|_{\pounds(n,m)} : \pounds(n,m) \to \mathbb{R}$ es una norma para el espacio $\pounds(n,m)$, vale la pena recordar una propiedad útil que se desprende de la desigualdad del triángulo.

Corolario 1.2 *Para cualesquiera $L_1, L_2 \in \pounds(n,m)$,*

$$\left|\|L_1\|_{\pounds(n,m)} - \|L_2\|_{\pounds(n,m)}\right| \leq \|L_1 - L_2\|_{\pounds(n,m)}$$

Su prueba como se mencionó, es inmediata de la desigualdad del triángulo.

1.2 El Teorema de la función inversa

En esta sección se desarrollarán con cuidado, las herramientas suficientes para dar una prueba del teorema de la función inversa. Una vez que hayamos demostrado el teorema de la función inversa, podremos concluir, como un corolario, el teorema de la función implícita.

A fin de evitar confusiones futuras, pongámonos un poco de acuerdo con la notación que usaremos para las derivadas; si $g : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$, con Ω un conjunto abierto, es diferenciable en $\hat{x}_0$, denotaremos a la diferencial de g en $\hat{x}_0$ como $Dg(\hat{x}_0)$. Es decir, $Dg(\hat{x}_0)$ representa a la transformación lineal que más aproxima a g cerca $\hat{x}_0$, y denotaremos por $g'(\hat{x}_0)$ a la derivada de g en $\hat{x}_0$ o también conocida como la matriz Jacobiana de g en $\hat{x}_0$, que no es otra cosa más que la matriz asociada a la transformación $Dg(\hat{x}_0)$ (respecto a la base canónica). En otras palabras lo que estamos diciendo es que, si $g : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ es diferenciable en $\hat{x}_0$ entonces para toda $\hat{y} \in \mathbb{R}^n$ $Dg(\hat{x}_0)(\hat{y}) = g'(\hat{x}_0)\hat{y}$. Así mismo para el determinante Jacobiano de una función g (o simplemente Jacobiano de g) usaremos la notación J_g, es decir, $J_g(\hat{x}) = \det g'(\hat{x})$.

Ya que se ha demostrado que $\|\cdot\|_{\pounds(n,m)}$ es una norma para $\pounds(n,m)$, este espacio se convierte (al igual que $\mathbb{R}^n$) en un espacio métrico y tiene una topología definida similarmente a la que se define en $\mathbb{R}^n$ con la norma euclideana, y nos permite pensar en conceptos como continuidad de funciones definidas en $\pounds(n,m)$. No profundizaremos mucho en este hecho pero enunciaremos un resultado que nos permitirá mostrar que el concepto de función continuamente diferenciable que se define en términos de la continuidad de las correspondientes derivadas parciales, es equivalente a la continuidad de una función que tiene el mismo dominio y como contradominio el espacio vectorial $\pounds(n,m)$. De manera más precisa, una función $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ es de clase C^1 en Ω si y sólo si la función $D_f : \Omega \subset \mathbb{R}^n \to \pounds(n,m)$, definida como $D_f(x) = Df(x)$, es continua con la norma $\|\cdot\|_{\pounds(n,m)}$.

Proposición 1.3 *Sea $f = (f_1, \ldots, f_m) : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ diferenciable en Ω; f es continuamente diferenciable en $\hat{x}_0 \in \Omega$ si y sólo si para todo $\varepsilon > 0$, existe $\delta > 0$ tal que para cualquier $\hat{x} \in \Omega$, si $\|\hat{x} - \hat{x}_0\| < \delta$ entonces $\|D_f(\hat{x}) - D_f(\hat{x}_0)\|_{\pounds(n,m)} < \varepsilon$.*

Dem. Supongamos que f es continuamente diferenciable en $\hat{x}_0$, es decir, las derivadas parciales de las funciones coordenadas f_i son continuas en $\hat{x}_0$.

Sea $\varepsilon > 0$, entonces para cada $\frac{\partial f_i}{\partial x_j}$ existe $\delta_{ij} > 0$ tal que para toda $x \in \Omega$ si $\|\hat{x} - \hat{x}_0\| < \delta_{ij}$, entonces

$$\left| \frac{\partial f_i}{\partial x_j}(\hat{x}) - \frac{\partial f_i}{\partial x_j}(\hat{x}_0) \right| < \frac{\varepsilon}{2\sqrt{nm}}$$

Sea $\delta = \min\{\delta_{ij} \mid \ i \in \{1, \ldots, m\}, j \in \{1, \ldots, n\}\}$ y tomemos $\hat{x} \in \Omega$ tal que $\|\hat{x} - \hat{x}_0\| < \delta$.

Si $\hat{y} \in \mathbb{R}^n$, $\hat{y} = (y_1, \ldots, y_n)$, se tiene que

$$D_f(\hat{x})(\hat{y}) - D_f(\hat{x}_0)(\hat{y}) = \left(\sum_{j=1}^{n} \left(\frac{\partial f_1}{\partial x_j}(\hat{x}) - \frac{\partial f_1}{\partial x_j}(\hat{x}_0) \right) y_j, \ldots, \sum_{j=1}^{n} \left(\frac{\partial f_m}{\partial x_j}(\hat{x}) - \frac{\partial f_m}{\partial x_j}(\hat{x}_0) \right) y_j \right)$$

Para estimar $\|D_f(\hat{x})(\hat{y}) - D_f(\hat{x}_0)(\hat{y})\|$, nos interesa acotar, para cada $i \in \{1, \ldots, m\}$, el valor de $\left| \sum_{j=1}^{n} \left(\frac{\partial f_i}{\partial x_j}(\hat{x}) - \frac{\partial f_i}{\partial x_j}(\hat{x}_0) \right) y_j \right|$. Para esto nótese que

$$\sum_{j=1}^{n} \left(\frac{\partial f_i}{\partial x_j}(\hat{x}) - \frac{\partial f_i}{\partial x_j}(\hat{x}_0) \right) y_j = \left(\frac{\partial f_i}{\partial x_1}(\hat{x}) - \frac{\partial f_i}{\partial x_1}(\hat{x}_0), \ldots, \frac{\partial f_i}{\partial x_n}(\hat{x}) - \frac{\partial f_i}{\partial x_n}(\hat{x}_0) \right) \cdot \hat{y}$$

y entonces

$$\begin{aligned} \left| \sum_{j=1}^{n} \left(\frac{\partial f_i}{\partial x_j}(\hat{x}) - \frac{\partial f_i}{\partial x_j}(\hat{x}_0) \right) y_j \right| &\leq \left\| \left(\frac{\partial f_i}{\partial x_1}(\hat{x}) - \frac{\partial f_i}{\partial x_1}(\hat{x}_0), \ldots, \frac{\partial f_i}{\partial x_n}(\hat{x}) - \frac{\partial f_i}{\partial x_n}(\hat{x}_0) \right) \right\| \|\hat{y}\| \\ &\leq \sqrt{n} \max \left\{ \left| \frac{\partial f_i}{\partial x_j}(\hat{x}) - \frac{\partial f_i}{\partial x_j}(\hat{x}_0) \right| \mid \ j = 1, \ldots, n \right\} \|\hat{y}\| \end{aligned}$$

Pero $\left| \frac{\partial f_i}{\partial x_j}(\hat{x}) - \frac{\partial f_i}{\partial x_j}(\hat{x}_0) \right| < \frac{\varepsilon}{2\sqrt{nm}}$, así que

$$\left| \sum_{j=1}^{n} \left(\frac{\partial f_i}{\partial x_j}(\hat{x}) - \frac{\partial f_i}{\partial x_j}(\hat{x}_0) \right) y_j \right| < \sqrt{n} \frac{\varepsilon}{2\sqrt{nm}} \|\hat{y}\| = \frac{\varepsilon}{2\sqrt{m}} \|\hat{y}\|$$

y esto ocurre $\forall i \in \{1, \ldots, m\}$, de donde

$$\begin{aligned} \|D_f(\hat{x})(\hat{y}) - D_f(\hat{x}_0)(\hat{y})\| &\leq \sqrt{m} \max \left\{ \left| \sum_{j=1}^{n} \left(\frac{\partial f_i}{\partial x_j}(\hat{x}) - \frac{\partial f_i}{\partial x_j}(\hat{x}_0) \right) y_j \right| \mid \ i \in \{1, \ldots, m\} \right\} \\ &< \sqrt{m} \frac{\varepsilon}{2\sqrt{m}} \|\hat{y}\| = \frac{\varepsilon}{2} \|\hat{y}\| \end{aligned}$$

Como $\hat{y}$ fue arbitrario, entonces por definición de $\|\cdot\|_{\pounds(n,m)}$ tenemos que

$$\|D_f(\hat{x}) - D_f(\hat{x}_0)\|_{\pounds(n,m)} \leq \frac{\varepsilon}{2} < \varepsilon$$

Ahora supongamos que la función D_f es continua en $\hat{x}_0$, es decir, que para todo $\varepsilon > 0$, existe $\delta > 0$ tal que, si $\|\hat{x} - \hat{x}_0\| < \delta$, entonces

$$\|D_f(\hat{x}) - D_f(\hat{x}_0)\|_{\pounds(n,m)} < \varepsilon$$

y veamos que $\frac{\partial f_i}{\partial x_j}$ es continua en $\hat{x}_0$. Sea $\varepsilon > 0$ y $\delta > 0$ tal que para toda $\hat{x} \in \Omega$ si $\|\hat{x} - \hat{x}_0\| < \delta$, entonces $\|D_f(\hat{x}) - D_f(\hat{x}_0)\|_{\pounds(n,m)} < \varepsilon$. Tomemos $\hat{x} \in \mathbb{R}^n$ tal que $\|\hat{x} - \hat{x}_0\| < \delta$ y consideremos al vector canónico $\hat{e}_j$, entonces

$$\|D_f(\hat{x})(\hat{e}_j) - D_f(\hat{x}_0)(\hat{e}_j)\| \leq \|D_f(\hat{x}) - D_f(\hat{x}_0)\|_{\pounds(n,m)} \|\hat{e}_j\| < \varepsilon$$

pero

$$D_f(\hat{x})(\hat{e}_j) - D_f(\hat{x}_0)(\hat{e}_j) = \left(\frac{\partial f_1}{\partial x_j}(\hat{x}) - \frac{\partial f_1}{\partial x_j}(\hat{x}_0), \ldots, \frac{\partial f_m}{\partial x_j}(\hat{x}) - \frac{\partial f_m}{\partial x_j}(\hat{x}_0)\right)$$

por lo que

$$\left|\frac{\partial f_i}{\partial x_j}(\hat{x}) - \frac{\partial f_i}{\partial x_j}(\hat{x}_0)\right| \leq \|D_f(\hat{x})(\hat{e}_j) - D_f(\hat{x}_0)(\hat{e}_j)\| < \varepsilon$$

es decir, $\frac{\partial f_i}{\partial x_j}$ es continua en $\hat{x}_0$.

■

Corolario 1.3 *$f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ es de clase C^1 en Ω si y sólo si la función $D_f : \Omega \subset \mathbb{R}^n \to \pounds(n,m)$ es continua en Ω con la norma $\|\cdot\|_{\pounds(n,m)}$.*

Con este último corolario podemos demostrar una propiedad importante de las funciones de clase C^1.

Si $f = (f_1, \ldots, f_n) : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ es diferenciable en Ω entonces dado $\hat{y} \in \mathbb{R}^n$, $\hat{y} = (y_1, \ldots, y_n)$ y $\hat{x} \in \Omega$

$$D_f(\hat{x})(\hat{y}) = \left(\sum_{j=1}^{n} \frac{\partial f_1}{\partial x_j}(\hat{x})\, y_j, \ldots, \sum_{j=1}^{n} \frac{\partial f_m}{\partial x_j}(\hat{x})\, y_j\right)$$

que, por cierto, usando la notación del gradiente de una función, lo anterior es equivalente a

$$D_f(\hat{x})(\hat{y}) = (\nabla f_1(\hat{x}) \cdot \hat{y}, \ldots, \nabla f_m(\hat{x}) \cdot \hat{y})$$

de modo que

$$\|D_f(\hat{x})(\hat{y})\| \leq \sum_{j=1}^{m} |\nabla f_j(\hat{x}) \cdot \hat{y}| \leq \left(\sum_{j=1}^{m} \|\nabla f_j(\hat{x})\|\right) \|\hat{y}\|$$

y por lo tanto

$$\|D_f(\hat{x})\|_{\pounds(n,m)} \leq \left(\sum_{j=1}^{m} \|\nabla f_j(\hat{x})\|\right)$$

Ahora bien, si f es una función de clase C^1 en Ω, entonces cada función coordenada lo es también, lo que convierte a cada gradiente en una función continua de Ω en $\mathbb{R}^n$, de manera que si $K \subset \Omega$ es compacto entonces $\sup \{\|\nabla f_j(\hat{x})\| \mid \hat{x} \in K\}$ existe para cada $j \in \{1, \ldots, m\}$ y en consecuencia $\sup \left\{ \|D_f(\hat{x})\|_{\pounds(n,m)} \mid \hat{x} \in K \right\}$ también existe. Sin embargo, si nos apoyamos del corolario 1.2, podemos demostrar algo un poco más fuerte.

Proposición 1.4 *Si $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ es de clase C^1 en Ω entonces para todo conjunto $K \subset \Omega$ compacto existe $\hat{x}_0 \in K$ tal que para toda $\hat{x} \in K$*

$$\|D_f(\hat{x})\|_{\pounds(n,m)} \leq \|D_f(\hat{x}_0)\|_{\pounds(n,m)}$$

Dem. Sea $\psi : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ definida como $\psi(x) = \|D_f(\hat{x})\|_{\pounds(n,m)}$. Veamos que ψ es una función continua. Sean $\hat{x} \in \Omega$ y $\varepsilon > 0$, como f es de clase C^1 en Ω entonces existe $\delta > 0$ tal que para toda $\hat{y} \in \Omega$ si $\|\hat{y} - \hat{x}\| < \delta$, entonces $\|D_f(\hat{y}) - D_f(\hat{x})\|_{\pounds(n,m)} < \varepsilon$. De manera que dado $\hat{y} \in \Omega$ tal que $\|\hat{y} - \hat{x}\| < \delta$; por el corolario 1.2 obtenemos que

$$\left| \|D_f(\hat{y})\|_{\pounds(n,m)} - \|D_f(\hat{x})\|_{\pounds(n,m)} \right| \leq \|D_f(\hat{y}) - D_f(\hat{x})\|_{\pounds(n,m)} < \varepsilon$$

es decir,

$$|\psi(\hat{y}) - \psi(\hat{x})| < \varepsilon$$

Por lo tanto ψ es continua como queríamos; por lo tanto si $K \subset \Omega$ es compacto entonces existe $\hat{x}_0 \in K$ tal que para toda $\hat{x} \in K$ $\psi(\hat{x}) \leq \psi(\hat{x}_0)$, es decir,

$$\|D_f(\hat{x})\|_{\pounds(n,m)} \leq \|D_f(\hat{x}_0)\|_{\pounds(n,m)}$$

■

Por supuesto que así como existe un máximo para los valores $\|D_f(\hat{x})\|_{\pounds(n,m)}$ en conjuntos compactos, también existe un mínimo. Para todo efecto práctico de este trabajo, nos es suficiente el hecho de que el supremo de estos valores exista.

Pasemos ahora a desarrollar las herramientas para poder demostrar el teorema de la función inversa.

Lema 1.2 *Sea $\Omega \subset \mathbb{R}^n$ abierto, si $f : \Omega \to \mathbb{R}^n$ es diferenciable en Ω y continuamente diferenciable en $\hat{x}_0$ y tal que $D_f(\hat{x}_0) : \mathbb{R}^n \to \mathbb{R}^n$ es uno a uno, entonces existen $\delta > 0$ y $M > 0$ tales que para todo $\hat{x} \in \Omega$ si $\hat{x} \in B_\delta(\hat{x}_0)$*[1]*, entonces $M\|\hat{y}\| \leq \|D_f(\hat{x})(\hat{y})\|$ para toda $\hat{y} \in \mathbb{R}^n$*

Dem. Por el lema 1.1 existe $m > 0$ tal que $\|D_f(\hat{x}_0)(\hat{y})\| \geq m\|\hat{y}\|$ para toda $\hat{y} \in \mathbb{R}^n$, esto porque $D_f(\hat{x}_0)$ es uno a uno. Como f es continuamente diferenciable en $\hat{x}_0$, existe $\delta > 0$ tal que si $\hat{x} \in B_\delta(\hat{x}_0)$ entonces $\|D_f(\hat{x}) - D_f(\hat{x}_0)\|_{\pounds} < \frac{m}{2}$. De manera que para toda $\hat{y} \in \mathbb{R}^n$

$$\|D_f(\hat{x})(\hat{y}) - D_f(\hat{x}_0)(\hat{y})\| \leq \|D_f(\hat{x}) - D_f(\hat{x}_0)\|_{\pounds} \|\hat{y}\| < \frac{m}{2}\|\hat{y}\|$$

[1] $B_\delta(\hat{x}_0) = \{\hat{x} \in \mathbb{R}^n \mid \|\hat{x} - \hat{x}_0\| < \delta\}$

de donde

$$m \|\hat{y}\| \leq \|D_f(\hat{x}_0)(\hat{y})\| \leq \|D_f(\hat{x})(\hat{y})\| + \|D_f(\hat{x})(\hat{y}) - D_f(\hat{x}_0)(\hat{y})\| < \|D_f(\hat{x})(\hat{y})\| + \frac{m}{2}\|\hat{y}\|$$

y por tanto $\frac{m}{2}\|\hat{y}\| \leq \|D_f(\hat{x})(\hat{y})\|$.

■

Corolario 1.4 *Sea $\Omega \subset \mathbb{R}^n$ abierto, si $f : \Omega \to \mathbb{R}^n$ es diferenciable en Ω, continuamente diferenciable en $\hat{x}_0$ y $D_f(\hat{x}_0)$ es uno a uno entonces $D_f(\hat{x})$ es uno a uno para toda $\hat{x}$ en una vecindad de $\hat{x}_0$.*

La prueba de este corolario es inmediata de los lemas 1.1 y 1.2.

El siguiente lema nos dice que si tenemos una función f, continuamente diferenciable en algún punto $\hat{x}_0$ en donde la derivada es uno a uno, entonces hay toda una vecindad del punto, donde f es uno a uno; este lema junto con el lema previo juegan un papel muy importante en la demostración del teorema de la función inversa.

Lema 1.3 *Sea $\Omega \subset \mathbb{R}^n$ abierto, si $f : \Omega \to \mathbb{R}^n$ es diferenciable en Ω, continuamente diferenciable en $\hat{x}_0$ y $D_f(\hat{x}_0)$ es uno a uno entonces existen números reales positivos δ y M tales que para toda $\hat{x} \in \Omega$*

$$M\|\hat{x}' - \hat{x}\| \leq \|f(\hat{x}') - f(\hat{x})\|$$

siempre que $\hat{x}, \hat{x}' \in B_\delta(\hat{x}_0)$.

Dem. Por el lema 1.1 existe $m > 0$ tal que $m\|\hat{y}\| \leq \|D_f(\hat{x}_0)(\hat{y})\|$ para toda $\hat{y} \in \mathbb{R}^n$. Sea $M = \frac{m}{2\sqrt{n}}$, entonces por la proposición 1.3 existe $\delta > 0$ tal que $B_\delta(\hat{x}_0) \subset \Omega$ y si $\|\hat{x} - \hat{x}_0\| < \delta$ entonces $\|D_f(\hat{x}) - D_f(\hat{x}_0)\|_{\pounds} < M$. Sean $\hat{x}, \hat{x}' \in \Omega$ tales que $\|\hat{x} - \hat{x}_0\| < \delta$ y $\|\hat{x}' - \hat{x}_0\| < \delta$, sea $\hat{z} = \hat{x}' - \hat{x}$; $\forall t \in [0,1]$ ocurre que

$$\begin{aligned} \|\hat{x} + t\hat{z} - \hat{x}_0\| &= \|t\hat{x}' + (1-t)\hat{x} - \hat{x}_0\| = \|t(\hat{x}' - \hat{x}_0) + (1-t)(\hat{x} - \hat{x}_0)\| \\ &\leq t\|\hat{x}' - \hat{x}_0\| + (1-t)\|\hat{x} - \hat{x}_0\| < t\delta + (1-t)\delta = \delta \end{aligned}$$

y por lo tanto

$$\|D_f(\hat{x} + t\hat{z}) - D_f(\hat{x}_0)\|_{\pounds} < M$$

de donde para toda $\hat{y} \in \mathbb{R}^n$ y para toda $t \in [0,1]$

$$\|D_f(\hat{x} + t\hat{z})(\hat{y}) - D_f(\hat{x}_0)(\hat{y})\| \leq M\|\hat{y}\|$$

Para cada $k = 1, \ldots, n$, sea $\pi_k : \mathbb{R}^n \to \mathbb{R}$ la proyección sobre la k-ésima coordenada $(\pi_k(x_1, \ldots, x_n) = x_k)$ y definamos $g_k : [0,1] \to \mathbb{R}$ como $g_k(t) = (\pi_k \circ f)(\hat{x} + t\hat{z})$, de esta manera cada g_k es derivable y

$$g_k'(t) = \pi_k(D_f(\hat{x} + t\hat{z})(\hat{z}))$$

entonces por el teorema del valor medio se sigue que existe $t_k \in (0,1)$ tal que

$$g_k(1) - g_k(0) = g_k'(t_k)$$

es decir,

$$(\pi_k \circ f)(\hat{x}') - (\pi_k \circ f)(\hat{x}) = \pi_k \left(D_f (\hat{x} + t_k \hat{z})(\hat{z}) \right)$$

Recordando que si $\hat{x} = (x_1, \ldots, x_n)$ entonces $\|\hat{x}\| \leq \sqrt{n} \max\{|x_i| \mid i = 1, \ldots, n\}$, tenemos que para alguna $k \in \{1, \ldots, n\}$

$$|\pi_k (D_f (\hat{x}_0)(\hat{z}))| \geq \frac{1}{\sqrt{n}} \|D_f (\hat{x}_0)(\hat{z})\|$$

y usando que para esta k

$$|\pi_k ((D_f(\hat{x} + t_k \hat{z}) - D_f(\hat{x}_0))(\hat{z}))| \leq \|(D_f (\hat{x} + t_k \hat{z}) - D_f (\hat{x}_0))(\hat{z})\|$$

obtenemos que

$$\begin{aligned} \|f(\hat{x}') - f(\hat{x})\| &\geq |(\pi_k \circ f)(\hat{x}') - (\pi_k \circ f)(\hat{x})| = |\pi_k (D_f (\hat{x} + t_k \hat{z})(\hat{z}))| \\ &\geq |\pi_k (D_f (\hat{x}_0)(\hat{z}))| - |\pi_k ((D_f (\hat{x} + t_k \hat{z}) - D_f (\hat{x}_0))(\hat{z}))| \\ &\geq \frac{1}{\sqrt{n}} \|D_f (\hat{x}_0)(\hat{z})\| - \|(D_f (\hat{x} + t_k \hat{z}) - D_f (\hat{x}_0))(\hat{z})\| \end{aligned}$$

pero $M = \frac{m}{2\sqrt{n}}$, $\|D_f (\hat{x}_0)(\hat{z})\| \geq m \|\hat{z}\|$ y $\|(D_f (\hat{x} + t_k \hat{z}) - D_f (\hat{x}_0))(\hat{z})\| \leq M \|\hat{z}\|$, por lo que obtenemos finalmente que

$$\|f(\hat{x}') - f(\hat{x})\| \geq 2M \|\hat{z}\| - M \|\hat{z}\| = M \|\hat{z}\| = M \|\hat{x}' - \hat{x}\|$$

■

Observación 1.1 *De acuerdo con el lema 1.3, si tenemos una función $f : \Omega \to \mathbb{R}^n$ diferenciable en Ω, continuamente diferenciable en $\hat{x}_0$ y $D_f(\hat{x}_0)$ es uno a uno entonces existen números reales positivos δ y M tales que para toda $\hat{x} \in \Omega$*

$$M \|\hat{x}' - \hat{x}\| \leq \|f(\hat{x}') - f(\hat{x})\|$$

siempre que $\hat{x}, \hat{x}' \in B_\delta(\hat{x}_0)$. De manera que si tomamos dos elementos $\hat{x}, \hat{x}' \in B_\delta(\hat{x}_0)$ tales que $\hat{x}' \neq \hat{x}$, entonces de la desigualdad se concluye que $f(\hat{x}') \neq f(\hat{x})$, es decir, la función f es uno a uno en $B_\delta(\hat{x}_0)$. Por lo tanto la función f tiene inversa, f^{-1}, definida en la imagen de la vecindad $B_\delta(\hat{x}_0)$ bajo f.

Teorema 1.2 *(Función inversa) Sea $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ de clase C^1 en Ω, con Ω un conjunto abierto, y tal que $D_f(\hat{x}_0)$ es uno a uno para algún $\hat{x}_0 \in \Omega$. Entonces existe un abierto $N \subseteq \Omega$ con $\hat{x}_0 \in N$ tal que*
i) f restringida a N es inyectiva.
ii) $f(N)$ es abierto.
iii) $f^{-1} : f(N) \to N$ es de clase C^1 en $f(N)$, y para todo $\hat{x} \in N$ se cumple que $D_{f^{-1}}(y) = (D_f(\hat{x}))^{-1}$ donde $y = f(\hat{x})$.

Dem. Por los lemas 1.2 y 1.3 existen $\delta > 0$ y $M > 0$ tales que $B_\delta(\hat{x}_0) \subset \Omega$, si $\hat{x}, \hat{x}' \in B_\delta(\hat{x}_0)$ entonces
$$\|f(\hat{x}') - f(\hat{x})\| \geq M \|\hat{x}' - \hat{x}\|$$
y para toda $\hat{y} \in \mathbb{R}^n$
$$\|D_f(\hat{x})(\hat{y})\| \geq M \|\hat{y}\|$$

Llamemos $N = B_\delta(\hat{x}_0)$; por la observación 1.1 f tiene inversa f^{-1} definida en $f(N)$.

Veamos que $f(N)$ es abierto; sea $\hat{y}_1 \in f(N)$ y sea $\hat{x}_1 \in N$ tal que $f(\hat{x}_1) = \hat{y}_1$, como N es abierto, $\exists \delta_1 > 0$ tal que si $\|\hat{x} - \hat{x}_1\| \leq \delta_1$, entonces $\hat{x} \in N$. Nombremos
$$B = \{\hat{x} \in \mathbb{R}^n \mid \|\hat{x} - \hat{x}_1\| \leq \delta_1\}$$
Sea $\hat{y}_2 \in \mathbb{R}^n$ tal que $\|\hat{y}_2 - \hat{y}_1\| < \frac{M\delta_1}{2}$. Si $\hat{y}_2 \in f(N)$ habremos terminado.

Sea $g : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ dada por
$$g(\hat{x}) = (f(\hat{x}) - \hat{y}_2) \cdot (f(\hat{x}) - \hat{y}_2) = \|f(\hat{x}) - \hat{y}_2\|^2$$
como B es compacto, g alcanza su valor mínimo en B. Sea $\hat{x}_2 \in B$ un punto donde g alcanza este mínimo, entonces $\|\hat{x}_2 - \hat{x}_1\| \leq \delta_1$.

Supongamos que $\|\hat{x}_2 - \hat{x}_1\| = \delta_1$, entonces
$$\|f(\hat{x}_2) - \hat{y}_1\| = \|f(\hat{x}_2) - f(\hat{x}_1)\| \geq M \|\hat{x}_2 - \hat{x}_1\| = M\delta_1$$
y por lo tanto:
$$\|f(\hat{x}_2) - \hat{y}_2\| \geq \|f(\hat{x}_2) - \hat{y}_1\| - \|\hat{y}_2 - \hat{y}_1\| > M\delta_1 - \frac{M\delta_1}{2} = \frac{M\delta_1}{2} > \|\hat{y}_1 - \hat{y}_2\| = \|f(\hat{x}_1) - \hat{y}_2\|$$
esto implica que $g(\hat{x}_2) > g(\hat{x}_1)$, contradiciendo que $\hat{x}_2$ es un punto en donde se alcanza el mínimo de g; por lo tanto $\|\hat{x}_2 - \hat{x}_1\| < \delta_1$.

Como g es derivable y alcanza su mínimo en $\hat{x}_2$ que es un punto del interior de su dominio tenemos que $D_g(\hat{x}_2)$ es la transformación idénticamente cero, es decir $D_g(\hat{x}_2)(\hat{z}) = \hat{0}$, para toda $\hat{z} \in \mathbb{R}^n$. Ahora, por la regla de la cadena sabemos que
$$Dg(\hat{x}_2) = 2(f(\hat{x}_2) - \hat{y}_2) \cdot Df(\hat{x}_2)$$
de modo que para toda $\hat{z} \in \mathbb{R}^n$ se tiene que
$$\hat{0} = D_g(\hat{x}_2)(\hat{z}) = Dg(\hat{x}_2)(\hat{z}) = 2[(f(\hat{x}_2) - \hat{y}_2) \cdot Df(\hat{x}_2)](\hat{z}) = 2(f(\hat{x}_2) - \hat{y}_2) \cdot (D_f(\hat{x}_2)(\hat{z}))$$
Como $\|D_f(\hat{x}_2)(\hat{y})\| \geq M\|\hat{y}\|$ entonces por el lema 1.1 concluimos que $D_f(\hat{x}_2)$ es uno a uno y por lo tanto su imagen es todo $\mathbb{R}^n$, de modo que existe $\hat{z} \in \mathbb{R}^n$ tal que $D_f(\hat{x}_2)(\hat{z}) = f(\hat{x}_2) - \hat{y}_2$, de donde
$$0 = 2(f(\hat{x}_2) - \hat{y}_2) \cdot (f(\hat{x}_2) - \hat{y}_2) = 2\|f(\hat{x}_2) - \hat{y}_2\|^2$$
es decir, $f(\hat{x}_2) - \hat{y}_2 = 0$ y por lo tanto $\hat{y}_2 \in f(N)$.

Sabemos que si $\hat{x} \in N$ entonces $\|D_f(\hat{x})(\hat{y})\| \geq M\|\hat{y}\|$ para toda $\hat{y} \in \mathbb{R}^n$, es decir, que por el lema 1.1 $D_f(\hat{x})$ es uno a uno. Demostremos ahora que f^{-1} es diferenciable en $f(N)$ y que $D_{f^{-1}}(\hat{y}) = (D_f(\hat{x}))^{-1}$ con $\hat{y} = f(\hat{x})$, es decir, queremos que si $\hat{y}' = f(\hat{x}')$ entonces

$$\lim_{\hat{y}\to f(\hat{x}')} \frac{\left\|f^{-1}(\hat{y}) - \hat{x}' - (D_f(\hat{x}'))^{-1}(\hat{y} - f(\hat{x}'))\right\|}{\|\hat{y} - f(\hat{x}')\|} = 0$$

Como las $\hat{y}$ que nos interesan se mueven en $f(N)$ y f^{-1} es continua (porque f lo es), entonces el límite que queremos probar es equivalente a demostrar que

$$\lim_{\hat{x}\to\hat{x}'} \frac{\left\|\hat{x} - \hat{x}' - (D_f(\hat{x}'))^{-1}(f(\hat{x}) - f(\hat{x}'))\right\|}{\|f(\hat{x}) - f(\hat{x}')\|} = 0$$

Para esto llamemos

$$s(\hat{x}) = \frac{\hat{x} - \hat{x}' - (D_f(\hat{x}'))^{-1}(f(\hat{x}) - f(\hat{x}'))}{\|f(\hat{x}) - f(\hat{x}')\|}$$

y

$$r(\hat{x}) = \frac{f(\hat{x}) - f(\hat{x}') - D_f(\hat{x}')(\hat{x} - \hat{x}')}{\|\hat{x} - \hat{x}'\|}$$

De la linealidad de $(D_f(\hat{x}'))^{-1}$ tenemos que

$$(D_f(\hat{x}'))^{-1}(r(\hat{x})) = \frac{(D_f(\hat{x}'))^{-1}(f(\hat{x}) - f(\hat{x}')) - (\hat{x} - \hat{x}')}{\|\hat{x} - \hat{x}'\|}$$

de manera que

$$s(\hat{x}) = \frac{-\|\hat{x} - \hat{x}'\|}{\|f(\hat{x}) - f(\hat{x}')\|}(D_f(\hat{x}'))^{-1}(r(\hat{x}))$$

Ahora bien, como $(D_f(\hat{x}'))^{-1}$ es continua entonces

$$\lim_{\hat{x}\to\hat{x}'}(D_f(\hat{x}'))^{-1}(r(\hat{x})) = (D_f(\hat{x}'))^{-1}\left(\lim_{\hat{x}\to\hat{x}'} r(\hat{x})\right) = (D_f(\hat{x}'))^{-1}(\hat{0}) = \hat{0}$$

Como $\|f(\hat{x}) - f(\hat{x}')\| \geq M\|\hat{x} - \hat{x}'\|$ entonces los valores $\frac{\|\hat{x}-\hat{x}'\|}{\|f(\hat{x})-f(\hat{x}')\|}$ están acotados, con lo que ya podemos concluir que $\lim_{\hat{x}\to\hat{x}'} s(\hat{x}) = \hat{0}$.

Finalmente debemos probar que f^{-1} es continuamente diferenciable en $f(N)$. Tomemos $\hat{y}_1 = f(\hat{x}_1)$ para algún $\hat{x}_1 \in N$, por la proposición 1.4 es suficiente probar que para todo $\varepsilon > 0$ existe $\delta > 0$ tal que si $\|\hat{y} - \hat{y}_1\| < \delta$, entonces $\|D_{f^{-1}}(\hat{y}) - D_{f^{-1}}(\hat{y}_1)\|_{\pounds} < \varepsilon$.

Para $\hat{x} \in N$ denotemos $L_{\hat{x}} = D_f(\hat{x})$; recordando que $\|D_f(\hat{x})(\hat{y})\| \geq M\|\hat{y}\|$ para toda $\hat{y} \in \mathbb{R}^n$ y sustituyendo $\hat{y} = L_{\hat{x}}^{-1}(\hat{z})$ tenemos que para toda $\hat{z} \in \mathbb{R}^n$, $\left\|L_{\hat{x}}^{-1}(\hat{z})\right\| \leq \frac{1}{M}\|\hat{z}\|$.

Sea $\varepsilon > 0$, como f es continuamente diferenciable en $\hat{x}_1$ tomemos $\delta > 0$ de manera que $B_\delta(\hat{x}_1) \subset N$ y para toda $\hat{x} \in B_\delta(\hat{x}_1)$

$$\|L_{\hat{x}} - L_{\hat{x}_1}\|_{\pounds} < \frac{\varepsilon}{2}M^2$$

entonces

$$\begin{aligned}
\left\| \left(L_{\hat{x}}^{-1} - L_{\hat{x}_1}^{-1} \right) (\hat{z}) \right\| &= \left\| L_{\hat{x}}^{-1} \left(L_{\hat{x}} - L_{\hat{x}_1} \right) L_{\hat{x}_1}^{-1} (\hat{z}) \right\| \\
&\leq \frac{1}{M} \left\| \left(L_{\hat{x}} - L_{\hat{x}_1} \right) L_{\hat{x}_1}^{-1} (\hat{z}) \right\| \leq \frac{1}{M} \left\| L_{\hat{x}} - L_{\hat{x}_1} \right\|_{\pounds} \left\| L_{\hat{x}_1}^{-1} (\hat{z}) \right\| \\
&< \frac{\varepsilon}{2} M \left\| L_{\hat{x}_1}^{-1} (\hat{z}) \right\| \leq \frac{\varepsilon}{2} \left\| \hat{z} \right\|
\end{aligned}$$

$\forall \hat{z} \in \mathbb{R}^n$, lo que por definición de $\|\cdot\|_{\pounds}$ implica que $\left\| L_{\hat{x}}^{-1} - L_{\hat{x}_1}^{-1} \right\|_{\pounds} \leq \frac{\varepsilon}{2} < \varepsilon$.

■

Como habíamos mencionado, el importante teorema de la función implícita se puede obtener como un "corolario" del teorema de la función inversa (de hecho, lo recíproco también es cierto, de donde se deduce que estos dos teoremas son en cierta forma "equivalentes").

Teorema 1.3 *(Función implícita) Sea $F : \Omega \subset \mathbb{R}^{n+m} \to \mathbb{R}^m$ de clase C^1 en Ω, con Ω abierto, supongamos que para algún $\hat{x}_0 \in \mathbb{R}^n$ y para algún $\hat{y}_0 \in \mathbb{R}^m$ tales que $(\hat{x}_0, \hat{y}_0) \in \Omega$ se cumple que:*

1. $F(\hat{x}_0, \hat{y}_0) = 0$

2. $F_y(\hat{x}_0, \hat{y}_0)$ *tiene inversa, donde*

$$F_y(\hat{x}_0, \hat{y}_0) = \begin{pmatrix} \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_m} \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial y_1} & \cdots & \frac{\partial F_m}{\partial y_m} \end{pmatrix}$$

y donde todas las derivadas parciales están evaluadas en $(\hat{x}_0, \hat{y}_0)$.

Entonces existe un abierto $N \subset \mathbb{R}^n$, con $\hat{x}_0 \in N$, y existe una función $f : N \subset \mathbb{R}^n \to \mathbb{R}^m$ de clase C^1 en N, tal que $f(\hat{x}_0) = \hat{y}_0$ y $F(\hat{x}, f(\hat{x})) = 0$ para toda $\hat{x} \in N$.

Dem. Sea $H : \Omega \subset \mathbb{R}^{n+m} \to \mathbb{R}^{n+m}$ definida por $H(\hat{x}, \hat{y}) = (\hat{x}, F(\hat{x}, \hat{y}))$, como F es continuamente diferenciable H también lo es y además

$$H'(\hat{x}_0, \hat{y}_0) = \begin{pmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 1 & 0 & \cdots & 0 \\ \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial F_m}{\partial x_1} & \cdots & \frac{\partial F_m}{\partial x_n} & \frac{\partial F_m}{\partial y_1} & \cdots & \frac{\partial F_m}{\partial y_m} \end{pmatrix}$$

donde nuevamente todas las derivadas parciales están evaluadas en $(\hat{x}_0, \hat{y}_0)$. Como $F_{\hat{y}}(\hat{x}_0, \hat{y}_0)$ tiene inversa sabemos que las últimas m columnas son linealmente independientes y por lo tanto las $n+m$ columnas lo son, es decir, que $H'(\hat{x}_0, \hat{y}_0)$ tiene inversa o dicho de otra forma $D_H(\hat{x}_0, \hat{y}_0)$ es uno a uno. Entonces por el teorema de la función inversa existe $N' \subseteq \Omega$ abierto tal que $(\hat{x}_0, \hat{y}_0) \in N'$, $H(N')$ es abierto y H restringido a N' tiene inversa continuamente diferenciable.

Sea $N = \left\{\hat{x} \in \mathbb{R}^n \mid \left(\hat{x}, \hat{0}\right) \in H(N')\right\}$; es claro que $\hat{x}_0 \in N$ pues $H(\hat{x}_0, \hat{y}_0) = (\hat{x}_0, F(\hat{x}_0, \hat{y}_0)) = \left(\hat{x}_0, \hat{0}\right)$. Además N es un conjunto abierto; para ver esto tomemos $\hat{x} \in N$, es decir, $\left(\hat{x}, \hat{0}\right) \in H(N')$ y como ya sabemos que éste es abierto, existe $\delta > 0$ tal que $B_\delta\left(\hat{x}, \hat{0}\right) \subseteq H(N')$; sea $\hat{x}' \in \mathbb{R}^n$ tal que $\|\hat{x} - \hat{x}'\| < \delta$, entonces $\left\|\left(\hat{x}, \hat{0}\right) - \left(\hat{x}', \hat{0}\right)\right\| = \|\hat{x} - \hat{x}'\| < \delta$ de modo que $\left(\hat{x}', \hat{0}\right) \in H(N')$, es decir, $\hat{x}' \in N$.

Sean $G_1 : \mathbb{R}^{n+m} \to \mathbb{R}^n$ y $G_2 : \mathbb{R}^{n+m} \to \mathbb{R}^m$ dadas por $G_1(\hat{x}, \hat{y}) = \hat{x}$ y $G_2(\hat{x}, \hat{y}) = \hat{y}$; nótese que $G_1 = G_1 \circ H^{-1}$ si nos restringimos a $H(N')$, ya que dado cualquier $H(\hat{x}, \hat{y}) \in H(N')$ se tiene que

$$G_1 \circ H^{-1}(H(\hat{x}, \hat{y})) = G_1(\hat{x}, \hat{y}) = \hat{x} = G_1(\hat{x}, F(\hat{x}, \hat{y})) = G_1(H(\hat{x}, \hat{y}))$$

Definamos entonces a $f : N \to \mathbb{R}^m$ como $f(\hat{x}) = G_2 \circ H^{-1}\left(\hat{x}, \hat{0}\right)$. Esta función f, es composición de tres funciones continuamente diferenciables, G_2, H^{-1} y la asignación $\hat{x} \mapsto \left(\hat{x}, \hat{0}\right)$, de manera que f resulta ser continuamente diferenciable; además

$$f(\hat{x}_0) = G_2 \circ H^{-1}\left(\hat{x}_0, \hat{0}\right) = G_2(\hat{x}_0, \hat{y}_0) = \hat{y}_0$$

Por último, como $H^{-1}\left(\hat{x}, \hat{0}\right) \in \mathbb{R}^{n+m}$, entonces lo podemos ver como

$$H^{-1}\left(\hat{x}, \hat{0}\right) = \left(G_1\left(H^{-1}\left(\hat{x}, \hat{0}\right)\right), G_2\left(H^{-1}\left(\hat{x}, \hat{0}\right)\right)\right) = (\hat{x}, f(\hat{x}))$$

por lo tanto

$$\left(\hat{x}, \hat{0}\right) = H(\hat{x}, f(\hat{x})) = (\hat{x}, F(\hat{x}, f(\hat{x})))$$

de donde $F(\hat{x}, f(\hat{x})) = \hat{0}$ para toda $\hat{x} \in N$.

■

Capítulo 2

Teorema de Lebesgue

En este capítulo se desarrollarán algunos elementos de la teoría de integración introduciendo un nuevo concepto para el estudiante de Cálculo en varias variables, el de conjunto con medida de Lebesgue cero. Veremos que significa esto, que propiedades tienen estos conjuntos y daremos una caracterización de las funciones Riemann integrables en términos de este concepto.

2.1 Funciones Riemann integrables

Comenzaremos este capítulo recordando lo que es un rectángulo en $\mathbb{R}^n$, el volumen de un rectángulo, una partición de un rectángulo, así como los conceptos de suma inferior y suma superior de una función acotada definida en un rectángulo.

Definición 2.1 *Si* $a_i, b_i \in \mathbb{R}$ *para* $i \in \{1, \ldots, n\}$ *con* $a_i \leq b_i$ *entonces*

$$\begin{aligned} R &= \{\hat{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n \mid a_i \leq x_i \leq b_i, i = 1, \ldots, n\} \\ &= [a_1, b_1] \times \cdots \times [a_n, b_n] \end{aligned}$$

es llamado un rectángulo cerrado en $\mathbb{R}^n$*.*

Definición 2.2 *Si* $a_i, b_i \in \mathbb{R}$ *para* $i \in \{1, \ldots, n\}$ *con* $a_i < b_i$*, entonces*

$$\begin{aligned} R &= \{\hat{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n \mid a_i < x_i < b_i, i = 1, \ldots, n\} \\ &= (a_1, b_1) \times \cdots \times (a_n, b_n) \end{aligned}$$

es llamado un rectángulo abierto en $\mathbb{R}^n$*.*

Por comodidad llamaremos a los rectángulos cerrados simplemente rectángulos; por definición para estos rectángulos queda abierta la posibilidad de que $a_i = b_i$ para alguna $i \in \{1, \ldots, n\}$, en cuyo caso diremos que se trata de un rectángulo degenerado. Es claro que los rectángulos son llamados abiertos o cerrados porque son en sí conjuntos abiertos o cerrados respectivamente, más aún, si R es un rectángulo abierto entonces $\overline{R}$ (la cerradura de R) es un rectángulo cerrado o viceversa si R es un rectángulo cerrado entonces $int\,(R)$ (el interior de R) es un rectángulo abierto; de manera implícita estamos diciendo que si R es un rectángulo entonces $Fr\,(R) = R \setminus int\,(R)$ y $\overline{int\,(R)} = R$ (donde $Fr\,(R)$ denota la frontera de R) detalle a tener muy en cuenta[1].

Ahora definiremos los conceptos de vértice y centro de un rectángulo; ambas definiciones son muy naturales. En el siguiente capítulo veremos como el concepto de centro de un rectángulo nos permitirá caracterizar a cierto tipo de retángulos, haciendo nuestro trabajo un poco más fácil.

Definición 2.3 *Sea $R \subset \mathbb{R}^n$ un rectángulo con $R = [a_1, b_1] \times \cdots \times [a_n, b_n]$; decimos que un punto $\hat{v} \in R$ es un vértice de R, si $\hat{v} = (x_1, \ldots, x_n)$ donde $x_i \in \{a_i, b_i\}$ para cada $i \in \{1, \ldots, n\}$.*

Definición 2.4 *Sea $R \subset \mathbb{R}^n$ un rectángulo con $R = [a_1, b_1] \times \cdots \times [a_n, b_n]$; al punto*

$$\hat{c} = \left(\frac{a_1 + b_1}{2}, \ldots, \frac{a_n + b_n}{2} \right) \in R$$

lo llamaremos el centro de R.

Una relación muy clara entre el centro de un rectángulo y los vértices del mismo es que el centro del rectángulo dista lo mismo de cualquier vértice.

Proposición 2.1 *Sea $R \subset \mathbb{R}^n$ un rectángulo con $R = [a_1, b_1] \times \cdots \times [a_n, b_n]$. Si $\hat{v}_1, \hat{v}_2 \in R$ son dos vértices arbitrarios y $\hat{c} \in R$ es el centro de R entonces $\|\hat{v}_1 - \hat{c}\| = \|\hat{v}_2 - \hat{c}\|$.*

Dem. La prueba de esto es muy sencilla; basta notar que para cualquier vértice $\hat{v} = (x_1, \ldots, x_n)$, con $x_i \in \{a_i, b_i\}$ para toda $i \in \{1, \ldots, n\}$, se tiene que

$$\left| x_i - \frac{a_i + b_i}{2} \right| = \frac{b_i - a_i}{2}$$

es decir, $\|\hat{v} - \hat{c}\| = \frac{1}{2}\sqrt{(b_1 - a_1)^2 + \cdots + (b_n - a_n)^2}$ ■

Definición 2.5 *Si $R \subset \mathbb{R}^n$ es un rectángulo con $R = [a_1, b_1] \times \cdots \times [a_n, b_n]$, definimos el volumen de R, $V(R)$, como*

$$V(R) = (b_1 - a_1)(b_2 - a_2) \cdots (b_n - a_n)$$

Si $R = (a_1, b_1) \times \cdots \times (a_n, b_n)$, definimos también $V(R) = V\left(\overline{R}\right)$.

Definición 2.6 *Sea $R \subset \mathbb{R}^n$ un rectángulo con $R = [a_1, b_1] \times \cdots \times [a_n, b_n]$, si $P_j = \left\{ x_0^{(j)}, \ldots, x_{k_j}^{(j)} \right\}$ donde $x_0^{(j)} = a_j$ y $x_{k_j}^{(j)} = b_j$, es una partición del intervalo $[a_j, b_j]$ entonces $P = P_1 \times \cdots \times P_n$ es llamada una partición del rectángulo R.*

[1] En general para $A \subset \mathbb{R}^n$ se definen los conjuntos interior de A y frontera de A como $int\,(A) = \{\hat{x} \in \mathbb{R}^n \mid \exists r > 0 \text{ tal que } B_r(\hat{x}) \subset A\}$ y $Fr\,(A) = \{\hat{x} \in \mathbb{R}^n \mid \forall r > 0,\ B_r(\hat{x}) \cap A \neq \emptyset \text{ y } B_r(\hat{x}) \cap A^c \neq \emptyset\}$, también se define la cerradura de A como $\bar{A} = A \cup Fr\,(A) = int\,(A) \cup Fr\,(A)$.

Así como en $\mathbb{R}$ una partición de un intervalo generaba subintervalos del mismo, en $\mathbb{R}^n$ una partición de un rectángulo genera subrectángulos, donde cada elemento de la partición es vértice de algunos subrectángulos inducidos por la misma.

Si $R \subset \mathbb{R}^n$ es un rectángulo entonces al conjunto de particiones de R lo denotaremos por $\mathcal{P}_R$.

Definición 2.7 *Sea $R \subset \mathbb{R}^n$ un rectángulo y $P, Q \in \mathcal{P}_R$, diremos que Q es un refinamiento de P si cada punto de P es punto de Q, es decir, $P \subset Q$.*

Si $P, Q \in \mathcal{P}_R$ y P genera subrectángulos $R_1, \ldots, R_k$ mientras que Q genera subrectángulos $R'_1, \ldots, R'_s$ y Q es un refinamiento de P entonces cada R'_i está contenido en algún R_j. Uno puede convencerse de este hecho sin mucho problema de manera geométrica, escribirlo se convierte en tarea difícil, más por la notación que por la dificultad que representa la afirmación. Para darnos una idea de como podríamos probar esto, supongamos que tenemos una partición $P = P_1 \times \cdots \times P_n$ y consideremos el refinamiento $Q = P_1 \cup \{c\} \times \cdots \times P_n$ donde $c \in [a_1, b_1] \setminus P_1$; los subrectángulos inducidos por cada una de estas particiones son casi los mismos, con excepción de algunos subrectángulos inducidos por P los cuales se pueden poner como la unión de dos subrectángulos inducidos por Q. Siendo un poco más exactos, si $P_1 = \left\{x_0^{(1)}, \ldots, x_{k_j}^{(1)}\right\}$ y $x_i^{(1)} < c < x_{i+1}^{(1)}$ para alguna $i \in \{0, \ldots, k_1\}$ entonces los subrectángulos a los que nos referimos son aquellos generados por los vértices que en su primer coordenada aparece $x_i^{(1)}$ junto con los vértices con primer coordenada $x_{i+1}^{(1)}$; dichos subrectángulos ahora en Q se subdividen en los rectángulos generados por los vértices con primer coordenada $x_i^{(1)}$ y los nuevos vértices con primer coordenada c y los rectángulos generados por estos nuevos vértices junto con los vértices de primer coordenada $x_{i+1}^{(1)}$.

Si en alguna medida hemos logrado convencernos de la relación que hay entre estas particiones P y Q, entonces para dar un argumento general, se puede considerar lo siguiente: supongamos que $P, Q \in \mathcal{P}_R$ y que Q refina a P; si $P = P_1 \times \cdots \times P_n$ y $Q = Q_1 \times \cdots \times Q_n$ entonces podemos construir una sucesión finita de particiones $Q^{(0)}, \ldots, Q^{(m)}$, tal que $Q^{(0)} = P$ y $Q^{(m)} = Q$ y con la propiedad de que cada $Q^{(i)}$ sólo tiene un punto más, en alguna de sus particiones coordenadas, que $Q^{(i-1)}$; dicha sucesión la podemos obetener comenzando con $Q^{(0)} = P$, $Q^{(1)} = P_1 \cup \{c\} \times \cdots \times P_n$ donde $c \in Q_1 \setminus P_1$; $Q^{(2)} = P_1 \cup \{c, d\} \times \cdots \times P_n$ donde $d \in Q_1 \setminus (P_1 \cup \{c\})$, y continuar así hasta agotar todos los puntos Q_1 que no pertenecen a P_1, estos son una cantidad finita, para después continuar con los puntos de Q_2 que no están en P_2, y una vez agotados éstos, continuar con el mismo procedimiento con las demás particiones coordenadas.

De esta manera se podrá concluir que cada subrectángulo inducido por P es la unión de subrectángulos inducidos por Q, o de manera equivalente, cada subrectángulo inducido por Q está contenido en algún subrectángulo inducido por P.

Un concepto, muy relacionado con la definición de partición de un rectángulo, y que estaremos usando frecuentemente, es el de subdivisión de un conjunto.

Definición 2.8 *Sea $A \subset \mathbb{R}^n$, diremos que los subconjuntos $A_1, \ldots, A_k \subset A$ forman una subdivisión de A si*

$$A = \bigcup_{i=1}^{k} A_i$$

y para cualesquiera $i \neq j$ se cumple que

$$int\,(A_i) \cap int\,(A_j) = \emptyset$$

Queda claro de la definición anterior que una partición de un rectángulo es una subdivisión del mismo, pero no toda subdivisión de un rectángulo es una partición.

El lector recordará que resulta muy útil, a la hora de trabajar con particiones y refinamientos, el concepto de longitud de la diagonal de un rectángulo.

Definición 2.9 *Si $R \subset \mathbb{R}^n$ es un rectángulo con $R = [a_1, b_1] \times \cdots \times [a_n, b_n]$, definimos la longitud de la diagonal de R, $d\,(R)$, como*

$$d\,(R) = \sqrt{(b_1 - a_1)^2 + \cdots + (b_n - a_n)^2}$$

Observación 2.1 *Dado un rectángulo $R = [a_1, b_1] \times \cdots \times [a_n, b_n]$ y $\hat{x}, \hat{y} \in R$ con $\hat{x} = (x_1, \ldots, x_n)$ y $\hat{y} = (y_1, \ldots, y_n)$ se tiene que $\|\hat{x} - \hat{y}\| \leq d\,(R)$ lo que se deduce del hecho de que $|x_i - y_i| \leq b_i - a_i \ \forall i \in \{1, \ldots, n\}$.*

Definición 2.10 *Si $R \subset \mathbb{R}^n$ es un rectángulo con $R = [a_1, b_1] \times \cdots \times [a_n, b_n]$; a las longitudes de cada intervalo $[a_i, b_i]$, las llamaremos las dimensiones de R.*

Lema 2.1 *Si $R \subset \mathbb{R}^n$ es un rectángulo entonces para todo $\varepsilon > 0$ existe $R' \subset \mathbb{R}^n$ rectángulo de dimensiones racionales tal que $R \subset R'$ y $V\,(R') - V\,(R) < \varepsilon$.*

Dem. Supongamos que $R = [a_1, b_1] \times \ldots \times [a_n, b_n]$ y sea $\varepsilon > 0$.

La idea para esta prueba es la construcción de n rectángulos $R_1, \ldots, R_n \subset \mathbb{R}^n$, siendo R_n el rectángulo buscado, en donde se cumplirán dos cosas

$$i) R \subset R_1 \subset \cdots \subset R_n$$

$$ii) V\,(R_j) - V\,(R_{j-1}) < \frac{\varepsilon}{n}$$

para $j = 1, \ldots, n$ donde $R_0 = R$. Más aún, conseguiremos que cada R_j tenga al menos j dimensiones racionales, con lo que R_n tendrá dimensiones racionales como queremos.

Para cada $i \in \{1, \ldots, n-1\}$ sea $c_i = \prod_{j=i+1}^{n} (b_j - a_j)$, y definamos $c_n = 1$. Nótese que $c_{i-1} = (b_i - a_i)\, c_i \ \forall i \in \{2, \ldots, n\}$.

La clave para la construcción de los R_j está en la elección de sus dimensiones racionales, mismas que haremos de forma recursiva; sea $s_1 \in \mathbb{Q}$ tal que

$$b_1 - a_1 < s_1 < b_1 - a_1 + \frac{\varepsilon}{nc_1}$$

y para $i \in \{2, \ldots, n\}$ tomemos $s_i \in \mathbb{Q}$ tal que

$$b_i - a_i < s_i < b_i - a_i + \frac{\varepsilon}{nc_i \prod_{j=1}^{i-1} s_j}$$

Para cada $i \in \{1, \ldots, n\}$ sean $r_i = s_i - (b_i - a_i)$, $a'_i = a_i - \frac{r_i}{2}$ y $b'_i = b_i + \frac{r_i}{2}$ y definamos

$$R_i = [a'_1, b'_1] \times \ldots \times [a'_i, b'_i] \times [a_{i+1}, b_{i+1}] \times \ldots \times [a_n, b_n]$$

De la construcción es claro que $R \subset R_1 \subset \cdots \subset R_n$, además es notable que a'_i y b'_i no necesariamente son racionales pero $b'_i - a'_i$ sí lo es ya que $b'_i - a'_i = s_i \in \mathbb{Q}$. Por lo tanto cada R_i tiene al menos i dimensiones racionales. Nos falta verificar que $V(R_j) - V(R_{j-1}) < \frac{\varepsilon}{n}$ para $j \in \{1, \ldots, n\}$ donde como ya se dijo llamamos $R_0 = R$.

Para $j = 1$, $V(R_1) = s_1 c_1$ y entonces

$$\begin{aligned} V(R_1) - V(R_0) &= s_1 c_1 - \prod_{i=1}^{n} (b_i - a_i) \\ &< \left(b_1 - a_1 + \frac{\varepsilon}{n c_1}\right) c_1 - \prod_{i=1}^{n} (b_i - a_i) \\ &= \prod_{i=1}^{n} (b_i - a_i) + \frac{\varepsilon}{n} - \prod_{i=1}^{n} (b_i - a_i) = \frac{\varepsilon}{n} \end{aligned}$$

En general $V(R_i) = c_i \prod_{j=1}^{i} s_j$ y entonces

$$\begin{aligned} V(R_i) - V(R_{i-1}) &= c_i \prod_{j=1}^{i} s_j - c_{i-1} \prod_{j=1}^{i-1} s_j \\ &< \left(b_i - a_i + \frac{\varepsilon}{n c_i \prod_{j=1}^{i-1} s_j}\right) c_i \prod_{j=1}^{i-1} s_j - c_{i-1} \prod_{j=1}^{i-1} s_j \\ &= c_{i-1} \prod_{j=1}^{i-1} s_j + \frac{\varepsilon}{n} - c_{i-1} \prod_{j=1}^{i-1} s_j = \frac{\varepsilon}{n} \end{aligned}$$

De manera que R_n es un rectángulo de dimensiones racionales tal que $R \subset R_n$ y

$$\begin{aligned} V(R_n) - V(R) &= V(R_n) + (V(R_{n-1}) - V(R_{n-1})) + \cdots + (V(R_1) - V(R_1)) - V(R) \\ &= (V(R_n) - V(R_{n-1})) + \cdots + (V(R_1) - V(R)) < \sum_{j=1}^{n} \frac{\varepsilon}{n} = \varepsilon \end{aligned}$$

■

Observación 2.2 *Cabe destacar que el resultado anterior es válido si en lugar de circunscribir un rectángulo a uno ya dado lo inscribimos, es decir, dado $R \subset \mathbb{R}^n$ rectángulo y una cantidad positiva ε, existe un rectángulo $R' \subset \mathbb{R}^n$ de dimensiones racionales tal que $R' \subset R$ y $V(R) - V(R') < \varepsilon$. Esto se puede probar de manera similar a lo hecho en el lema y ya no lo demostraremos (pero sí lo usaremos). También vale la pena destacar que por la demostración realizada en el lema se concluye que R no sólo está contenido en R' sino que no toca a la frontera de R', es decir $R \subset int(R')$ y por supuesto algo similar se puede asegurar para el caso en el que se inscribe un rectángulo.*

El lema anterior en realidad se cumple sin importar cómo sean las dimensiones del rectángulo encontrado, éstas pueden ser racionales o no, no importa, siempre se puede conseguir un R' tal que $R \subset R'$ y $V(R') - V(R) < \varepsilon$ e incluso la prueba de esto es un poco más sencilla, pero para nuestros propósitos nos es de gran ayuda poder asegurar además que las dimensiones de este R' sean racionales.

Definición 2.11 *Sea $R \subset \mathbb{R}^n$ un rectángulo y $f : R \to \mathbb{R}$ una función acotada. Dada $P \in \mathcal{P}_R$ que genera subrectángulos $R_1, \ldots R_k$, definimos la suma superior de f respecto a P, $\overline{S}(f,P)$, como*

$$\overline{S}(f,P) = \sum_{j=1}^{k} M_j(f)\, V(R_j)$$

en donde $M_j(f) = \sup\{f(\hat{x}) \mid \hat{x} \in R_j\}$ y definimos la suma inferior de f respecto a P, $\underline{S}(f,P)$, como

$$\underline{S}(f,P) = \sum_{j=1}^{k} m_j(f)\, V(R_j)$$

en donde $m_j(f) = \inf\{f(\hat{x}) \mid \hat{x} \in R_j\}$.

Al ser f acotada en R las sumas inferiores y superiores siempre existen.

La siguiente proposición nos presenta algunas propiedades, que no demostaremos, (ver [3]) de las sumas inferiores y las superiores .

Proposición 2.2 *Sea $R \subset \mathbb{R}^n$ rectángulo, si $f : R \to \mathbb{R}$ es una función acotada entonces*

i) $\underline{S}(f,P) \leq \overline{S}(f,P)$ *para cualquier $P \in \mathcal{P}_R$.*

ii) Si $P, Q \in \mathcal{P}_R$ y Q refina a P, entonces $\underline{S}(f,P) \leq \underline{S}(f,Q)$ *y* $\overline{S}(f,Q) \leq \overline{S}(f,P)$.

iii) $\underline{S}(f,P) \leq \overline{S}(f,Q)$ *para cualesquiera particiones P y Q.*

iv) $\sup\{\underline{S}(f,P) \mid P \in \mathcal{P}_R\}$, $\inf\{\overline{S}(f,P) \mid P \in \mathcal{P}_R\}$ *existen y*

$$\sup\{\underline{S}(f,P) \mid P \in \mathcal{P}_R\} \leq \inf\{\overline{S}(f,P) \mid P \in \mathcal{P}_R\}$$

Definición 2.12 *Sea $R \subset \mathbb{R}^n$ rectángulo y $f : R \to \mathbb{R}$ una función acotada; decimos que f es una función Riemann integrable sobre R si y sólo si*

$$\sup\{\underline{S}(f,P) \mid P \in \mathcal{P}_R\} = \inf\{\overline{S}(f,P) \mid P \in \mathcal{P}_R\}$$

y se define la integral de Riemann de f sobre R como el número

$$\int_R f = \sup\{\underline{S}(f,P) \mid P \in \mathcal{P}_R\} = \inf\{\overline{S}(f,P) \mid P \in \mathcal{P}_R\}$$

Un ejemplo de una función Riemann integrable es la función constante 1, aunque es un ejemplo bastante trivial ofrece una buena relación entre una integral y el volumen de un rectángulo.

Ejemplo 2.1 *Sea $R \subset \mathbb{R}^n$ rectángulo y $f : R \to \mathbb{R}$ la función constante 1, es decir, $f(\hat{x}) = 1$ $\forall \hat{x} \in R$, entonces f es Riemann integrable sobre R y*

$$V(R) = \int_R f = \int_R 1$$

Al ser $f = 1$ en R

$$\overline{S}(f, P) = \sum_{j=1}^{k} M_j(f) V(R_j) = \sum_{j=1}^{k} V(R_j) = V(R)$$

y

$$\underline{S}(f, P) = \sum_{j=1}^{k} m_j(f) V(R_j) = \sum_{j=1}^{k} V(R_j) = V(R)$$

para cualquier partición P de R, de manera que

$$\sup\{\underline{S}(f, P) \mid P \in \mathcal{P}_R\} = V(R) = \inf\{\overline{S}(f, P) \mid P \in \mathcal{P}_R\}$$

es decir, f es integrable en R y $V(R) = \int_R 1$.

El siguiente teorema establece una condición necesaria y suficiente para que una función sea Riemann integrable, dicha condición es conocida también como condición de Riemann.

Teorema 2.1 *Sean $R \subset \mathbb{R}^n$ rectángulo y $f : R \to \mathbb{R}$ una función acotada; f es Riemann integrable sobre R si y sólo si para cada $\varepsilon > 0$ hay una partición P de R tal que*

$$\overline{S}(f, P) - \underline{S}(f, P) < \varepsilon$$

Dem. Supongamos primero que f es Riemann integrable sobre R, es decir,

$$\sup\{\underline{S}(f, P) \mid P \in \mathcal{P}_R\} = \inf\{\overline{S}(f, P) \mid P \in \mathcal{P}_R\}$$

Sea $\varepsilon > 0$, por definición de supremo e ínfimo se sabe que existen $P_1, P_2 \in \mathcal{P}_R$ tales que

$$\sup\{\underline{S}(f, P) \mid P \in \mathcal{P}_R\} - \frac{\varepsilon}{2} < \underline{S}(f, P_1)$$

y

$$\overline{S}(f, P_2) < \inf\{\overline{S}(f, P) \mid P \in \mathcal{P}_R\} + \frac{\varepsilon}{2}$$

por lo tanto

$$\overline{S}(f, P_2) - \underline{S}(f, P_1) < \inf_{P \in \mathcal{P}_R}\{\overline{S}(f, P)\} + \frac{\varepsilon}{2} - \sup_{P \in \mathcal{P}_R}\{\underline{S}(f, P)\} + \frac{\varepsilon}{2} = \varepsilon$$

Si tomamos $P \in \mathcal{P}_R$ un refinamiento tanto de P_1 como de P_2 por la proposición 2.2 tenemos que

$$\overline{S}(f, P) - \underline{S}(f, P) \leq \overline{S}(f, P_2) - \underline{S}(f, P_1) < \varepsilon$$

es decir, la condición de Riemann se satisface.

Ahora demostraremos por contrapositiva que la condición de Riemann implica la integrabilidad. Supongamos que f no es Riemann integrable sobre R, esto significa por la proposición 2.2 que

$$\sup_{P\in\mathcal{P}_R}\left\{\underline{S}(f,P)\right\}<\inf_{P\in\mathcal{P}_R}\left\{\overline{S}(f,P)\right\}$$

entonces hay una cantidad positiva ε tal que

$$\sup_{P\in\mathcal{P}_R}\left\{\underline{S}(f,P)\right\}+\varepsilon=\inf_{P\in\mathcal{P}_R}\left\{\overline{S}(f,P)\right\}$$

y por lo tanto para toda partición P de R

$$\underline{S}(f,P)+\varepsilon\leq\sup_{P\in\mathcal{P}_R}\left\{\underline{S}(f,P)\right\}+\varepsilon=\inf_{P\in\mathcal{P}_R}\left\{\overline{S}(f,P)\right\}\leq\overline{S}(f,P)$$

de donde

$$\overline{S}(f,P)-\underline{S}(f,P)\geq\varepsilon$$

es decir, la condición de Riemann no se satisface. ■

Gracias a este criterio de integrabilidad se puede demostrar, de manera más sencilla que de la sola definición, algunas de las propiedades de la integral de Riemann y de funciones Riemann integrables. Nosotros sólo nos limitaremos a enunciar algunas de éstas (ver [3]).

Proposición 2.3 *Sea $R\subset\mathbb{R}^n$ rectángulo y $f,g:R\to\mathbb{R}$ funciones acotadas y Riemann integrables sobre R. Entonces*

i) $f+g:R\to\mathbb{R}$ es Riemann integrable sobre R y $\int_R(f+g)=\int_R f+\int_R g$

ii) Para toda $c\in\mathbb{R}$, $cf:R\to\mathbb{R}$ es Riemann integrable sobre R y $\int_R cf=c\int_R f$

iii) $fg:R\to\mathbb{R}$ es Riemann integrable sobre R.

iv) Si $R_1,\ldots,R_m\subset\mathbb{R}^n$ son todos los subrectángulos inducidos por una partición de R, entonces f es integrable sobre cada R_k y $\int_R f=\sum_{i=1}^m\left(\int_{R_i}f\right)$

2.2 Medida de Lebesgue cero

Como se mencionó al inicio de este capítulo definiremos un nuevo concepto, el de medida de Lebesgue cero para un conjunto en $\mathbb{R}^n$, así que pasaremos a ver que significa que un conjunto tenga medida de Lebesgue cero y algunas propiedades importantes de estos conjuntos.

Definición 2.13 *Un conjunto $S\subset\mathbb{R}^n$ tiene medida de Lebesgue cero, $\lambda(S)=0$, si y sólo si para todo $\varepsilon>0$ hay una colección numerable de rectángulos $\{R_j\}_{j=1}^{\infty}$ tales que*

$$S\subset\bigcup_{j=1}^{\infty}R_j$$

y

$$\sum_{j=1}^{\infty} V(R_j) < \varepsilon$$

Observación 2.3 *Un punto muy interesante es, que si un conjunto S satisface que para todo $\varepsilon > 0$ hay una colección finita de rectángulos $\{R_j\}_{j=1}^{k}$ tal que $S \subset \bigcup_{j=1}^{k} R_j$ y $\sum_{j=1}^{k} V(R_j) < \varepsilon$, entonces S es un conjunto de medida de Lebesgue cero. Para tratar de ver esto con algo de rigor simplemente podemos decir que ya que se tiene a la colección finita de rectángulos, a estos les podemos agregar una cantidad numerable de rectángulos degenerados (arbitrarios); al tener volumen cero estos rectángulos, no alteran nuestra suma total de volúmenes. Nótese que para que un conjunto pueda ser cubierto por una cantidad finita de rectángulos éste debe ser acotado.*

Aunque en la definición no es necesario pedir que un conjunto sea acotado para que tenga medida de Lebesgue cero, nosotros nos quedaremos con estos conjuntos; para ilustrar lo fuerte que es la condición mencionada en la observación previa, más adelante trataremos un ejemplo de un conjunto que pese a ser acotado y de medida de Lebesgue cero no se puede cubrir con una cantidad finita de rectángulos de manera que su suma de volúmenes sea suficientemente pequeña.

Analicemos ahora algunas propiedades de los conjuntos que tienen medida de Lebesgue cero.

Proposición 2.4 *Si $S \subset \mathbb{R}^n$ es un conjunto de medida de Lebesgue cero entonces todo subconjunto de S tiene medida de Lebesgue cero.*

Esto es inmediato de la definición de medida de Lebesgue cero.

Proposición 2.5 *Todo conjunto a lo más numerable tiene medida de Lebesgue cero.*

Dem. Sea $S \subset \mathbb{R}^n$ un conjunto infinito numerable, es decir, $S = \{\hat{a}_j \in \mathbb{R}^n \mid j \in \mathbb{N}\}$, sea $R_k \subset \mathbb{R}^n$ rectángulo tal que $\hat{a}_k \in R_k$ y

$$V(R_k) = \frac{\varepsilon}{2^{k+1}}$$

Es sencillo convencernos de la existencia de cada R_k: si $\hat{a}_k = (x_1, \ldots, x_n)$, podemos considerar al rectángulo

$$R_k = \left[x_1 - \frac{1}{2}c^{\frac{1}{n}}, x_1 + \frac{1}{2}c^{\frac{1}{n}}\right] \times \cdots \times \left[x_n - \frac{1}{2}c^{\frac{1}{n}}, x_n + \frac{1}{2}c^{\frac{1}{n}}\right]$$

donde

$$c = \frac{\varepsilon}{2^{k+1}}$$

Entonces $\hat{a}_k \in R_k$ y $V(R_k) = c = \frac{\varepsilon}{2^{k+1}}$.

Dado que

$$\sum_{k=1}^{\infty} \frac{\varepsilon}{2^{k+1}} = \frac{\varepsilon}{2}$$

podemos concluir que $\sum_{k=1}^{\infty} V(R_k)$ existe y además

$$\sum_{k=1}^{\infty} V(R_k) = \frac{\varepsilon}{2} < \varepsilon$$

es decir, $\lambda(S) = 0$.

Para el caso en el que tenemos un conjunto finito se puede imitar la demostración previa y aplicar la observación 2.3. ■

Proposición 2.6 *Si $R \subset \mathbb{R}^n$ es un rectángulo entonces $Fr(R)$ tiene medida de Lebesgue cero.*

Dem. Si R es un rectángulo degenerado entonces el resultado se sigue de forma inmediata, ya que $Fr(R) \subset R$ y $V(R) = 0$.

Supongamos que R es un rectángulo no degenerado; sea $\varepsilon > 0$, y sean $R', R'' \subset \mathbb{R}^n$ rectángulos tales que $R' \subset int(R)$, $R \subset int(R'')$ y

$$\begin{aligned} V(R) - V(R') &< \frac{\varepsilon}{2} \\ V(R'') - V(R) &< \frac{\varepsilon}{2} \end{aligned}$$

tal y como asegura que existen el lema 2.1 y la observación 2.2. Sea $P \in \mathcal{P}_{R''}$ una partición generada a partir de los vértices de R' y sean $R_1, \ldots R_k$ los subrectángulos generados por P, nótese que R' es uno de estos subrectángulos, supongamos sin pérdida de generalidad que $R' = R_k$. Ahora bien, como $R' \subset int(R)$ entonces $Fr(R) \cap R' = \emptyset$, lo que implica que $Fr(R) \subset \bigcup_{j=1}^{k-1} R_j$ y como

$$V(R'') = \sum_{i=1}^{k} V(R_i) = V(R') + \sum_{i=1}^{k-1} V(R_i)$$

tenemos que

$$\sum_{i=1}^{k-1} V(R_i) = V(R'') - V(R') = (V(R'') - V(R)) + (V(R) - V(R')) < \varepsilon$$

De acuerdo con la observación 2.3 podemos concluir que $\lambda(Fr(R)) = 0$. ■

Proposición 2.7 *Si $R \subset \mathbb{R}^n$ es un rectángulo no degenerado ($V(R) > 0$) entonces R no tiene medida de Lebesgue cero.*

Esto es porque toda colección numerable de rectángulos $\{R_j\}_{j=1}^{\infty}$ tal que $R \subset \bigcup_{j=1}^{\infty} R_j$ satisface que

$$0 < V(R) \leq \sum_{k=1}^{\infty} V(R_k)$$

y por lo tanto la suma de volúmenes no se puede hacer menor que $V(R) > 0$.

La proposición 2.5 deja ver que no hace falta que un conjunto sea acotado para tener medida de Lebesgue cero y como un caso particular de este hecho se tiene que el conjunto de puntos de $\mathbb{R}^n$ cuyas coordenadas son racionales forman un conjunto de medida de Lebesgue cero. La proposición 2.6 jugará un papel importante a la hora de caracterizar las funciones Riemann integrables como estamos buscando hacerlo.

Proposición 2.8 *Sea $\{S_i \subset \mathbb{R}^n \mid i \in \mathbb{N}\}$ una colección de conjuntos de medida de Lebesgue cero, si definimos $S = \bigcup_{i=1}^{\infty} S_i$ entonces S tiene medida de Lebesgue cero.*

Dem. Sea $\varepsilon > 0$, como $\lambda(S_k) = 0$ para toda $k \in \mathbb{N}$, entonces para cada $k \in \mathbb{N}$ existe una colección numerable de rectángulos $\left\{R_i^{(k)} \mid i \in \mathbb{N}\right\}$ tales que $S_k \subset \bigcup_{i=1}^{\infty} R_i^{(k)}$ y

$$\sum_{i=1}^{\infty} V\left(R_i^{(k)}\right) < \frac{\varepsilon}{2^{k+1}}$$

entonces

$$S \subset \bigcup_{k=1}^{\infty}\left[\bigcup_{i=1}^{\infty} R_i^{(k)}\right]$$

Como sabemos que la unión numerable de conjuntos numerables es numerable[2], se tiene que la colección de rectángulos $\left\{R_i^{(k)} \mid i \in \mathbb{N}, k \in \mathbb{N}\right\}$ es numerable de manera que la podemos renombrar para que formen una sucesión de rectángulos. Ahora definamos

$$a_k = \sum_{i=1}^{\infty} V\left(R_i^{(k)}\right)$$

para cada $k \in \mathbb{N}$, entonces $0 < a_k < \frac{\varepsilon}{2^{k+1}}$ y como

$$\sum_{k=1}^{\infty} \frac{\varepsilon}{2^{k+1}} = \frac{\varepsilon}{2}$$

se tiene que $\sum_{k=1}^{\infty} a_k$ existe y

$$\sum_{k=1}^{\infty} a_k \leq \frac{\varepsilon}{2} < \varepsilon$$

Es decir,

$$\sum_{k=1}^{\infty}\left[\sum_{i=1}^{\infty} V\left(R_i^{(k)}\right)\right] < \varepsilon$$

■

Por definición para un conjunto que tiene medida de Lebesgue de cero y una cantidad positiva ε se puede encontrar una colección numerable de rectángulos que cubren al conjunto

[2] Una manera de demostrar que la unión numerable de conjuntos numerables es numerable es considerar la función $f(k,i) = (2i-1)\,2^{k-1}$, la cual es una biyección de $\mathbb{N} \times \mathbb{N}$ en $\mathbb{N}$

y cuya suma de volúmenes no excede ε, debería resultar normal preguntarnos si también existe una colección numerable de rectángulos abiertos que satisfacen la misma condición. Y claro, gracias al lema 2.1 la respuesta es afirmativa.

Proposición 2.9 *$S \subset \mathbb{R}^n$ tiene medida de Lebesgue cero si y sólo si para todo $\varepsilon > 0$ existe una colección numerable de rectángulos abiertos $\{R_j\}_{j=1}^{\infty}$ tales que $S \subset \bigcup_{j=1}^{\infty} R_j$ y*

$$\sum_{j=1}^{\infty} V(R_j) < \varepsilon$$

Dem. La necesidad es bastante trivial pues dada una colección numerable de rectángulos abiertos $\{R_j\}_{j=1}^{\infty}$, $\left\{\overline{R_j}\right\}_{j=1}^{\infty}$ es una colección numerable de rectángulos tales que $V(R_j) = V\left(\overline{R_j}\right)$.

Para la suficiencia, dada $\varepsilon > 0$ tomemos $\{Q_j\}_{j=1}^{\infty}$ colección de rectángulos tales que

$$S \subset \bigcup_{j=1}^{\infty} Q_j$$

y

$$\sum_{j=1}^{\infty} V(Q_j) < \frac{\varepsilon}{2}$$

por el lema 2.1 para cada $j \in \mathbb{N}$ existe un rectángulo R_j tal que $Q_j \subset int(R_j)$ y

$$V(R_j) < \frac{\varepsilon}{2^{j+1}} + V(Q_j)$$

entonces $\sum_{j=1}^{\infty} V(R_j)$ existe y

$$\sum_{j=1}^{\infty} V(R_j) \leq \sum_{j=1}^{\infty} \left[\frac{\varepsilon}{2^{j+1}} + V(Q_j)\right] < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

de manera que $\{int(R_j)\}_{j=1}^{\infty}$ es la colección de rectángulos abiertos deseada ■

Ya estamos listos para dar la caracterización de las funciones Riemann integrables de la que tanto hemos hablado. Para ello, establezcamos la siguiente notación: si $f : A \subset \mathbb{R}^n \to \mathbb{R}$, denotaremos por $D_{f,A} = \{\hat{x} \in A \mid f \text{ es discontinua en } \hat{x}\}$, es decir, $D_{f,A}$ es el conjunto de discontinuidades de f en A.

Teorema 2.2 *(Lebesgue) Sean $R \subset \mathbb{R}^n$ rectángulo y $f : R \to \mathbb{R}$ una función acotada. f es Riemann integrable sobre R si y sólo si el conjunto de discontinuidades de f en R tiene medida de Lebesgue cero.*

Dem. Primero supongamos que $D_{f,R}$ tiene medida de Lebesgue cero y demostremos que f debe ser Riemann integrable sobre R.

Sean $\varepsilon > 0$ y $\{R_j\}_{j=1}^{\infty}$ una colección de rectángulos tales que $D_{f,R} \subset \bigcup_{j=1}^{\infty} int\,(R_j)$ y

$$\sum_{j=1}^{\infty} V\,(R_j) < \frac{\varepsilon}{2\,(M-m)}$$

en donde $M = \sup\{f\,(\hat{x}) \mid \hat{x} \in R\}$ y $m = \inf\{f\,(\hat{x}) \mid \hat{x} \in R\}$. No está de más decir que si $M - m = 0$ entonces f sería una función constante en R y por la proposición 2.3 f es Riemann integrable, así que podemos suponer que $M - m > 0$. También nótese que de acuerdo con la proposición 2.9 podemos pedir que $D_{f,R} \subset \bigcup_{j=1}^{\infty} int\,(R_j)$.

Si $\hat{x} \in R \setminus D_{f,R}$, entonces f es continua en $\hat{x}$, de modo que existe $\delta_{\hat{x}} > 0$ tal que para cualquier $\hat{\xi} \in B_{\delta_{\hat{x}}}\,(\hat{x}) \cap R$

$$\left|f\left(\hat{\xi}\right) - f\,(\hat{x})\right| < \frac{\varepsilon}{8V\,(R)}$$

La importancia de poder pedir que $D_{f,R} \subset \bigcup_{j=1}^{\infty} int\,(R_j)$ es que el conjunto

$$U = \{int\,(R_j) \mid j \in \mathbb{N}\} \cup \left\{B_{\frac{\delta_{\hat{x}}}{2}}\,(\hat{x}) \mid \hat{x} \in R \setminus D_{f,R}\right\}$$

es una cubierta abierta para R y de la compacidad de R tenemos que existen $k_1, \ldots, k_r \in \mathbb{N}$ y $\hat{x}_1, \ldots, \hat{x}_s \in R \setminus D_{f,R}$ tales que

$$R \subset \left[\bigcup_{j=k_1}^{k_r} int\,(R_j)\right] \cup \left[\bigcup_{j=1}^{s} B_{\frac{\delta_j}{2}}\,(\hat{x}_j)\right]$$

en donde $\delta_j = \delta_{\hat{x}_j}$. Llamemos $R'_j = R \cap R_{k_j}$ para cada $j \in \{1, \ldots, r\}$, sea $P_0 \in \mathcal{P}_R$ una partición que contenga a todos los vértices de cada R'_j y $P \in \mathcal{P}_R$ un refinamiento de P_0 de forma tal que si $Q_1, \ldots, Q_m$ son los subrectángulos generados por P, entonces $d\,(Q_j) < \frac{\delta}{2}$ para cada $j \in \{1, \ldots, m\}$, en donde $\delta = \min\{\delta_1, \ldots, \delta_s\}$. De la discusión hecha tras la definición 2.7 sabemos que cada $Q_j \subset R'_i$ para alguna $i \in \{1, \ldots, r\}$ o $int\,(Q_j) \cap int\,(R'_i) = \emptyset$ para toda $i \in \{1, \ldots, r\}$.

Sean

$$A_1 = \{k \in \{1, \ldots, m\} \mid \exists i \in \{1, \ldots, r\} \text{ tal que } Q_k \subset R'_i\}$$

y $A_2 = \{1, \ldots, m\} \setminus A_1$. Por construcción $\{1, \ldots, m\} = A_1 \cup A_2$ y $A_1 \cap A_2 = \emptyset$, de modo que

$$\overline{S}\,(f,P) - \underline{S}\,(f,P) = \sum_{j\in A_1} (M_j\,(f) - m_j\,(f))\,V\,(Q_j) + \sum_{j\in A_2} (M_j\,(f) - m_j\,(f))\,V\,(Q_j)$$

Queremos comprobar que cada suma sea menor que $\frac{\varepsilon}{2}$. Para la primera se tiene que

$$\begin{aligned}
\sum_{j\in A_1} (M_j\,(f) - m_j\,(f))\,V\,(Q_j) &\leq (M-m)\sum_{j\in A_1} V\,(Q_j) \\
&\leq (M-m)\sum_{j=1}^{r} V\left(R'_j\right) \leq (M-m)\sum_{j=1}^{\infty} V\,(R_j) < \frac{\varepsilon}{2}
\end{aligned}$$

Para la segunda suma se tiene lo siguiente, dada $j \in A_2$ existen $\hat{x}, \hat{y} \in Q_j$ tales que

$$M_j(f) - \frac{\varepsilon}{8V(R)} < f(\hat{x})$$

y

$$f(\hat{y}) < m_j(f) + \frac{\varepsilon}{8V(R)}$$

de donde

$$M_j(f) - m_j(f) < f(\hat{x}) - f(\hat{y}) + \frac{\varepsilon}{4V(R)}$$

Como por definición para $j \in A_2$, $int(Q_j) \cap int(R'_i) = \emptyset$ para toda $i \in \{1, \ldots, r\}$ y $\hat{x} \in Q_j$, entonces $\hat{x} \in B_{\frac{\delta i}{2}}(\hat{x}_i)$ para alguna $i \in \{1, \ldots, s\}$ y por lo tanto

$$|f(\hat{x}) - f(\hat{x}_i)| < \frac{\varepsilon}{8V(R)}$$

Ahora bien, al tomar a la partición P, se pidió que $d(Q_j) < \frac{\delta}{2}$ lo que implica que $\|\hat{x} - \hat{y}\| < \frac{\delta}{2}$ y entonces

$$\|\hat{y} - \hat{x}_i\| \leq \|\hat{y} - \hat{x}\| + \|\hat{x} - \hat{x}_i\| < \delta_i$$

por lo tanto

$$\left|f(\hat{y}) - f\left(\hat{\xi}_i\right)\right| < \frac{\varepsilon}{8V(R)}$$

de donde

$$|f(\hat{x}) - f(\hat{y})| < \frac{\varepsilon}{4V(R)}$$

y

$$M_j(f) - m_j(f) < f(\hat{x}) - f(\hat{y}) + \frac{\varepsilon}{4V(R)} < \frac{\varepsilon}{2V(R)}$$

de manera que

$$\sum_{j \in A_2} (M_j(f) - m_j(f)) V(Q_j) < \frac{\varepsilon}{2V(R)} \sum_{j \in A_2} V(Q_j) < \frac{\varepsilon}{2V(R)} V(R) = \frac{\varepsilon}{2}$$

concluyéndose así que $\overline{S}(f, P) - \underline{S}(f, P) < \varepsilon$, es decir, que f es Riemann integrable sobre R.

Comenzamos la demostración de este teorema con la condición de suficiencia (suficiencia en términos de la integrabilidad de la función) por ser más técnica que la condición de necesidad, en esta parte de la prueba hay más aproximaciones y cálculos que hacer. En cambio para probar la condición de necesidad simplemente vamos a valernos de la proposición 2.8 para demostrar que $\lambda(D_{f,R}) = 0$, es decir, si logramos ver a $D_{f,R}$ como una unión numerable de conjuntos de medida de Lebesgue cero entonces por la proposición 2.8 podremos concluir que $D_{f,R}$ tiene también medida de Lebesgue cero.

Supongamos ahora que f es Riemann integrable sobre R, y para cada $m \in \mathbb{N}$ definamos al conjunto

$$E_m = \left\{ \hat{\xi} \in R \mid \ \forall \delta > 0 \exists \hat{x} \in B_\delta\left(\hat{\xi}\right) \cap R \ \text{ tal que } \left| f(\hat{x}) - f\left(\hat{\xi}\right) \right| \geq \tfrac{1}{m} \ \right\}$$

Es claro que $E_m \subset D_{f,R}$ y por lo tanto $\bigcup_{m=1}^{\infty} E_m \subset D_{f,R}$. Por otra parte si $\hat{x} \in D_{f,R}$ entonces $\exists \varepsilon > 0$ tal que para cada $\delta > 0$ se puede encontrar un $\hat{y} \in B_\delta(\hat{x}) \cap R$ que cumpla que $|f(\hat{x}) - f(\hat{y})| \geq \varepsilon$, por lo que si tomamos $m \in \mathbb{N}$ con $\frac{1}{m} < \varepsilon$ entonces $x \in E_m$, es decir, $D_{f,R} \subset \bigcup_{m=1}^{\infty} E_m$ y por lo tanto $D_{f,R} = \bigcup_{m=1}^{\infty} E_m$. De acuerdo con lo que ya se dijo, si probamos que $\lambda(E_m) = 0$ para toda $m \in \mathbb{N}$ entonces $\lambda(D_{f,R}) = 0$.

Tomemos $m \in \mathbb{N}$ fija y sea $\varepsilon > 0$, por hipótesis existe $P \in \mathcal{P}_R$ tal que

$$\overline{S}(f,P) - \underline{S}(f,P) < \frac{\varepsilon}{2m}$$

Sean $R_1, \ldots, R_k$ los subrectángulos generados por P y definamos

$$A_1 = \{j \in \{1, \ldots, k\} \mid int(R_j) \cap E_m \neq \emptyset\}$$

y

$$A_2 = \{j \in \{1, \ldots, k\} \mid int(R_j) \cap E_m = \emptyset\}$$

Nuevamente por construcción se tiene que $\{1, \ldots, k\} = A_1 \cup A_2$ y $A_1 \cap A_2 = \emptyset$, de manera que

$$\overline{S}(f,P) - \underline{S}(f,P) = \sum_{j \in A_1} (M_j(f) - m_j(f)) V(R_j) + \sum_{j \in A_2} (M_j(f) - m_j(f)) V(R_j)$$

y al tratarse de sumandos no negativos se tiene que

$$\sum_{j \in A_1} (M_j(f) - m_j(f)) V(R_j) < \frac{\varepsilon}{2m}$$

Por definición para $j \in A_1$ podemos encontrar un $\hat{\xi} \in int(R_j) \cap E_m$, que por pertenecer al interior de R_j $\exists \delta > 0$ tal que $B_\delta\left(\hat{\xi}\right) \subset R_j$ y por pertenecer a E_m $\exists \hat{x} \in B_\delta\left(\hat{\xi}\right)$ tal que $\left| f(\hat{x}) - f\left(\hat{\xi}\right) \right| \geq \frac{1}{m}$. Pero

$$\left| f(\hat{x}) - f\left(\hat{\xi}\right) \right| \leq M_j(f) - m_j(f)$$

porque $\hat{x}, \hat{\xi} \in R_j$, entonces $M_j(f) - m_j(f) \geq \frac{1}{m}$ y esto para cada $j \in A_1$, de donde

$$\frac{1}{m} \sum_{j \in A_1} V(R_j) \leq \sum_{j \in A_1} (M_j(f) - m_j(f)) V(R_j) < \frac{\varepsilon}{2m}$$

es decir,

$$\sum_{j \in A_1} V(R_j) < \frac{\varepsilon}{2}$$

Cabe destacar que

$$E_m = \left[\bigcup_{j \in A_1} \left(E_m \cap int\left(R_j \right) \right) \right] \cup \left[\bigcup_{j=1}^{k} \left(E_m \cap Fr\left(R_j \right) \right) \right]$$

y lo que podemos decir hasta este momento es que el conjunto $\bigcup_{j \in A_1} \left(E_m \cap int\left(R_j \right) \right)$ está contenido en una unión finita de rectángulos cuya suma de volúmenes no excede $\frac{\varepsilon}{2}$.

Por otro lado, de la proposición 2.6 sabemos que $\lambda\left(Fr\left(R_j\right)\right) = 0$ para toda $j \in \{1, \ldots, k\}$ y como $E_m \cap Fr\left(R_j\right) \subset Fr\left(R_j\right)$ entonces de la proposición 2.4 concluimos que $\lambda\left(E_m \cap Fr\left(R_j\right)\right) = 0$ $\forall j \in \{1, \ldots, k\}$ y por la proposición 2.8

$$\lambda \left(\bigcup_{j=1}^{k} \left(E_m \cap Fr\left(R_j \right) \right) \right) = 0$$

De manera que existe una colección numerable de rectángulos $\{Q_i\}_{i=1}^{\infty}$ tales que

$$\bigcup_{j=1}^{k} \left(E_m \cap Fr\left(R_j \right) \right) \subset \bigcup_{i=1}^{\infty} Q_i$$

y

$$\sum_{i=1}^{\infty} V\left(Q_i\right) < \frac{\varepsilon}{2}$$

Es decir, que $\{R_j \mid j \in A_1\} \cup \{Q_i\}_{i=1}^{\infty}$ es una colección numerable de rectángulos tales que

$$E_m \subset \left(\bigcup_{j \in A_1} R_j \right) \cup \left(\bigcup_{i=1}^{\infty} Q_i \right)$$

y

$$\sum_{j \in A_1} V\left(R_j\right) + \sum_{i=1}^{\infty} V\left(Q_i\right) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

Lo que prueba que $\lambda\left(E_m\right) = 0$ y por lo tanto $\lambda\left(D_{f,R}\right) = 0$. ■

Observación 2.4 *Una inmediata consecuencia de este teorema es el hecho de que toda función continua sobre un rectángulo es Riemann integrable ahí.*

Ya que contamos con este poderoso teorema, podemos demostrar parte de la proposición 2.3, es decir, el teorema nos permite dar un argumento muy sencillo de porque si f y g son Riemann integrables sobre un rectángulo R entonces $f+g$ y fg también son Riemann integrables sobre R: Para ello simplemente nótese que si $\hat{x} \notin D_{f,R} \cup D_{g,R}$ entonces f y g son continuas en x, por lo que las funciones $f+g$ y fg son continuas en $\hat{x}$, es decir, $\hat{x} \notin D_{f+g,R}$ y $\hat{x} \notin D_{fg,R}$. Esto prueba que $D_{f+g,R} \subset D_{f,R} \cup D_{g,R}$ y $D_{fg,R} \subset D_{f,R} \cup D_{g,R}$.De manera que si f y g son integrables sobre R entonces por el teorema 2.2 $\lambda\left(D_{f,R}\right) = 0 = \lambda\left(D_{g,R}\right)$ y por la proposición 2.8 se concluye que $\lambda\left(D_{f,R} \cup D_{g,R}\right) = 0$, de donde $\lambda\left(D_{f+g,R}\right) = 0$ y $\lambda\left(D_{fg,R}\right) = 0$. Es decir, $f+g$ y fg son Riemann integrables sobre R.

Capítulo 3

Medida de Jordan

En este capítulo se hablará un poco de la medida de Jordan y su relación con los conjuntos de medida de Lebesgue cero así como de funciones definidas sobre conjuntos Jordan medibles y la integrabilidad de estas funciones. No olvidemos que en este trabajo se espera que el lector tenga ciertos conocimientos del cálculo integral en varias variables, razón por la cual se enunciarán sin demostrar algunos resultados y en ocasiones simplemente se hará uso de ellos.

3.1 Conjuntos Jordan medibles

Comencemos recordando que dado un conjunto $D \subset \mathbb{R}^n$ se define a la función característica de D, $\mathcal{X}_D : \mathbb{R}^n \to \mathbb{R}$ como

$$\mathcal{X}_D(\hat{y}) = \begin{cases} 1 & \hat{y} \in D \\ 0 & \hat{y} \notin D \end{cases}$$

Definición 3.1 *Sea $A \subset \mathbb{R}^n$ un conjunto acotado, diremos que A es Jordan medible si y sólo si la función característica de A, $\mathcal{X}_A : \mathbb{R}^n \to \mathbb{R}$, es Riemann integrable sobre R para algún rectángulo $R \subset \mathbb{R}^n$ tal que $A \subset R$, y se define la medida de Jordan de A como*

$$J(A) = \int_R \mathcal{X}_A$$

Como se puede ver en la definición anterior, es importante pedir que el conjunto al que llamamos Jordan medible sea acotado para poder hablar de la existencia de un rectángulo R que contenga a nuestro conjunto así como de la integrablidad de la función característica; pero por supuesto la parte fundamental de la definición es que la integrabilidad de la función característica no dependa del rectángulo que se tome, ya que si un conjunto está acotado entonces está contenido en una infinidad de rectángulos.

Uno puede demostrar que la medida de Jordan está bien definida a partir de la definición de integrabilidad, nosotros lo haremos apoyándonos en el teorema 2.2. Lo que se quiere demostrar es que si $A \subset \mathbb{R}^n$ es un conjunto acotado, $R \subset \mathbb{R}^n$ un rectángulo tal que $A \subset R$ y $\mathcal{X}_A : \mathbb{R}^n \to \mathbb{R}$ es Riemann integrable sobre R, entonces $\mathcal{X}_A$ es Riemann integrable sobre todo rectángulo $R' \subset \mathbb{R}^n$ que contenga al conjunto A y además

$$\int_R \mathcal{X}_A = \int_{R'} \mathcal{X}_A$$

Gracias al teorema 2.2, la primera parte se traduce en que si

$$\lambda\left(D_{\mathcal{X}_A,R}\right) = 0$$

para algún rectángulo $R \subset \mathbb{R}^n$ tal que $A \subset R$ entonces

$$\lambda\left(D_{\mathcal{X}_A,R'}\right) = 0$$

para todo rectángulo $R' \subset \mathbb{R}^n$ que contenga a A. Supongamos que $\lambda\left(D_{\mathcal{X}_A,R}\right) = 0$ para algún rectángulo $R \subset \mathbb{R}^n$ tal que $A \subset R$, sea $R' \subset \mathbb{R}^n$ un rectángulo tal que $A \subset R'$ entonces $A \subset (R \cap R') \subset R$ y por tanto

$$D_{\mathcal{X}_A,R\cap R'} \subset D_{\mathcal{X}_A,R}$$

Como

$$\lambda\left(D_{\mathcal{X}_A,R}\right) = 0$$

entonces por la proposición 2.4 se tiene que

$$\lambda\left(D_{\mathcal{X}_A,R\cap R'}\right) = 0$$

Ahora bien, demostremos que

$$D_{\mathcal{X}_A,R'} \subset D_{\mathcal{X}_A,R\cap R'} \cup Fr\left(R'\right) \cup Fr\left(R\right)$$

esto nos permitirá concluir que $\lambda\left(D_{\mathcal{X}_A,R'}\right) = 0$ puesto que ya sabemos que la frontera de todo rectángulo tiene medida de Lebesgue cero y acabamos de mostrar que $\lambda\left(D_{\mathcal{X}_A,R\cap R'}\right) = 0$.

Tomemos $\hat{x}_0 \in D_{\mathcal{X}_A,R'}$ y supongamos que $\hat{x}_0 \notin D_{\mathcal{X}_A,R\cap R'}$ y $\hat{x}_0 \notin Fr\left(R'\right)$; queremos que $\hat{x}_0 \in Fr\left(R\right)$.

Como $\hat{x}_0 \notin Fr\left(R'\right)$ y $\hat{x}_0 \in R'$ entonces $\exists \delta_0 > 0$ tal que $B_{\delta_0}\left(\hat{x}_0\right) \subset R'$. Observemos que para cualquier $\delta > 0$,con $\delta \leq \delta_0$,se tiene que $B_\delta\left(\hat{x}_0\right) \cap R \neq \emptyset$, de otra manera tendríamos que $B_\delta\left(\hat{x}_0\right) \cap A = \emptyset$ y entonces la función $\mathcal{X}_A$ restringida a R' se convierte en la constante cero en el abierto $B_\delta\left(\hat{x}_0\right)$ y por tanto sería continua en $\hat{x}_0$ contradiciendo nuestra elección de $\hat{x}_0$. De hecho, podemos asegurar que $\hat{x}_0 \in B_\delta\left(\hat{x}_0\right) \cap R$, pues al ser R un conjunto cerrado, si $\hat{x}_0 \notin R$ entonces existe $\delta > 0$ con $\delta \leq \delta_0$ tal que $B_\delta\left(\hat{x}_0\right) \cap R = \emptyset$ contrario a lo que acabamos de probar. De manera que $\hat{x}_0 \in R$ y por tanto $\hat{x}_0 \in R \cap R'$.

Como $\hat{x}_0 \in D_{\mathcal{X}_A,R'}$, entonces existe $\varepsilon_0 > 0$, tal que para toda $\delta > 0$ existe $\hat{x} \in B_\delta\left(\hat{x}_0\right) \cap R'$ y $\left|\mathcal{X}_A\left(\hat{x}\right) - \mathcal{X}_A\left(\hat{x}_0\right)\right| \geq \varepsilon_0$; pero como $\hat{x}_0 \notin D_{\mathcal{X}_A,R\cap R'}$ y $\hat{x}_0 \in R \cap R'$ entonces $\mathcal{X}_A$ restringida a $R \cap R'$ es continua en $\hat{x}_0$, es decir, existe $\delta_1 > 0$, con $\delta_1 \leq \delta_0$, tal que para toda $\hat{x} \in B_{\delta_1}\left(\hat{x}_0\right) \cap \left(R \cap R'\right)$ se tiene que $\left|\mathcal{X}_A\left(\hat{x}\right) - \mathcal{X}_A\left(\hat{x}_0\right)\right| < \varepsilon_0$.

De modo que dado $\delta > 0$, con $\delta \leq \delta_1$, tenemos que existe $\hat{y} \in B_\delta\left(\hat{x}_0\right) \cap R'$ tal que $\left|\mathcal{X}_A\left(\hat{y}\right) - \mathcal{X}_A\left(\hat{x}_0\right)\right| \geq \varepsilon_0$ y como

$$B_\delta\left(\hat{x}_0\right) \cap \left(R \cap R'\right) \subset B_{\delta_1}\left(\hat{x}_0\right) \cap \left(R \cap R'\right)$$

entonces $\hat{y} \notin R$, por lo tanto $B_\delta(\hat{x}_0) \cap R^c \neq \emptyset$, y como ya habíamos dicho que $\forall\delta > 0$ con $\delta \leq \delta_0$, se tiene que $B_\delta(\hat{x}_0) \cap R \neq \emptyset$ y $\delta \leq \delta_1 \leq \delta_0$ entonces hemos demostrado que para toda $\delta > 0$, con $\delta \leq \delta_1$, se cumple que $B_\delta(\hat{x}_0) \cap R \neq \emptyset$ y $B_\delta(\hat{x}_0) \cap R^c \neq \emptyset$, es decir, $\hat{x}_0 \in Fr(R)$, de modo que $\lambda(D_{\mathcal{X}_A,R'}) = 0$ como queríamos.

Para la segunda parte simplemente notemos que $R \cap R'$ es un rectángulo contenido tanto en R como R', de manera que si tomamos una partición $P \in \mathcal{P}_R$ que genere subrectángulos $R_1, \ldots, R_m, R \cap R'$ y tomamos también una partición $Q \in \mathcal{P}_{R'}$ que genere subrectángulos $R'_1, \ldots, R'_k, R \cap R'$, entonces por la proposición 2.3 se tiene que

$$\int_R \mathcal{X}_A = \sum_{j=1}^{m} \int_{R_j} \mathcal{X}_A + \int_{R\cap R'} \mathcal{X}_A$$

y

$$\int_{R'} \mathcal{X}_A = \sum_{j=1}^{k} \int_{R'_j} \mathcal{X}_A + \int_{R\cap R'} \mathcal{X}_A$$

Como $A \subset R \cap R'$, entonces $\mathcal{X}_A(\hat{x}) = 0$ para toda $\hat{x} \in R \setminus (R \cap R')$ y para toda $\hat{x} \in R' \setminus (R \cap R')$, de donde $\int_{R_j} \mathcal{X}_A = 0$ para toda $j \in \{1, \ldots, m\}$ y $\int_{R'_j} \mathcal{X}_A = 0$ para toda $j \in \{1, \ldots, k\}$, es decir, que

$$\int_R \mathcal{X}_A = \int_{R\cap R'} \mathcal{X}_A = \int_{R'} \mathcal{X}_A$$

En el ejemplo 2.1 vimos que la función constante 1 era integrable sobre cualquier rectángulo R y $V(R) = \int_R 1$ y como para cualquier rectángulo R, $\mathcal{X}_R$ es la constante 1 en R entonces por definición se tiene que todo rectángulo es Jordan medible y además

$$J(R) = \int_R \mathcal{X}_R = \int_R 1 = V(R)$$

Con esto lo que estamos diciendo es que la medida de Jordan generaliza el concepto de volumen de un rectángulo llevándolo a conjuntos acotados. Ahora, hay que recordar que lamentablemente no todo conjunto acotado es Jordan medible.

Ejemplo 3.1 *Sean $R = [0,1] \times [0,1]$ y $A = ([0,1] \times [0,1]) \cap (\mathbb{Q} \times \mathbb{Q})$. Por la densidad de los racionales y los irracionales se tiene que $\mathcal{X}_A$ es discontinua en todo R y sabemos que un rectángulo no tiene medida de Lebesgue cero (proposición 2.7), entonces por el teorema 2.2 $\mathcal{X}_A$ no es integrable en R y por lo tanto A no es Jordan medible a pesar de ser un conjunto acotado. Sin embargo cabe destacar que este conjunto tiene medida de Lebesgue cero por ser un conjunto numerable.*

Si un conjunto $A \subset \mathbb{R}^n$ es Jordan medible entonces, por definición de medida de Jordan, para todo rectángulo $R \subset \mathbb{R}^n$ que contenga a A se tiene que

$$\sup\left\{\underline{S}(\mathcal{X}_A, P) \mid P \in \mathcal{P}_R\right\} = \inf\left\{\bar{S}(\mathcal{X}_A, P) \mid P \in \mathcal{P}_R\right\}$$

Sin embargo sabemos que en general estos valores no son iguales; se puede ver en el caso de nuestro ejemplo anterior que

$$\sup\left\{\underline{S}\left(\mathcal{X}_A, P\right) \mid P \in \mathcal{P}_R\right\} = 0$$

mientras que

$$\inf\left\{\bar{S}\left(\mathcal{X}_A, P\right) \mid P \in \mathcal{P}_R\right\} = 1$$

En el contexto de la medida de Jordan llamaremos

$$\bar{J}\left(A\right) = \inf\left\{\bar{S}\left(\mathcal{X}_A, P\right) \mid P \in \mathcal{P}_R\right\}$$

la medida exterior de A y

$$\underline{J}\left(A\right) = \sup\left\{\underline{S}\left(\mathcal{X}_A, P\right) \mid P \in \mathcal{P}_R\right\}$$

la medida interior de A. A diferencia de la medida de Jordan, la medida exterior y la medida interior de un conjunto acotado siempre existe porque la función característica es acotada, y queda claro que un conjunto es Jordan medible si y sólo si $\bar{J}\left(A\right) = \underline{J}\left(A\right)$. Aunque vale la pena recordar el concepto de medida interior, la realidad es que no nos ocuparemos de ella, no nos hace falta para los objetivos que se tienen planteados; sin embargo la medida exterior sí nos será de gran utilidad a lo largo de este trabajo.

Gracias a la sencillez de la función característica de un conjunto, se demuestran algunas propiedades básicas de la medida de Jordan, y aunque la función característica parece muy simple, posee propiedades bastante interesantes, una de éstas es que $D_{\mathcal{X}_A,\mathbb{R}^n} = Fr\left(A\right)$, y si en particular A es un conjunto acotado y $R \subset \mathbb{R}^n$ un rectángulo tal que $\overline{A} \subset int\left(R\right)$ entonces $D_{\mathcal{X}_A,R} = Fr\left(A\right)$; la prueba de ambas relaciones es análoga y vale la pena enunciarla en una pequeña proposición.

Proposición 3.1 *Sea $A \subset \mathbb{R}^n$ acotado. Si $R \subset \mathbb{R}^n$ es un rectángulo tal que $\overline{A} \subset int\left(R\right)$ entonces $D_{\mathcal{X}_A,R} = Fr\left(A\right)$.*

Dem. El argumento para demostrar que $D_{\mathcal{X}_A,R} \subset Fr\left(A\right)$ es similar a lo mencionado unos párrafos atrás, pues dado $\hat{x} \in R$ si $\hat{x} \notin Fr\left(A\right)$ entonces habría $\delta > 0$ tal que $B_\delta\left(\hat{x}\right) \subset A$ o $B_\delta\left(\hat{x}\right) \subset A^c$, pero esto implicaría la continuidad de $\mathcal{X}_A$ en $\hat{x}$ porque sería constante en el abierto $B_\delta\left(\hat{x}\right)$; por lo tanto si $\hat{x} \in D_{\mathcal{X}_A,R}$ entonces $\hat{x} \in Fr\left(A\right)$.

Para la contención contraria, tomamos $\hat{x} \in Fr\left(A\right)$ y consideremos el valor $\varepsilon = \frac{1}{2}$; como $\overline{A} \subset int\left(R\right)$ entonces existe $\delta_0 > 0$ tal que $B_{\delta_0}\left(\hat{x}\right) \subset R$ y por tanto para toda $\delta > 0$ tal que $\delta \leq \delta_0$, se cumple que $B_\delta\left(\hat{x}\right) \subset R$. De modo que dado $\delta > 0$, con $\delta \leq \delta_0$, como $\hat{x} \in Fr\left(A\right)$ se tiene que $B_\delta\left(\hat{x}\right) \cap A \neq \emptyset$ y $B_\delta\left(\hat{x}\right) \cap A^c \neq \emptyset$; de manera que si $\hat{x} \in A$ podemos tomar $\hat{y} \in B_\delta\left(\hat{x}\right) \cap A^c$ y entonces

$$\left|\mathcal{X}_A\left(\hat{x}\right) - \mathcal{X}_A\left(\hat{y}\right)\right| = \left|1 - 0\right| = 1 > \varepsilon$$

y si $x \notin A$ podemos tomar $\hat{y} \in B_\delta\left(\hat{x}\right) \cap A$ y entonces

$$\left|\mathcal{X}_A\left(\hat{x}\right) - \mathcal{X}_A\left(\hat{y}\right)\right| = \left|0 - 1\right| = 1 > \varepsilon$$

En ambos casos el punto $\hat{y} \in B_\delta(\hat{x})$ vive en R, puesto que $B_\delta(\hat{x}) \subset R$; lo que prueba que $\hat{x} \in D_{\mathcal{X}_A, R}$.

$\blacksquare$

Observación 3.1 *Como consecuencia de esta propiedad y el teorema 2.2 se tiene que un conjunto acotado es Jordan medible si y sólo si su frontera tiene medida de Lebesgue cero.*

En esta observación, toma mucho valor (en la parte de la suficiencia) el hecho de que la definición de conjunto Jordan medible no dependa del rectángulo que se tome para integrar a la función característica del conjunto; ya que a la cerradura de todo conjunto acotado lo podemos meter en el interior de un rectángulo.

La observación anterior muestra algo muy interesante; sabemos que un conjunto acotado es Jordan medible si y sólo si su frontera tiene medida de Jordan cero (demostraremos este criterio un poco más adelante), de manera que gracias a lo dicho en la observación, cuando trabajamos con la frontera de un conjunto acotado, es equivalente que ésta tenga medida de Lebesgue cero a que tenga medida de Jordan cero; y es interesante porque no todo conjunto con medida de Lebesgue cero es Jordan medible, como nos lo demuestra el ejemplo 3.1. Ahora bien, daremos una condición suficiente bajo la cual un conjunto de medida de Lebesgue cero es Jordan medible, condición que por supuesto satisface la frontera de todo conjunto acotado; y para ello a continuación veremos un resultado que deja ver lo cercanos que son los conceptos de medida de Jordan cero con medida de Lebesgue cero.

Lema 3.1 *Sea $A \subset \mathbb{R}^n$ un conjunto acotado, A tiene medida de Jordan cero si y sólo si para toda $\varepsilon > 0$ hay una cantidad finita de rectángulos $R_1, \ldots, R_k$ tales que*

$$A \subset \bigcup_{j=1}^{k} R_j$$

y

$$\sum_{j=1}^{k} V(R_j) < \varepsilon$$

Dem. Si $J(A) = 0$ entonces dado $\varepsilon > 0$ existe una partición $P \in \mathcal{P}_R$, donde $R \subset \mathbb{R}^n$ es un rectángulo que contiene a A, tal que $\bar{S}(\mathcal{X}_A, P) < \varepsilon$. Sean $R_1, \ldots, R_m$ los rectángulos inducidos por P y nótese que

$$\bar{S}(\mathcal{X}_A, P) = \sum_{j=1}^{m} M_j(\mathcal{X}_A) V(R_j) = \sum_{R_j \cap A \neq \emptyset} V(R_j)$$

(con frecuencia estaremos utilizando esta identidad). En efecto, para toda $\hat{x} \in R_j \cap A$ se tiene que $\mathcal{X}_A(\hat{x}) = 1$ y para toda $\hat{x} \in R_j \setminus A$ ocurre que $\mathcal{X}_A(\hat{x}) = 0$, de modo que $M_j(\mathcal{X}_A) = 1$ si $R_j \cap A \neq \emptyset$ y $M_j(\mathcal{X}_A) = 0$ si $R_j \cap A = \emptyset$, por lo tanto si $\{j \in \{1, \ldots, m\} \mid R_j \cap A \neq \emptyset\} = \{k_1, \ldots, k_s\}$ entonces

$$A \subset \bigcup_{j=1}^{s} R_{k_j}$$

y

$$\sum_{j=1}^{s} V\left(R_{k_j}\right) = \bar{S}\left(\mathcal{X}_A, P\right) < \varepsilon$$

Recíprocamente si $R \subset \mathbb{R}^n$ es un rectángulo tal que $A \subset int\left(R\right)$ y $\varepsilon > 0$, entonces por hipótesis hay una cantidad finita de rectángulos $R_1, \ldots, R_k$ tales que

$$A \subset \bigcup_{j=1}^{k} R_j$$

y

$$\sum_{j=1}^{k} V\left(R_j\right) < \frac{\varepsilon}{2}$$

Como $R \cap R_j$ es un rectángulo tal que $V\left(R \cap R_j\right) \leq V\left(R_j\right)$ y $A \subset \bigcup_{j=1}^{k} R \cap R_j$, entonces podemos suponer sin pérdida de generalidad que $R_j \subset R$ para toda $j \in \{1, \ldots, k\}$. Por una cuestión totalmente técnica nos gustaría que $A \subset \bigcup_{j=1}^{k} int\left(R_j\right)$, pero en realidad esto es algo que por el momento no podemos asegurar. Afortunadamente por el lema 2.1 esto no va a ser un problema, de manera más precisa, por el lema 2.1 sabemos que para cada $j \in \{1, \ldots, k\}$ existe un rectángulo $R'_j \subset \mathbb{R}^n$ tal que $R_j \subset int\left(R'_j\right)$ y

$$V\left(R'_j\right) - V\left(R_j\right) < \frac{\varepsilon}{2k}$$

de donde $A \subset \bigcup_{j=1}^{k} int\left(R'_j\right)$ y

$$\sum_{j=1}^{k} V\left(R'_j\right) < \sum_{j=1}^{k} \left[V\left(R_j\right) + \frac{\varepsilon}{2k}\right] < \varepsilon$$

Como $A \subset int\left(R\right)$ entonces

$$A \subset \bigcup_{j=1}^{k} \left[int\left(R\right) \cap int\left(R'_j\right)\right] = \bigcup_{j=1}^{k} int\left(R \cap R'_j\right)$$

Con lo que nuevamente sin pérdida de generalidad podemos suponer que $R'_j \subset R$ para toda $j \in \{1, \ldots, k\}$. Sea $P \in \mathcal{P}_R$ una partición que contenga todos los vértices de cada R'_j, esto es posible porque $R'_j \subset R \; \forall j \in \{1, \ldots, k\}$; cabe recordar nuevamente que de la discusión hecha tras la definición 2.7 si Q es un subrectángulo generado por P entonces $Q \subset R'_j$ para alguna $j \in \{1, \ldots, k\}$ o $int\left(Q\right) \cap int\left(R'_j\right) = \emptyset$ para toda $j \in \{1, \ldots, k\}$. Ahora bien, nos era importante asegurar que $A \subset \bigcup_{j=1}^{k} int\left(R'_j\right)$ porque de esta manera si Q es un subrectángulo generado por P tal que $Q \cap A \neq \emptyset$ entonces $Q \subset R'_j$ para alguna $j \in \{1, \ldots, k\}$. Por lo tanto si $Q_1, \ldots, Q_m$ son los rectángulos inducidos por P entonces

$$\bar{S}\left(\mathcal{X}_A, P\right) = \sum_{Q_j \cap A \neq \emptyset} V\left(R_j\right) \leq \sum_{j=1}^{k} V\left(R'_j\right) < \varepsilon$$

de donde $\overline{J}(A) = 0$ y como $0 \leq \underline{J}(A) \leq \overline{J}(A)$ concluímos que $\underline{J}(A) = \overline{J}(A)$ y por lo tanto A es Jordan medible y tiene medida de Jordan cero. ■

Aunque este resultado ya nos era familiar, valía la pena demostrarlo por el razonamiento empleado en la parte de la necesidad. A modo de preparación para el teorema del cambio de variable, sería bueno tener presente este truco de poder "inflar" rectángulos de manera que sus volúmenes no se nos escapen de las manos y que nos permita asegurar que determinado conjunto se queda contenido en el interior de estos rectángulos inflados, con la única intención de que al tomar un refinamiento de una partición generada a partir de dichos rectángulos, los nuevos rectángulos que cubran a nuestro conjunto no generen más volumen del que ya se tenía.

Observación 3.2 *Gracias a esta caracterización de los conjuntos de medida de Jordan cero podemos concluir que la unión finita de conjuntos de medida de Jordan cero también es de medida de Jordan cero, así como también podemos concluir que todo subconjunto de un conjunto de medida de Jordan cero tiene medida de Jordan cero.*

Retomando la discusión previa al lema 3.1, de éste mismo y de la observación 2.3 se concluye que todo conjunto de medida de Jordan cero tiene medida de Lebesgue cero. Es claro que, si un conjunto se puede cubrir con un número finito de rectángulos cuya suma de volúmenes es pequeña, entonces este mismo conjunto se puede cubrir por una cantidad numerable de rectángulos cuya suma de volúmenes también es pequeña. Sin embargo, lo recíproco no es necesariamente cierto. En este sentido, es más restrictivo pedir que un conjunto tenga medida de Jordan cero que pedir que tenga medida de Lebesgue cero. Para concluir esta discusión regresemos al ejemplo 3.1, en él se demuestra que no todo conjunto de medida de Lebesgue cero es Jordan medible y ahora con el lema 3.1 podemos agregar que éste es un ejemplo de un conjunto acotado de medida de Lebesgue cero que no puede ser cubierto por una cantidad finita de rectángulos cuya suma de volúmenes sea tan pequeña como queramos. Sin embargo el lema 3.1 también deja ver claramente cuáles son las condiciones suficientes para que un conjunto con medida de Lebesgue cero sea Jordan medible y que además tenga medida de Jordan cero.

Proposición 3.2 *Sea $A \subset \mathbb{R}^n$ un conjunto compacto. Si A tiene medida de Lebesgue cero entonces A tiene medida de Jordan cero.*

Dem. Por supuesto que la idea es aplicar el lema anterior para concluir que A tiene medida de Jordan cero. Sea $\varepsilon > 0$, por la proposición 2.9 existe una colección numerable de rectángulos abiertos $\{R_j\}_{j=1}^{\infty}$ tales que $A \subset \bigcup_{j=1}^{\infty} R_j$ y $\sum_{j=1}^{\infty} V(R_j) < \varepsilon$, pero entonces la colección de rectángulos es una cubierta abierta para A y éste es compacto. Por lo tanto existen $R_{k_1}, \ldots, R_{k_m}$ en $\{R_j\}_{j=1}^{\infty}$ tales que $A \subset \bigcup_{j=1}^{m} R_{k_j} \subset \bigcup_{j=1}^{m} \overline{R_{k_j}}$ y como

$$\sum_{j=1}^{m} V\left(\overline{R_{k_j}}\right) = \sum_{j=1}^{m} V\left(R_{k_j}\right) \leq \sum_{j=1}^{\infty} V(R_j) < \varepsilon$$

Entonces por el lema anterior A tiene medida de Jordan cero. ■

Con base en los resultados anteriores, ahora estamos en condiciones de enunciar el siguiente teorema, el cual nos da condiciones necesarias y suficientes para que un conjunto sea Jordan medible.

Teorema 3.1 *Sea $A \subset \mathbb{R}^n$ un conjunto acotado. A es Jordan medible si y sólo si $Fr(A)$ es Jordan medible y $J(Fr(A)) = 0$.*

Dem. Supongamos que A es Jordan medible, por la observación 3.1 sabemos que $Fr(A)$ tiene medida de Lebesgue cero, y como la frontera de todo conjunto acotado es compacta entonces por la proposición 3.2 $Fr(A)$ es Jordan medible y $J(Fr(A)) = 0$.

De manera recíproca si $J(Fr(A)) = 0$, entonces $\lambda(Fr(A)) = 0$. Ahora bien, si $R \subset \mathbb{R}^n$ es un rectángulo tal que $\overline{A} \subset int(R)$, entonces por la proposición 3.1, sabemos que $D_{\mathcal{X}_A,R} = Fr(A)$, de manera que $\lambda(D_{\mathcal{X}_A,R}) = 0$ y esto implica, por el teorema 2.2, que A es Jordan medible. ■

Como ya se dijo, el lema 3.1 deja ver lo parecido que son los conjuntos de medida de Jordan cero con los conjuntos de medida de Lebesgue cero, concretamente sabemos que todo conjunto de medida de Jordan cero es de medida de Lebesgue cero, partiendo de esto y de la caracterización dada en el lema 3.1 uno podría preguntarse si se pueden caracterizar a las funciones Riemann integrables con los conjuntos de medida de Jordan cero tal y como lo hicimos en el teorema 2.2, y la respuesta es no. Si bien es cierto que toda función definida en un rectángulo, cuyo conjunto de discontinuidades forme un conjunto de medida de Jordan cero, es Riemann integrable sobre este rectángulo, el recíproco no es cierto. Para ver que esto es así, consideremos el siguiente ejemplo.

Ejemplo 3.2 *Sean $R = [0,1] \times [0,1]$ y*

$$A = \{(x,y) \in (0,1) \times (0,1) \mid x \in \mathbb{Q} \text{ ó } y \in \mathbb{Q}\} \cup (Fr(R) \cap \mathbb{Q} \times \mathbb{Q}) \setminus \{(0,0),(0,1),(1,0),(1,1)\}$$

Sea $f : R \subset \mathbb{R}^2 \to \mathbb{R}$ definida como

$$f(x,y) = \begin{cases} \frac{1}{n} + \frac{1}{q} & x = \frac{m}{n},\ y = \frac{p}{q} \\ 0 & \text{en otro caso} \end{cases}$$

en donde $m < n$, con n y m primos relativos, y lo mismo para p y q. Veremos que $D_{f,R} = A$, y que A no es Jordan medible pero tiene medida de Lebesgue cero, con lo que concluiremos que f es Riemann integrable a pesar de que su conjunto de discontinuidades no es Jordan medible.

Primero notemos que

$$R \setminus A = \{(x,y) \in [0,1] \times [0,1] \mid x \in (\mathbb{R} \setminus \mathbb{Q}) \cup \{0,1\} \text{ y } y \in (\mathbb{R} \setminus \mathbb{Q}) \cup \{0,1\}\}$$

Como este conjunto es denso en R y A también lo es, se tiene que $\mathcal{X}_A$ es discontinua en todo R y por lo tanto A no es Jordan medible.

Ahora veamos que A tiene medida de Lebesgue cero; para esto, aprovechemos que $(0,1) \cap \mathbb{Q}$ es un conjunto numerable, es decir, $(0,1) \cap \mathbb{Q} = \{q_n \mid n \in \mathbb{N}\}$. Para cada $n \in \mathbb{N}$ definamos a los conjuntos $A_n = \{(q_n, y) \in [0,1] \times [0,1] \mid y \in \mathbb{R}\}$ y $B_n = \{(x, q_n) \in [0,1] \times [0,1] \mid x \in \mathbb{R}\}$, claramente

$$A = \bigcup_{n=1}^{\infty} A_n \cup \bigcup_{n=1}^{\infty} B_n$$

de manera que si probamos que cada A_n y cada B_n tiene medida de Lebesgue cero entonces A tendrá medida de Lebesgue cero.

Dado $\varepsilon > 0$, definamos al rectángulo $R_n = \left[q_n - \frac{\varepsilon}{4}, q_n + \frac{\varepsilon}{4}\right] \times [0,1]$, R_n contiene a A_n y $V(R_n) = \frac{\varepsilon}{2} < \varepsilon$, lo que prueba que $\lambda(A_n) = 0$; de manera análoga se tiene que $\lambda(B_n) = 0$, por lo tanto $\lambda(A) = 0$.

Nos resta probar que $D_{f,R} = A$. Sea $(x_0, y_0) \in A$ y supongamos, para empezar, que $x_0 = \frac{m}{n}$ y $y_0 = \frac{p}{q}$ con n y m primos relativos así como p y q, sea $\varepsilon = \frac{1}{2} f(x_0, y_0)$, dado $\delta > 0$ existe $x \in [0,1] \cap (\mathbb{R} \setminus \mathbb{Q})$tal que si $|x - x_0| < \delta$, entonces $(x, y_0) \in R$ es tal que

$$\|(x, y_0) - (x_0, y_0)\| = |x - x_0| < \delta$$

Pero

$$|f(x, y_0) - f(x_0, y_0)| = f(x_0, y_0) > \varepsilon$$

es decir, f no es continua en (x_0, y_0).

Supongamos ahora, sin pérdida de generalidad, que $x_0 \in (\mathbb{R} \setminus \mathbb{Q}) \cup \{0,1\}$ y $y_0 \in \mathbb{Q} \setminus \{0,1\}$, con $y_0 = \frac{p}{q}$. Sea $\{a_n\}_{n=1}^{\infty} \subset [0,1] \cap \mathbb{Q}$ una sucesión que converja a x_0, y consideremos a la sucesión $\{(a_n, y_0)\}_{n=1}^{\infty} \subset R \cap (\mathbb{Q} \times \mathbb{Q})$, esta sucesión converge a (x_0, y_0). Si f fuera continua en (x_0, y_0) entonces la sucesión de imágenes, $\{f(a_n, y_0)\}_{n=1}^{\infty}$, debería converger al valor $f(x_0, y_0) = 0$, sin embargo $f(a_n, y_0) \geq \frac{1}{q}$ para toda $n \in \mathbb{N}$ y por tanto no puede converger a cero, de modo que f no es continua en (x_0, y_0); con lo que concluimos que $A \subset D_{f,R}$.

Para la contención contraria tomemos $(x_0, y_0) \in R \setminus A$ y veamos que f es continua en este punto. Sea $\varepsilon > 0$ y supongamos que $\varepsilon < 1$, sea $n \in \mathbb{N}$ tal que

$$\frac{1}{n} < \frac{\varepsilon}{2} \leq \frac{1}{n-1}$$

n existe porque $\frac{\varepsilon}{2} < \frac{1}{2}$. Para toda $m < n$,

$$\frac{\varepsilon}{2} \leq \frac{1}{n-1} \leq \frac{1}{m}$$

mientras que para toda $m > n$

$$\frac{1}{m} < \frac{1}{n} < \frac{\varepsilon}{2}$$

de manera que sólo hay un cantidad finita de números de la forma $\frac{p}{q}$, con $p < q$ primos relativos tales que $\frac{\varepsilon}{2} \leq \frac{1}{q}$, entonces

$$\delta_1 = \min\left\{\left|\frac{p}{q} - x_0\right| \mid \frac{1}{q} \geq \frac{\varepsilon}{2}, \frac{p}{q} \neq 1, (p,q) = 1\right\}$$

y

$$\delta_2 = \min\left\{\left|\frac{p}{q} - y_0\right| \mid \frac{1}{q} \geq \frac{\varepsilon}{2}, \frac{p}{q} \neq 1, (p,q) = 1\right\}$$

son dos valores bien definidos, donde $(p,q) = 1$ representa que p y q son primos relativos. Además $\delta_i > 0$, $i = 1, 2$, porque $x_0, y_0 \in \mathbb{R}\backslash\mathbb{Q}\cup\{0,1\}$ y los números $\frac{p}{q}$ son racionales distintos de cero y uno. Sean $\delta = \min\{\delta_1, \delta_2\}$ y $(x,y) \in R \cap B_\delta(x_0, y_0)$, si $x \in (\mathbb{R} \setminus \mathbb{Q}) \cup \{0,1\}$ ó $y \in (\mathbb{R} \setminus \mathbb{Q}) \cup \{0,1\}$ entonces

$$|f(x,y) - f(x_0,y_0)| = 0 < \varepsilon$$

mientras que si $(x,y) = \left(\frac{r}{m}, \frac{p}{q}\right)$ entonces

$$\left|\frac{r}{m} - x_0\right| < \delta \leq \delta_1$$

y

$$\left|\frac{p}{q} - y_0\right| < \delta \leq \delta_2$$

entonces $\frac{1}{m} < \frac{\varepsilon}{2}$ y $\frac{1}{q} < \frac{\varepsilon}{2}$, de donde

$$|f(x,y) - f(x_0,y_0)| = \frac{1}{m} + \frac{1}{q} < \varepsilon$$

es decir, f es continua en (x_0, y_0); lo que implica que $D_{f,R} \subset A$, es decir, $D_{f,R} = A$.

Este ejemplo demuestra que existen funciones Riemann integrables cuyo conjunto de discontinuidades no es Jordan medible, sin embargo, si sabemos que el conjunto de discontinuidades de una función es Jordan medible, entonces podemos formular el siguiente teorema.

Teorema 3.2 *Sea $f : R \subset \mathbb{R}^n \to \mathbb{R}$ acotada tal que $D_{f,R}$ es Jordan medible. f es integrable sobre R si y sólo si $J(D_{f,R}) = 0$.*

Dem. Primero supongamos que f es integrable sobre R. Bajo esta hipótesis, $D_{f,R}$ debe tener interior vacío. De lo contrario existe $R' \subset \mathbb{R}^n$ rectángulo no degenerado tal que $R' \subset D_{f,R}$; como f es integrable sobre R entonces $D_{f,R}$ tiene medida de Lebesgue cero y por lo tanto también R' tiene medida de Lebesgue cero, pero esto contradice el hecho de que todo rectángulo no degenerado no tiene medida de Lebesgue cero. De manera que $int(D_{f,R}) = \emptyset$, esto implica que $D_{f,R} \subset Fr(D_{f,R})$ y como $D_{f,R}$ es Jordan medible entonces $J(Fr(D_{f,R})) = 0$ y por lo tanto $J(D_{f,R}) = 0$.

La implicación contraria es inmediata del teorema 2.2. ■

Ya que nos hemos familiarizado un poco más con la medida de Jordan y tenemos buenos criterios para saber cuándo un conjunto es Jordan medible, nos resta enunciar algunas propiedades básicas que tienen dichos conjuntos. Como se dijo al principio de este capítulo estamos asumiendo que ya se conoce a la medida de Jordan y sus propiedades básicas, de manera que sólo enunciaremos dichas propiedades, siendo la más natural que la unión e intersección finita de conjuntos Jordan medibles son Jordan medibles, y cuya prueba, por cierto, se reduce a un simple hecho de topología de $\mathbb{R}^n$.

Teorema 3.3 *Sean $A, B \subset \mathbb{R}^n$ acotados y Jordan medibles entonces $A \cup B$ y $A \cap B$ son Jordan medibles y*
$$J(A \cup B) = J(A) + J(B) - J(A \cap B)$$

Corolario 3.1 *Sean $A, B \subset \mathbb{R}^n$ acotados y Jordan medibles, si $int(A) \cap int(B) = \emptyset$ entonces $J(A \cup B) = J(A) + J(B)$*

Corolario 3.2 *Sean $A, B \subset \mathbb{R}^n$ acotados y Jordan medibles, entonces $A \setminus B$ es Jordan medible y $J(A \setminus B) = J(A) - J(A \cap B)$.*

La lista de propiedades podría seguir y seguir, pero vamos a enunciar las que nos son indispensables para nuestra labor, y a las que estaremos recurriendo frecuentemente.

Proposición 3.3 *Sean $A, B \subset \mathbb{R}^n$ acotados y Jordan medibles,*

i) Si $B \subset A$ entonces $J(A \setminus B) = J(A) - J(B)$.

ii) Si $B \subset A$ entonces $J(B) \leq J(A)$.

iii) $\overline{A}$ e $int(A)$ son Jordan medibles y $J(A) = J(\overline{A}) = J(int(A))$.

La prueba de esto es una consecuencia fácil del teorema 3.3 y los corolarios 3.1 y 3.2.

Vale la pena mencionar también que la medida exterior de un conjunto satisface una condición análoga al inciso *ii*) de esta proposición, partiendo del hecho de que si $B \subset A$ entonces $\mathcal{X}_B \leq \mathcal{X}_A$.

Hay otra propiedad más, a la que sí le daremos prueba, que nos dice que todo conjunto Jordan medible puede aproximarse tanto como se quiera por un conjunto compacto Jordan medible. Esta propiedad va encaminada hacia el teorema de cambio de variable.

Teorema 3.4 *Sea $A \subset \mathbb{R}^n$ un conjunto Jordan medible; entonces para toda $\varepsilon > 0$ existe $K \subset A$ compacto Jordan medible, tal que $J(A) - J(K) < \varepsilon$.*

Dem. Sea $R \subset \mathbb{R}^n$ rectángulo tal que $A \subset R$. Dado $\varepsilon > 0$, por hipótesis existe $P \in \mathcal{P}_R$ tal que
$$J(A) - \underline{S}(\mathcal{X}_A, P) < \varepsilon$$

Pero
$$\underline{S}(\mathcal{X}_A, P) = \sum_{R_j \subset A} V(R_j)$$

donde los rectángulos R_j son generados por la partición P; definimos entonces
$$K = \bigcup_{R_j \subset A} R_j$$

K es compacto por ser unión finita de rectángulos, y por la misma razón es Jordan medible. Además como los rectángulos R_j forman una partición de R entonces $int(R_j) \cap int(R_i) = \emptyset$ para $i \neq j$, de modo que
$$J(K) = \sum_{R_j \subset A} V(R_j) = \underline{S}(\mathcal{X}_A, P)$$

y por lo tanto K es un compacto tal que $K \subset A$ y $J(A) - J(K) < \varepsilon$. ■

Observación 3.3 *El teorema anterior lo podemos refinar un poco más, es decir, si tenemos un conjunto Jordan medible A y $\varepsilon > 0$, entonces podemos conseguir un conjunto compacto $K \subset int(A)$ y Jordan medible tal que $J(A) - J(K) < \varepsilon$, ya que por la proposición 3.3 $int(A)$ es Jordan medible y $J(A) = J(int(A))$, entonces aplicando el teorema 3.4 al $int(A)$, existe $K \subset int(A)$ compacto Jordan medible, tal que $J(A) - J(K) = J(int(A)) - J(K) < \varepsilon$.*

3.2 Integración sobre conjuntos Jordan medibles

Para finalizar este capítulo, hablaremos de integrales sobre conjuntos Jordan medibles; el lector recordará que podemos extender la integral sobre rectángulos a conjuntos más complicados, es decir, si $A \subset \mathbb{R}^n$ es un conjunto acotado y $f : A \subset \mathbb{R}^n \to \mathbb{R}$ es acotada entonces podemos preguntarnos ¿qué significa que f sea integrable sobre A? La clave estaba en poder extender f a todo $\mathbb{R}^n$ de forma que resultara integrable en $R \subset \mathbb{R}^n$ rectángulo que contuviera a A, y la manera más natural de hacerlo era mandando a cero a todo punto que no estuviese en A, es decir, se definía a la función característica de f en A, $f_A : \mathbb{R}^n \to \mathbb{R}$ como

$$f_A(\hat{x}) = \begin{cases} f(\hat{x}) & \hat{x} \in A \\ 0 & \hat{x} \notin A \end{cases}$$

Si esta función resultaba integrable en R decíamos que f era integrable sobre A y que $\int_A f = \int_R f_A$. Así como se demuestra que la medida de Jordan no depende del rectángulo que se tome, también se prueba que cuando hablamos de la integral de f sobre un conjunto acotado arbitrario no importa el rectángulo que se tome.

El problema ahora es ¿cómo saber cuando f_A es integrable? Definitivamente f_A no siempre resultará integrable (ejemplo 3.1 y tomando a f como la constante 1). Notemos que, como $f_A(\hat{x}) = 0$ para toda $\hat{x} \in \left(\overline{A}\right)^c$ y este es un conjunto abierto entonces f_A es continua en este conjunto, de manera que si f_A es discontinua en algún punto $\hat{x} \in \mathbb{R}^n$ entonces $\hat{x} \in D_{f,A}$ o $\hat{x} \in Fr(A)$, es decir, $D_{f_A,\mathbb{R}^n} \subset (D_{f,A} \cup Fr(A))$ y por lo tanto para todo rectángulo $R \subset \mathbb{R}^n$ que contenga a A, $D_{f_A,R} \subset (D_{f,A} \cup Fr(A))$; esto ya deja ver que es lo que necesitamos, porque sabemos que si $D_{f_A,R}$ tiene medida de Lebesgue cero entonces f_A es Riemann integrable sobre R, por lo que si podemos asegurar que $D_{f,A} \cup Fr(A)$ es un conjunto de medida de Lebesgue cero entonces f_A será Riemann integrable sobre R. Ahora bien, en el caso de los conjuntos Jordan medibles sabemos que su frontera tiene medida de Jordan cero y por tanto medida de Lebesgue cero, de manera que si A es Jordan medible y $f : A \subset \mathbb{R}^n \to \mathbb{R}$ una función acotada cuyo conjunto de discontinuidades tiene medida de Lebesgue cero entonces f será integrable sobre A, formulemos todo esto con un poco más de rigor.

Definición 3.2 *Si $A \subset \mathbb{R}^n$ es un conjunto acotado y $f : A \subset \mathbb{R}^n \to \mathbb{R}$ es acotada en A decimos que f es integrable sobre A si y sólo si la función característica de f en A, $f_A : \mathbb{R}^n \to \mathbb{R}$ dada por*

$$f_A(\hat{x}) = \begin{cases} f(\hat{x}) & \hat{x} \in A \\ 0 & \hat{x} \notin A \end{cases}$$

es integrable sobre $R \subset \mathbb{R}^n$ rectángulo que contiene A y definimos la integral de f sobre A como

$$\int_A f = \int_R f_A$$

Teorema 3.5 *Si $A \subset \mathbb{R}^n$ es un conjunto Jordan medible y $f : A \subset \mathbb{R}^n \to \mathbb{R}$ es acotada en A entonces f es integrable sobre A si y sólo si $\lambda(D_{f,A}) = 0$.*

Como ya se dijo, la prueba de este teorema se sigue totalmente del teorema 2.2

Pasemos a ver algunas propiedades de funciones integrables sobre conjuntos Jordan medibles; como se ha hecho a lo largo de este trabajo, sólo enunciaremos las que nos son indispensables para nuestra labor.

Teorema 3.6 *Si $f, g : A \subset \mathbb{R}^n \to \mathbb{R}$ son integrables en el conjunto Jordan medible A, entonces*
i) $f + g : A \subset \mathbb{R}^n \to \mathbb{R}$ es integrable sobre A y $\int_A (f+g) = \int_A f + \int_A g$

ii) Para cualquier $c \in \mathbb{R}$, $cf : A \subset \mathbb{R}^n \to \mathbb{R}$ es integrable sobre D y $\int_A cf = c\int_A f$

iii) $fg : A \subset \mathbb{R}^n \to \mathbb{R}$ es integrable sobre A.

Teorema 3.7 *Si $f : A \subset \mathbb{R}^n \to \mathbb{R}$ es integrable en el Jordan medible $A \subset \mathbb{R}^n$ y $f(\hat{x}) \geq 0$ para toda $\hat{x} \in A$ entonces $\int_A f \geq 0$; más aún, si existe $\hat{x}_0 \in int(A)$ tal que f es continua en este punto y $f(\hat{x}_0) > 0$ entonces $\int_A f > 0$.*

Corolario 3.3 *Si $f, g : A \subset \mathbb{R}^n \to \mathbb{R}$ son integrables en el Jordan medible $A \subset \mathbb{R}^n$ y para toda $\hat{x} \in A$ $f(\hat{x}) \leq g(\hat{x})$, entonces $\int_A f \leq \int_A g$.*

Corolario 3.4 *Si $f : A \subset \mathbb{R}^n \to \mathbb{R}$ es integrable en el Jordan medible $A \subset \mathbb{R}^n$, entonces $|f| : A \subset \mathbb{R}^n \to \mathbb{R}$ es integrable sobre A y $\left|\int_A f\right| \leq \int_A |f|$.*

Otra propiedad, bastante natural, es que si una función definida sobre un conjunto Jordan medible y con medida de Jordan cero es acotada entonces la función es automáticamente integrable sobre este conjunto y además la integral es cero.

Proposición 3.4 *Sea $A \subset \mathbb{R}^n$ Jordan medible con $J(A) = 0$. Si $f : A \subset \mathbb{R}^n \to \mathbb{R}$ es acotada en A entonces f es integrable sobre A y $\int_A f = 0$.*

Dem. Como $D_{f,A} \subset A$ y $J(A) = 0$ entonces $J(D_{f,A}) = 0$ y por lo tanto f es integrable sobre A. Ahora bien, si $M > 0$ es tal que $|f(\hat{x})| \leq M$ para toda $\hat{x} \in A$, entonces $-M \leq f(\hat{x}) \leq M$ para toda $\hat{x} \in A$ y por el corolario 3.3 tenemos que

$$\int_A -M \leq \int_A f \leq \int_A M$$

es decir,

$$-M \cdot J(A) \leq \int_A f \leq M \cdot J(A)$$

pero $J(A) = 0$, por lo tanto $\int_A f = 0$. ■

El siguiente teorema, que sí demostraremos, es conocido como el teorema del valor intermedio para integrales. El lector sin duda recordará el resultado análogo para el cálculo de una variable, cuya prueba tenía como base el teorema del valor intermedio para funciones continuas. La prueba del teorema del valor intermedio para integrales, ahora para funciones de varias variables, también tiene como base el teorema del valor intermedio para funciones continuas, así como los teoremas 3.6 y 3.7 y el corolario 3.3.

Teorema 3.8 *(Valor intermedio para integrales) Si $f, g : A \subset \mathbb{R}^n \to \mathbb{R}$ son continuas y acotadas en el conjunto Jordan medible A, $g(\hat{x}) \geq 0$ para toda $\hat{x} \in A$ y A es conexo, entonces existe $\hat{y} \in A$ tal que*

$$\int_A fg = f(\hat{y}) \int_A g$$

Dem. Por el teorema 3.6 fg es integrable sobre A. Si $g(\hat{x}) = 0$ para toda $\hat{x} \in int(A)$, entonces el resultado es trivialmente cierto, de manera que podemos suponer que existe $\hat{x}_0 \in int(A)$ tal que $g(\hat{x}_0) > 0$ y entonces por el teorema 3.7 $\int_A g > 0$.

Sean

$$m = \inf\{f(\hat{x}) \mid \hat{x} \in A\}$$

y

$$M = \sup\{f(\hat{x}) \mid \hat{x} \in A\}$$

Como g es una función no negativa entonces $mg(\hat{x}) \leq f(\hat{x})g(\hat{x}) \leq Mg(\hat{x})$ para toda $\hat{x} \in A$. Por el corolario 3.3 y el teorema 3.6 se tiene que

$$m \int_A g \leq \int_A fg \leq M \int_A g$$

y como $\int_A g > 0$

$$m \leq \frac{\int_A fg}{\int_A g} \leq M$$

La idea ahora es aplicar el teorema del valor intermedio, mismo que es válido porque A es conexo y f continua en A, y entonces asegurar que existe $\hat{y} \in A$ tal que

$$f(\hat{y}) = \frac{\int_A fg}{\int_A g}$$

es decir,

$$\int_A fg = f(\hat{y}) \int_A g$$

pero esto hay que hacerlo con cuidado, porque de momento sólo hemos dicho que

$$m \leq \frac{\int_A fg}{\int_A g} \leq M$$

dejando abierta la posibilidad de que

$$m = \frac{\int_A fg}{\int_A g}$$

o

$$\frac{\int_D fg}{\int_D g} = M$$

y si f no alcanza a su ínfimo o a su supremo entonces no podemos asegurar la existencia del punto $\hat{y} \in A$. Sin embargo lo que en realidad sucede, es que si f no alcanza a su ínfimo o a su supremo entonces el cociente de integrales no puede ser ni m ni M, es decir, si $f(\hat{x}) > m$ para toda $\hat{x} \in A$ o $f(\hat{x}) < M$ para toda $\hat{x} \in A$, entonces en particular tendríamos que

$$f(\hat{x}_0)\, g(\hat{x}_0) - mg(\hat{x}_0) > 0$$

o

$$Mg(\hat{x}_0) - f(\hat{x}_0)\, g(\hat{x}_0) > 0$$

según sea el caso, lo que por el teorema 3.7 implica que $m\int_A g < \int_A fg$ o $\int_A fg < M\int_A g$, y por tanto,

$$m < \frac{\int_A fg}{\int_A g}$$

o

$$\frac{\int_A fg}{\int_A g} < M$$

y entonces podemos aplicar con toda libertad el teorema del valor intermedio. ■

Si tomamos a la función g como la constante 1 en el teorema anterior, podemos concluir que $\int_A f = f(\hat{y})\, J(A)$ para algún $\hat{y} \in A$.

Por último, demostraremos un resultado análogo al inciso *iv*) de la proposición 2.3, pero ahora por supuesto para funciones definidas sobre conjuntos Jordan medibles.

Teorema 3.9 *Sean $A, A_1, A_2 \subset \mathbb{R}^n$ conjuntos Jordan medibles tales que $A = A_1 \cup A_2$ y $J(A_1 \cap A_2) = 0$. $f : A \subset \mathbb{R}^n \to \mathbb{R}$ es integrable sobre A si y sólo si es integrable sobre A_1 y A_2, además*

$$\int_A f = \int_{A_1} f + \int_{A_2} f$$

Dem. Supongamos primero que f es integrable sobre A. Por el teorema 3.5 se tiene que $\lambda(D_{f,A}) = 0$ y por lo tanto $\lambda(D_{f,A_1}) = 0$ y $\lambda(D_{f,A_2}) = 0$, ya que $D_{f,A_1} \cup D_{f,A_2} \subset D_{f,A_1 \cup A_2} = D_{f,A}$. De modo que f es integrable tanto en A_1 como en A_2.

Supongamos ahora que f es integrable en A_1 y A_2. Como $J(A_1 \cap A_2) = 0$ entonces por la proposición 3.4 f también es integrable en $A_1 \cap A_2$. Sea $R \subset \mathbb{R}^n$ rectángulo tal que $A \subset R$ y sean $f_A, f_{A_1}, f_{A_2}, f_{A_1 \cap A_2} : \mathbb{R}^n \to \mathbb{R}$ las funciones características de f en los conjuntos A, A_1,A_2 y $A_1 \cap A_2$ respectivamente como en la definición 3.2. Nótese que $f_A(\hat{x}) = f_{A_1}(\hat{x}) + f_{A_2}(\hat{x}) - f_{A_1 \cap A_2}(\hat{x})$; de acuerdo con la proposición 2.3 como las funciones

f_{A_1}, f_{A_2} y $f_{A_1 \cap A_2}$ son integrables sobre R podemos concluir que f_A es integrable sobre R, es decir, f es integrable sobre A.

Para la igualdad de integrales consideremos una vez más a un rectángulo $R \subset \mathbb{R}^n$ tal que $A \subset R$ y a las funciones características $f_A, f_{A_1}, f_{A_2}, f_{A_1 \cap A_2} : \mathbb{R}^n \to \mathbb{R}$ de f en los conjuntos A, A_1,A_2 y $A_1 \cap A_2$. De la identidad $f_A(\hat{x}) = f_{A_1}(\hat{x}) + f_{A_2}(\hat{x}) - f_{A_1 \cap A_2}(\hat{x})$ y el hecho de que cada función es integrable sobre R, obtenemos que

$$\int_R f_A = \int_R f_{A_1} + \int_R f_{A_2} - \int_R f_{A_1 \cap A_2}$$

es decir,

$$\int_A f = \int_{A_1} f + \int_{A_2} f - \int_{A_1 \cap A_2} f$$

pero por la proposición 3.4 sabemos que $\int_{A_1 \cap A_2} f = 0$, por lo tanto, $\int_A f = \int_{A_1} f + \int_{A_2} f$ ■

Corolario 3.5 *Si $A, A_1 \subset \mathbb{R}^n$ son Jordan medibles tales que $A_1 \subset A$ y $f : A \to \mathbb{R}$ es integrable sobre A entonces f es integrable sobre $A \setminus A_1$ y además*

$$\int_{A \setminus A_1} f = \int_A f - \int_{A_1} f$$

Dem. Como A_1 es Jordan medible entonces $A \setminus A_1$ es Jordan medible. Al ser $A = A_1 \cup (A \setminus A_1)$ y $J(A_1 \cap (A \setminus A_1)) = J(\emptyset) = 0$, del teorema 3.9 se concluye que f es integrable en $A \setminus A_1$ y A_1 y además $\int_A f = \int_{A_1} f + \int_{A \setminus A_1} f$, es decir,

$$\int_{A \setminus A_1} f = \int_A f - \int_{A_1} f$$

■

Capítulo 4

Teorema de cambio de variable

En este último capítulo, daremos una demostración del teorema de cambio de variable en varias variables, teorema que nos es absolutamente familiar. Su prueba es bastante larga y nos llevará algo de tiempo desarrollar las herramientas necesarias para realizar dicha prueba.

Para comenzar, veamos lo que dice el teorema de cambio de variable:

Sea $\Omega \subset \mathbb{R}^n$ un conjunto abierto y $D \subset \mathbb{R}^n$ Jordan medible tal que $D \cup Fr(D) \subset \Omega$. Sea $g : \Omega \to \mathbb{R}^n$ de clase C^1 en Ω, inyectiva en $int(D)$ y tal que $J_g(\hat{x}) \neq 0$ para toda $\hat{x} \in int(D)$. Si $f : g(D) \to \mathbb{R}$ es acotada y continua entonces

$$\int_{g(D)} f = \int_D (f \circ g) |J_g|$$

Analícemos muy a grandes rasgos por qué se da esta identidad y tratemos de ver algunas complicaciones que se nos presentarán al momento de demostrar el teorema.

Comencemos suponiendo que queremos integrar una función f sobre un conjunto Jordan medible $A \subset \mathbb{R}^n$ y que sabemos que este conjunto A se puede ver como la imagen bajo una cierta función g de otro conjunto Jordan medible $D \subset \mathbb{R}^n$, es decir, $g(D) = A$.

Sea $R \subset \mathbb{R}^n$ un rectángulo que contiene a D y Q una partición de R; para cada subrectángulo R_i inducido por Q que intersecte a D elijamos un $\hat{\eta}_i \in R_i \cap D$.

Lo primero que nos gustaría conseguir es que sumas de la forma

$$\sum_{R_i \cap D \neq \emptyset} f(g(\hat{\eta}_i)) J(g(R_i))$$

se aproximen a $\int_A f$ y que esta aproximación sea mejor en la medida de que Q sea una partición más fina de R. Aquí surgen de inmediato dos detalles. El primero es poder asegurar que los conjuntos $g(R_i)$ son Jordan medibles. El segundo es que no sumemos de más, es decir, quitar la posibilidad de que para dos subrectángulos distintos R_i y R_j ocurra que

$$int(g(R_i) \cap g(R_j)) \neq \emptyset$$

El primer detalle requiere un poco de trabajo, la suposición de que la función g es de clase C^1 es fundamental; el segundo detalle es más sencillo de cubrir, basta pedir inyectividad de la función g.

También sería muy bueno encontrar una forma de expresar la medida de Jordan de $g(R_i)$ en términos del volumen de R_i y de la función g. En este punto, las transformaciones lineales serán muy útiles. Demostraremos que, si $R \subset \mathbb{R}^n$ es un rectángulo y $T : \mathbb{R}^n \to \mathbb{R}^n$ una transformación lineal, entonces

$$J(T(R)) = |\det L| \, V(R)$$

en donde L es la representación de T respecto a la base canónica.

Con esto en mente y considerando que trabajamos con particiones muy finas (y por tanto con rectángulos muy pequeños), si g es una función tal que en cada punto $\hat{x}$ de su dominio existe su mejor aproximación lineal, es decir, su derivada, y R_i es un rectángulo "muy pequeño", parece razonable esperar que

$$J(g(R_i)) \approx J(D_g(\hat{\eta}_i)(R_i)) = |J_g(\hat{\eta}_i)| \, V(R_i)$$

en donde $\hat{\eta}_i$ es algún elemento de $R_i \cap D$. Con esta aproximación obtenemos que

$$\int_A f \approx \sum_{R_i \cap D \neq \emptyset} f(g(\hat{\eta}_i)) \, J(g(R_i)) \approx \sum_{R_i \cap D \neq \emptyset} f(g(\hat{\eta}_i)) \, |J_g(\hat{\eta}_i)| \, V(R_i)$$

Si se observa con cuidado, la segunda suma es una suma de Riemann asociada a la función $(f \circ g) |J_g|$ y correspondiente a la partición Q de R. En este sentido, si dicha partición es muy fina, deberá ser cierto que

$$\sum_{R_i \cap D \neq \emptyset} f(g(\hat{\eta}_i)) \, |J_g(\hat{\eta}_i)| \, V(R_i) \approx \int_D (f \circ g) \, |J_g|$$

concluyendo así que

$$\int_A f \approx \int_D (f \circ g) \, |J_g|$$

y conforme tengamos particiones más finas, podremos asegurar que esta aproximación es en realidad una igualdad.

Con este pequeño esbozo de lo que será la prueba del teorema de cambio de variable, podemos decir que el camino a seguir para demostrar este teorema consiste de tres partes. Primero daremos una condición suficiente para que la imagen bajo una función de un conjunto Jordan medible continúe siendo Jordan medible; para esto será necesario introducir el concepto de uniformidad de Lipschitz, concepto posiblemente nuevo para el lector; también hará falta hablar un poco de cubos, que no son otra cosa más que una clase especial de rectángulos (trabajar con cubos nos facilitará las cosas en más de una ocasión). El segundo paso será demostrar un caso particular del teorema de cambio de variable, concretamente, demostraremos la fórmula propuesta unos renglones arriba

$$J(T(R)) = |\det L| \, V(R)$$

en donde T es una transformación lineal y L su representación matricial (respecto a la base canónica); pero no nos limitaremos a los rectángulos, sino que hallaremos una manera de calcular la medida de Jordan de la imagen de un conjunto Jordan medible bajo una transformación lineal. Finalmente la tercera parte será demostrar el caso general del teorema de cambio de variable.

4.1 Preliminares

En esta breve sección desarrollaremos algunas herramientas necesarias para nuestra labor, en la primera parte veremos algunos resultados que involucran exclusivamente a nuestros conocimentos del cálculo diferencial. De manera un poco más precisa, recurriremos constantemente al teorema del valor medio para funciones de $\mathbb{R}$ en $\mathbb{R}$. Veremos un par de lemas, fundamentales para cuando entremos de lleno en la demostración del teorema de cambio de variable, y finalizaremos esta sección hablando un poco de la "distancia" entre conjuntos en $\mathbb{R}^n$. Cabe advertitr que las demostraciones de los resultados en esta sección son muy técnicas.

Proposición 4.1 *Sea $\sigma : [a,b] \to \mathbb{R}^n$ una función continua. Si σ es diferenciable en (a,b) y existe $M > 0$ tal que $\|\sigma'(t)\| \leq M$ para toda $t \in (a,b)$, entonces $\|\sigma(b) - \sigma(a)\| \leq M(b-a)$.*

Dem. Consideremos a la función $f : [a,b] \to \mathbb{R}$ definida por

$$f(t) = (\sigma(b) - \sigma(a)) \cdot \sigma(t)$$

f resulta continua en $[a,b]$ y derivable en (a,b) de manera que por el Teorema del valor medio existe $\xi \in (a,b)$ tal que

$$f(b) - f(a) = f'(\xi)(b-a)$$

Por definición de f obtenemos que

$$\begin{aligned} f(b) - f(a) &= (\sigma(b) - \sigma(a)) \cdot \sigma(b) - (\sigma(b) - \sigma(a)) \cdot \sigma(a) \\ &= (\sigma(b) - \sigma(a)) \cdot (\sigma(b) - \sigma(a)) = \|\sigma(b) - \sigma(a)\|^2 \end{aligned}$$

mientras que

$$f'(\xi) = (\sigma(b) - \sigma(a)) \cdot \sigma'(\xi)$$

De manera que por hipótesis y por la desigualdad de Cauchy - Schwarz se tiene que

$$\begin{aligned} \|\sigma(b) - \sigma(a)\|^2 &= (b-a)(\sigma(b) - \sigma(a)) \cdot \sigma'(\xi) \\ &\leq (b-a)\|\sigma(b) - \sigma(a)\| \|\sigma'(\xi)\| \\ &\leq M(b-a)\|\sigma(b) - \sigma(a)\| \end{aligned}$$

Si $\sigma(b) = \sigma(a)$ entonces la desigualdad que queremos se sigue de manera trivial y si $\sigma(b) \neq \sigma(a)$ entonces podemos cancelar el término $\|\sigma(b) - \sigma(a)\|$ de la desigualdad anterior con lo que concluimos la proposición ■

Una consecuencia del resultado anterior es la siguiente "generalización" del mismo para $\mathbb{R}^n$.

Proposición 4.2 *Si $\Omega \subset \mathbb{R}^n$ es abierto, $g : \Omega \to \mathbb{R}^m$ es de clase C^1 en Ω y $\hat{u}_0, \hat{u}_1 \in \Omega$ son tales que $\hat{u}_t = (1-t)\hat{u}_0 + t\hat{u}_1$ pertenece a Ω para toda $t \in [0,1]$, entonces*

$$\|g(\hat{u}_1) - g(\hat{u}_0)\| \leq c\|\hat{u}_1 - \hat{u}_0\|$$

donde

$$c = \sup_{t\in[0,1]} \|D_g(\hat{u}_t)\|_{\pounds(n,m)}$$

Dem. Como $g \in C^1(\Omega)$ y $\{\hat{u}_t \in \mathbb{R}^n \mid t \in [0,1]\} \subset \Omega$ es un conjunto compacto entonces por la proposición 1.4

$$c = \sup_{t \in [0,1]} \|D_g(\hat{u}_t)\|_{\pounds(n,m)}$$

existe.

Sea $\sigma : [0,1] \to \mathbb{R}^m$ definida como $\sigma(t) = g(\hat{u}_t)$; σ es continua y diferenciable y por la regla de la cadena $\sigma'(t) = D_g(\hat{u}_t)(\hat{u}_1 - \hat{u}_0)$ de modo que

$$\|\sigma'(t)\| = \|D_g(\hat{u}_t)(\hat{u}_1 - \hat{u}_0)\| \leq \|D_g(u_t)\|_{\pounds(n,m)} \|\hat{u}_1 - \hat{u}_0\| \leq c\|\hat{u}_1 - \hat{u}_0\|$$

y por la proposición anterior

$$\|g(\hat{u}_1) - g(\hat{u}_0)\| = \|\sigma(1) - \sigma(0)\| \leq c\|\hat{u}_1 - \hat{u}_0\|(1-0) = c\|\hat{u}_1 - \hat{u}_0\|$$

■

A continuación veremos una especie de teorema de valor medio para funciones de $\mathbb{R}^n$ en $\mathbb{R}$.

Teorema 4.1 *Sea $f : D \subset \mathbb{R}^n \to \mathbb{R}$ y $\hat{a}, \hat{b} \in int(D)$. Si $(1-t)\hat{a} + t\hat{b} \in int(D)$ para toda $t \in [0,1]$ y f es diferenciable en cada uno de estos puntos entonces existe $t_0 \in (0,1)$ tal que*

$$f\left(\hat{b}\right) - f(\hat{a}) = \nabla f\left((1-t_0)\hat{a} + t_0\hat{b}\right) \cdot \left(\hat{b} - \hat{a}\right)$$

Dem. Sea $\sigma : [0,1] \to \mathbb{R}$ definida como $\sigma(t) = f\left((1-t)\hat{a} + t\hat{b}\right)$. σ es una función diferenciable en $[0,1]$, y por la regla de la cadena $\sigma'(t) = \nabla f\left((1-t)\hat{a} + t\hat{b}\right) \cdot \left(\hat{b} - \hat{a}\right)$.

Por el teorema del valor medio para funciones reales se tiene que $\sigma(1) - \sigma(0) = \sigma'(t_0)$ para algún $t_0 \in (0,1)$, es decir, $f\left(\hat{b}\right) - f(\hat{a}) = \nabla f\left((1-t_0)\hat{a} + t_0\hat{b}\right) \cdot \left(\hat{b} - \hat{a}\right)$. ■

Corolario 4.1 *Sea $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ y $\hat{a}, \hat{b} \in int(D)$. Si $(1-t)\hat{a} + t\hat{b} \in int(D)$ para toda $t \in [0,1]$ y f es diferenciable en cada uno de estos puntos entonces $\forall \hat{u} \in \mathbb{R}^m$ $\exists t_{\hat{u}} \in (0,1)$ tal que*

$$\left(f\left(\hat{b}\right) - f(\hat{a})\right) \cdot \hat{u} = D_f(c_{\hat{u}})\left(\hat{b} - \hat{a}\right) \cdot \hat{u}$$

donde $c_{\hat{u}} = (1-t_{\hat{u}})\hat{a} + t_{\hat{u}}\hat{b}$.

Dem. Dado $\hat{u} \in \mathbb{R}^m$ definamos a la función $g : D \subset \mathbb{R}^n \to \mathbb{R}$ como $g(\hat{x}) = f(\hat{x}) \cdot \hat{u}$; g resulta una función diferenciable en $(1-t)\hat{a} + t\hat{b} \in int(D)$ para toda $t \in [0,1]$. Por el teorema 4.1 se tiene que existe $t_{\hat{u}} \in (0,1)$ tal que

$$g\left(\hat{b}\right) - g(\hat{a}) = \nabla g(c_{\hat{u}}) \cdot \left(\hat{b} - \hat{a}\right)$$

donde $c_{\hat{u}} = (1-t_{\hat{u}})\hat{a} + t_{\hat{u}}\hat{b}$, es decir,

$$\left(f\left(\hat{b}\right) - f(\hat{a})\right) \cdot \hat{u} = D_f(c_{\hat{u}})\left(\hat{b} - \hat{a}\right) \cdot \hat{u}$$

■

Sabemos que si una función $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ es derivable en un punto $\hat{x}_0 \in D$ entonces $D_f(\hat{x}_0)$ es la transformación lineal que más se parece a f cerca de $\hat{x}_0$. En términos de la definición de derivabilidad, esto significa que para cualquier cantidad $\varepsilon > 0$ se cumple que $\|f(\hat{x}) - f(\hat{x}_0) - D_f(\hat{x}_0)(\hat{x} - \hat{x}_0)\| \leq \varepsilon \|\hat{x} - \hat{x}_0\|$ para $\hat{x}$ suficientemente cercano a $\hat{x}_0$. Ahora bien, nosotros trataremos de encontrar condiciones suficientes para que lo anterior se siga cumpliendo con cierta uniformidad, es decir, trataremos de hallar condiciones suficientes bajo las cuales dada cualquier cantidad positiva ε se cumpla que $\|f(\hat{x}) - f(\hat{a}) - D_f(\hat{a})(\hat{x} - \hat{a})\| \leq \varepsilon \|\hat{x} - \hat{a}\|$ para cualquier $\hat{a} \in D$ y para cualquier $\hat{x} \in D$ suficientemente cercano a $\hat{a}$. En realidad no demostraremos que la desigualdad anterior se cumple para cualquier $\hat{a} \in D$, nos basta con que se cumpla para cualquier $\hat{a} \in K$, donde K es un subconjunto compacto de D (el lector estará de acuerdo en que normalmente es en los conjuntos compactos donde se consigue uniformidad). Nuestro siguiente lema nos acerca un poco a lo que buscamos, y cabe decir que la prueba de éste, se apoya en el corolario previo; de hecho, la razón principal por la que el teorema 4.1 y el corolario 4.1 fueron incluidos en este trabajo fue para poder aterrizar en este lema.

Lema 4.1 *Sea $D \subset \mathbb{R}^n$ abierto y $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$. Si f es de clase C^1 en D y tomamos $\hat{a} \in D$, entonces para toda $\varepsilon > 0$, existe $\delta > 0$ tal que $B_\delta(\hat{a}) \subset D$ y para cualesquiera $\hat{x}, \hat{y} \in B_\delta(\hat{a})$*

$$\|f(\hat{x}) - f(\hat{y}) - D_f(\hat{a})(\hat{x} - \hat{y})\| \leq \varepsilon \|\hat{x} - \hat{y}\|$$

Dem. Sea $\varepsilon > 0$, como f es de clase C^1 en D, existe $\delta > 0$ tal que $B_\delta(\hat{a}) \subset D$ y para toda $\hat{x} \in B_\delta(\hat{a})$

$$\|D_f(\hat{x}) - D_f(\hat{a})\|_{\mathcal{L}(n,m)} < \varepsilon$$

Tomemos $\hat{x}, \hat{y} \in B_\delta(\hat{a})$; de la convexidad de $B_\delta(\hat{a})$, se sigue que $(1-t)\hat{x} + t\hat{y} \in B_\delta(\hat{a})$ para toda $t \in [0,1]$. Llamemos $\hat{u} = f(\hat{x}) - f(\hat{y}) - D_f(\hat{a})(\hat{x} - \hat{y})$; aplicando el corolario anterior, se sigue que existe $t_{\hat{u}} \in (0,1)$ tal que

$$(f(\hat{x}) - f(\hat{y})) \cdot \hat{u} = D_f(\hat{x}_{\hat{u}})(\hat{x} - \hat{y}) \cdot \hat{u}$$

con $\hat{x}_{\hat{u}} = (1 - t_{\hat{u}})\hat{x} + t_{\hat{u}}\hat{y}$. De modo que

$$\begin{aligned}
\|f(\hat{x}) - f(\hat{y}) - D_f(\hat{a})(\hat{x} - \hat{y})\|^2 &= (f(\hat{x}) - f(\hat{y}) - D_f(\hat{a})(\hat{x} - \hat{y})) \cdot \hat{u} \\
&= (f(\hat{x}) - f(\hat{y})) \cdot \hat{u} - D_f(\hat{a})(\hat{x} - \hat{y}) \cdot \hat{u} \\
&= D_f(\hat{x}_{\hat{u}})(\hat{x} - \hat{y}) \cdot \hat{u} - D_f(\hat{a})(\hat{x} - \hat{y}) \cdot \hat{u} \\
&= (D_f(\hat{x}_{\hat{u}})(\hat{x} - \hat{y}) - D_f(\hat{a})(\hat{x} - \hat{y})) \cdot \hat{u} \\
&\leq \|D_f(\hat{x}_{\hat{u}})(\hat{x} - \hat{y}) - D_f(\hat{a})(\hat{x} - \hat{y})\| \|\hat{u}\| \\
&\leq \|D_f(\hat{x}_{\hat{u}}) - D_f(\hat{a})\|_{\mathcal{L}(n,m)} \|\hat{x} - \hat{y}\| \|\hat{u}\| \\
&< \varepsilon \|\hat{x} - \hat{y}\| \|\hat{u}\|
\end{aligned}$$

Nótese que en la segunda desigualdad se está aplicando la proposición 1.2.

Por lo tanto

$$\|f(\hat{x}) - f(\hat{y}) - D_f(\hat{a})(\hat{x} - \hat{y})\| \|\hat{u}\| = \|f(\hat{x}) - f(\hat{y}) - D_f(\hat{a})(\hat{x} - \hat{y})\|^2 < \varepsilon \|\hat{x} - \hat{y}\| \|\hat{u}\|$$

y suponiendo que $\|\hat{u}\| \neq 0$ (de lo contrario no habría nada que probar), podemos cancelar de ambos lados de la desigualdad $\|\hat{u}\|$ y concluimos que

$$\|f(\hat{x}) - f(\hat{y}) - D_f(\hat{a})(\hat{x} - \hat{y})\| \leq \varepsilon \|\hat{x} - \hat{y}\|$$

■

Lema 4.2 *Sea $D \subset \mathbb{R}^n$ abierto y $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$. Si f es de clase C^1 en D y $K \subset D$ es un conjunto compacto, entonces para toda $\varepsilon > 0$, existe $\delta > 0$ tal que para toda $\hat{a} \in K$*

$$\|f(\hat{x}) - f(\hat{a}) - D_f(\hat{a})(\hat{x} - \hat{a})\| \leq \varepsilon \|\hat{x} - \hat{a}\|$$

siempre que $\hat{x} \in B_\delta(\hat{a}) \cap K$.

Dem. Sea $\varepsilon > 0$, por el Lema 4.1 y como f es de clase C^1, para cada $\hat{a} \in K$ podemos hallar $\delta_{\hat{a}} > 0$ tal que $B_{\delta_{\hat{a}}}(\hat{a}) \subset D$, para cualesquiera $\hat{x}, \hat{y} \in B_{\delta_{\hat{a}}}(\hat{a})$

$$\|f(\hat{x}) - f(\hat{y}) - D_f(\hat{a})(\hat{x} - \hat{y})\| \leq \frac{\varepsilon}{2} \|\hat{x} - \hat{y}\|$$

y para toda $\hat{x} \in B_{\delta_{\hat{a}}}(\hat{a})$

$$\|D_f(\hat{x}) - D_f(\hat{a})\|_{\mathcal{L}(n,m)} < \frac{\varepsilon}{2}$$

Entonces

$$\left\{B_{\frac{\delta_{\hat{a}}}{2}}(\hat{a}) \mid \hat{a} \in K\right\}$$

es una cubierta abierta para K y, por tanto, existen $\hat{a}_1, \ldots, \hat{a}_l \in K$ tales que

$$K \subset \bigcup_{j=1}^{l} B_{\delta_j}(\hat{a}_j)$$

en donde

$$\delta_j = \frac{\delta_{\hat{a}_j}}{2}$$

Sean $\delta = \min\{\delta_1, \ldots, \delta_l\}$ y $\hat{a} \in K$ y tomemos $\hat{x} \in B_\delta(\hat{a}) \cap K$. Por la última contención sabemos que existe $\hat{a}_j \in \{\hat{a}_1, \ldots, \hat{a}_l\}$ tal que $\hat{x} \in B_{\delta_j}(\hat{a}_j)$; entonces

$$\|\hat{a} - \hat{a}_j\| \leq \|\hat{a} - \hat{x}\| + \|\hat{x} - \hat{a}_j\| < \delta + \delta_j \leq 2\delta_j = \delta_{\hat{a}_j}$$

es decir,

$$\hat{a} \in B_{\delta_{\hat{a}_j}}(\hat{a}_j)$$

de donde,

$$\|f(\hat{x}) - f(\hat{a}) - D_f(\hat{a}_j)(\hat{x} - \hat{a})\| \leq \frac{\varepsilon}{2} \|\hat{x} - \hat{a}\|$$

y

$$\left\| D_f(\hat{a})(\hat{x}-\hat{a}) - D_f(\hat{a}_j)(\hat{x}-\hat{a}) \right\| \leq \left\| D_f(\hat{x}) - D_f(\hat{a}) \right\|_{\mathcal{L}(n,m)} \|\hat{x}-\hat{a}\| < \frac{\varepsilon}{2} \|\hat{x}-\hat{a}\|$$

Aplicando la desigualdad del triángulo y las dos desigualdades anteriores se tiene que

$$\| f(\hat{x}) - f(\hat{a}) - D_f(\hat{a})(\hat{x}-\hat{a}) \| \leq \frac{\varepsilon}{2} \|\hat{x}-\hat{a}\| + \frac{\varepsilon}{2} \|\hat{x}-\hat{a}\| = \varepsilon \|\hat{x}-\hat{a}\|$$

■

Del lema 1.1 sabemos que toda transformación lineal $L : \mathbb{R}^n \to \mathbb{R}^n$ es uno a uno si y sólo si existe $\mu > 0$ tal que para toda $\hat{x} \in \mathbb{R}^n$

$$\mu \|\hat{x}\| \leq \|L(\hat{x})\|$$

y en particular, para una función $f : D \subset \mathbb{R}^n \to \mathbb{R}^n$, diferenciable en algún punto $\hat{a} \in D$ para el cual $J_f(\hat{a}) \neq 0$, se tiene que existe $\mu > 0$ tal que para toda $\hat{x} \in \mathbb{R}^n$

$$\mu \|\hat{x}\| \leq \|D_f(\hat{a})(\hat{x})\|$$

Al igual que hicimos en el lema anterior, ahora buscamos que esta desigualdad se siga cumpliendo con cierta uniformidad, es decir, queremos que exista una cantidad positiva μ tal que para cualquier punto $\hat{a} \in D$ para el cual $J_f(\hat{a}) \neq 0$, se cumpla que $\mu \|\hat{x}\| \leq \|D_f(\hat{a})(\hat{x})\|$ para toda $\hat{x} \in \mathbb{R}^n$. Nuevamente, no podremos garantizar que la $\mu > 0$ nos sirve para cualquier $\hat{a} \in D$, pero sí demostraremos que sirve si nos restringimos a un subconjunto compacto de D (con eso nos basta).

Lema 4.3 *Sean $D \subset \mathbb{R}^n$ un conjunto abierto y $f : D \subset \mathbb{R}^n \to \mathbb{R}^n$. Si f es de clase C^1 en D y $K \subset D$ compacto y además $J_f(\hat{a}) \neq 0$ para toda $\hat{a} \in K$, entonces existe $\mu > 0$ tal que para toda $\hat{a} \in K$ y para toda $\hat{x} \in \mathbb{R}^n$*

$$\mu \|\hat{x}\| \leq \|D_f(\hat{a})(\hat{x})\|$$

Dem. El razonamiento para esta prueba es muy similar al lema previo. Al ser f de clase C^1 en D y $J_f(\hat{a}) \neq 0$ para toda $\hat{a} \in K$ se tiene por el lema 1.2 que para cada $\hat{a} \in K$ existen $\delta_{\hat{a}} > 0$ y $M_{\hat{a}} > 0$ tales que para toda $\hat{x} \in B_{\delta_{\hat{a}}}(\hat{a})$ y para toda $\hat{y} \in \mathbb{R}^n$

$$M_{\hat{a}} \|\hat{y}\| \leq \|D_f(\hat{x})(\hat{y})\|$$

De manera que tomando la cubierta abierta de K

$$\{B_{\delta_{\hat{a}}}(\hat{a}) \mid \hat{a} \in K\}$$

existen $\hat{a}_1, \ldots, \hat{a}_l \in K$ tales que

$$K \subset \bigcup_{j=1}^{l} B_{\delta_{\hat{a}_j}}(\hat{a}_j)$$

Sea $\mu = \min\{M_{\hat{a}_1}, \ldots, M_{\hat{a}_l}\}$, dado $\hat{a} \in K$, $\hat{a} \in B_{\delta_{\hat{a}_j}}(\hat{a}_j)$, para alguna j, y entonces para toda $\hat{x} \in \mathbb{R}^n$

$$\mu\|\hat{x}\| \leq M_{\hat{a}_j}\|\hat{x}\| \leq \|D_f(\hat{a})(\hat{x})\|$$

■

Como dijimos al principio de esta sección, los lemas 4.2 y 4.3 serán muy importantes para desarrollar la prueba del teorema del cambio de variable, y podemos agregar que la proposición 4.2 nos será de utilidad cuando tratemos de ver cuándo la imagen de un conjunto Jordan medible bajo alguna función continúa siendo Jordan medible.

Hablemos ahora de la "distancia" entre conjuntos de $\mathbb{R}^n$.

Definición 4.1 *Sean $A, B \subset \mathbb{R}^n$ no vacíos, definimos la "distancia" de A a B como el número*

$$d(A,B) = \inf\{\|\hat{x} - \hat{y}\| \mid \hat{x} \in A,\ \hat{y} \in B\}$$

Hemos puesto entre comillas la palabra "distancia", porque el número $d(A,B)$ no es en sí una métrica para los subconjuntos de $\mathbb{R}^n$. De la definición resulta muy claro que si A es un subconjunto propio de B entonces $d(A,B) = 0$ a pesar de que $A \neq B$; y no sólo eso, sino que tampoco es una pseudométrica, porque no se satisface la desigualdad del triángulo: es posible conseguir tres conjuntos tales que $d(A,C) + d(C,B) < d(A,B)$. Por ejemplo, si $d(A,B) > 0$ entonces se tiene que $d(A, A \cup B) + d(A \cup B, B) = 0 < d(A,B)$.

A pesar de estos detalles, al número que acabamos de definir vale la pena llamarlo "distancia" entre conjuntos porque de alguna manera nos dice qué tan separados están dos conjuntos A y B cuando son ajenos.

Veamos una propiedad muy importante que tiene esta "distancia".

Proposición 4.3 *Sean $A, B \subset \mathbb{R}^n$ no vacíos tales que A es compacto y B es cerrado. Si $A \cap B = \emptyset$ entonces $d(A,B) > 0$.*

Dem. Sunpongamos que $A \cap B = \emptyset$. Sea $f : \mathbb{R}^n \to \mathbb{R}^n$ definida como

$$f(\hat{x}) = \inf\left\{\left\|\hat{x} - \hat{b}\right\| \mid \hat{b} \in B\right\}$$

Nótese que $d(A,B) = \inf\{f(\hat{x}) \mid \hat{x} \in A\}$. Además si $\hat{x} \notin B$ entonces $f(\hat{x}) > 0$; esto sucede ya que al ser B un conjunto cerrado, si tomamos $\hat{x} \notin B$ existe $\delta > 0$ tal que $B_\delta(\hat{x}) \cap B = \emptyset$, de modo que, $\left\|\hat{x} - \hat{b}\right\| \geq \delta$ para toda $\hat{b} \in B$, es decir, $f(\hat{x}) \geq \delta > 0$.

Demostremos que f es continua. Sean $\hat{x}, \hat{y} \in \mathbb{R}^n$ arbitrarios; por la desigualdad del triángulo tenemos que para toda $\hat{b} \in B$

$$\left\|\hat{x} - \hat{b}\right\| \leq \|\hat{x} - \hat{y}\| + \left\|\hat{y} - \hat{b}\right\|$$

pero por definición de f sabemos que $f(\hat{x}) \leq \left\|\hat{x} - \hat{b}\right\|$, por lo tanto,

$$f(\hat{x}) \leq \|\hat{x} - \hat{y}\| + \left\|\hat{y} - \hat{b}\right\|$$

y como esta desigualdad se cumple para toda $\hat{b} \in B$ obtenemos que

$$f(\hat{x}) \leq \|\hat{x} - \hat{y}\| + f(\hat{y})$$

de donde concluimos que

$$|f(\hat{x}) - f(\hat{y})| \leq \|\hat{x} - \hat{y}\|$$

y por tanto f es una función continua (de hecho es uniformemente continua en todo $\mathbb{R}^n$). Ahora bien, como f es continua y A es compacto entonces f alcanza un valor mínimo en A, es decir, existe $\hat{x}_0 \in A$ tal que $f(\hat{x}) \geq f(\hat{x}_0)$ para todo $\hat{x} \in A$, y por tanto $d(A, B) \geq f(\hat{x}_0)$, pero como $\hat{x}_0 \notin B$ entonces $f(\hat{x}_0) > 0$, y por lo tanto $d(A, B) > 0$ ■

Para concluir esta sección, demostraremos algo que geométricamente es muy claro: si tenemos un conjunto compacto y un rectángulo contenido en el interior del compacto entonces podemos meter un rectángulo de dimensiones racionales entre el compacto y el rectángulo inicial de manera que la diferencia de volúmenes entre los rectángulos sea tan pequeña como queramos. Para esto haremos uso de la proposición anterior por supuesto.

Proposición 4.4 *Sean $K \subset \mathbb{R}^n$ un conjunto compacto y $\delta > 0$. Si $R \subset int(K)$ es un rectángulo de dimensiones menores a δ entonces para toda $\varepsilon > 0$ existe $R' \subset int(K)$ rectángulo de dimensiones racionales menores a δ y tal que $R \subset int(R')$ y $V(R') - V(R) < \varepsilon$.*

Dem. Sea $\varepsilon > 0$. El lema 2.1 nos asegura la existencia de un rectángulo $R' \subset \mathbb{R}^n$ con dimensiones racionales tal que $R \subset int(R')$ y cuyo volumen difiera del volumen de R en menos que ε. Ahora, si se recuerda la prueba hecha en este lema, la clave estaba en la elección de las dimensiones del rectángulo R' y de acuerdo con la construcción realizada en esa prueba y el hecho de que la dimensiones R son menores a δ, podemos decir que se puede elegir a las dimensiones de R' también menores a δ.

La parte que hay que analizar con cuidado es que $R' \subset int(K)$. Llamemos $B = (int(K))^c$, B es un conjunto cerrado y como R es un conjunto compacto tal que $R \cap B = \emptyset$ entonces $d(R, B) > 0$.

Supongamos que $R = [a_1, b_1] \times \ldots \times [a_n, b_n]$, para cada $r > 0$ el rectángulo

$$R_r = [a_1 - r, b_1 + r] \times \ldots \times [a_n - r, b_n + r]$$

satisface que $R \subset int(R_r)$ y la diagonal R_r tiene longitud

$$\begin{aligned} d(R_r) &= \left[\sum_{i=1}^{n} (b_i - a_i + 2r)^2\right]^{\frac{1}{2}} = \|(b_1 - a_1, \ldots, b_n - a_n) + (2r, \ldots, 2r)\| \\ &\leq \|(b_1 - a_1, \ldots, b_n - a_n)\| + \|(2r, \ldots, 2r)\| = d(R) + 2r\sqrt{n} \end{aligned}$$

Si tomamos $r > 0$ suficientemente pequeña tal que $2r\sqrt{n} < d(R, B)$, entonces en vista de la desigualdad anterior, parece que debería ocurrir que $R_r \subset int(K)$. Y en efecto; dado $\hat{x} \in R_r \setminus R$ con $\hat{x} = (x_1, \ldots, x_n)$, se tiene que existen $i_1, \ldots, i_k \in \{1, \ldots, n\}$ tales que

$$x_{i_m} \in [a_{i_m} - r, b_{i_m} + r] \setminus [a_{i_m}, b_{i_m}]$$

Es decir, $a_{i_m} - r \leq x_{i_m} < a_{i_m}$ o $b_{i_m} < x_{i_m} \leq b_{i_m} + r$. Sea $\hat{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$ donde $y_j = x_j$ si $j \notin \{i_1, \ldots, i_k\}$ y para $j \in \{i_1, \ldots, i_k\}$

$$y_j = \begin{cases} a_j & si \quad a_j - r \leq x_j < a_j \\ \\ b_j & si \quad b_j < x_j \leq b_j + r \end{cases}$$

Hemos tomado $\hat{y}$ de manera que $\hat{y} \in R$, pero además para cada $j \in \{i_1, \ldots, i_k\}$, $|x_j - y_j| \leq r$, de donde

$$\|\hat{x} - \hat{y}\| = \left[\sum_{m=1}^{k} (x_{i_m} - y_{i_m})^2\right]^{\frac{1}{2}} \leq \left[\sum_{m=1}^{k} r^2\right]^{\frac{1}{2}} = r\sqrt{k} \leq 2r\sqrt{n} < d(R, B)$$

y como $\forall \hat{z} \in B$ se tiene que $d(R, B) \leq \|\hat{z} - \hat{y}\|$, conluimos que $\hat{x} \notin B = (int\,(K))^c$, es decir, $\hat{x} \in int\,(K)$. Recurriendo una vez más a la construcción hecha en el lema 2.1, si pedimos que las dimensiones del rectángulo R' que estamos buscando sean menores no sólo que δ sino además menores que las dimensiones del rectángulo R_r considerado hace un momento entonces R' cumplirá con todas las condiciones buscadas. ■

4.2 Cubos en $\mathbb{R}^n$

Definición 4.2 *Un cubo en $\mathbb{R}^n$ es un rectángulo $C \subset \mathbb{R}^n$ con todas las dimensiones iguales, es decir, si $C = [a_1, b_1] \times \ldots \times [a_n, b_n]$ entonces $b_i - a_i = b_j - a_j$ para cualesquiera $i, j \in \{1, \ldots, n\}$; al valor $s = b_i - a_i$ lo llamaremos la dimensión del cubo.*

Proposición 4.5 *Si $C = [a_1, b_1] \times \ldots \times [a_n, b_n] \subset \mathbb{R}^n$ es un cubo de dimensión s y $\hat{c} = (c_1, \ldots c_n) \in C$ es el centro del cubo, entonces $c_i = a_i + \frac{s}{2}$ para toda i y la diagonal de C tiene longintud $s\sqrt{n}$, es decir, $d(C) = s\sqrt{n}$.*

Dem. Sabemos que el centro $\hat{c}$ de C tiene por coordenadas

$$c_i = \frac{b_i + a_i}{2}$$

y como $s = b_i - a_i$ y

$$\frac{b_i + a_i}{2} = a_i + \frac{b_i - a_i}{2}$$

concluimos que

$$c_i = a_i + \frac{b_i - a_i}{2} = a_i + \frac{s}{2}$$

Por otro lado

$$d(C) = \sqrt{(b_1 - a_1)^2 + \cdots + (b_n - a_n)^2} = \sqrt{ns^2} = s\sqrt{n}$$

■

Observación 4.1 *De acuerdo con la proposición anterior, si $C \subset \mathbb{R}^n$ es un cubo con centro en $\hat{x} = (x_1, \ldots, x_n)$ y dimensión s entonces*

$$C = \left[x_1 - \frac{s}{2}, x_1 + \frac{s}{2}\right] \times \ldots \times \left[x_n - \frac{s}{2}, x_n + \frac{s}{2}\right]$$

y por tanto para toda $\hat{y} \in C$, con $\hat{y} = (y_1, \ldots, y_n)$, se tiene que $|y_i - x_i| \leq \frac{s}{2}$ para toda $i \in \{1, \ldots, n\}$, donde la igualdad se da sólo si $\hat{y}$ es un elemento de la frontera de C. De manera más precisa: $|y_i - x_i| = \frac{s}{2}$ para alguna $i \in \{1, \ldots, n\}$ si y sólo si $\hat{y} \in Fr(C)$.

De esto que acabamos de decir también se sigue que si $\hat{y} \in C$ entonces $\|\hat{x} - \hat{y}\| \leq \frac{s}{2}\sqrt{n}$. De manera recíproca si $\hat{y} \in \mathbb{R}^n$ es tal que $\|\hat{x} - \hat{y}\| \leq s$ para alguna $s \in \mathbb{R}$ entonces y vive en un cubo de dimensión $2s$ y con centro en x. Esto es claro pues si $\hat{x} = (x_1, \ldots, x_n)$ entonces basta tomar el cubo

$$C = [x_1 - s, x_1 + s] \times \ldots \times [x_n - s, x_n + s]$$

mismo que evidentemente tiene centro en $\hat{x}$ y dimensión $2s$. Además como para cada $i \in \{1, \ldots, n\}$

$$|x_i - y_i| \leq \|\hat{x} - \hat{y}\| \leq s$$

entonces $y_i \in [x_i - s, x_i + s]$, es decir, $\hat{y} \in C$. Como podemos ver, con todo lo dicho hace un momento, el hecho de que a diferencia de los rectángulos arbitrarios en un cubo todas sus dimensiones coinciden, nos permite caracterizar a los cubos a partir de su centro.

Lema 4.4 *Si $R = [a_1, b_1] \times \ldots \times [a_n, b_n] \subset \mathbb{R}^n$ es un rectángulo de dimensiones racionales, es decir, $b_i - a_i \in \mathbb{Q}$ para toda $i \in \{1, \ldots, n\}$, entonces existe una partición de R en cubos, todos con la misma dimensión y ésta además es racional.*

Dem. Por hipótesis cada $b_i - a_i$ es de la forma $\frac{p_i}{q_i}$ con p_i y q_i enteros positivos y primos relativos, sea $r = mcm\{q_1, \ldots, q_n\}$, entonces $r\frac{p_i}{q_i} \in \mathbb{Z}$ para toda $i \in \{1, \ldots, n\}$; de manera que

$$P_i = \left\{x_j^{(i)} = a_i + \frac{j}{r} \mid j \in \left\{0, 1, \ldots, r\frac{p_i}{q_i}, \right\}\right\}$$

es una partición del intervalo $[a_i, b_i]$ pero además ésta es uniforme, ya que

$$x_j^{(i)} - x_{j-1}^{(i)} = a_i + \frac{j}{r} - a_i - \frac{j-1}{r} = \frac{1}{r}$$

Esto implica que $P = P_1 \times \cdots \times P_n$ es una partición de R en cubos todos de dimensión $\frac{1}{r} \in \mathbb{Q}$
■

Observación 4.2 *Nótese que el razonamiento utilizado en el lema 4.4 nos permite demostrar algo todavía más fuerte. Si $R \subset \mathbb{R}^n$ tiene dimensiones racionales, entonces siempre se puede encontrar una partición de R en cubos de igual dimensión, tan pequeña como se quiera, y racional. En otras palabras, dada $\delta > 0$, existe una partición de R en cubos todos de dimensión $\frac{1}{r} \in \mathbb{Q}$ y con $\frac{1}{r} < \delta$. Para ver esto, en lugar de tomar $r = mcm\{q_1, \ldots, q_n\}$,* *tomamos $r = mcm\{q_1, \ldots, q_n\} k$, donde $k \in \mathbb{N}$ es tal que $\frac{1}{k} < mcm\{q_1, \ldots, q_n\} \delta$, entonces nuevamente se tiene que $r\frac{p_i}{q_i} \in \mathbb{Z}$ y $P_i = \left\{x_j^{(i)} = a_i + \frac{j}{r} \mid j \in \left\{0, 1, \ldots, r\frac{p_i}{q_i}, \right\}\right\}$ es una partición uniforme de $[a_i, b_i]$, lo que hace que $P = P_1 \times \cdots \times P_n$ sea una partición de R como se buscaba.*

Lema 4.5 *Si $R_1, \ldots, R_k \subset \mathbb{R}^n$ son rectángulos entonces para toda $\varepsilon > 0$ existe una cantidad finita de cubos $C_1, \ldots, C_l \subset \mathbb{R}^n$ con dimensiones racionales tales que*

$$\bigcup_{j=1}^{k} R_j \subset \bigcup_{i=1}^{l} C_i$$

y

$$0 \leq \sum_{i=1}^{l} V(C_i) - \sum_{j=1}^{k} V(R_j) < \varepsilon$$

Dem. Sea $\varepsilon > 0$, por el lema 2.1 para cada R_j existe un $R_j' \subset \mathbb{R}^n$ rectángulo de dimensiones racionales tal que $R_j \subset R_j'$ y $V\left(R_j'\right) - V(R_j) < \frac{\varepsilon}{k}$. Por el lema 3.1 para cada R_j' existen $C_1^{(j)}, \ldots, C_{l(j)}^{(j)} \subset \mathbb{R}^n$ cubos con la misma dimensión y además racional tales que forman una partición de R_j', de donde $V\left(R_j'\right) = \sum_{i=1}^{l(j)} V\left(C_i^{(j)}\right)$. De manera que

$$\bigcup_{j=1}^{k} R_j \subset \bigcup_{j=1}^{k} \left(\bigcup_{i=1}^{l(j)} C_i^{(j)}\right)$$

y

$$\begin{aligned} \sum_{j=1}^{k} \left[\sum_{i=1}^{l(j)} V\left(C_i^{(j)}\right)\right] - \sum_{j=1}^{k} V(R_j) &= \sum_{j=1}^{k} V\left(R_j'\right) - \sum_{j=1}^{k} V(R_j) \\ = \sum_{j=1}^{k} \left[V\left(R_j'\right) - V(R_j)\right] &< \sum_{j=1}^{k} \frac{\varepsilon}{k} = \varepsilon \end{aligned}$$

Como cada sumando es no negativo se tiene que

$$0 \leq \sum_{j=1}^{k} \left[\sum_{i=1}^{l(j)} V\left(C_i^{(j)}\right)\right] - \sum_{j=1}^{k} V(R_j) < \varepsilon$$

■

Observación 4.3 *Retomando la observación 4.2 podemos decir que dada una colección finita de rectángulos $R_1, \ldots, R_k \subset \mathbb{R}^n$ se tiene que para toda $\varepsilon > 0$ existe una cantidad finita de* *cubos $C_1, \ldots, C_l \subset \mathbb{R}^n$, con la misma dimensión $s \in \mathbb{Q}$, y tan pequeña como se quiera, tales que*

$$\bigcup_{j=1}^{k} R_j \subset \bigcup_{j=1}^{l} C_j$$

y

$$\sum_{i=1}^{l} V(C_i) - \sum_{j=1}^{k} V(R_j) < \varepsilon$$

Esto ya no lo vamos a probar pues resultaría en una copia exacta de la prueba hecha en la observación 4.2.

4.3 Uniformidad de Lipschitz

Otro concepto que necesitaremos es el que se refiere a la llamada condición de uniformidad de Lipschitz de una función de $\mathbb{R}^n$ en $\mathbb{R}^m$. Específicamente, tenemos la siguiente definición.

Definición 4.3 *Decimos que una función $g : A \subset \mathbb{R}^n \to \mathbb{R}^m$ satisface la condición de uniformidad de Lipschitz en A si y sólo si existen $L > 0$ y $\delta > 0$ tales que para cualesquiera $\hat{x}, \hat{y} \in A$ si $\|\hat{x} - \hat{y}\| < \delta$, entonces*

$$\|g(\hat{x}) - g(\hat{y})\| \leq L \|\hat{x} - \hat{y}\|$$

Como el lector habrá notado, el concepto de uniformidad de Lipschitz es muy cercano al concepto de que una función sea de Lipschitz, con la diferencia sutil de que para poder asegurar que $\|g(\hat{x}) - g(\hat{y})\| \leq L \|\hat{x} - \hat{y}\|$ necesitamos que los puntos $\hat{x}$ y $\hat{y}$ sean suficientemente cercanos, cosa que no importa cuando una función sólo es de Lipschitz, convirtiendo a la uniformidad de Lipschitz en una condición más débil; no obstante, sigue siendo un concepto muy útil.

Proposición 4.6 *Si $g : A \subset \mathbb{R}^n \to \mathbb{R}^m$ satisface la condición de uniformidad de Lipschitz en A con $L > 0$ y $\delta > 0$ entonces g es uniformemente continua en A.*

Dem. Dado $\varepsilon > 0$ consideramos $\delta_0 = \min\left\{\delta, \frac{\varepsilon}{L}\right\}$, de manera que dados $\hat{x}, \hat{y} \in A$ tales que $\|\hat{x} - \hat{y}\| < \delta_0$ entonces

$$\|g(\hat{x}) - g(\hat{y})\| \leq L \|\hat{x} - \hat{y}\| < L\frac{\varepsilon}{L} = \varepsilon$$

■

El recíproco de la proposición anterior, no necesariamente es cierto. Veamos un ejemplo.

Ejemplo 4.1 *Sea $f : [0, \infty) \subset \mathbb{R} \to \mathbb{R}$, definida como $f(x) = \sqrt{x}$. f es una función continua en su dominio y por tanto uniformemente continua en el intervalo $[0, 1]$, sin embargo f no cumple la condición de uniformidad de Lipschitz en este intervalo. Para $x, y \in [0, \infty)$, con $x \neq y$, la expresión*

$$|f(x) - f(y)| \leq L |x - y|$$

en donde $L > 0$, es equivalente a la expresión

$$\left| \frac{\sqrt{x} - \sqrt{y}}{x - y} \right| \leq L$$

que a su vez es equivalente a

$$\frac{1}{\sqrt{x} + \sqrt{y}} \leq L$$

Ahora bien, si tomamos a $x \in [0,1]$ tan cercano de $y = 0$ como queramos, entonces los valores $\frac{1}{\sqrt{x}}$ no son valores acotados. Por lo tanto no podemos tener uniformidad de Lipschitz en el intervalo $[0,1]$.

Nuestro próximo paso tiene como objetivo probar el siguiente teorema: Si $g : \Omega \to \mathbb{R}^n$ satisface la condición de uniformidad de Lipschitz en un conjunto compacto $K \subset \Omega$ y $J(K) = 0$, entonces $J(g(K)) = 0$. Para ello, se demostrará un lema, técnico pero muy útil, que nos permite aproximar la medida exterior de la imagen de un conjunto a partir de la medida exterior del conjunto. Aunque parecerá que le estaremos pidiendo demasiadas condiciones al conjunto, éstas sólo dependen de la función con la que estemos trabajando. La importancia de este lema se verá reflejada en todo momento a lo largo de nuestro camino.

Lema 4.6 *Sean $\Omega \subset \mathbb{R}^n$ un conjunto abierto y $K \subset \Omega$ compacto. Si $g : \Omega \to \mathbb{R}^n$ satisface la condición de uniformidad de Lipschitz en K con $L > 0$ y $\delta > 0$, y para alguna $a > 0$*

$$\bar{J}(K) < \frac{a}{4\left(2L\sqrt{n}\right)^n}$$

entonces

$$\bar{J}(g(K)) < a.$$

Dem. Primero nótese que, por la proposición 4.6 $g(K)$ es compacto, de modo que tiene sentido hablar de $\bar{J}(K)$ y $\bar{J}(g(K))$.

Sea $R \subset \mathbb{R}^n$ un rectángulo tal que $K \subset R$, por definición $\bar{J}(K) = \inf\left\{\bar{S}(\mathcal{X}_K, P) \mid P \in \mathcal{P}_R\right\}$. Entonces $\exists P \in \mathcal{P}_R$ tal que

$$\bar{S}(\mathcal{X}_K, P) < \bar{J}(K) + \frac{a}{4\left(2L\sqrt{n}\right)^n} < \frac{a}{2\left(2L\sqrt{n}\right)^n}$$

Sean $R_1, \ldots, R_k \subset R$ los subrectángulos inducidos por P, tales que

$$\bar{S}(\mathcal{X}_K, P) = \sum_{j=1}^{k} V(R_j)$$

es decir, que los R_j son aquellos subrectángulos inducidos por P cuya intersección con K es no vacía; de modo que $K \subset \bigcup_{j=1}^{k} R_j$ y

$$\sum_{j=1}^{k} V(R_j) < \frac{a}{2\left(2L\sqrt{n}\right)^n}$$

Por la observación 4.3 existen $C_1, \ldots, C_l \subset \mathbb{R}^n$ cubos con la misma dimensión $s \in \mathbb{Q}$ tales que $s < \frac{\delta}{\sqrt{n}}$, $\bigcup_{j=1}^{k} R_j \subset \bigcup_{j=1}^{l} C_j$ y

$$0 \leq \sum_{j=1}^{l} V(C_j) - \sum_{j=1}^{k} V(R_j) < \frac{a}{2\left(2L\sqrt{n}\right)^n}$$

de modo que

$$\sum_{j=1}^{l} V(C_j) < \sum_{j=1}^{k} V(R_j) + \frac{a}{2\left(2L\sqrt{n}\right)^n} < \frac{a}{\left(2L\sqrt{n}\right)^n}$$

Nos interesan aquellos cubos cuya intersección con K sea no vacía; supongamos, sin pérdida de generalidad, que estos son $C_1, \ldots, C_r$, en donde $r \leq l$. Nótese que al tratarse posiblemente de una cantidad menor de los l cubos que ya se tenían entonces se cumple que

$$\sum_{j=1}^{r} V(C_j) \leq \sum_{j=1}^{l} V(C_j) < \frac{a}{\left(2L\sqrt{n}\right)^n}$$

Para cada $i \in \{1, \ldots, r\}$ sean $\hat{c}_i \in C_i \cap K$ y $\hat{x} \in C_i \cap K$ arbitrarios; como $\hat{x}, \hat{c}_i \in C_i$ entonces de acuerdo con la proposición 4.5

$$\|\hat{c}_i - \hat{x}\| \leq d(C_i) = s\sqrt{n}$$

pero $s\sqrt{n} < \delta$ así que $\|\hat{c}_i - \hat{x}\| < \delta$ y por tanto $\|g(\hat{c}_i) - g(\hat{x})\| \leq L\|\hat{c}_i - \hat{x}\| \leq Ls\sqrt{n}$; de la observación 4.1 tenemos que $g(\hat{x})$ pertenece a un cubo Γ_i de dimensión $2Ls\sqrt{n}$ y centro $g(\hat{c}_i)$ y como esto ocurre para cada $\hat{x} \in C_i \cap K$, y sabemos que todo cubo queda determinado por su centro y su dimensión, hemos probado que $g(C_i \cap K) \subset \Gamma_i$.

Como los cubos C_i tienen la misma dimensión entonces $V(C_i) = V(C_j)$para cualesquiera $i, j \in \{1, \ldots, r\}$ y de la expresión

$$\sum_{j=1}^{r} V(C_j) < \frac{a}{\left(2L\sqrt{n}\right)^n}$$

se tiene que

$$V(C_j) < \frac{a}{r\left(2L\sqrt{n}\right)^n}$$

y por tanto

$$s^n = V(C_j) < \frac{a}{r\left(2L\sqrt{n}\right)^n}$$

Como Γ_i tiene dimensión $2Ls\sqrt{n}$ tenemos que

$$V(\Gamma_i) = \left(2Ls\sqrt{n}\right)^n = \left(2L\sqrt{n}\right)^n s^n < \left(2L\sqrt{n}\right)^n \frac{a}{r\left(2L\sqrt{n}\right)^n} = \frac{a}{r}$$

y por lo tanto

$$\sum_{i=1}^{r} V(\Gamma_i) < a$$

Finalmente, como $K \subset \bigcup_{i=1}^{r} C_i$ entonces $g(K) \subset \bigcup_{i=1}^{r} \Gamma_i$, de donde

$$\bar{J}(g(K)) \leq \sum_{i=1}^{r} V(\Gamma_i) < a$$

■

Teorema 4.2 *Sean $\Omega \subset \mathbb{R}^n$ un conjunto abierto y $K \subset \Omega$ compacto. Si $g : \Omega \to \mathbb{R}^n$ satisface la condición de uniformidad de Lipschitz en K y $J(K) = 0$ entonces $J(g(K)) = 0$.*

Dem. Si $J(K) = 0$ entonces $\forall \varepsilon > 0$

$$\bar{J}(K) < \frac{\varepsilon}{4\left(2L\sqrt{n}\right)^n}$$

donde $L > 0$ es la constante de la uniformidad de Lipschitz de g sobre K, por el lema 4.6

$$\bar{J}(g(K)) < \varepsilon$$

es decir, $J(g(K)) = 0$ ■

Como se dijo al inicio de este capítulo, en esta primera parte, estamos en busca de dar una condición suficiente para que la imagen de un conjunto Jordan medible siga siendo Jordan medible, y el teorema anterior nos da una pequeña aproximación a la misma. Sabemos que para que un conjunto sea Jordan medible, nos basta con que su frontera tenga medida de Jordan cero, de manera que dado un conjunto Jordan medible D, si podemos encontrar las condiciones que debe cumplir una función g para que $Fr(g(D)) \subset g(Fr(D))$, entonces por el teorema 4.2 prácticamente podremos concluir que $g(D)$ será Jordan medible. Ahora bien, para poder hacer válido el teorema 4.2, de entrada necesitaremos que nuestra función g satisfaga la condición de uniformidad de Lipschitz, y lo que nos gustaría saber entonces, es cuándo una función satisface esta condición. Quizá esto es algo que no podamos decidir de manera general, sin embargo podremos demostrar que sobre conjuntos compactos una buena clase de funciones sí satisface la condición de uniformidad de Lipschitz.

Lema 4.7 *Sean $\Omega \subset \mathbb{R}^n$ un conjunto abierto y $K \subset \Omega$ compacto. Si $g : \Omega \to \mathbb{R}^n$ es de clase C^1 en Ω entonces g satisface la condición de uniformidad de Lipschitz en K.*

Dem. Para cada $\hat{a} \in \Omega$ existe $\delta_{\hat{a}} > 0$ tal que $\overline{B_{\delta_{\hat{a}}}(\hat{a})} \subset \Omega$ pues Ω es abierto. Como $g \in C^1(\Omega)$ entonces por la proposición 1.4

$$c_{\hat{a}} = \sup\left\{\left\|D_g(\hat{z})\right\|_{\pounds} \mid \hat{z} \in \overline{B_{\delta_{\hat{a}}}(\hat{a})}\right\}$$

existe. Consideremos

$$\mathcal{U} = \left\{B_{\frac{\delta_{\hat{a}}}{2}}(\hat{a}) \mid \hat{a} \in K\right\},$$

$\mathcal{U}$ resulta ser una cubierta abierta para K, por lo que existen $\hat{a}_1, \ldots, \hat{a}_l \in K$ tales que $K \subset \bigcup_{i=1}^{l} B_{\delta_i}(\hat{a}_i)$ con $\delta_i = \frac{\delta_{\hat{a}_i}}{2}$, sean $\delta = \min\{\delta_1, \ldots, \delta_l\}$ y $L = \max\{c_{\hat{a}_j} \mid j \in \{1, \ldots, l\}\}$.

Dados $\hat{x}, \hat{y} \in K$ tales que $\|\hat{x} - \hat{y}\| < \delta$ existe $\hat{a}_j \in K$ tal que $\hat{x} \in B_{\delta_j}(\hat{a}_j)$, entonces

$$\|\hat{y} - \hat{a}_j\| \leq \|\hat{y} - \hat{x}\| + \|\hat{x} - \hat{a}_j\| < \delta + \delta_j \leq 2\delta_j = \delta_{\hat{a}_j}$$

es decir, $\hat{x}, \hat{y} \in B_{\delta\hat{a}_j}(\hat{a}_j)$. De la convexidad de $B_{\delta\hat{a}_j}(\hat{a}_j)$ se tiene que para toda $t \in [0,1]$

$$\hat{x}_t = (1-t)\,\hat{x} + t\hat{y}$$

pertenece a $B_{\delta\hat{a}_j}(\hat{a}_j) \subset \Omega$, de modo que por la proposición 4.2

$$\|g(\hat{x}) - g(\hat{y})\| \leq c\,\|\hat{x} - \hat{y}\|$$

donde $c = \sup\limits_{t\in[0,1]} \|D_g(\hat{x}_t)\|_{\mathcal{L}}$, y como $\hat{x}_t \in B_{\delta\hat{a}_j}(\hat{a}_j)$ para toda $t \in [0,1]$ se tiene que $c \leq c_{\hat{a}_j}$ y $c_{\hat{a}_j} \leq L$, por tanto

$$\|g(\hat{x}) - g(\hat{y})\| \leq L\,\|\hat{x} - \hat{y}\|$$

■

Dentro de la prueba del lema 4.7 se nota la importancia de la hipótesis de que la función g sea de clase C^1, ya que gracias a esta hipótesis es que se consigue a una constante de Lipschitz. Sin embargo, una pregunta interesante es: ¿qué tan necesaria es dicha hipótesis?

En el ejemplo 4.1 vimos que, no es suficiente la continuidad de una función para poder asegurar que se cumple la condición de uniformidad de Lipschitz. Ahora veamos un ejemplo de una función derivable en todo su dominio, pero no de clase C^1, para la cual tampoco se cumple la condición de uniformidad de Lipschitz. Para ello, demostremos antes una propiedad que nos ayudará a exhibir lo que queremos.

Proposición 4.7 *Sean $A \subset \mathbb{R}$ un conjunto abierto y $B \subset A$. Si $f : A \subset \mathbb{R} \to \mathbb{R}$ es una función derivable en A y cumple la condición de uniformidad de Lipschitz en B, entonces f' es acotada en B.*

Dem. Sean $L > 0$ y $\delta > 0$ tales que para cualesquiera $x, y \in B$ se cumpla que si $|x - y| < \delta$, entonces

$$|f(x) - f(y)| \leq L\,|x - y|$$

Sea $x_0 \in B$, como para toda $x \in (x_0 - \delta, x_0 + \delta) \cap B$, se cumple que

$$|f(x) - f(x_0)| \leq L\,|x - x_0|$$

entonces para toda $x \in (x_0 - \delta, x_0 + \delta) \cap B$, con $x \neq x_0$, se tiene que

$$\left|\frac{f(x) - f(x_0)}{x - x_0}\right| \leq L$$

de donde

$$\lim_{x\to x_0}\left|\frac{f(x) - f(x_0)}{x - x_0}\right| \leq L$$

es decir, $|f'(x_0)| \leq L$. Por lo tanto f' está acotada en B por el valor positivo L ∎

Con la proposición anterior, la tarea de encontrar una función dervible pero para la cual no se cumpla la condición de uniformidad de Lipschitz, se vuelve un poco más sencilla. Sólo debemos encontrar un ejemplo de una función derivable pero cuya derivada no sea acotada.

Ejemplo 4.2 *Sea $f : \mathbb{R} \to \mathbb{R}$ definida como*

$$f(x) = \begin{cases} x\sqrt{|x|}sen\left(\frac{1}{x}\right) & si \quad x \neq 0 \\ \\ 0 & si \quad x = 0 \end{cases}$$

f es claramente derivable para toda $x \neq 0$ por involucrar a un producto de funciones derivables; mientras que para $x = 0$ la derivabilidad la obtenemos por definición, es decir,

$$f'(0) = \lim_{x \to 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x \to 0} \sqrt{|x|}sen\left(\frac{1}{x}\right) = 0$$

Por lo tanto f es derivable en todo $\mathbb{R}$. Ahora ubiquémonos en el intervalo $[0, 1]$, en este intervalo tenemos que

$$f'(x) = \begin{cases} \sqrt{x}sen\left(\frac{1}{x}\right) - \frac{1}{\sqrt{x}}\cos\left(\frac{1}{x}\right) & \text{si} \quad x \neq 0 \\ \\ 0 & \text{si} \quad x = 0 \end{cases}$$

Para cada $n \in \mathbb{N}$, tomemos a $x_n = \frac{1}{(2n+1)\pi} \in [0, 1]$, y notemos que

$$f'(x_n) = \sqrt{x_n}sen\left((2n+1)\pi\right) - \frac{1}{\sqrt{x_n}}\cos\left((2n+1)\pi\right) = \sqrt{(2n+1)\pi}$$

de manera que el conjunto de valores $\left\{ \acute{f}'(x_n) \mid n \in \mathbb{N} \right\}$ no es un conjunto acotado. De modo que f' no es acotada en el intervalo $[0, 1]$. De acuerdo con la proposición anterior, concluimos que f no satisface la condición de uniformidad de Lipschitz en el conjunto compacto $[0, 1]$.

Con este ejemplo, se puede ver con más fuerza la importancia de trabajar con una función de clase C^1.

Regresemos al problema que nos interesa. Como ya se había discutido; para que la imagen de un conjunto Jordan medible siga siendo Jordan medible, nos es suficiente que la función satisfaga la condición de uniformidad de Lipschitz junto con el hecho de que $Fr(g(D)) \subset g(Fr(D))$, el lema 4.7 nos resuelve el primer punto, quedando entonces el detalle de poder asegurar que $Fr(g(D)) \subset g(Fr(D))$. Afortunadamente el teorema de la función inversa nos dará solución a este problema.

Lema 4.8 *Sea $\Omega \subset \mathbb{R}^n$ un conjunto abierto, sea $D \subset \mathbb{R}^n$ un conjunto acotado tal que $D \cup Fr(D) \subset \Omega$. Si $g : \Omega \to \mathbb{R}^n$ es de clase C^1 en Ω y $J_g(\hat{x}) \neq 0$ para toda $\hat{x} \in int(D)$ entonces $Fr(g(D)) \subset g(Fr(D))$.*

Dem. Dado que $g(D) \subset g(D \cup Fr(D))$ y $g(D \cup Fr(D))$ es compacto por la continuidad de g, tenemos que $Fr(g(D)) \subset g(D \cup Fr(D))$. Sea $\hat{\xi} \in Fr(g(D))$, entonces existe $\hat{x}_0 \in D \cup Fr(D)$ tal que $g(\hat{x}_0) = \hat{\xi}$, como $D \cup Fr(D) = int(D) \cup Fr(D)$ entonces $\hat{x}_0 \in int(D) \cup Fr(D)$; supongamos que $\hat{x}_0 \in int(D)$, por hipótesis $g \in C^1(\Omega)$ y $J_g(\hat{x}) \neq 0$ para toda $\hat{x} \in int(D)$, entonces por el Teorema de la función inversa existe $N \subset D$ abierto tal que $\hat{x}_0 \in N$ y $g(N)$ es abierto, pero $\hat{\xi} = g(\hat{x}_0) \in g(N) \subset g(D)$, es decir, $\hat{\xi}$ es un punto interior de $g(D)$ lo que contradice nuestra elección de $\hat{\xi}$.

Por lo tanto $\hat{x}_0 \in Fr(D)$ y por lo tanto $\hat{\xi} \in g(Fr(D))$, lo que prueba que $Fr(g(D)) \subset g(Fr(D))$. ■

Teorema 4.3 *Sea $\Omega \subset \mathbb{R}^n$ abierto, sea $D \subset \mathbb{R}^n$ un conjunto Jordan medible tal que $D \cup Fr(D) \subset \Omega$. Si $g : \Omega \to \mathbb{R}^n$ es de clase C^1 en Ω y $J_g(\hat{x}) \neq 0$ para toda $\hat{x} \in int(D)$ entonces $g(D)$ es Jordan medible.*

Dem. Primero veamos que $J(g(Fr(D))) = 0$. Como D es Jordan medible entonces $J(Fr(D)) = 0$, pero $Fr(D) \subset \Omega$ es compacto, así que por el lema 4.7 g satisface la condición de uniformidad de Lipschitz en $Fr(D)$ y por el Teorema 4.2 $J(g(Fr(D))) = 0$.

Ahora, $D \cup Fr(D)$ es compacto y por tanto $g(D \cup Fr(D))$ también lo es, de manera que $g(D)$ resulta ser acotado ya que $g(D) \subset g(D \cup Fr(D))$, por lo que resulta válido preguntarnos si $g(D)$ es Jordan medible.

Sabemos que todo subconjunto de un conjunto de medida de Jordan cero tiene medida Jordan cero, y como por el lema anterior sabemos que $Fr(g(D)) \subset g(Fr(D))$ entonces $J(Fr(g(D))) = 0$ y por lo tanto $g(D)$ es Jordan medible. ■

4.4 El caso lineal

Lo siguiente en nuestro camino es poder demostrar que el teorema de cambio de variable es válido para transformaciones lineales, en el siguiente sentido: si $D \subset \mathbb{R}^n$ es un conjunto Jordan medible y $T : \mathbb{R}^n \to \mathbb{R}^n$ es una transformación lineal, entonces $T(D)$ es Jordan medible y

$$J(T(D)) = |\det L| \, J(D)$$

donde L es la representación matricial de T respecto a la base canónica. La fórmula anterior es válida sin importar la transformación lineal; recordemos que una transformación lineal de $\mathbb{R}^n$ en $\mathbb{R}^n$ puede ser singular o no singular, lo que en términos de su representación matricial, significa que el determinante de la misma es cero o diferente de cero respectivamente. La base para probar el resultado para una transformación lineal no singular está en el hecho de que este tipo de transformaciones puede descomponerse en una cantidad finita de transformaciones elementales (ver [2]), por lo que demostrando el resultado para transformaciones elementales la prueba para una transformación no singular arbitraria es casi inmediata. Sabemos que hay dos tipos de transformaciones elementales:

1) La multiplicación por escalares en una coordenada $E_1 : \mathbb{R}^n \to \mathbb{R}^n$

$$E_1(x_1, \ldots, x_n) = (x_1, \ldots, \lambda x_k, \ldots, x_n)$$

para alguna $\lambda \in \mathbb{R} \setminus \{0\}$.

2) Suma de un múltiplo de una coordenada a otra $E_2 : \mathbb{R}^n \to \mathbb{R}^n$

$$E_2(x_1, \ldots, x_n) = (x_1, \ldots, x_k + \lambda x_j, \ldots, x_n)$$

para alguna $\lambda \in \mathbb{R} \setminus \{0\}$.

Por comodidad, de aquí en adelante cuando hablemos de la representación matricial de una transformación lineal, consideraremos que ésta estará dada respecto a la base canónica; por lo que simplemente la llamaremos la representación matricial de la transformación.

Lema 4.9 *Si $R \subset \mathbb{R}^n$ es un rectángulo, con $R = [a_1, b_1] \times \ldots \times [a_n, b_n]$, y $T : \mathbb{R}^n \to \mathbb{R}^n$ es una transformación lineal elemental, entonces $T(R)$ es Jordan medible y*

$$J(T(R)) = |\det L| \, V(R)$$

donde L es la representación matricial de T.

Dem. Primero supongamos que T es una transformación elemental del tipo uno, es decir, $T(x_1, \ldots, x_n) = (x_1, \ldots, \lambda x_k, \ldots, x_n)$ para alguna $\lambda \in \mathbb{R}$. Entonces

$$L = \begin{pmatrix} \hat{e}_1 & \cdots & \hat{e}_{k-1} & \lambda \hat{e}_k & \hat{e}_{k+1} & \cdots & \hat{e}_n \end{pmatrix}$$

donde $\hat{e}_j$ es el j-ésimo vector canónico (visto como columna), de donde $\det L = \lambda$. Por otro lado, si $\lambda > 0$ entonces

$$T(R) = [a_1, b_1] \times \ldots \times [\lambda a_k, \lambda b_k] \times \ldots \times [a_n, b_n]$$

mientras que si $\lambda < 0$ entonces

$$T(R) = [a_1, b_1] \times \ldots \times [\lambda b_k, \lambda a_k] \times \ldots \times [a_n, b_n]$$

en ambos casos $T(R)$ es nuevamente un rectángulo y por lo tanto es Jordan medible y

$$J(T(R)) = V(T(R)) = \begin{cases} \lambda V(R) & \lambda > 0 \\ -\lambda V(R) & \lambda < 0 \end{cases}$$

es decir, $J(T(R)) = |\lambda| \, V(R)$.

Supongamos ahora que T es una transformación elemental del tipo dos, es decir,

$$T(x_1, \ldots, x_n) = (x_1, \ldots, x_k + \lambda x_j, \ldots, x_n)$$

si $\hat{x} = (x_1, \ldots, x_n)$ entonces otra manera de ver a T es

$$T(\hat{x}) = \hat{x} + \lambda x_j \hat{e}_k$$

de modo que su representación matricial es

$$L = \begin{pmatrix} \hat{e}_1 & \cdots & \hat{e}_{j-1} & \hat{e}_j + \lambda \hat{e}_k & \hat{e}_{j+1} & \cdots & \hat{e}_n \end{pmatrix}$$

donde nuevamente los vectores están vistos como columnas. Entonces L es una matriz triangular superior si $j > k$ y una matriz triangular inferior si $j < k$; en ambos casos se tiene que $\det L = 1$. A diferencia de una transformación elemental tipo uno, la imagen de un rectángulo bajo una transformación lineal tipo dos ya no es un rectángulo (en $\mathbb{R}^3$ se puede ver geométricamente que, esta imagen es un paralelepípedo); la imagen la podemos describir de la siguiente manera, si $\alpha, \beta : [a_j, b_j] \to \mathbb{R}$ están dadas por $\alpha(t) = a_k + \lambda t$ y $\beta(t) = b_k + \lambda t$ entonces

$$T(R) = \left\{ (x_1, \ldots, x_n) \in \mathbb{R}^n \mid \ a_i \leq x_i \leq b_i \quad \forall i \neq k, \quad \alpha(x_j) \leq x_k \leq \beta(x_j) \ \right\}$$

Entonces por el teorema de Fubini, la función constante 1 es integrable sobre $T(R)$ (es decir, $T(R)$ es Jordan medible) y

$$\begin{aligned} J(T(R)) &= \int_{T(R)} 1 = \left(\int_{a_j}^{b_j} \left(\int_{\alpha(x_j)}^{\beta(x_j)} 1 dx_k \right) dx_j \right) \prod_{i \neq k,j} (b_i - a_i) \\ &= \left(\int_{a_j}^{b_j} (b_k - a_k) \, dx_j \right) \prod_{i \neq k,j} (b_i - a_i) \\ &= \prod_{i=1}^{n} (b_i - a_i) = V(R) \end{aligned}$$

■

Nuestro objetivo es poder demostrar que $J(T(D)) = |\det L| \, J(D)$ para un conjunto Jordan medible D arbitrario y una transformación elemental T, para después poder generalizarlo a una transformación no singular cualquiera, partiendo del hecho de que una transformación no singular T es la composición de transformaciones elementales, es decir, $T = T_1 \circ T_2 \circ \cdots \circ T_k$. En el lema anterior pudimos describir sin mayor problema como era la imagen de un rectángulo bajo una transformación elemental y auxiliarnos de ello para concluir el resultado, sin embargo para un conjunto Jordan medible arbitrario no es tan sencillo ver qué efecto tiene una de estas transformaciones (su imagen puede ser difícil de describir) así que no podremos actuar de la misma manera, razón por la que tendremos que darle la vuelta a este incoveniente.

Lema 4.10 *Si $D \subset \mathbb{R}^n$ es un conjunto Jordan medible, y $T : \mathbb{R}^n \to \mathbb{R}^n$ es una transformación lineal elemental, entonces $T(D)$ es Jordan medible y*

$$J(T(D)) = |\det L| \, J(D)$$

donde L es la representación matricial de T.

Dem. Sabemos que una transformación lineal es de clase C^1, y al ser elemental tenemos que $\det L \neq 0$, entonces por el teorema 4.3 ya se tiene que $T(D)$ es un conjunto Jordan medible. Tomemos $R \subset \mathbb{R}^n$ rectángulo tal que $D \subset R$. A fin de demostrar la igualdad, demostraremos que para toda $\varepsilon > 0$

$$|J(T(D)) - |\det L| \, J(D)| < \varepsilon$$

Sea $\varepsilon > 0$; como D es Jordan medible, existe $P \in \mathcal{P}_R$ tal que

$$\bar{S}(\mathcal{X}_R, P) - J(D) < \frac{\varepsilon}{|\det L|}$$

y

$$J(D) - \underline{S}(\mathcal{X}_R, P) < \frac{\varepsilon}{|\det L|}$$

Como dijimos, al tratarse de una transformación elemental tenemos que $\det L \neq 0$, por lo que es válido dividir entre él.

Sean $Q_1, \ldots, Q_k$ los subrectángulos generados por P, y sean $A_1 = \{j \in \{1, \ldots, k\} \mid Q_j \subset D\}$ y $A_2 = \{j \in \{1, \ldots, k\} \mid Q_j \cap D \neq \emptyset\}$, llamemos

$$G = \bigcup_{j \in A_1} Q_j$$

y

$$B = \bigcup_{j \in A_2} Q_j$$

Como T es elemental, es inyectiva y entonces los conjuntos $T(Q_1), \ldots, T(Q_k)$ forman una subdivisión de $T(R)$ en conjuntos Jordan medibles, es decir, $T(R) = \bigcup_{j=1}^{k} T(Q_j)$ y que además satisfacen que

$$int(T(Q_j)) \cap int(T(Q_i)) = \emptyset$$

para $i \neq j$, de manera que por el corolario 3.1 $T(G)$ y $T(B)$ son conjuntos Jordan medibles y

$$J(T(G)) = \sum_{j \in A_1} J(T(Q_j))$$

y

$$J(T(B)) = \sum_{j \in A_2} J(T(Q_j))$$

pero por el lema 4.9 para cada Q_j se cumple que

$$J(T(Q_j)) = |\det L| \, J(Q_j)$$

de modo que

$$\begin{aligned} J(T(G)) &= \sum_{j \in A_1} J(T(Q_j)) = \sum_{j \in A_1} |\det L| \, J(Q_j) = |\det L| \sum_{j \in A_1} J(Q_j) \\ &= |\det L| \, \underline{S}(\mathcal{X}_R, P) > |\det L| \left(J(D) - \frac{\varepsilon}{|\det L|} \right) \end{aligned}$$

es decir, $J(T(G)) > |\det L| \, J(D) - \varepsilon$, así mismo

$$\begin{aligned} J(T(B)) &= \sum_{j \in A_2} J(T(Q_j)) = |\det L| \sum_{j \in A_2} J(Q_j) \\ &= |\det L| \, \bar{S}(\mathcal{X}_R, P) < |\det L| \left(J(D) + \frac{\varepsilon}{|\det L|} \right) \end{aligned}$$

es decir, $J(T(B)) < |\det L| J(D) + \varepsilon$.

Por construcción se cumple que $G \subset D \subset B$ y por lo tanto $T(G) \subset T(D) \subset T(B)$, y como cada conjunto es Jordan medible, entonces

$$J(T(G)) \leq J(T(D)) \leq J(T(B))$$

de donde

$$|\det L| J(D) - \varepsilon < J(T(D)) < |\det L| J(D) + \varepsilon$$

es decir, $|J(T(D)) - |\det L| J(D)| < \varepsilon$ ■

Teorema 4.4 *Si $D \subset \mathbb{R}^n$ es un conjunto Jordan medible, y $T : \mathbb{R}^n \to \mathbb{R}^n$ es una transformación lineal no singular, entonces $T(D)$ es Jordan medible y*

$$J(T(D)) = |\det L| J(D)$$

donde L es la representación matricial de T.

Dem. Por el teorema 4.3 ya sabemos que $T(D)$ es Jordan medible. Como T es no singular existen $T_1, \ldots, T_k : \mathbb{R}^n \to \mathbb{R}^n$ transformaciones elementales tales que $T = T_k \circ T_{k-1} \circ \cdots \circ T_1$ entonces $T(R) = T_k(T_{k-1}(\cdots(T_1(D))\cdots))$ donde cada $T_j(\cdots(T_1(D))\cdots)$ es Jordan medible; sean $L_1, \ldots, L_k$ las representaciones matriciales de cada T_j, por el lema 4.10 tenemos que para cada $j \in \{1, \ldots, k\}$

$$J(T_j(\cdots(T_1(D))\cdots)) = |\det L_j| J(T_{j-1}(\cdots(T_1(D))\cdots))$$

de donde $J(T(D)) = |\det L_k| \cdots |\det L_1| J(D)$, pero $|\det L| = |\det L_k| \cdots |\det L_1|$, por lo tanto $J(T(D)) = |\det L| J(D)$ ■

Previo al lema 4.9 mencionamos que la fórmula $J(T(D)) = |\det L| J(D)$ es válida para cualquier transformación lineal, lo que incluye a las transformaciones singulares, y aunque en la práctica (de la integración) uno difícilmente se cruza con este tipo de transformaciones lineales, vale la pena argumentar por qué la fórmula propuesta sigue siendo válida. Con este fin probaremos primero el siguiente resultado.

Proposición 4.8 *Si para cualquier $m < n$ definimos el conjunto*

$$P_m = \left\{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid \ x_i = 0 \quad \forall i \in \{m+1, \ldots, n\} \ \right\}$$

entonces P_m tiene medida de Lebesgue cero.

Dem. Cabe notar que P_m no es otra cosa más que un subespacio vectorial de $\mathbb{R}^n$ isomorfo a $\mathbb{R}^m$, en otras palabras lo que esta proposición afirma es que $\mathbb{R}^m$ visto como subespacio de $\mathbb{R}^n$ tiene medida de Lebesgue cero en $\mathbb{R}^n$. La prueba es en realidad muy sencilla, consideremos al conjunto $P_m \cap \mathbb{Z}^n$, es decir, el conjunto de puntos de P_m con coordenadas enteras, este conjunto es numerable, es decir, $P_m \cap \mathbb{Z}^n = \left\{\hat{y}^{(i)} \mid i \in \mathbb{N}\right\}$.

Dado $\varepsilon > 0$, para cada $\hat{y}^{(i)} \in P_m \cap \mathbb{Z}^n$ con $\hat{y}^{(i)} = \left(y_1^{(i)}, \ldots, y_m^{(i)}, 0, \ldots, 0\right)$, definimos el rectángulo

$$R_i = \left[y_1^{(i)}, y_1^{(i)} + 1\right] \times \cdots \times \left[y_m^{(i)}, y_m^{(i)} + 1\right] \times [0, \delta_i] \times \underset{n-m}{\cdots} \times [0, \delta_i]$$

en donde

$$\delta_i = \left(\frac{\varepsilon}{2^{i+1}}\right)^{\frac{1}{n-m}}$$

Esta colección de rectangulos R_i es numerable, porque el conjunto $P_m \cap \mathbb{Z}^n$ lo es, además

$$V(R_i) = (\delta_i)^{n-m} = \frac{\varepsilon}{2^{i+1}}$$

y por tanto $\sum_{i=1}^{\infty} V(R_i) = \frac{\varepsilon}{2} < \varepsilon$, de manera que si

$$P_m \subset \bigcup_{i=1}^{\infty} R_i$$

entonces habremos probado que P_m tiene medida de Lebesgue cero.

Dado $\hat{x} = (x_1, \ldots, x_m, 0, \ldots, 0) \in P_m$, existen $n_1, \ldots, n_m \in \mathbb{Z}$ tales que $n_j \leq x_j \leq n_j + 1$ para toda $j \in \{1, \ldots, m\}$. Sea $i \in \mathbb{N}$ tal que $\hat{y}^{(i)} = (n_1, \ldots, n_m, 0, \ldots, 0) \in P_m \cap \mathbb{Z}^n$ y R_i el rectángulo correspondiente a este punto. Como $n_j \leq x_j \leq n_j + 1$ para toda $j \in \{1, \ldots, m\}$, es claro que $\hat{x} \in R_i$ y por tanto que

$$P_m \subset \bigcup_{i=1}^{\infty} R_i$$

de donde concluimos que $\lambda(P_m) = 0$ ■

Lo hecho en esta proposición, es la clave para demostar que la identidad $J(T(D)) = |\det L|\, J(D)$ es válida aún cuando T es singular. En este caso, al ser T singular, se tiene que $\det L = 0$, de forma que lo que se tratará de demostrar es que si T es singular entonces $J(T(D)) = 0$; es aquí donde entra la proposición anterior, porque $T(\mathbb{R}^n)$ es un subespacio de $\mathbb{R}^n$ de dimensión $m < n$, y por lo tanto $T(\mathbb{R}^n)$ será isomorfo al conjunto P_m; como este conjunto tiene medida de Lebesgue cero, entonces de acuerdo con la proposición 3.2, todo subconjunto compacto tendrá medida de Jordan cero. Hagamos esto con un poco más de cuidado.

Lema 4.11 *Si $T : \mathbb{R}^n \to \mathbb{R}^n$ es una transformación lineal singular tal que $T(\mathbb{R}^n)$ es un subespacio de $\mathbb{R}^n$ de dimensión $m < n$, entonces existe un isomorfismo $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ tal que*

$$\varphi(T(\mathbb{R}^n)) = P_m$$

Dem. Si $\{\hat{v}_1, \ldots, \hat{v}_m\}$ es una base para $T(\mathbb{R}^n)$, entonces podemos extenderla a una base para $\mathbb{R}^n$, $\{\hat{v}_1, \ldots, \hat{v}_m, \hat{v}_{m+1}, \ldots, \hat{v}_n\}$; como toda transformación lineal queda determinada a partir de la base, definamos $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ como

$$\varphi(\hat{v}_j) = \hat{e}_j$$

donde $\hat{e}_j$, como siempre, representa al j-ésimo vector canónico. Entonces φ definida de esta manera es un isomorfismo de $\mathbb{R}^n$ en sí mismo, y dado que $\{\hat{e}_1, \ldots, \hat{e}_m\}$ es una base para P_m, se concluye que $\varphi(T(\mathbb{R}^n)) = P_m$. ■

Proposición 4.9 *Si $D \subset \mathbb{R}^n$ es un conjunto Jordan medible y $T : \mathbb{R}^n \to \mathbb{R}^n$ es una transformación lineal singular, entonces $T(D)$ es Jordan medible y*

$$J(T(D)) = 0$$

Dem. Al ser T singular, como se dijo, significa que $\det L = 0$. De manera que $T(\mathbb{R}^n)$ es un subespacio de $\mathbb{R}^n$ de dimensión $m < n$ y por tanto existe un isomorfismo $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ tal que

$$\varphi(T(\mathbb{R}^n)) = P_m$$

Como D es Jordan medible entonces el compacto $D \cup Fr(D)$ también es Jordan medible. Ahora bien, $\varphi(T(D \cup Fr(D)))$ es un subconjunto compacto de P_m y como éste tiene medida de Lebesgue cero, se tiene por la proposición 3.2 que $\varphi(T(D \cup Fr(D)))$ es Jordan medible en $\mathbb{R}^n$ y además

$$J(\varphi(T(D \cup Fr(D)))) = 0$$

Por otra parte, como φ es un isomorfismo de $\mathbb{R}^n$, eso significa que es una transformación lineal no singular, lo que implica que $\varphi^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ también es una transformación lineal no singular y por lo tanto $\varphi^{-1}(\varphi(T(D \cup Fr(D))))$ es un conjunto Jordan medible, es decir, $T(D \cup Fr(D))$ es Jordan medible; pero $\varphi(T(D \cup Fr(D)))$ es un conjunto compacto de medida de Jordan cero, entonces por el teorema 4.2 concluímos que

$$J\left(\varphi^{-1}(\varphi(T(D \cup Fr(D))))\right) = 0$$

es decir, $J(T(D \cup Fr(D))) = 0$ y por lo tanto $J(T(D)) = 0$ ■

Con esto damos por terminada la segunda parte de este largo camino para demostrar el teorema de cambio de variable, y abriremos paso ahora sí, a este importante teorema.

4.5 El caso general

El teorema de cambio de variable, como ya todos sabemos, nos da una poderosa herramienta para calcular integrales en varias variables, cuando éstas están definidas sobre conjuntos no tan simples como los rectángulos. Si nuestra región de integración puede verse como la imagen de un conjunto Jordan medible bajo una función inyectiva y de clase C^1 en este conjunto, entonces podemos cambiar nuestra región de integración a este conjunto Jordan medible, que en principio, debería resultar más accesible para trabajar. Como mencionamos al principio de este capítulo, bajo estas condiciones el teorema afirma que

$$\int_{g(D)} f = \int_D (f \circ g)\,|J_g|$$

donde D es el conjunto Jordan medible, g es la función inyectiva y de clase C^1 en D, y $f : g(D) \to \mathbb{R}$ es una función continua y acotada en $g(D)$.

Al igual que lo sucedido cuando tratamos de deducir la fórmula para transformaciones lineales en donde nos fue útil comenzar con rectángulos, aquí también tendremos que trabajar con rectángulos, es decir, como todo conjunto Jordan medible puede aproximarse mediante particiones de rectángulos, analizaremos la medida de Jordan de $g(R)$, con R rectángulo suficientemente pequeño, a partir de la medida de Jordan de $D_g(\hat{x})(R)$ para alguna $\hat{x} \in R$. La idea central de este razonamiento es que, entre más pequeño sea el rectángulo R, si $\hat{x} \in R$ entonces $D_g(\hat{x})$ se "parece" mucho a la función g en R, de tal forma que la medida de Jordan del conjunto $D_g(\hat{x})(R)$ se "parece" mucho a la medida de Jordan del conjunto $g(R)$. Para lograr este objetivo, primero lo haremos para una clase especial de rectángulos, los cubos; por su sencillez y fácil caracterización, nos resultará muy conveniente iniciar con ellos.

Lema 4.12 *Sea $\Omega \subset \mathbb{R}^n$ un conjunto abierto y $K \subset \Omega$ compacto. Si $g : \Omega \to \mathbb{R}^n$ es de clase C^1 en Ω, $J_g(\hat{x}) \neq 0$ para toda $\hat{x} \in K$ y además g es inyectiva en $int(K)$ entonces para toda $\varepsilon \in (0,2)$ existe $\delta > 0$ tal que para cualquier $C \subset int(K)$ cubo de dimensión $s < \delta$ y centro en $\hat{a}$ se tiene que*

$$\left(1 - \frac{\varepsilon}{2}\right) |J_g(\hat{a})| V(C) \leq J(g(C)) \leq \left(1 + \frac{\varepsilon}{2}\right) |J_g(\hat{a})| V(C)$$

Dem. Antes de comenzar la prueba vale la pena detenernos un poco a analizar la desigualdad propuesta en este lema. Primero nótese que la desigualdad puede ser escrita como

$$\left| \frac{J(g(C))}{|J_g(\hat{a})| V(C)} - 1 \right| < \frac{\varepsilon}{2}$$

Esto nos dice que, para cubos suficientemente pequeños, las cantidades $J(g(C))$ y $|J_g(\hat{a})| V(C)$ se parecen mucho en proporción; lo que nos da un poco de luz en cuanto a la fórmula de cambio de variable si consideramos que,

$$\int_{g(D)} f \approx \sum_i (f \circ g)(\hat{x}_i) J(g(R_i))$$

y

$$\int_D (f \circ g) |J_g| \approx \sum_i (f \circ g)(\hat{x}_i) |J_g(\hat{x}_i)| V(R_i)$$

en donde los rectángulos R_i forman una partición de algún rectángulo que contenga a D. Ahora bien, de acuerdo con el teorema 4.4

$$|J_g(\hat{a})| V(C) = J(D_g(\hat{a})(C))$$

de manera que la desigualdad también la podemos ver como

$$\left| \frac{J(g(C))}{J(D_g(\hat{a})(C))} - 1 \right| < \frac{\varepsilon}{2}$$

Desigualdad que, como dijimos previo al lema, muestra que para cubos suficientemente pequeños centrados en $\hat{a}$, la medida de Jordan de $D_g(\hat{a})(C)$ se "parece" mucho a la medida de Jordan del conjunto $g(C)$. Una manera equivalente de exhibir esta idea es considerar que, al tratarse de un cubo pequeño centrado en $\hat{a}$ y el hecho de que la función $D_g(\hat{a})$ se parece mucho a la función g en puntos póximos a $\hat{a}$, parece natural que la función $[D_g(\hat{a})]^{-1} \circ g$ no altere mucho a la medida de Jordan de C. En otras palabras, apoyados del teorema 4.4 obtenemos que

$$J\left([D_g(\hat{a})]^{-1} \circ g(C)\right) = \frac{1}{|J_g(\hat{a})|} J(g(C))$$

de manera que la desigualdad

$$\left| \frac{J(g(C))}{J(D_g(\hat{a})(C))} - 1 \right| < \frac{\varepsilon}{2}$$

también puede ser vista como

$$\left| \frac{J\left([D_g(\hat{a})]^{-1} \circ g(C)\right)}{V(C)} - 1 \right| < \frac{\varepsilon}{2}$$

Planteamos esta nueva expresión para propósitos de la prueba, ya que nos resultará un poco sencillo deducir esta última desigualdad y a partir de ésta concluir lo que queremos.

Ahora sí iniciemos la prueba. Comencemos notando que por el teorema 4.3, tiene sentido hablar de $J(g(C))$.

Como $J_g(\hat{x}) \neq 0$ para toda $\hat{x} \in K$ entonces por el lema 4.3 existe $\mu > 0$ tal que para toda $\hat{x} \in K$ y para toda $\hat{u} \in \mathbb{R}^n$

$$\mu \|\hat{u}\| \leq \|D_g(\hat{x})(\hat{u})\|$$

Dado $\hat{x} \in K$ y $\hat{v} \in \mathbb{R}^n$ arbitrario, podemos hacer $\hat{u} = [D_g(\hat{x})]^{-1}(\hat{v})$ (la inversa existe porque $J_g(\hat{x}) \neq 0$); entonces

$$\mu \left\| [D_g(\hat{x})]^{-1}(\hat{v}) \right\| \leq \|\hat{v}\|$$

De manera que si hacemos $M = \frac{1}{\mu}$, se tiene que para toda $\hat{x} \in K$ y para toda $\hat{v} \in \mathbb{R}^n$

$$\left\| [D_g(\hat{x})]^{-1}(\hat{v}) \right\| \leq M \|\hat{v}\| \tag{4.1}$$

Sea $\varepsilon \in (0, 2)$ y consideremos

$$0 < \varepsilon_1 < \frac{1}{M\sqrt{n}} \min\left\{1 - \left(1 - \frac{\varepsilon}{2}\right)^{\frac{1}{n}}, \left(1 + \frac{\varepsilon}{2}\right)^{\frac{1}{n}} - 1\right\}$$

en particular $\varepsilon_1 < \frac{1}{M\sqrt{n}}$.

De nuestra elección de ε_1 también se tiene que

$$\left(1 - M\varepsilon_1\sqrt{n}\right)^n > 1 - \frac{\varepsilon}{2} \tag{4.2}$$

y

$$\left(1 + M\varepsilon_1\sqrt{n}\right)^n < 1 + \frac{\varepsilon}{2} \tag{4.3}$$

Por el lema 4.2 existe $\delta_1 > 0$ tal que para toda $\hat{a} \in K$

$$\|g(\hat{x}) - g(\hat{a}) - D_g(\hat{a})(\hat{x} - \hat{a})\| \leq \varepsilon_1 \|\hat{x} - \hat{a}\| \tag{4.4}$$

siempre que $\hat{x} \in B_{\delta_1}(\hat{a}) \cap K$. Definamos $\delta > 0$ como

$$\delta = \frac{2\delta_1}{\sqrt{n}}$$

y tomemos $s < \delta$; sea $C \subset int(K)$ un cubo con centro en $\hat{a}$ y dimensión s, tenemos que para toda $\hat{x} \in C$

$$\|\hat{x} - \hat{a}\| \leq \frac{s}{2}\sqrt{n} < \delta_1$$

y por lo tanto para toda $\hat{x} \in C$

$$\|g(\hat{x}) - g(\hat{a}) - D_g(\hat{a})(\hat{x} - \hat{a})\| \leq \varepsilon_1 \|\hat{x} - \hat{a}\|$$

Sea $h : \Omega \to \mathbb{R}^n$ definida por $h(x) = [D_g(\hat{a})]^{-1}(g(\hat{x}))$; nótese que h es inyectiva en $int(K)$ por la inyectividad de g ahí; además, como g es de clase C^1 en Ω entonces h también lo es y finalmente como $J_g(\hat{x}) \neq 0$ para toda $\hat{x} \in K$ tenemos que $J_h(\hat{x}) \neq 0$ para toda $\hat{x} \in K$. La idea ahora es conseguir dos cubos γ y Γ tales que $\gamma \subset h(C) \subset \Gamma$ y cuyos volúmenes puedan ser calculados en términos del volumen de C.

Por definición de h y la ecuación (4.1)

$$\begin{aligned} \|h(\hat{x}) - h(\hat{a}) - (\hat{x} - \hat{a})\| &= \left\|[D_g(\hat{a})]^{-1}(g(\hat{x}) - g(\hat{a}) - D_g(\hat{a})(\hat{x} - \hat{a}))\right\| \\ &\leq M\|g(\hat{x}) - g(\hat{a}) - D_g(\hat{a})(\hat{x} - \hat{a})\| \end{aligned}$$

cabe destacar que para que esta última desigualdad sea cierta, es importante el hecho de que $\left\|[D_g(\hat{x})]^{-1}(\hat{v})\right\| \leq M\|\hat{v}\|$ es válida para toda $\hat{x} \in K$.

Supongamos que h tiene funciones coordenadas $h_1, \ldots, h_n : \Omega \to \mathbb{R}$ y el centro de C tiene coordenadas $a_1, \ldots, a_n \in \mathbb{R}$, es decir, $\hat{a} = (a_1, \ldots, a_n)$. Dado $\hat{x} = (x_1, \ldots, x_n) \in C$, por (4.4) tenemos que para cada $j \in \{1, \ldots, n\}$

$$\begin{aligned} |h_j(\hat{x}) - h_j(\hat{a}) - (x_j - a_j)| &\leq \|h(\hat{x}) - h(\hat{a}) - (\hat{x} - \hat{a})\| \\ &\leq M\|g(\hat{x}) - g(\hat{a}) - D_g(\hat{a})(\hat{x} - \hat{a})\| \\ &\leq M\varepsilon_1\|\hat{x} - \hat{a}\| \end{aligned}$$

Aquí también vale la pena recalcar la importancia de conseguir un $\delta_1 > 0$ tal que la expresión (4.4) es válida para cualquier $\hat{a} \in K$; entonces de la última desigualdad se concluye que

$$|h_j(\hat{x}) - h_j(\hat{a})| \leq |x_j - a_j| + M\varepsilon_1\|\hat{x} - \hat{a}\|$$

De la observación 4.1, hecha casi al principio de este capítulo, tenemos que $|x_j - a_j| \leq \frac{s}{2}$ y $\|\hat{x} - \hat{a}\| \leq \frac{s}{2}\sqrt{n}$, por lo tanto

$$|h_j(\hat{x}) - h_j(\hat{a})| \leq \frac{s}{2}\left(1 + M\varepsilon_1\sqrt{n}\right)$$

y esto ocurre para toda $j \in \{1, \ldots, n\}$, de modo que, nuevamente por la observación 4.1, $h(\hat{x})$ pertenece a un cubo Γ con centro $h(\hat{a})$ y dimensión $s(1 + M\varepsilon_1\sqrt{n})$.Esto significa que $h(C) \subset \Gamma$ pues la $\hat{x} \in C$ que tomamos fue arbitraria.

Por otro lado, dado que h es de clase C^1 en Ω, $J_h(\hat{x}) \neq 0$ para toda $\hat{x} \in K$ y hemos tomado a C contenido en el interior de K, entonces por el lema 4.8 podemos concluir que $Fr(h(C)) \subset h(Fr(C))$; más aún, gracias a la inyectividad de h en $int(K)$, podremos demostrar la contención contraria.

Sea $\hat{x} \in h(Fr(C))$, es decir, $\hat{x} = h(\hat{y})$ para algún $\hat{y} \in Fr(C)$, y supongamos que $\hat{x} \notin Fr(h(C))$; como C es cerrado entonces $Fr(C) \subset C$ y por tanto $\hat{x} \in h(Fr(C)) \subset h(C)$, de modo que existe $r > 0$ tal que $B_r(\hat{x}) \subset h(C)$. Como h es continua en $\hat{y}$, existe $r' > 0$ tal que $B_{r'}(\hat{y}) \subset int(K)$ y tal que si $\hat{z} \in B_{r'}(\hat{y})$,entonces $\|h(\hat{z}) - h(\hat{y})\| < r$. Ahora, como $\hat{y} \in Fr(C)$, entonces existe $\hat{z} \in B_{r'}(\hat{y})$ tal que $\hat{z} \notin C$, pero entonces

$$\|h(\hat{z}) - \hat{x}\| = \|h(\hat{z}) - h(\hat{y})\| < r$$

por lo tanto $h(\hat{z}) \in h(C)$, lo que por la inyectividad de h implica que $\hat{z} \in C$, contradiciendo nuestra elección de $\hat{z}$. Por lo tanto $\hat{x} \in Fr(h(C))$ y por lo tanto

$$Fr(h(C)) = h(Fr(C))$$

Retomando la desigualdad

$$|h_j(\hat{x}) - h_j(\hat{a}) - (x_j - a_j)| \leq M\varepsilon_1 \|\hat{x} - \hat{a}\|$$

concluimos que

$$|h_j(\hat{x}) - h_j(a)| \geq |x_j - a_j| - M\varepsilon_1 \|\hat{x} - a\|$$

Una vez más, de la observación 4.1, sabemos que si $\hat{x} \in Fr(C)$ entonces $|x_j - a_j| = \frac{s}{2}$ para alguna $j \in \{1, \ldots, n\}$, de manera que

$$|h_j(\hat{x}) - h_j(\hat{a})| \geq \frac{s}{2} - M\varepsilon_1 \|\hat{x} - \hat{a}\| \geq \frac{s}{2}\left(1 - M\varepsilon_1\sqrt{n}\right)$$

Esto quiere decir que los puntos de $h(Fr(C))$, o lo que es lo mismo como ya lo demostramos, los puntos de $Fr(h(C))$, viven afuera o en la frontera del cubo γ con centro en $h(\hat{a})$ y dimensión $s(1 - M\varepsilon_1\sqrt{n})$, la cual es una cantidad positiva porque $\varepsilon_1 < \frac{1}{M\sqrt{n}}$.

Nos interesa mostrar que $\gamma \subset h(C)$, y de la compacidad de $h(C)$ y el hecho de que γ es un cubo, es suficiente mostrar que $int(\gamma) \subset h(C)$. Sabemos que

$$\mathbb{R}^n = int(h(C)) \cup ext(h(C)) \cup Fr(h(C))$$

y como

$$Fr(h(C)) \subset \gamma^c \cup Fr(\gamma)$$

entonces $int(\gamma) \cap Fr(h(C)) = \emptyset$, de donde

$$int(\gamma) = (int(h(C)) \cap int(\gamma)) \cup (ext(h(C)) \cap int(\gamma))$$

y estos dos conjuntos son ajenos, de manera que si ambos fueran diferentes del vacío se tendría una disconexión de $int(\gamma)$, lo que es absurdo, por lo tanto uno de éstos es vacío. Sin embargo como h es de clase C^1 en Ω y $J_h(\hat{a}) \neq 0$, por el teorema de la función inversa tenemos que $h(\hat{a}) \in int(h(C))$, y este punto es el centro de γ por lo que $int(h(C)) \cap int(\gamma) \neq \emptyset$ y entonces $ext(h(C)) \cap int(\gamma) = \emptyset$, es decir, $int(\gamma) \subset int(h(C))$ y por tanto $\gamma \subset h(C)$ como queríamos.

Resumiendo un poco lo que se ha logrado hasta este momento, hemos encontrado un cubo γ con centro en $h(\hat{a})$ y dimensión $s(1 - M\varepsilon_1\sqrt{n})$, así como un cubo Γ con centro en $h(\hat{a})$ y dimensión $s(1 + M\varepsilon_1\sqrt{n})$, tales que $\gamma \subset h(C) \subset \Gamma$.

Por el teorema 4.3 $h(C)$ es Jordan medible, y por lo tanto

$$V(\gamma) \leq J(h(C)) \leq V(\Gamma)$$

en donde

$$V(\gamma) = s^n \left(1 - M\varepsilon_1\sqrt{n}\right)^n$$

y

$$V(\Gamma) = s^n \left(1 + M\varepsilon_1\sqrt{n}\right)^n$$

De las ecuaciones (4.2) y (4.3) y como $V(C) = s^n$ se tiene que

$$\left(1 - \frac{\varepsilon}{2}\right) V(C) \leq V(\gamma) \leq J(h(C)) \leq V(\Gamma) \leq \left(1 + \frac{\varepsilon}{2}\right) V(C)$$

Por otra parte, $[D_g(\hat{a})]^{-1}$ es una transformación lineal no singular, con matriz asociada $[g'(\hat{a})]^{-1}$ y

$$\det\left([g'(\hat{a})]^{-1}\right) = \frac{1}{\det g'(\hat{a})} = \frac{1}{J_g(\hat{a})}$$

Como $h(C) = \left([D_g(\hat{a})]^{-1} \circ g\right)(C)$ entonces por el teorema 4.4

$$J(h(C)) = J\left([D_g(\hat{a})]^{-1} \circ g(C)\right) = \frac{1}{|J_g(\hat{a})|} J(g(C))$$

concluyendo finalmente que

$$\left(1 - \frac{\varepsilon}{2}\right) V(C)\,|J_g(\hat{a})| \leq J(g(C)) \leq \left(1 + \frac{\varepsilon}{2}\right) V(C)\,|J_g(\hat{a})|$$

■

Lo siguiente que haremos será, generalizar el lema 4.12 para rectángulos, haciendo uso de la observación 3.2 y aplicando el lema anterior.

Lema 4.13 *Sea $\Omega \subset \mathbb{R}^n$ un conjunto abierto y $K \subset \Omega$ compacto. Si $g : \Omega \to \mathbb{R}^n$ es de clase C^1 en Ω, $J_g(\hat{x}) \neq 0$ para toda $\hat{x} \in K$ y además g es inyectiva en $int(K)$ entonces para toda $\varepsilon \in (0,1)$ existe $\delta > 0$ tal que para cualquier $R \subset int(K)$ rectángulo de dimensiones menores que δ y para toda $\hat{x} \subset R$*

$$(1 - \varepsilon)\,|J_g(\hat{x})|\,V(R) \leq J(g(R)) \leq (1 + \varepsilon)\,|J_g(\hat{x})|\,V(R)$$

Dem. Este lema nos dice algo similar que el lema previo, es decir, que para rectángulos suficientemente pequeños las cantidades $|J_g(\hat{x})| V(R)$ y $J(g(R))$ son muy parecidos en proporción, con la diferencia sutil de que en el lema 4.12 el Jacobiano de g está evaluado en el centro del cubo, mientras que aquí la desigualdad es válida sin importar el punto del rectángulo que se considere para el Jacobiano de g.

La clave para este prueba, es encontrar la manera de aplicar el lema 4.12; para esto, encerraremos a R entre dos rectángulos $R_1, R_2 \subset \mathbb{R}^n$ de dimensiones racionales tales que $R_1 \subset R \subset R_2$ y cuyo volumen difiera del de R por una cantidad tan pequeña como sea necesario, para después, de acuerdo con el lema 4.1, poder hacer una partición de estos nuevos rectángulos en cubos y poder aplicar el lema 4.12.

Comencemos notando que, al ser g de clase C^1 en Ω, se satisfacen dos cosas:

1) J_g es una función continua en Ω, pues es producto y suma de funciones continuas.

2) g satisface la condición de uniformidad de Lipschitz en K.

Del primer hecho y dado que $J_g(\hat{x}) \neq 0$ para toda $\hat{x} \in K$, existe $\mu > 0$ tal que $|J_g(\hat{x})| \geq \mu$ para toda $\hat{x} \in K$, y además J_g es uniformemente continua en K.

Sea $\varepsilon \in (0, 1)$ y tomemos un $\varepsilon_1 > 0$ tal que

$$\varepsilon_1 < \mu \frac{\varepsilon}{2 + \varepsilon}$$

por lo que también se cumple que

$$\varepsilon_1 < \mu \frac{\varepsilon}{2 - \varepsilon}$$

Como

$$\frac{\varepsilon}{2 + \varepsilon} = \frac{2(1 + \varepsilon)}{2 + \varepsilon} - 1 = \frac{1 + \varepsilon}{1 + \frac{\varepsilon}{2}} - 1$$

y

$$\frac{\varepsilon}{2 - \varepsilon} = 1 - \frac{2(1 - \varepsilon)}{2 - \varepsilon} = 1 - \frac{1 - \varepsilon}{1 - \frac{\varepsilon}{2}}$$

entonces las condiciones que cumple ε_1 implican que

$$\left(1 + \frac{\varepsilon}{2}\right)\left(1 + \frac{\varepsilon_1}{\mu}\right) < 1 + \varepsilon \tag{4.5}$$

y

$$1 - \varepsilon < \left(1 - \frac{\varepsilon}{2}\right)\left(1 - \frac{\varepsilon_1}{\mu}\right) \tag{4.6}$$

De la continuidad uniforme de J_g en K, sabemos que existe $\delta_1 > 0$ tal que $|J_g(\hat{x}) - J_g(\hat{y})| < \varepsilon_1$ para cualesquiera $\hat{x}, \hat{y} \in K$ tales que $\|\hat{x} - \hat{y}\| < \delta_1$, de donde

$$|J_g(\hat{x})| < \varepsilon_1 + |J_g(\hat{y})|$$

y

$$|J_g(\hat{x})| - \varepsilon_1 < |J_g(\hat{y})|$$

De la primera desigualdad y en vista de que $|J_g(\hat{x})| \geq \mu$ para toda $\hat{x} \in K$, se tiene

$$\frac{|J_g(\hat{x})|}{|J_g(\hat{y})|} < \frac{\varepsilon_1}{|J_g(\hat{y})|} + 1 \leq \frac{\varepsilon_1}{\mu} + 1$$

o de manera equivalente

$$|J_g(\hat{x})| < \left(\frac{\varepsilon_1}{\mu} + 1\right) |J_g(\hat{y})| \tag{4.7}$$

así mismo, de la segunda desigualdad se desprende que

$$\frac{|J_g(\hat{y})|}{|J_g(\hat{x})|} > 1 - \frac{\varepsilon_1}{|J_g(\hat{x})|} \geq 1 - \frac{\varepsilon_1}{\mu}$$

es decir,

$$|J_g(\hat{y})| > \left(1 - \frac{\varepsilon_1}{\mu}\right) |J_g(\hat{x})| \tag{4.8}$$

Sea $\delta \in \left(0, \frac{\delta_1}{\sqrt{n}}\right)$ lo suficientemente pequeña para que el lema 4.12 aplique en el compacto K y para la ε con la que partimos; pidámosle además que esta δ sea útil para la condición de uniformidad de Lipschitz de g en K, es decir, que para alguna $L > 0$ si $\|\hat{x} - \hat{y}\| < \delta$ entonces

$$\|g(\hat{x}) - g(\hat{y})\| \leq L \|\hat{x} - \hat{y}\|$$

con $\hat{x}, \hat{y} \in K$.

Sea $R \subset int(K)$ rectángulo de dimensiones menores que δ; pedimos que $\delta < \frac{\delta_1}{\sqrt{n}}$, porque de esta forma dados $\hat{x}, \hat{y} \in R$, con $\hat{x} = (x_1, \ldots, x_n)$ y $\hat{y} = (y_1, \ldots, y_n)$, obtenemos que

$$\|\hat{x} - \hat{y}\| = \left(\sum_{j=1}^{n} (x_j - y_j)^2\right)^{\frac{1}{2}} \leq \left(\sum_{j=1}^{n} \delta^2\right)^{\frac{1}{2}} = \delta\sqrt{n} < \delta_1$$

y entonces las desigualdades (4.7) y (4.8)son válidas para $\hat{x}$ y $\hat{y}$.

Sea $\varepsilon_2 > 0$, demostraremos que para toda $\hat{x} \in R$

$$(1 - \varepsilon) |J_g(\hat{x})| V(R) - \varepsilon_2 \leq J(g(R)) \leq (1 + \varepsilon) |J_g(\hat{x})| V(R) + \varepsilon_2$$

Por el lema 2.1, la observación 2.2 y la proposición 4.4 sabemos que existen rectángulos $R_1, R_2 \subset \mathbb{R}^n$ con dimensiones racionales y menores a δ tales que $R_1 \subset R \subset R_2 \subset int(K)$ y

$$V(R) - V(R_1) < \frac{\varepsilon_2}{4(2L\sqrt{n})^n}$$

y

$$V(R_2) - V(R) < \frac{\varepsilon_2}{4(2L\sqrt{n})^n}$$

donde la $L > 0$, como ya se dijo, es la constante de Lipschitz. Es muy importante poder asegurar que $R_2 \subset int(K)$ porque necesitaremos que g sea inyectiva en este rectángulo.

Al ser R_1 un rectángulo, se tiene que $V(R_1) = V(int(R_1))$, por lo que

$$J(R \setminus int(R_1)) = V(R) - V(int(R_1)) = V(R) - V(R_1) < \frac{\varepsilon_2}{4(2L\sqrt{n})^n}$$

en donde la primera igualdad es válida ya que $R_1 \subset R$. Entonces $R \setminus int(R_1)$ es un conjunto compacto tal que

$$J(R \setminus int(R_1)) < \frac{\varepsilon_2}{4(2L\sqrt{n})^n}$$

aplicando el lema 4.6 se concluye que $J(g(R \setminus int(R_1))) < \varepsilon_2$.

De la inyectividad de g en $int(K)$ se sigue la inyectividad de g en R, de donde $g(R \setminus int(R_1)) = g(R) \setminus g(int(R_1))$ y obtenemos entonces

$$J(g(R \setminus int(R_1))) = J(g(R)) - J(g(int(R_1)))$$

y por tanto

$$J(g(R)) < \varepsilon_2 + J(g(int(R_1)))$$

De manera análoga se tiene que

$$J(g(R_2)) < \varepsilon_2 + J(g(int(R)))$$

vale destacar que es aquí donde usamos que $R_2 \subset int(K)$ pues como ya se dijo, aseguramos la inyectividad de g en R_2.

Ahora bien, como $g(int(R_1)) \subset g(R)$ y $g(int(R)) \subset g(R)$ y todos son Jordan medibles, se tiene que

$$J(g(int(R_1))) \leq J(g(R_1))$$

y

$$J(g(int(R))) \leq J(g(R))$$

de modo que

$$J(g(R)) < \varepsilon_2 + J(g(R_1)) \tag{4.9}$$

y

$$J(g(R_2)) - \varepsilon_2 < J(g(R)) \tag{4.10}$$

Tomemos una partición de R_1 en k cubos $C_1, \ldots, C_k \subset \mathbb{R}^n$ y una partición de R_2 en l cubos $C'_1, \ldots, C'_l \subset \mathbb{R}^n$, esto es posible por el lema 4.1; sean $\hat{a}_1, \ldots, \hat{a}_k$ los centros de cada C_j y $\hat{b}_1, \ldots, \hat{b}_l$ los centros de cada C'_j. Nuevamente de la inyectividad de g tanto en R_1 como en R_2, se sigue que $g(C_1), \ldots, g(C_k)$ es una subdivisión del conjunto $g(R_1)$, es decir,

$$g(R_1) \subset \bigcup_{j=1}^{k} g(C_j)$$

y además $int(C_j) \cap int(C_i) = \emptyset$ para $i \neq j$; similarmente $g(C'_1), \ldots, g(C'_l)$ es una subdivisión de $g(R_2)$, de modo que por el corolario 3.1

$$J(g(R_1)) = \sum_{j=1}^{k} J(g(C_j))$$

y

$$J(g(R_2)) = \sum_{j=1}^{l} J\left(g\left(C'_j\right)\right)$$

Como las dimensiones de cada C_j y cada C'_j son menores que δ, por estar contenidos en R_1 y R_2 respectivamente, podemos aplicar el lema 4.12 a cada uno de estos cubos y entonces usaremos el hecho de que para C_j se cumple que

$$J(g(C_j)) \leq \left(1 + \frac{\varepsilon}{2}\right) |J_g(\hat{a}_j)| V(C_j) \tag{4.11}$$

y para C_j' se cumple que

$$\left(1-\frac{\varepsilon}{2}\right)\left|J_g\left(\hat{b}_j\right)\right|V\left(C_j'\right)\leq J\left(g\left(C_j'\right)\right) \tag{4.12}$$

Como $R_1 \subset R \subset R_2$ y R_2 tiene dimensiones menores a δ, entonces para toda $\hat{x} \in R$, $\|\hat{x}-\hat{a}_j\| < \delta_1$ para toda $j \in \{1,\ldots,k\}$ y $\left\|\hat{x}-\hat{b}_j\right\| < \delta_1$ para toda $j \in \{1,\ldots,l\}$, de modo que por las ecuaciones (4.7) y (4.8)

$$\left|J_g\left(\hat{a}_j\right)\right| < \left(1+\frac{\varepsilon_1}{\mu}\right)\left|J_g\left(\hat{x}\right)\right|$$

y

$$\left(1-\frac{\varepsilon_1}{\mu}\right)\left|J_g\left(\hat{x}\right)\right| < \left|J_g\left(\hat{b}_j\right)\right|$$

utilizando estas dos desigualdades en las ecuaciones (4.11) y (4.12) respectivamente, se tiene que

$$J\left(g\left(C_j\right)\right) < \left(1+\frac{\varepsilon}{2}\right)\left(1+\frac{\varepsilon_1}{\mu}\right)\left|J_g\left(\hat{x}\right)\right|V\left(C_j\right)$$

y

$$\left(1-\frac{\varepsilon}{2}\right)\left(1-\frac{\varepsilon_1}{\mu}\right)\left|J_g\left(\hat{x}\right)\right|V\left(C_j'\right) < J\left(g\left(C_j'\right)\right)$$

y por (4.5) y (4.6) conluimos que

$$J\left(g\left(C_j\right)\right) < \left(1+\varepsilon\right)\left|J_g\left(\hat{x}\right)\right|V\left(C_j\right) \tag{4.13}$$

y

$$\left(1-\varepsilon\right)\left|J_g\left(\hat{x}\right)\right|V\left(C_j'\right) < J\left(g\left(C_j'\right)\right) \tag{4.14}$$

Finalmente con las ecuaciones (4.10) y (4.14) obtenemos

$$\begin{aligned}
\left(1-\varepsilon\right)\left|J_g\left(\hat{x}\right)\right|V\left(R\right)-\varepsilon_2 &< \left(1-\varepsilon\right)\left|J_g\left(\hat{x}\right)\right|V\left(R_2\right)-\varepsilon_2 \\
&= \left(1-\varepsilon\right)\left|J_g\left(\hat{x}\right)\right|\sum_{j=1}^{l}V\left(C_j'\right)-\varepsilon_2 \\
&< \sum_{j=1}^{l}J\left(g\left(C_j'\right)\right)-\varepsilon_2 \\
&= J\left(g\left(R_2\right)\right)-\varepsilon_2 < J\left(g\left(R\right)\right)
\end{aligned}$$

y del mismo modo con las ecuaciones (4.9) y (4.13)

$$\begin{aligned}
J\left(g\left(R\right)\right) &< J\left(g\left(R_1\right)\right)+\varepsilon_2 = \sum_{j=1}^{k}J\left(g\left(C_j\right)\right)+\varepsilon_2 \\
&< \varepsilon_2+\left(1+\varepsilon\right)\left|J_g\left(\hat{x}\right)\right|\sum_{j=1}^{k}V\left(C_j\right) \\
&= \varepsilon_2+\left(1+\varepsilon\right)\left|J_g\left(\hat{x}\right)\right|V\left(R_1\right) \\
&< \varepsilon_2+\left(1+\varepsilon\right)\left|J_g\left(\hat{x}\right)\right|V\left(R\right)
\end{aligned}$$

es decir

$$(1-\varepsilon)\left|J_g(\hat{x})\right| V(R)-\varepsilon_2<J(g(R))<\varepsilon_2+(1+\varepsilon)\left|J_g(\hat{x})\right| V(R)$$

y como esto ocurre para toda $\varepsilon_2>0$, se tiene que

$$(1-\varepsilon)\left|J_g(\hat{x})\right| V(R) \leq J(g(R)) \leq (1+\varepsilon)\left|J_g(\hat{x})\right| V(R)$$

■

Con este último lema por fin tenemos todas las herramientas necesarias para demostrar el teorema de cambio de variable.

Teorema 4.5 *(Cambio de variable) Sea $\Omega \subset \mathbb{R}^n$ un conjunto abierto y $D \subset \mathbb{R}^n$ Jordan medible tal que $D \cup Fr(D) \subset \Omega$. Sea $g: \Omega \rightarrow \mathbb{R}^n$ de clase C^1 en Ω, inyectiva en $int(D)$ y tal que $J_g(\hat{x}) \neq 0$ para toda $\hat{x} \in int(D)$. Si $f: g(D) \rightarrow \mathbb{R}$ es acotada y continua entonces*

$$\int_{g(D)} f=\int_D(f \circ g)\left|J_g\right|$$

Dem. Primero nótese que $g(D)$ es Jordan medible por el teorema 4.3 y entonces $\int_{g(D)} f$ existe, y por la continuidad de $(f \circ g)\left|J_g\right|$ se tiene que $\int_D(f \circ g)\left|J_g\right|$ existe.

Supongamos primero que $f(\hat{y}) \geq 0$ para toda $\hat{y} \in g(D)$ y probemos el resultado en este caso; más adelante nos ocuparemos del caso general.

Sean $M>0$ y $S>0$ tales que para toda $\hat{y} \in g(D)$

$$f(\hat{y}) \leq M \tag{4.15}$$

y para toda $\hat{x} \in D$

$$\left|J_g(\hat{x})\right| \leq S \tag{4.16}$$

M existe por hipótesis y S existe porque $\left|J_g\right|$ es continua en el compacto $D \cup Fr(D)$ y por tanto acotada ahí.

Como Ω es abierto, para cada $\hat{a} \in D \cup Fr(D)$ existe $r_{\hat{a}}>0$ tal que

$$\overline{B_{r_{\hat{a}}}(\hat{a})} \subset \Omega$$

y como $D \cup Fr(D)$ es compacto, existen $\hat{a}_1, \ldots, \hat{a}_l \in D \cup Fr(D)$ tales que

$$D \cup Fr(D) \subset \bigcup_{j=1}^l B_{r_{\hat{a}_j}}\left(\hat{a}_j\right)$$

Llamemos

$$A=\bigcup_{j=1}^l B_{r_{\hat{a}_j}}\left(\hat{a}_j\right)$$

A es un conjunto abierto cuya cerradura es compacta, pues

$$\overline{\bigcup_{j=1}^l B_{r_{\hat{a}_j}}\left(\hat{a}_j\right)} \subset \bigcup_{j=1}^l \overline{B_{r_{\hat{a}_j}}\left(\hat{a}_j\right)}$$

y este segundo es un conjunto compacto y como además el conjunto está contenido en Ω, se tiene entonces que $\overline{A} \subset \Omega$.

Tomemos $\varepsilon > 0$; demostraremos que

$$\left| \int_D (f \circ g) |J_g| - \int_{g(D)} f \right| < \varepsilon$$

con lo que concluiremos la igualdad deseada.

Al ser g de clase C^1 en Ω y $\overline{A}$ compacto, tenemos que g satisface la condición de uniformidad de Lipschitz en $\overline{A}$, así que tomemos $L > 0$ la constante de Lipschitz de g en $\overline{A}$. Como D es Jordan medible entonces $J(Fr(D)) = 0$ y por el lema 3.1 existen $J_1, \ldots, J_k \subset \mathbb{R}^n$ rectángulos tales que $Fr(D) \subset \bigcup_{i=1}^{k} J_i$ y

$$\sum_{j=1}^{k} V(J_i) < \delta_1$$

en donde

$$\delta_1 = \frac{1}{2} \min \left\{ \frac{\varepsilon}{5MS}, \frac{\varepsilon}{60M(2L\sqrt{n})^n} \right\} \tag{4.17}$$

Recordando el "truco" de "inflar" rectángulos como lo hicimos en el lema 3.1, "inflaremos" los réctangulos J_i de manera que cada punto de la frontera de D pertenezca al interior de algún rectángulo inflado; por el lema 2.1 y la observación 2.2 sabemos que para cada $i \in \{1, \ldots, k\}$ existe $J_i' \subset \mathbb{R}^n$ rectángulo tal que $J_i \subset int(J_i')$ y

$$V(J_i') - V(J_i) < \frac{\delta_1}{k}$$

entonces para toda $\hat{x} \in Fr(D)$ existe $i \in \{1, \ldots, k\}$ tal que $\hat{x} \in int(J_i')$ y

$$\sum_{j=1}^{k} V(J_i') = \sum_{j=1}^{k} (V(J_i') - V(J_i)) + \sum_{j=1}^{k} V(J_i) < 2\delta_1$$

Parte de la idea para realizar esta prueba es la aplicación del lema 4.13, pero para eso necesitamos un compacto $K \subset \Omega$ tal que para toda $\hat{x} \in K$ $J_g(\hat{x}) \neq 0$ y además g deberá ser inyectiva en el interior de este compacto.

Como D es Jordan medible, por la observación 3.3 existe $K \subset int(D)$ compacto y Jordan medible tal que

$$J(D) - J(K) < \delta_2 \tag{4.18}$$

en donde

$$\delta_2 = \min \left\{ \frac{\varepsilon}{10MS}, \frac{\varepsilon}{60M(2L\sqrt{n})^n} \right\} \tag{4.19}$$

de esta manera para toda $\hat{x} \in K$ $J_g(\hat{x}) \neq 0$ y g resulta inyectiva en K, por lo que estaremos en la libertad de usar el lema 4.13 sobre K.

De manera análoga a lo hecho con $Fr\,(D)$, podemos decir que existen $I_1, \ldots, I_{k'} \subset \mathbb{R}^n$ rectángulos tales que $Fr\,(K) \subset \bigcup_{j=1}^{k'} int\,(I_j)$ y

$$\sum_{j=1}^{k'} V\,(I_i) < \delta_2$$

Sea $R \subset \mathbb{R}^n$ rectángulo, suficientemente grande, de manera que

$$\overline{A}, \bigcup_{j=1}^{k'} I_j, \bigcup_{i=1}^{k} J'_i \subset R$$

Estamos en busca de una partición de R lo bastante buena para que nos permita estimar el valor de

$$\left| \int_D (f \circ g)\,|J_g| - \int_{g(D)} f \right|$$

Vale la pena recalcar la importancia de que los puntos tanto de $Fr\,(D)$ como de $Fr\,(K)$ sean puntos interiores de algún rectángulo; construiremos una partición de R a partir de los vértices de los J'_i y los I_j y queremos tener control del volumen de los rectángulos generados por la partición que intersectan a $Fr\,(D)$ o a $Fr\,(K)$, en otras palabras, si tenemos una partición de R que tenga a todos los vértices de cada J'_i y Q es un subrectángulo generado por esta partición tal que $Q \cap Fr\,(D) \neq \emptyset$ entonces por la discusión hecha tras la definición 2.7 y por el hecho de que cada punto de $Fr\,(D)$ vive en el interior de algún J'_i, se tiene que $Q \subset J'_i$ de modo que

$$\sum_{Q \cap Fr(D) \neq \emptyset} V\,(Q) \leq \sum_{j=1}^{k} V\,(J'_i) < 2\delta_1$$

Sea $P_0 \in \mathcal{P}_R$ tal que contiene a todos los vértices de cada J'_i y cada I_j, por lo ya argumentado sabemos que si $Q_1, \ldots, Q_r$ son los subrectángulos generados por P_0 entonces

$$\sum_{Q_i \cap Fr(D) \neq \emptyset} V\,(Q_i) < 2\delta_1$$

y

$$\sum_{Q_i \cap Fr(K) \neq \emptyset} V\,(Q_i) < \delta_2$$

de hecho, esta propiedad se preserva bajo cualquier refinamiento de P_0, una vez más, por la discusión tras la definición2.7, dado cualquier refinamiento de P_0 que genere subrectángulos $R_1, \ldots, R_m$, se tiene que para cada $i \in \{1, \ldots, m\}$ existe $j \in \{1, \ldots, r\}$ tal que $R_i \subset Q_j$, de donde

$$\sum_{R_i \cap Fr(D) \neq \emptyset} V\,(R_i) \leq \sum_{Q_i \cap Fr(D) \neq \emptyset} V\,(Q_i) < 2\delta_1$$

y

$$\sum_{R_i \cap Fr(K) \neq \emptyset} V\,(R_i) \leq \sum_{Q_i \cap Fr(K) \neq \emptyset} V\,(Q_i) < \delta_2$$

Por otro lado, usando la proposición 4.3 tenemos que $d\left(A^c,\overline{D}\right) > 0$, ya que A es abierto y $\overline{D} \subset A$ es compacto. Así mismo $d\left((int\,(D))^c, K\right) > 0$ pues $(int\,(D))^c$ es cerrado y $K \subset int\,(D)$ es compacto.

Ahora bien, si tomamos un rectángulo R' cuya longitud de la diagonal, $d\,(R')$, sea menor que $d\left(A^c,\overline{D}\right)$ y tal que $R' \cap (D \cup Fr\,(D)) \neq \emptyset$ entonces $R' \subset \Omega$. Esto es sencillo de ver, pues dado $\hat{x} \in R'$, si $\hat{y} \in R' \cap (D \cup Fr\,(D))$ entonces

$$\|\hat{x} - \hat{y}\| \leq d\,(R') < d\left(A^c,\overline{D}\right)$$

y por lo tanto $\hat{x} \notin A^c$, es decir, $\hat{x} \in A \subset \Omega$. Algo similar sucede si tomamos un rectángulo R' tal que $d\,(R') < d\left((int\,(D))^c, K\right)$ y $R' \cap K \neq \emptyset$ en este caso lo que sucede es que $R' \subset int\,(D)$.

Finalmente, antes de construir la partición deseada, tomemos $\delta > 0$ con

$$\delta < \min\left\{d\left(A^c,\overline{D}\right), d\left((int\,(D))^c, K\right)\right\}$$

para la cual el lema 4.13 sea válido en el compacto K y para el valor positivo

$$\frac{\varepsilon}{5T}$$

donde

$$T = V\,(R)\,MS \tag{4.20}$$

Sea $P \in \mathcal{P}_R$ un refinamiento de P_0, con subrectángulos generados $R_1, \ldots, R_m$, tal que

$$\begin{array}{ll} i) & \bar{S}\left((f\circ g)\,|J_g|\,\mathcal{X}_D, P\right) - \underline{S}\left((f\circ g)\,|J_g|\,\mathcal{X}_D, P\right) < \frac{\varepsilon}{5} \\ ii) & \text{para toda } j \in \{1,\ldots,m\}, \quad d\,(R_j) < \delta \end{array}$$

La primera condición la podemos conseguir porque $(f\circ g)\,|J_g|$ es integrable sobre D, y la segunda se puede pedir porque cualquier refinamiento de una partición que satisfaga $i)$ no altera la validez del mismo.

Sean

$$\begin{aligned} \beta &= \{j \in \{1,\ldots,m\} \mid R_j \cap Fr\,(D) \neq \emptyset\}, \\ \gamma_1 &= \{j \in \{1,\ldots,m\} \mid R_j \subset int\,(K)\}, \\ \gamma_2 &= \{j \in \{1,\ldots,m\} \mid R_j \cap Fr\,(K) \neq \emptyset\}, \\ \gamma_3 &= \{j \in \{1,\ldots,m\} \mid R_j \subset int\,(D) \setminus K\}, \\ \gamma &= \gamma_1 \cup \gamma_2 \cup \gamma_3 = \{j \in \{1,\ldots,m\} \mid R_j \subset int\,(D)\} \end{aligned}$$

y sean también,

$$\begin{aligned} B_1 &= \bigcup_{j\in\beta} R_j \\ B_2 &= \bigcup_{j\in\gamma_2} R_j \\ B_3 &= \bigcup_{j\in\gamma_3} R_j \\ G &= \bigcup_{j\in\gamma} R_j \end{aligned}$$

Por todas las condiciones pedidas anteriormente, se tiene que

1) $G \cup B_1 \subset \Omega$

2) Los conjuntos B_1 y B_2 son jordan medibles con

$$J(B_1) = \sum_{j \in \beta} V(R_j) < 2\delta_1 \tag{4.21}$$

y

$$J(B_2) = \sum_{j \in \gamma_2} V(R_j) < \delta_2 \tag{4.22}$$

3) Para toda $j \in \gamma_1$, $R_j \subset int(K)$ y $d(R_j) < \delta$, de modo que aplicando el lema 4.13, se tiene que para toda $\hat{x} \in R_j$

$$\left(1 - \frac{\varepsilon}{5T}\right) |J_g(\hat{x})| V(R_j) \leq J(g(R_j)) \leq \left(1 + \frac{\varepsilon}{5T}\right) |J_g(\hat{x})| V(R_j)$$

como $f \geq 0$ entonces $\forall \hat{x} \in R_j$

$$\left(1 - \frac{\varepsilon}{5T}\right) (f \circ g)(\hat{x}) |J_g(\hat{x})| V(R_j) \leq (f \circ g)(\hat{x}) J(g(R_j)) \tag{4.23}$$

y

$$(f \circ g)(\hat{x}) J(g(R_j)) \leq \left(1 + \frac{\varepsilon}{5T}\right) (f \circ g)(\hat{x}) |J_g(\hat{x})| V(R_j) \tag{4.24}$$

Si para cada $j \in \{1, \ldots, m\}$ tomamos $\hat{x}_j \in R_j$ arbitrario (por el momento) entonces podemos estimar $\left| \int_D (f \circ g) |J_g| - \int_{g(D)} f \right|$ de la siguiente manera

$$\begin{aligned}
& \left| \int_D (f \circ g) |J_g| - \int_{g(D)} f \right| \qquad (4.25) \\
\leq & \left| \int_D (f \circ g) |J_g| - \sum_{j=1}^{m} (f \circ g)(\hat{x}_j) |J_g(\hat{x}_j)| \mathcal{X}_D(\hat{x}_j) V(R_j) \right| \\
& + \left| \sum_{j=1}^{m} (f \circ g)(\hat{x}_j) |J_g(\hat{x}_j)| \mathcal{X}_D(\hat{x}_j) V(R_j) - \sum_{j \in \gamma} (f \circ g)(\hat{x}_j) |J_g(\hat{x}_j)| V(R_j) \right| \\
& + \left| \sum_{j \in \gamma} (f \circ g)(\hat{x}_j) |J_g(\hat{x}_j)| V(R_j) - \sum_{j \in \gamma_1} (f \circ g)(\hat{x}_j) |J_g(\hat{x}_j)| V(R_j) \right| \\
& + \left| \sum_{j \in \gamma_1} (f \circ g)(\hat{x}_j) |J_g(\hat{x}_j)| V(R_j) - \sum_{j \in \gamma_1} (f \circ g)(\hat{x}_j) J(g(R_j)) \right| \\
& + \left| \sum_{j \in \gamma_1} (f \circ g)(\hat{x}_j) J(g(R_j)) - \int_{g(D)} f \right|
\end{aligned}$$

La idea desde luego, es conseguir que cada sumando sea menor que $\frac{\varepsilon}{5}$, siendo el primero el más claro pues $\int_D (f \circ g) |J_g|$ y

$$\sum_{j=1}^{m} (f \circ g) (\hat{x}_j) |J_g (x_j)| \mathcal{X}_D (\hat{x}_j) V (R_j)$$

son dos valores entre $\bar{S} ((f \circ g) |J_g| \mathcal{X}_D, P)$ y $\underline{S} ((f \circ g) |J_g| \mathcal{X}_D, P)$, así que por la condición $i)$

$$\left| \int_D (f \circ g) |J_g| - \sum_{j=1}^{m} (f \circ g) (\hat{x}_j) |J_g (\hat{x}_j)| \mathcal{X}_D (\hat{x}_j) V (R_j) \right| < \frac{\varepsilon}{5}$$

Cabe destacar que para $j \notin \beta \cup \gamma$, $\hat{x}_j \notin D$, de modo que $\mathcal{X}_D (\hat{x}_j) = 0$, entonces

$$\sum_{j=1}^{m} (f \circ g) (\hat{x}_j) |J_g (\hat{x}_j)| \mathcal{X}_D (\hat{x}_j) V (R_j) = \sum_{j \in \beta \cup \gamma} (f \circ g) (\hat{x}_j) |J_g (\hat{x}_j)| \mathcal{X}_D (\hat{x}_j) V (R_j)$$

De hecho, es posible que para alguna $j \in \beta$, $\hat{x}_j \notin D$, anulándose también ese sumando y permitiéndonos concluir junto con las ecuaciones (4.15), (4.16), (4.21) y (4.17) que

$$\sum_{j \in \beta} (f \circ g) (\hat{x}_j) |J_g (\hat{x}_j)| \mathcal{X}_D (\hat{x}_j) V (R_j) \leq MS \sum_{j \in \beta} V (R_j) < MS (2\delta_1) \leq MS \left(\frac{\varepsilon}{5MS} \right) = \frac{\varepsilon}{5}$$

de donde

$$\begin{aligned} & \left| \sum_{j \in \beta \cup \gamma} (f \circ g) (\hat{x}_j) |J_g (\hat{x}_j)| \mathcal{X}_D (\hat{x}_j) V (R_j) - \sum_{j \in \gamma} (f \circ g) (\hat{x}_j) |J_g (\hat{x}_j)| V (R_j) \right| \\ = & \sum_{j \in \beta} (f \circ g) (\hat{x}_j) |J_g (\hat{x}_j)| \mathcal{X}_D (\hat{x}_j) V (R_j) < \frac{\varepsilon}{5} \end{aligned}$$

Para el tercer sumando usando nuevamente (4.15) y (4.16) así como (4.22) y (4.19) se tiene que

$$\begin{aligned} & \left| \sum_{j \in \gamma} (f \circ g) (\hat{x}_j) |J_g (\hat{x}_j)| V (R_j) - \sum_{j \in \gamma_1} (f \circ g) (\hat{x}_j) |J_g (\hat{x}_j)| V (R_j) \right| \\ = & \sum_{j \in \gamma_2} (f \circ g) (\hat{x}_j) |J_g (\hat{x}_j)| V (R_j) + \sum_{j \in \gamma_3} (f \circ g) (\hat{x}_j) |J_g (\hat{x}_j)| V (R_j) \\ \leq & MS \sum_{j \in \gamma_2} V (R_j) + MS \sum_{j \in \gamma_3} V (R_j) < MS \left(\frac{\varepsilon}{10MS} \right) + MS \sum_{j \in \gamma_3} V (R_j) \end{aligned}$$

Pero $\bigcup_{j \in \gamma_3} R_j \subset int (D) \setminus K$, y como por (4.18) y (4.19) tenemos que

$$J (int (D) \setminus K) = J (int (D)) - J (K) < \delta_2 \leq \frac{\varepsilon}{10MS}$$

entonces

$$\sum_{j\in\gamma_3} V(R_j) < \frac{\varepsilon}{10MS}$$

y por tanto

$$\left|\sum_{j\in\gamma} (f\circ g)(\hat{x}_j)\left|J_g(\hat{x}_j)\right| V(R_j) - \sum_{j\in\gamma_1} (f\circ g)(\hat{x}_j)\left|J_g(\hat{x}_j)\right| V(R_j)\right| < \frac{\varepsilon}{5}$$

El cuarto sumando es problamente el más importante, porque es de alguna manera, el sumando que nos dice que las integrales $\int_D (f\circ g)|J_g|$ y $\int_{g(D)} f$ se parecen (como lo habíamos mencionado en el lema 4.12). Es aquí donde hechamos mano del lema 4.13.

Sabemos por (4.15), (4.16) y (4.20) que para toda $\hat{x}\in D$

$$(f\circ g)(\hat{x})\left|J_g(\hat{x})\right| \leq MS = \frac{T}{V(R)}$$

de manera que

$$(f\circ g)(\hat{x})\left|J_g(\hat{x})\right| V(R_j)\frac{\varepsilon}{5} \leq \frac{T}{V(R)}V(R_j)\frac{\varepsilon}{5}$$

dividiendo ambos lados de la desigualdad entre T y sumando a cada lado el término

$$(f\circ g)(\hat{x})\left|J_g(\hat{x})\right| V(R_j)$$

obtenemos

$$(f\circ g)(\hat{x})\left|J_g(\hat{x})\right| V(R_j)\left(1+\frac{\varepsilon}{5T}\right) \leq (f\circ g)(\hat{x})\left|J_g(\hat{x})\right| V(R_j) + \frac{\varepsilon}{5V(R)}V(R_j) \tag{4.26}$$

equivalentemente

$$(f\circ g)(\hat{x})\left|J_g(\hat{x})\right| V(R_j) - \frac{\varepsilon}{5V(R)}V(R_j) \leq \left(1-\frac{\varepsilon}{5T}\right)(f\circ g)(\hat{x})\left|J_g(\hat{x})\right| V(R_j) \tag{4.27}$$

Ahora bien, como por (4.23) y (4.24) para $j\in\gamma_1$ se cumple que

$$\left(1-\frac{\varepsilon}{5T}\right)(f\circ g)(\hat{x}_j)\left|J_g(\hat{x}_j)\right| V(R_j) \leq (f\circ g)(\hat{x}_j)\, J(g(R_j))$$

y

$$(f\circ g)(\hat{x}_j)\, J(g(R_j)) \leq \left(1+\frac{\varepsilon}{5T}\right)(f\circ g)(\hat{x}_j)\left|J_g(\hat{x}_j)\right| V(R_j)$$

entonces aplicando las desigualdades (4.26) y (4.27) obtenemos que

$$(f\circ g)(\hat{x}_j)\left|J_g(\hat{x}_j)\right| V(R_j) - \frac{\varepsilon}{5V(R)}V(R_j) \leq (f\circ g)(\hat{x}_j)\, J(g(R_j))$$

y

$$(f\circ g)(\hat{x}_j)\, J(g(R_j)) \leq (f\circ g)(\hat{x}_j)\left|J_g(\hat{x}_j)\right| V(R_j) + \frac{\varepsilon}{5V(R)}V(R_j)$$

sumando respecto a los $j \in \gamma_1$ en estas últimas dos desigualdades, y usando que

$$\frac{\varepsilon}{5V(R)}\sum_{j\in\gamma_1} V(R_j) < \frac{\varepsilon}{5V(R)}\sum_{j=1}^{m} V(R_j) = \frac{\varepsilon}{5V(R)}V(R) = \frac{\varepsilon}{5}$$

se tiene que

$$\begin{aligned} &\sum_{j\in\gamma_1}(f\circ g)(\hat{x}_j)\left|J_g(\hat{x}_j)\right|V(R_j) - \frac{\varepsilon}{5} < \sum_{j\in\gamma_1}(f\circ g)(\hat{x}_j)J(g(R_j)) \\ < &\sum_{j\in\gamma_1}(f\circ g)(\hat{x}_j)\left|J_g(\hat{x}_j)\right|V(R_j) + \frac{\varepsilon}{5} \end{aligned}$$

y por lo tanto

$$\left|\sum_{j\in\gamma_1}(f\circ g)(\hat{x}_j)\left|J_g(\hat{x}_j)\right|V(R_j) - \sum_{j\in\gamma_1}(f\circ g)(\hat{x}_j)J(g(R_j))\right| < \frac{\varepsilon}{5}$$

Es importante observar que ninguna de las desigualdades demostradas depende de nuestra elección de $\hat{x}_j \in R_j$; esto resulta bastante conveniente porque para probar que el último sumando de (4.25) también es menor que $\frac{\varepsilon}{5}$ sí tendremos que usar a unos $\hat{x}_j \in R_j$ específicos y esto no alterará lo ya logrado con los sumandos previos.

Como $G \subset int(D)$ es Jordan medible entonces por las hipótesis iniciales y el teorema 4.3, $g(G)$ también es Jordan medible y por lo tanto $g(D)\setminus g(G)$ es Jordan medible y entonces por el corolario 3.5

$$\int_{g(D)\setminus g(G)} f = \int_{g(D)} f - \int_{g(G)} f$$

Pero $D \subset B_1 \cup G$, entonces $g(D) \subset g(B_1) \cup g(G)$, de modo que $g(D)\setminus g(G) \subset g(B_1)$, de donde

$$\int_{g(D)\setminus g(G)} f \leq MJ(g(D)\setminus g(G)) = M\bar{J}(g(D)\setminus g(G)) \leq M\bar{J}(g(B_1))$$

Nótese que no podemos asegurar que $g(B_1)$ sea un conjunto Jordan medible pues no sabemos nada de $J_g(\hat{x})$ para $\hat{x} \in B_1$; afortunadamente no es algo que necesitemos saber. Por el lema 4.6 y dado que

$$J(B_1) < \frac{\varepsilon}{60M\left(2L\sqrt{n}\right)^n}$$

se tiene que

$$\bar{J}(g(B_1)) < \frac{\varepsilon}{15M}$$

(aquí estamos usando fuertemente que $B_1 \subset \Omega$ así como la compacidad del mismo) y entonces

$$\int_{g(D)} f - \int_{g(G)} f = \int_{g(D)\setminus g(G)} f \leq \frac{\varepsilon}{15} \tag{4.28}$$

Por otro lado; sabemos que para toda $j \in \gamma$, $R_j \subset int(D)$ y g es inyectiva ahí, de manera que $\{g(R_j) \mid j \in \gamma\}$ forma una subdivisión para el conjunto $g(G)$, del mismo modo

$\{g(R_j) \mid j \in \gamma_2\}$ y $\{g(R_j) \mid j \in \gamma_3\}$ forman una subdivisión de los conjuntos $g(B_2)$ y $g(B_3)$ respectivamente.

El hecho de que para $j \neq i$, con $i, j \in \gamma$, se tenga que $int(g(R_i)) \cap int(g(R_j)) = \emptyset$, implica que $J(g(R_i) \cap g(R_j)) = 0$, entonces por el teorema 3.9

$$\int_{g(G)} f = \sum_{j\in\gamma} \left[\int_{g(R_j)} f \right]$$

Por el teorema del valor intermedio para integrales

$$\int_{g(R_j)} f = f(\hat{y}_j) J(g(R_j))$$

para alguna $\hat{y}_j \in g(R_j)$ y $j \in \gamma$ (nótese que el teorema del valor intermedio para integrales es válido porque el conjunto $g(R_j)$ es conexo). De la inyectividad de g se sigue que existe un único $\hat{x}_j \in R_j$ tal que $\hat{y}_j = g(\hat{x}_j)$, así que

$$\int_{g(R_j)} f = (f \circ g)(\hat{x}_j) J(g(R_j))$$

de donde

$$\int_{g(G)} f = \sum_{j\in\gamma} (f \circ g)(\hat{x}_j) J(g(R_j))$$

Entonces para estas $\hat{x}_j \in R_j$ específicas, auxiliándonos de (4.28), el último sumando que nos falta analizar lo podemos aproximar como

$$\begin{aligned}
&\left| \int_{g(D)} f - \sum_{j\in\gamma_1} (f \circ g)(\hat{x}_j) J(g(R_j)) \right| \\
= &\left| \int_{g(D)} f - \sum_{j\in\gamma} (f \circ g)(\hat{x}_j) J(g(R_j)) + \sum_{j\in\gamma_2\cup\gamma_3} (f \circ g)(\hat{x}_j) J(g(R_j)) \right| \\
= &\left| \left(\int_{g(D)} f - \int_{g(G)} f \right) + \sum_{j\in\gamma_2\cup\gamma_3} (f \circ g)(\hat{x}_j) J(g(R_j)) \right| \\
< &\frac{\varepsilon}{15} + \sum_{j\in\gamma_2\cup\gamma_3} (f \circ g)(\hat{x}_j) J(g(R_j)) \leq \frac{\varepsilon}{15} + M \sum_{j\in\gamma_2\cup\gamma_3} J(g(R_j))
\end{aligned}$$

y como ya se dijo que $\{g(R_j) \mid j \in \gamma_2\}$ y $\{g(R_j) \mid j \in \gamma_3\}$ forman una subdivisión de los conjuntos $g(B_2)$ y $g(B_3)$ respectivamente, entonces por el corolario 3.1

$$J(g(B_2)) = \sum_{j\in\gamma_2} J(g(R_j))$$

y

$$J(g(B_3)) = \sum_{j\in\gamma_3} J(g(R_j))$$

Ahora, por (4.22) y (4.19) B_2 es un compacto tal que

$$J(B_2) < \frac{\varepsilon}{60M(2L\sqrt{n})^n}$$

así que por el lema 4.6

$$J(g(B_2)) < \frac{\varepsilon}{15M}$$

Si logramos que $J(g(B_3))$ también sea menor que $\frac{\varepsilon}{15M}$, prácticamente habremos terminado. Recordemos que K es un compacto que tomamos de manera que

$$J(D) - J(K) < \frac{\varepsilon}{60M(2L\sqrt{n})^n}$$

esto junto con la proposición 3.3 implican que el compacto $\overline{D} \setminus int(K)$ es un conjunto Jordan medible tal que $J\left(\overline{D} \setminus int(K)\right) = J(D) - J(K)$, es decir,

$$J\left(\overline{D} \setminus int(K)\right) < \frac{\varepsilon}{60M(2L\sqrt{n})^n}$$

entonces por el lema 4.6

$$\bar{J}\left(g\left(\overline{D} \setminus int(K)\right)\right) < \frac{\varepsilon}{15M}$$

y como por definición, para $j \in \gamma_3$ se cumple que $R_j \subset int(D) \setminus K \subset \overline{D} \setminus int(K)$ entonces $g(B_3) \subset g\left(\overline{D} \setminus int(K)\right)$ y por lo tanto

$$J(g(B_3)) \leq \bar{J}\left(g\left(\overline{D} \setminus int(K)\right)\right) < \frac{\varepsilon}{15M}$$

logrando así que

$$\left| \int_{g(D)} f - \sum_{j \in \gamma_1} (f \circ g)(\hat{x}_j) J(g(R_j)) \right| < \frac{\varepsilon}{15} + M\left(\frac{\varepsilon}{15M} + \frac{\varepsilon}{15M}\right) = \frac{\varepsilon}{5}$$

como queríamos. Con esto hemos mostrado que cada sumando de la expresión (4.25) es menor que $\frac{\varepsilon}{5}$ y por lo tanto

$$\left| \int_D (f \circ g)|J_g| - \int_{g(D)} f \right| < \varepsilon$$

es decir,

$$\int_{g(D)} f = \int_D (f \circ g)|J_g|$$

Algo que no se dijo, pero que fue muy importante para esta prueba, fue el poder meter al compacto $\overline{A}$ en Ω y tomar la constante de Lipschitz de g en $\overline{A}$, misma que se convirtió en constante de Lipschitz para los compactos B_1, B_2 y $\overline{D} \setminus int(K)$, permitiéndonos para estos compactos, aplicar el lema 4.6.

Nos resta ver el caso en que f tome valores negativos sobre $g(D)$. En este caso definimos $f_1, f_2 : g(D) \to \mathbb{R}$ por

$$f_1(\hat{y}) = \frac{1}{2}\left(|f(\hat{y})| + f(\hat{y})\right)$$

y

$$f_2(\hat{y}) = \frac{1}{2}\left(|f(\hat{y})| - f(\hat{y})\right)$$

entonces $f_1(\hat{y}), f_2(\hat{y}) \geq 0$ para toda $\hat{y} \in g(D)$, f_1 y f_2 son funciones continuas y acotadas en $g(D)$ y además $f = f_1 - f_2$; por lo que ya demostramos tenemos que

$$\begin{aligned}\int_{g(D)} f &= \int_{g(D)} f_1 - \int_{g(D)} f_2 = \int_D (f_1 \circ g)|J_g| - \int_D (f_2 \circ g)|J_g| \\ &= \int_D (f_1 \circ g - f_2 \circ g)|J_g| = \int_D (f \circ g)|J_g|\end{aligned}$$

■

Para finalizar este capítulo (y este trabajo), demostraremos una pequeña extensión del teorema del cambio de variable para el caso en el que $J_g(\hat{x}) = 0$ en un subconjunto del dominio de g con medida de Jordan cero. Vale la pena que el lector recuerde la importancia que tuvo el hecho de que $J_g(\hat{x}) \neq 0$ en el teorema 4.3 así como en los lemas 4.12 y 4.13, resultados esenciales para la prueba del teorema del cambio de variable.

Teorema 4.6 *Sea $\Omega \subset \mathbb{R}^n$ abierto y $D \subset \mathbb{R}^n$ Jordan medible tal que $D \cup Fr(D) \subset \Omega$. Sea $g : \Omega \to \mathbb{R}^n$ de clase C^1 en Ω, inyectiva en $int(D)$ y tal que $J_g(\hat{x}) \neq 0$ para toda $\hat{x} \in int(D) \setminus A$ donde $J(A) = 0$. Si $f : g(D) \to \mathbb{R}$ es acotada y continua entonces*

$$\int_{g(D)} f = \int_D (f \circ g)|J_g|$$

Dem. Lo primero que hay que demostrar, es que bajo estas condiciones el conjunto $g(D)$ continúa siendo Jordan medible, para que de este modo tenga sentido hablar de $\int_{g(D)} f$; para esto, demostraremos que los conjuntos $g(A)$ y $g(D \setminus A)$ son Jordan medibles.

Como $J(A) = 0$ entonces $J(\bar{A}) = 0$, pero $\bar{A}$ es un compacto tal que $\bar{A} \subset \bar{D} \subset \Omega$ y $g \in C^1(\Omega)$, entonces $J(g(\bar{A})) = 0$ y por tanto $g(A)$ es Jordan medible y $J(g(A)) = 0$. Como $g \in C^1(\Omega)$ y $D \setminus A$ es Jordan medible; de acuerdo con el teorema 4.3, para ver que $g(D \setminus A)$ es Jordan medible nos faltaría demostrar que $J_g(\hat{x}) \neq 0$ para toda $\hat{x} \in int(D \setminus A)$, pero esto es inmediato pues $int(D \setminus A) \subset int(D) \setminus A$ y por hipótesis $J_g(\hat{x}) \neq 0$ para toda $\hat{x} \in int(D) \setminus A$, por lo tanto $g(D \setminus A)$ es Jordan medible, y como $g(D) = g(D \setminus A) \cup g(A)$ entonces $g(D)$ también es Jordan medible. Además como $J(A) = 0$ y $J(g(A)) = 0$ entonces por el teorema 3.9

$$\int_D (f \circ g)|J_g| = \int_{(D\setminus A)\cup A} (f \circ g)|J_g| = \int_{D\setminus A} (f \circ g)|J_g| + \int_A (f \circ g)|J_g| = \int_{D\setminus A} (f \circ g)|J_g|$$

y

$$\int_{g(D)} f = \int_{g(D\setminus A)\cup g(A)} f = \int_{g(D\setminus A)} f + \int_{g(A)} f = \int_{g(D\setminus A)} f$$

Como las hipótesis del teorema de cambio de variable se cumplen para el conjunto $D \setminus A$, tenemos que $\int_{g(D\setminus A)} f = \int_{D\setminus A} (f \circ g) \, |J_g|$ de donde

$$\int_{g(D)} f = \int_D (f \circ g) \, |J_g|$$

$\blacksquare$

Bibliografía

[1] Trotter HF, Williamson RE, Crowell RH. Calculus of vector functions. 2nd ed. United States of America: Prentice Hall; 1968

[2] Friedberg HS, Insel AJ, Spence LE. Linear Algebra. 4th ed. United States of America: Prentice Hall; 2003

[3] Sagan H. Advanced Calculus. United States of America: Houghton Mifflin Company; 1974

[4] Bartle RG. The Elements of Integration and Lebesgue Measure. United States of America: John Wiley & Sons; 1995

[5] Bartle RG. The Elements of Real Analysis. 2nd ed. United States of America: John Wiley & Sons; 1976

[6] Bartle RG, Sherbet DR. Introduction to Real Analysis. 2nd ed. United States of America: John Wiley & Sons; 1992

[7] Spivak M. Cálculos en variedades. España: Editorial Reverté; 1970

[8] Spivak M. Cálculo Infinitesimal. 2a ed. España: Editorial Reverté; 1992

[9] Rudin W. Principles of Mathematical Analysis. 3rd ed. Singapur: McGraw-Hill Kogakusha; 1976

[10] Marsden JE, Tromba AJ. Cálculo Vectorial. 5a ed. España: Pearson Educación; 2004

[11] Bachman G, Narici L. Functional Analysis. United States of America: Dover; 2000

[12] Halmos PR. Measure Theory. United States of America: Springer-Verlag; 1974

[13] Grabinsky G. Teoría de la medida. México: Las Prensas de Ciencias; 2009

[14] Dugundji J. Topology. United States of America: McGraw-Hill; 1996

[15] Casarrubias F, Tamariz A. Elementos de topología general. México; 2007

[16] Willard S. General Topology. United States of America: Dover; 2004