Universidad de Guanajuato

Teoría métrica de punto fijo, algunos resultados fundamentales

TESIS
Que para obtener el título de
Licenciado en Matemáticas

PRESENTA:
Alejandro Méndez Rojas

Director de tesis:
Dr. Fernando Núñez Medina

GUANAJUATO, GTO, MAYO 2019

Índice

Prólogo

Este trabajo es un compendio de algunos de los resultados principales de la teoría métrica del punto fijo, entre ellos: el teorema de contracción de Banach, el teorema de Kirk y los teoremas de Brouwer y de Schauder. La teoría del punto fijo es una de las herramientas más poderosas de la matemática moderna. Es una mezcla de análisis, topología y geometría, y sus resultados pueden ser aplicados en campos como la física, economía, teoría de juegos, biología, química, etcétera. En 1886, Poincaré fue uno de los primeros en trabajar en este campo. Más tarde, en 1912 Brouwer probó teoremas de punto fijo para el cuadrado, la esfera y sus contrapartes en n-dimensiones. Mientras tanto, en 1920 apareció el principio de contracción de Banach, el cual fue considerado uno de los principios fundamentales del análisis funcional. En 1930, el teorema de Brouwer fue extendido por Schauder a espacios de Banach. Más tarde, en 1965 Kirk probó su teorema para mapeos no expansivos.

Este trabajo está dividido en cuatro capítulos. En el primer capítulo damos resultados y definiciones que se usarán a lo largo de la tesis. En la sección 1.1 se introducen los preliminares. En la subsección 1.1.1 introducimos notación y resultados generales que serán usados más adelante. En la sección 1.1.2 damos una breve introducción a la topología débil, junto a unos resultados que serán de ayuda para la lectura de esta tesis. En la sección 1.2 damos una introducción breve a la teoría del punto fijo, que es el tema principal de esta tesis. Indicamos cuál es su objeto de estudio y para qué sirve.

En el capítulo 2 enunciamos el teorema de contracción de Banach, el cual dio inicio a la teoría del punto fijo en espacios métricos. Dicho teorema, a pesar de parecer muy sencillo es un resultado muy útil, por lo que es una herramienta muy popular para resolver problemas de existencia de puntos fijos en análisis. Una de las aplicaciones más conocidas del teorema de contracción de Banach es la solución del problema de valor inicial de Cauchy, la cual mostramos es este capítulo. También introducimos la propiedad del punto fijo (PPF) en espacios de Banach.

Uno de los resultados que veremos en el capítulo 2 es que, en un espacio de Banach X, los subconjuntos no vacíos, convexos y compactos tienen la propiedad del punto fijo. Sin embargo, la compacidad en un espacio de Banach suele ser una propiedad muy fuerte, por lo que es deseable sustituir esta hipótesis por una menos restrictiva, por ejemplo la compacidad débil. En el capítulo 3 veremos que si K es un subconjunto con estructura normal (concepto que definiremos en este capítulo), $\omega-$compacto, no vacío y convexo de un espacio de Banach, entonces K tiene la propiedad del punto fijo. El resultado anterior es conocido como el teorema de Kirk, y como ya mencionamos, fue probado en 1965. Para probar el teorema de Kirk estudiamos antes varias propiedades de los conjuntos con estructura normal.

En el capítulo 4 estudiamos los teoremas de Brouwer y de Schauder. Para ello, en la sección 4.1 desarrollamos la teoría del grado en $\mathbb{R}^n$, la cual señala la existencia de una función d, definida en el conjunto

$$\{f : \overline{\Omega} \to \mathbb{Z} |\ f \text{ es continua, } \Omega \subset \mathbb{R}^n \text{ es abierto y acotado, y } y \in \mathbb{R}^n \backslash f(\partial\Omega)\},$$

tal que $d(f, \Omega, y)$ nos da información sobre las soluciones en Ω de la ecuación $f(x) = y$. Cabe destacar que en este trabajo estudiaremos la teoría del grado de manera puramente analítica, y no topológica. Además de utilizar la teoría del grado en $\mathbb{R}^n$ para probar los teoremas de Brouwer y de Schauder, como otro ejemplo de su utilidad, mostramos el teorema del erizo. En la sección 4.2 probamos algunas propiedades del grado que son de utilidad para probar el teorema de Brouwer, y finalmente demostramos el teorema de Schauder. También probamos el teorema de Perron-Frobenius como una aplicación del teorema de Brouwer.

Agradecimientos

A mi familia que siempre me ha apoyado cuando lo he requerido.

A mi asesor de tesis, el Dr. Fernando Núñez Medina, por su paciencia durante este proceso y por todo lo que me enseñó durante el mismo y a lo largo de la licenciatura.

Al CIMAT y la Universidad de Guanajuato por los apoyos que recibí durante la licenciatura.

A todas aquellas personas que me impartieron clases, de quienes aprendí, con quienes me divertí y quienes me guiaron, porque soy quien soy gracias a todas aquellas experiencias.

Al proyecto SEP-Conacyt 243722, "Propiedad de Punto fijo en espacios de Banach".

Capítulo 1

Introducción

1.1 Preliminares

En esta sección veremos los preliminares necesarios para entender los resultados de esta tesis, así como la notación que se empleará en la misma.

1.1.1 Notación, definiciones y algunos resultados

Sea (X, d) un espacio métrico. Dados $x \in X$ y $r > 0$, denotamos la bola abierta con centro en x y radio r como $B_r(x)$. Además, si X es un espacio normado, B_r denotará la bola abierta con centro en 0 y radio r.

Los espacios vectoriales considerados en este trabajo serán reales. Sean X un espacio vectorial de sucesiones y $x \in X$. Dado $i \in \mathbb{N}$ denotaremos por $(x)_i$ la i-ésima coordenada de x.

Antes de enunciar el lema siguiente, recordemos el concepto de diámetro de un conjunto.

Definición 1 (Diámetro de un conjunto) Sea X un subconjunto acotado de un espacio normado. Definimos el diámetro de X, denotado por $diam(X)$, como

$$diam(X) = \sup\{||u - v|| : u, v \in X\}.$$

Definición 2 (Conjunto convexo) Sea A un subconjunto de un espacio vectorial. Decimos que A es convexo si para todo $x, y \in A$ y $\lambda \in [0, 1]$,

$$\lambda x + (1 - \lambda)y \in A.$$

Lema 3 Sean X un espacio de normado y $A \subset X$. Entonces

(1) $diam(A) = diam(\overline{A})$;
(2) Si A es convexo entonces $\overline{A}$ es convexo.

Demostración (1) Es claro que $diam(A) \leq diam(\overline{A})$. Sean $x, y \in \overline{A}$. Entonces existen (x_n) y (y_n) en A tales que $x_n \to x$ y $y_n \to y$. Notemos que

$$||x_n - y_n|| \leq diam(A) \text{ para todo } n \in \mathbb{N},$$

de donde

$$||x - y|| = \lim_{n\to\infty} ||x_n - y_n|| \leq diam(A).$$

Así $diam(\overline{A}) \leq diam(A)$ y en consecuencia $diam(A) = diam(\overline{A})$.

(2) Sean $x, y \in \overline{A}$ y $\lambda \in [0,1]$. Tomemos (x_n) y (y_n) en A tales que $x_n \to x$ y $y_n \to y$. Como $x_n, y_n \in A$ para todo $n \in \mathbb{N}$ y A es convexo, $\lambda x_n + (1-\lambda)y_n \in A$ para todo $n \in \mathbb{N}$, de donde

$$\lambda x + (1-\lambda)y = \lim_{n\to\infty} \lambda x_n + (1-\lambda)y_n \in \overline{A},$$

que es lo que queríamos demostrar. ■

Definición 4 Sean X un espacio de Banach y $A \subset X$. Denotaremos por $conv(A)$ (respectivamente $\overline{conv}(A)$) a la intersección de todos los subconjuntos convexos (respectivamente convexos y cerrados) de X que contienen a A.

Lema 5 Sean X un espacio vectorial y $A \subset X$. Entonces

$$conv(A) = \Big\{ \sum_{i=1}^{n} \alpha_i x_i : n \in \mathbb{N}, \sum_{i=1}^{n} \alpha_i = 1, \alpha_i \geq 0 \text{ y } x_i \in A, i = 1, ..., n \Big\}.$$

Demostración Para simplificar la notación, para cada $n \in \mathbb{N}$ definamos

$$C_n = \Big\{ \sum_{i=1}^{n} \alpha_i x_i : \sum_{i=1}^{n} \alpha_i = 1, \alpha_i \geq 0 \text{ y } x_i \in A, i = 1, ..., n \Big\}$$

y

$$C = \Big\{ \sum_{i=1}^{n} \alpha_i x_i : n \in \mathbb{N}, \sum_{i=1}^{n} \alpha_i = 1, \alpha_i \geq 0 \text{ y } x_i \in A, i = 1, ..., n \Big\}.$$

Notemos que $C = \bigcup_{n=1}^{\infty} C_n$. Para probar que $C \subset conv(A)$, probaremos que para todo B subconjunto convexo de X con $A \subset B$, se tiene que $C \subset B$. Sea pues B un subconjunto convexo de X que contiene a A. Probaremos por inducción que $C_n \subset B$, y en consecuencia se tendrá que $C \subset B$.

El resultado es claro para $n = 1$. Supongamos que para m se tiene el resultado. Sea $c \in C_{m+1}$. Entonces

$$c = \sum_{i=1}^{m+1} \alpha_i x_i, \text{ donde } \alpha_i \geq 0, \sum_{i=1}^{m+1} \alpha_i = 1, x_i \in A, i = 1, ..., m+1.$$

Notemos que

$$\begin{aligned} c = \sum_{i=1}^{m+1} \alpha_i x_i &= \sum_{i=1}^{m} \alpha_i x_i + \alpha_{m+1} x_{m+1} \\ &= (1 - \alpha_{m+1}) \sum_{i=1}^{m} \frac{\alpha_i}{1 - \alpha_{m+1}} x_i + \alpha_{m+1} x_{m+1}. \end{aligned}$$

Por hipótesis, $\sum_{i=1}^{m} \frac{\alpha_i}{1-\alpha_{m+1}} x_i \in B$ y como $x_{m+1} \in B$, por convexidad de B, $c \in B$. Por lo tanto $C \subset conv(A)$.

Por otro lado, para ver que $conv(A) \subset C$, notemos que $A \subset C$ y por lo tanto $conv(A) \subset conv(C)$, por lo que basta probar que C es convexo. Sean $c, c' \in C$. Entonces

$$c = \sum_{i=1}^{n} \alpha_i x_i \text{ y } c' = \sum_{i=1}^{n'} \alpha_i' x_i',$$

con

$$\alpha_i \geq 0, \sum_{i=1}^{n} \alpha_i = 1, x_i \in A, i = 1, ..., n$$

y

$$\alpha_i' \geq 0, \sum_{i=1}^{n'} \alpha_i' = 1, x_i' \in A, i = 1, ..., n'.$$

Notemos que si $t \in [0,1]$, entonces

$$\begin{aligned} tc + (1-t)c' &= t \sum_{i=1}^{n} \alpha_i x_i + (1-t) \sum_{i=1}^{n'} \alpha_i' x_i' \\ &= \sum_{i=1}^{n} t\alpha_i x_i + \sum_{i=1}^{n'} (1-t)\alpha_i' x_i', \end{aligned}$$

donde

$$\sum_{i=1}^{n} t\alpha_i + \sum_{i=1}^{n'} (1-t)\alpha_i' = t + (1-t) = 1.$$

Por lo tanto $tc + (1-t)c' \in C$ y C es convexo. $\blacksquare$

Lema 6 Sean X un espacio de Banach y $A \subset X$ no vacío y acotado. Entonces

(1) $diam(A) = diam(conv(A))$;

(2) $diam(A) = diam(\overline{conv}(A))$.

Demostración (1) Recordemos que

$$A \subset conv(A),$$

por lo cual

$$diam(A) \leq diam(conv(A)).$$

Definamos la función $f_x : X \to \mathbb{R}$ como $f_x(y) = ||x - y||$. Es claro que $f_x(y) = f_y(x)$. Notemos que si $y, z \in X$ y $\lambda \in [0,1]$,

$$\begin{aligned} f_x(\lambda y + (1-\lambda)z) &= ||x - (\lambda y + (1-\lambda)z)|| \\ &= ||\lambda(x-y) + (1-\lambda)(x-z)|| \\ &\leq \lambda||x-y|| + (1-\lambda)||x-z||, \end{aligned}$$

de donde

$$f_x(\lambda y + (1-\lambda)z) \le \lambda f_x(y) + (1-\lambda) f_x(z). \tag{1.1}$$

En general, si $\sum_{i=1}^{n} \lambda_i = 1$, $\lambda_i \in [0,1]$, $x_i \in X$ para $i = 1, ..., n$, entonces

$$f_x\left(\sum_{i=1}^{n} \lambda_i x_i\right) \le \sum_{i=1}^{n} \lambda_i f_x(x_i). \tag{1.2}$$

Probaremos ahora que $diam(A) \ge diam(conv(A))$. Sean $l = diam(A)$ y $x, y \in conv(A)$. Existen entonces $N_1, N_2 \in \mathbb{N}$ tales que

$$x = \sum_{i=1}^{N_1} \alpha_i x_i \text{ y } y = \sum_{i=1}^{N_2} \beta_i y_i,$$

con $x_i, y_i \in A$ y $\alpha_i, \beta_i \in [0,1]$ con $\sum_{i=1}^{N_1} \alpha_i = \sum_{i=1}^{N_2} \beta_i = 1$. Notemos que

$$\begin{aligned}
||x-y|| &= f_x(y) \\
&= f_x\left(\sum_{j=1}^{N_2} \beta_j y_j\right) \\
&\le \sum_{j=1}^{N_2} \beta_j f_x(y_j) \\
&= \sum_{j=1}^{N_2} \beta_j f_{y_j}\left(\sum_{i=1}^{N_1} \alpha_i x_i\right) \text{ (por (1.2))} \\
&\le \sum_{j=1}^{N_2} \beta_j \sum_{i=1}^{N_1} \alpha_i f_{y_j}(x_i) \\
&\le \sum_{j=1}^{N_2} \beta_j \sum_{i=1}^{N_1} \alpha_i l = l,
\end{aligned}$$

de donde $diam(conv(A)) \le diam(A)$ y por lo tanto $diam(conv(A)) = diam(A)$.

(2) Por el inciso (1) del lema 3 se tiene lo buscado. ■

Acontinuación recordaremos el criterio M de Weierstrass, el cual usaremos en la proposición siguiente.

Teorema 7 (Criterio M de Weierstrass) Sea $\{f_n\}$ una sucesión de funciones real-valuadas definidas en un espacio normado X y supongamos que

$$|f_n(x)| \le M_n \; x \in X, \, n \in \mathbb{N}.$$

Entonces $\sum_{n=1}^{\infty} f_n$ converge uniformemente en X si $\sum_{n=1}^{\infty} M_n$ converge.

Sean (X, d) un espacio métrico, $A \subset X$ y $x \in X$. Recordemos que la distancia del punto x al conjunto A se define como

$$d(x, A) = \inf\{d(x, y) : y \in A\}.$$

La proposición siguiente nos será de utilidad más adelante para probar que dada una función continua definida en un subconjunto compacto de $\mathbb{R}^n$, existe una función de clase C^∞ tan "cercana" a ella como se requiera.

Proposición 8 Sean $A \subset \mathbb{R}^n$ no vacío y compacto, $\{a^1, a^2, ...\}$ un subconjunto denso de A y $f : A \to \mathbb{R}^n$ continua. Entonces la función $\widetilde{f} : \mathbb{R}^n \to \mathbb{R}^n$ dada por

$$\widetilde{f}(x) = \begin{cases} f(x) & \text{si } x \in A, \\ \left(\sum_{i=1}^{\infty} 2^{-i}\varphi_i(x)\right)^{-1} \sum_{i=1}^{\infty} 2^{-i}\varphi_i(x) f(a^i) & \text{si } x \notin A \end{cases}$$

define una extensión continua de f, donde

$$\varphi_i(x) = \max\left\{2 - \frac{\|x - a^i\|}{d(x, A)}, 0\right\} \text{ para } x \notin A.$$

Demostración Notemos que como A es compacto, entonces dicho conjunto $\{a^1, a^2, ...\}$ denso en A existe. Claramente cada función $\varphi_i : \mathbb{R}^n \backslash A \to \mathbb{R}$ está bien definida ya que $d(x, A) \neq 0$ para $x \notin A$, y además es continua. Veremos ahora que la función $\widetilde{f}$ está bien definida. Como $d(x, A) \leq \|x - a^i\|$, se tiene que $0 \leq \varphi_i \leq 1$ para toda $i \in \mathbb{N}$. Luego por el criterio M de Weierstrass con $M_i = 2^{-i}$ la serie $\sum_{i=1}^{\infty} 2^{-i}\varphi_i(x)$ converge uniformemente en $\mathbb{R}^n \backslash A$. Además, como la función f es continua y A es compacto, existe $M \in \mathbb{N}$ tal que $|f(a)| < M$ para todo $a \in A$, de donde nuevamente por el criterio M de Weierstrass la serie $\sum_{i=1}^{\infty} 2^{-i}\varphi_i(x) f(a^i)$ converge uniformemente en $\mathbb{R}^n \backslash A$. Por lo tanto, la función $\widetilde{f}$ está bien definida y además $\widetilde{f}\Big|_{\mathbb{R}^n \backslash A}$ es continua por estar definida como producto de funciones continuas. Por definición $\widetilde{f}$ es extensión de f. Resta probar que $\widetilde{f}$ es continua. Sea $x_0 \in \mathbb{R}^n$. Si $x_0 \in \mathbb{R}^n \backslash A$, puesto que $\mathbb{R}^n \backslash A$ es abierto y $\widetilde{f}\Big|_{\mathbb{R}^n \backslash A}$ es continua, entonces $\tilde{f}$ es continua en x_0. Supongamos ahora que $x_0 \in A$. Sea $\varepsilon > 0$. Como $f\big|_A$ es continua, existe $\delta > 0$ tal que si $x \in A \cap B_\delta(x_0)$, entonces

$$\left\|\widetilde{f}(x) - \widetilde{f}(x_0)\right\| = \|f(x) - f(x_0)\| < \varepsilon.$$

Notemos que $\varphi_i(x) = 0$ si $2d(x, A) \leq \|x - a^i\|$. Así, si $x \in (\mathbb{R}^n \backslash A) \cap B_{\delta/4}(x_0)$, entonces $\varphi_i(x) = 0$ para toda i tal que $a^i \notin B_\delta(x_0)$, pues para tales x e i,

$$\|x - a^i\| \geq \|x_0 - a^i\| - \|x - x_0\| \geq \delta - \frac{\delta}{4} > \frac{\delta}{2} > 2\|x - x_0\| \geq 2d(x, A).$$

Por simplicidad definamos

$$\psi_i(x) = \frac{2^{-i}\varphi_i(x)}{\sum_{i=1}^{\infty} 2^{-i}\varphi_i(x)}, \quad x \notin A, \quad i \in \mathbb{N}.$$

Así, para $x \notin A$, $\widetilde{f}(x) = \sum_{i=1}^{\infty} \psi_i(x) f(a^i)$. Notemos que para toda $i \in \mathbb{N}$ y $x \notin A$ se tiene que

$$0 \leq \psi_i(x) \leq 1 \text{ y } \sum_{i=1}^{\infty} \psi_i(x) = 1.$$

Luego, si $x \in (\mathbb{R}^n \backslash A) \cap B_{\delta/4}(x_0)$, entonces

$$\begin{aligned}
\left\| \widetilde{f}(x) - \widetilde{f}(x_0) \right\| &= \left\| \sum_{i=1}^{\infty} \psi_i(x) \left[f\left(a^i\right) - f(x_0) \right] \right\| \\
&\leq \sum_{i=1}^{\infty} \psi_i(x) \left\| f\left(a^i\right) - f(x_0) \right\| \\
&< \varepsilon.
\end{aligned}$$

Por lo que hemos probado, si $x \in B_{\delta/4}(x_0)$, entonces $\left\| \widetilde{f}(x) - \widetilde{f}(x_0) \right\| < \varepsilon$. Así $\widetilde{f}$ es continua en x_0. ■

Definición 9 (Propiedad de la intersección finita) Sea X un conjunto. Se dice que una colección $\mathfrak{C}$ de subconjuntos de X satisface la propiedad de la intersección finita si para toda subcolección finita

$$\{C_1, C_2, ..., C_n\}$$

de $\mathfrak{C}$, la intersección $\cap_{i=1}^n C_i$ es no vacía.

La demostración del siguiente resultado se puede encontrar en [9].

Teorema 10 Sea X un espacio topológico. Entonces X es compacto si y sólo si para toda colección $\mathfrak{C}$ de conjuntos cerrados en X con la propiedad de la intersección finita, la intersección $\cap_{C \in \mathfrak{C}} C$ de todos los elementos de $\mathfrak{C}$ es no vacía.

Antes de enunciar el siguiente resultado es necesario introducir la siguiente notación.

Definición 11 Sean $E \subset \mathbb{R}^n$ abierto, $f : E \to \mathbb{R}^m$ y $x \in E$.

(a) Si existe una transformación lineal $A : \mathbb{R}^n \to \mathbb{R}^m$ tal

$$\lim_{h \to 0} \frac{||f(x+h) - f(x) - Ah||}{||h||} = 0,$$

decimos que f es diferenciable en x y escribimos $f'(x) = A$.

(b) Al determinante del operador lineal $f'(x)$ se le llama jacobiano de f en x, y se denota como $J_f(x) = \det f'(x)$.

Definición 12 Sean $E \subset \mathbb{R}^n$ abierto y $f : E \to \mathbb{R}^m$ una función. Sean $\{e_1, ..., e_n\}$ y $\{u_1, ..., u_m\}$ las bases estándar de $\mathbb{R}^n$ y $\mathbb{R}^m$ respectivamente. Las componentes de f son las funciones reales $f_1, ..., f_m$ definidas como

$$f_i(x) = f(x) \cdot u_i \text{ para } x \in E.$$

Por simplicidad en ocasiones escribiremos de manera más simple $\partial_j f_i$ en vez de $\partial f_i / \partial x_j$.

Definición 13 Sean Ω un subconjunto abierto de $\mathbb{R}^n$ y $k \in \mathbb{N}$.

(a) Denotamos por $C(\Omega, \mathbb{R}^m)$ al conjunto de funciones $f : \Omega \to \mathbb{R}^m$ que son continuas.

(b) Denotamos por $C^k(\Omega, \mathbb{R}^m)$ al conjunto de funciones $f : \Omega \to \mathbb{R}^m$ cuyas funciones componentes tienen derivadas parciales continuas hasta el orden k.

(c) Denotaremos por $\overline{C}^k(\Omega, \mathbb{R}^m)$ al conjunto de funciones continuas $f : \overline{\Omega} \to \mathbb{R}^m$ tales que $f \in C^k(\Omega, \mathbb{R}^n)$.

(d) Definimos $C^\infty(\Omega, \mathbb{R}^m)$ y $\overline{C}^\infty(\Omega, \mathbb{R}^m)$ como la intersección de todos los $C^k(\Omega, \mathbb{R}^m)$ y $\overline{C}^k(\Omega, \mathbb{R}^m)$ respectivamente.

Para simplificar la notación, si $n = m$, denotaremos a $C(\Omega, \mathbb{R}^n)$ por $C(\Omega)$, a $C^k(\Omega, \mathbb{R}^n)$ por $C^k(\Omega)$, etcétera.

Ahora enunciaremos los teoremas de la función inversa, cambio de variable y Arzelá-Ascoli, los cuales usaremos en el capítulo 4, y pueden encontrarse en [11].

Teorema 14 (de la función inversa) Sean Ω abierto en $\mathbb{R}^n$, $f \in C^1(\Omega)$ y $J_f(x_0) \neq 0$ para algún $x_0 \in \Omega$. Entonces existe una vecindad U de x_0 tal que $f\big|_U$ es un homeomorfismo sobre una vecindad de $f(x_0)$.

Recordemos que el soporte de una función $f : \mathbb{R}^n \to \mathbb{R}$ es la cerradura del conjunto de todos los puntos $x \in \mathbb{R}^n$ en los que $f(x) \neq 0$.

Teorema 15 (de cambio de variable) Sean $E \subset \mathbb{R}^n$ un conjunto abierto y $T \in C^1(E)$ un mapeo inyectivo tal que $J_T(x) \neq 0$ para todo $x \in E$. Si f es una función continua en $\mathbb{R}^n$ con soporte compacto en $T(E)$, entonces

$$\int_{\mathbb{R}^n} f(y)dy = \int_{\mathbb{R}^n} f(T(x))|J_T(x)|dx.$$

Las siguientes dos definiciones nos ayudarán a enunciar el teorema de Arzelá-Ascoli.

Definición 16 Sea $\{f_n\}$ una sucesión de funciones reales definidas en un conjunto E.

(a) Decimos que $\{f_n\}$ es acotada puntualmente en E si la sucesión $\{f_n(x)\}$ es acotada para todo $x \in E$, esto es, si existe una función ϕ en E con valores reales tal que

$$|f_n(x)| < \phi(x) \ (x \in E, n = 1, 2, 3, ...).$$

(b) Decimos que $\{f_n\}$ es uniformemente acotada en E si existe un número M tal que

$$|f_n(x)| < M \ (x \in E, n = 1, 2, 3, ...).$$

Definición 17 (Equicontinuidad) Sea $\mathfrak{F}$ una familia de funciones complejas definida en un subconjunto E de un espacio métrico. Diremos que $\mathfrak{F}$ es equicontinua si para cada $\varepsilon > 0$, existe $\delta > 0$ tal que

$$|f(x) - f(y)| < \varepsilon \text{ si } x, y \in E,\ d(x, y) < \delta \text{ y } f \in \mathfrak{F}.$$

Teorema 18 (Arzelá-Ascoli) Sean X un espacio métrico, $K \subset X$ compacto y $\{f_n\}$ una sucesión en $C(K, \mathbb{R})$. Si $\{f_n\}$ es puntualmente acotada y equicontinua entonces

(a) $\{f_n\}$ es uniformemente acotada;
(b) $\{f_n\}$ tiene una subsucesión que converge uniformemente.

La siguiente proposición nos será de utilidad en la sección de la unicidad del grado en $\mathbb{R}^n$, y puede encontrarsse en [1].

Proposición 19 Sean $\Omega \subset \mathbb{R}^n$ abierto, $f \in C^2(\Omega)$ y $d_{ij}(x)$ el cofactor de $\frac{\partial f_j(x)}{\partial x_i}$ en $J_f(x)$, es decir, $(-1)^{i+j}$ veces el determinante obtenido de $J_f(x)$ cancelando el j-ésimo renglón y la i-ésima columna. Entonces,

$$\sum_{i=1}^{n} \frac{\partial d_{ij}(x)}{\partial x_i} = 0 \text{ para } j = 1, ..., n.$$

Demostración Fijemos $1 \leq j \leq n$ y para $1 \leq k \leq n$ denotemos por f_{x_k} a la columna

$$(\partial_k f_1(x), ..., \partial_k f_{j-1}(x), \partial_k f_{j+1}(x), ..., \partial_k f_n(x))^T.$$

Entonces

$$d_{ij}(x) = (-1)^{i+j} \det(f_{x_1}, ..., f_{x_{i-1}}, \tilde{f}_{x_i}, f_{x_{i+1}}, ..., f_{x_n}),$$

donde la tilde indica que se omite la columna señalada. Notemos que, de la fórmula de Leibniz para el determinante,

$$\begin{aligned}
\partial_i d_{ij}(x) &= \partial_i \Big[(-1)^{i+j} \sum_{\sigma \in \tilde{P}_n} (Sgn(\sigma) \prod_{l=1, l \neq i}^{n} \partial_{\sigma_l} f_l(x)) \Big] \\
&= (-1)^{i+j} \sum_{\sigma \in \tilde{P}_n} Sgn(\sigma) \partial_i \Big[\prod_{l=1, l \neq i}^{n} \partial_{\sigma_l} f_l(x) \Big] \\
&= (-1)^{i+j} \sum_{\sigma \in \tilde{P}_n} Sgn(\sigma) \Big(\sum_{k=1, k \neq i}^{n} \partial_i \partial_{\sigma_k} f_k(x) \prod_{l=1, l \neq i, l \neq k}^{n} \partial_{\sigma_l} f_l(x) \Big) \\
&= (-1)^{i+j} \sum_{k=1, k \neq i}^{n} \Big(\sum_{\sigma \in \tilde{P}_n} Sgn(\sigma) \partial_i \partial_{\sigma_k} f_k(x) \prod_{l=1, l \neq i, l \neq k}^{n} \partial_{\sigma_l} f_l(x) \Big) \\
&= (-1)^{i+j} \sum_{k=1, k \neq i}^{n} \det(f_{x_1}, ..., \tilde{f}_{x_i}, ... f_{x_{k-1}}, \partial_i f_{x_k}, f_{x_{k+1}}, ..., f_{x_n}),
\end{aligned}$$

donde $\tilde{P}_n$ denota el conjunto de todas las permutaciones de los números $1, 2, ...,$ $i-1, i+1, ..., n$, $Sgn(\sigma)$ es la signatura de σ y σ_l denota la posición del número l en la permutación σ. Para $1 \leq i < k \leq n$, definamos

$$c_{ki} = \det(\partial_i f_{x_k}, f_{x_1}, ..., \tilde{f}_{x_i}, ..., \tilde{f}_{x_k}, ..., f_{x_n})$$

y

$$c_{ik} = \det(\partial_k f_{x_i}, f_{x_1}, ..., \tilde{f}_{x_i}, ..., \tilde{f}_{x_k}, ..., f_{x_n}).$$

Notemos que como $f \in C^2(\Omega)$, $c_{ik} = c_{ki}$, y como el determinante cambia de signo si intercambiamos dos columnas consecutivas, entonces

$$\begin{aligned}(-1)^{i+j}\partial_i d_{ij}(x) &= \sum_{k=1,k\neq i}^{n} \det(f_{x_1}, ..., \tilde{f}_{x_i}, ...f_{x_{k-1}}) \\ &= \sum_{k<i}(-1)^{k-1}c_{ki} + \sum_{k>i}(-1)^{k-2}c_{ki} \\ &= \sum_{k=1,k\neq i}^{n} (-1)^{k-1}\gamma_{ki}c_{ki},\end{aligned}$$

donde $\gamma_{ki} = 1$ para $k < i$, y $\gamma_{ki} = -1$ para $k > i$. Entonces

$$(-1)^j \sum_{i=1}^{n} \partial_i d_{ij}(x) = \sum_{i,k=1,k\neq i}^{n} (-1)^{k-1+i}\gamma_{ki}c_{ki} = 0,$$

pues en la última suma aparece el inverso aditivo de cada sumando. ■

1.1.2 La topología débil

Una de las primeras cosas que notamos al comenzar a trabajar con espacios normados de dimensión arbitraria (no necesariamente finita) es la reducción de subconjuntos compactos del espacio en cuestión. Uno de los ejemplos más clásicos al respecto es que en un espacio de Banach de dimensión infinita X, la bola cerrada unitaria B_X no es compacta. Con el fin de tener una mayor cantidad de conjuntos compactos consideraremos una topología distinta para X, la cual, al tener una menor cantidad de abiertos acepta una mayor cantidad de compactos que la topología inducida por la norma de X. Tal topología es la llamada topología débil, la cual definieremos a continuación. La demostración de los resultados siguientes puede encontrarse en [7].

Definición 20 (Topología débil) Sean X un espacio de Banach y X^* su espacio dual. La topología débil es la topología generada por la base que consiste de los conjuntos

$$B(x_0, f_1, ..., f_n, \varepsilon) = \{x \in X : |f_i(x - x_0)| < \varepsilon, i = 1, ..., n\},$$

donde $x_0 \in X$, $f_1, ..., f_n \in X^*$, $n \in \mathbb{N}$ y $\varepsilon > 0$.

La topología débil es la mínima topología que hace continuos a los elementos de X^*. Denotaremos la topología débil de X como τ_ω, mientras que la topología inducida por la norma de X la denotaremos por $\tau_{||\cdot||}$. Escribiremos ω-abierto para indicar que un conjunto es abierto con respecto a la topología débil. De manera análoga con ω-compacto, ω-continuo, etcétera.

Como consecuencia del teorema de Hahn Banach, dado $x \in X$ con $x \neq 0$, existe $f \in X^*$ tal que $f(x) \neq 0$. En consecuencia la topología débil es Hausdorff, y como ya comentamos, $\tau_\omega \subset \tau_{||\cdot||}$.

Definición 21 Sea X un espacio de Banach.

(a) Diremos que $A \subset X$ es débilmente acotado (ω-acotado) si para toda $f \in X^*$, $f(A)$ es acotado.
(b) Diremos que una sucesión (x_n) en X es débilmente convergente (ω-convergente) a $x \in X$ si para toda $f \in X^*$, la sucesión $(f(x_n))$ es convergente a $f(x)$.

Teorema 22 Sean X un espacio de Banach y $A \subset X$. Entonces A es acotado si y sólo si A es ω-acotado.

Teorema 23 (Mazur) Sean X un espacio de Banach y $A \subset X$. Si A es ω-cerrado, entonces A es cerrado. Si A es convexo y cerrado, entonces es ω-cerrado.

1.2 ¿Qué es la teoría del punto fijo?

En esta sección veremos una breve introducción a la teoría del punto fijo en espacios métricos, en particular en espacios de Banach.

Es natural hacernos dos preguntas, **¿qué es la teoría del punto fijo?**, y, **¿para qué sirve?** Para responder estas preguntas comencemos con la definición de punto fijo.

Definición 24 Sean X un conjunto y $f : X \to X$ una función. Se dice que $x \in X$ es un punto fijo de f si se cumple que $f(x) = x$.

Sabiendo ahora lo que es un *punto fijo*, podemos dar una respuesta a la primer pregunta. La teoría del punto fijo consiste en buscar propiedades sobre un conjunto X y una familia de funciones $\mathfrak{F}$ de X en X para así poder asegurar la existencia de algún punto fijo para cualquier $f \in \mathfrak{F}$.

Ahora, ¿para qué sirve la teoría del punto fijo? Como todos sabemos, muchos problemas que se encuentran en las llamadas ciencias exactas y en otras áreas del conocimiento podemos plantearlos de forma matemática. Por tal motivo, es común el buscar resolver una ecuación o sistema de ecuaciones de la forma $f(x) = 0$, donde $f : X \to X$ es una fucnción y X es un espacio vectorial. Consideremos a la función $g : X \to X$ definida como $g(x) = f(x) + x$. Notemos que x es solución de la ecuación $f(x) = 0$ si y sólo si x es un punto fijo de la función g. De esta forma, pasamos de buscar un cero para f a buscar un punto fijo de la función g. Si bien puede que el replanteamiento de buscar un punto fijo en lugar de un cero de una función no parezca cambiar realmente el problema, el cambiar nuestro punto de vista sobre el problema hace que nos resulte más natural utilizar una gran cantidad de herramientas, de las cuales disponemos gracias a la teoría del punto fijo.

Debido al replanteamiento anterior, la teoría del punto fijo nos permite dar respuesta a una gran cantidad de problemas que pueden surgir de manera natural en muchas áreas del conocimiento, siendo una de las razones por las que resulta ser tan importante.

Como comentamos anteriormente, la teoría del punto fijo consiste en determinar qué propiedades sobre un conjunto X y una familia de funciones $\mathfrak{F}$ de X en X aseguran la existencia de algún punto fijo para cualquier f en $\mathfrak{F}$. Si nos restringimos al caso en que el conjunto X es un espacio topológico y la familia $\mathfrak{F}$ es el conjunto de todas las funciones continuas de X en X, entonces hablamos de la teoría topológica del punto fijo. En particular, si X es un espacio métrico hablamos de la teoría métrica del punto fijo. Como comentamos en el prólogo de este trabajo, los pilares de la teoría métrica del punto fijo son los teoremas de Brouwer, Schauder, Banach, y Kirk. En los capitulos siguientes estudiaremos estos teoremas y algunas de sus consecuencias, además de proporcionar algunas de sus aplicaciones.

Capítulo 2

El teorema de contracción de Banach y la propiedad del punto fijo en espacios métricos

En su tesis doctoral (1920, publicada en 1922), el matemático polaco Stefan Banach estableció un importante teorema de punto fijo conocido como *Teorema de Contracción de Banach*, el cual es uno de los resultados más importantes en análisis y considerado el inicio de la teoría del punto fijo en espacios métricos. El teorema de contracción de Banach fue primeramente enunciado y probado por Banach para las llamadas funciones contractivas en espacios vectoriales normados y completos, nombrados ahora espacios de Banach. Los resultados de este capítulo pueden encontrarse en [3].

Definición 25 Sean (M, d) un espacio métrico y $T : M \to M$. Se dice que T es de Lipschitz si existe $k \geq 0$ tal que

$$d(Tx, Ty) \leq kd(x, y),\ x, y \in M.$$

El menor k que cumple con lo anterior es llamado constante de Lipschitz de T. A las funciones con constante de Lipschitz menor a 1 les llamaremos contracciones.

Lema 26 Sean (M, d) un espacio métrico y $T, S : M \to M$ funciones de Lipschitz. Entonces

$$k(T \circ S) \leq k(T)k(S).$$

Demostración Dados x, $y \in M$,

$$\begin{aligned} d((T \circ S)x, (T \circ S)y) &= d(T(Sx), T(Sy)) \\ &\leq k(T)d(S(x), S(y)) \\ &\leq k(T)k(S)d(s, y), \end{aligned}$$

por lo tanto $k(T \circ S) \leq k(T)k(S)$. ■

Notemos que dados (M, d) un espacio métrico y $T : M \to M$ una función de Lipschitz, del lema anterior se tiene que $k(T^n) \leq (k(T))^n$, donde T^n representa la n-ésima iterada de T, por lo cual T^n es de Lipschitz para todo $n \in \mathbb{N}$.

Es momento de enunciar y demostrar el teorema de contracción de Banach.

Teorema 27 (de contracción de Banach) Sean (M, d) un espacio métrico, no vacío y completo y $T : M \to M$ una contracción. Entonces T tiene un único punto fijo $z \in M$ y además $\lim_{n\to\infty} T^n x_0 = z$ para todo $x_0 \in M$.

Demostración Sean k la constante de Lipschitz de T y $x_0 \in M$. Definamos la sucesión (x_n) como $x_n = T^n x_0$, $n \geq 0$. Entonces, dados $n, p \in \mathbb{N}$,

$$\begin{aligned}
d(x_n, x_{n+p}) &= d(T^n x_0, T^{n+p} x_0) \\
&= d(T^n x_0, T^n(T^p x_0)) \\
&\leq k(T^n) d(x_0, T^p x_0) \\
&\leq k^n (d(x_0, Tx_0) + d(Tx_0, T^2 x_0) + ... + d(T^{p-1} x_0, T^p x_0)) \\
&\leq k^n (1 + k + ... + k^{p-1}) d(x_0, Tx_0) \\
&= k^n \Big(\frac{1 - k^p}{1 - k}\Big) d(x_0, Tx_0).
\end{aligned}$$

Como $k < 1$, la sucesión (x_n) es de Cauchy, y puesto que M es completo, entonces converge, digamos a un punto z. Por lo tanto, por la continuidad de T,

$$z = \lim_{n\to\infty} x_n = \lim_{n\to\infty} T^n x_0 = \lim_{n\to\infty} T^{n+1} x_0 = T(\lim_{n\to\infty} T^n x_0) = Tz.$$

Así, z es un punto fijo de T. Por último, supongamos que existe $x \in (M, d)$ otro punto fijo de T. Entonces $d(x, z) = d(Tx, Tz) \leq kd(x, z)$, por lo cual $d(x, z) = 0$ y en consecuencia $x = z$. Por lo tanto z es el único punto fijo de T. ■

Una de las aplicaciones más conocidas del teorema de contracción de Banach es la solución al problema de valor inicial de Cauchy sobre la existencia y unicidad de ecuaciones diferenciales. Dicho ejemplo refleja la importancia del teorema de contracción de Banach, siendo el problema de valor inicial de Cauchy uno de los problemas más importantes en el área de ecuaciones diferenciales.

Ejemplo 28: Sea $f(t, x)$ una función continua real-valuada definida para $t \in [0, T]$, y $x \in \mathbb{R}$. El problema de valor inicial de Cauchy consiste en encontrar una función x continuamente diferenciable en $[0, T]$ que satisfaga la siguiente ecuación diferencial

$$\begin{cases} x'(t) = f(t, x(t)), t \in [0, T]; \\ x(0) = \xi. \end{cases} \tag{2.1}$$

El teorema de existencia y unicidad nos dice que si f es de Lipschitz con respecto a x, es decir, si existe $L > 0$ tal que para todo $x, y \in \mathbb{R}$,

$$|f(t, x) - f(t, y)| \leq L|x - y|, t \in [0, T], \tag{2.2}$$

entonces la solución (2.1) existe y es única. Supongamos que $x(t)$ cumple las hipótesis del problema de Cauchy (2.1) y que f cumple (2.2). Integrando ambos lados de (2.1), y utilizando que $x(0) = \xi$ obtenemos

$$x(t) = \xi + \int_0^t f(s, x(s))ds. \tag{2.3}$$

Recíprocamente si $x(t)$ satisface (2.3), entonces es solución al problema de Cauchy (2.1). Por lo tanto, resolver el problema (2.1) es equivalente a resolver la ecuación (2.3). Consideremos el espacio $C[0,T]$ de las funciones continuas real-valuadas dotado de la norma del supremo

$$||f|| = \sup_{0\leq t\leq T} \{|f(t)|\},$$

para $f \in C[0,T]$. El espacio $C[0,T]$ es un espacio de Banach, y por lo tanto es un espacio métrico completo. Definamos la función Fx como

$$(Fx)(t) = \xi + \textstyle\int_0^t f(s, x(s))ds.$$

Es claro que $Fx \in C[0,T]$, por lo cual $F : C[0,T] \to C[0,T]$. Notemos que $x(t)$ es solución de (2.3) y en consecuencia de (2.1) si y sólo si $x(t)$ es un punto fijo F. Notemos que para todo $x, y \in C[0,T]$,

$$\begin{aligned}
|(Fx)(t) - (Fy)(t)| &= \left| \int_0^t f(s, x(s))ds - \int_0^t f(s, y(s))ds \right| \\
&\leq \int_0^t |f(s, x(s)) - f(s, y(s))|ds \\
&\leq \int_0^t L|x(s) - y(s)|ds \\
&\leq Lt||x - y||.
\end{aligned}$$

Se sigue que

$$||Fx - Fy|| \leq LT||x - y||,$$

es decir, $k(F) \leq LT$. Mediante un procedimiento análogo al anterior se tiene que, dada $t \in [0,T]$,

$$\begin{aligned}
|(F^2x)(t) - (F^2y)(t)| &= \left| \int_0^t f(s, (Fx)(s))ds - \int_0^t f(s, (Fy)(s))ds \right| \\
&\leq \int_0^t |f(s, (Fx)(s)) - f(s, (Fy)(s))|ds \\
&\leq \int_0^t L|(Fx)(s) - (Fy)(s)|ds \\
&\leq L \int_0^t Ls||x - y||ds \\
&\leq \frac{(Lt)^2}{2}||x - y||.
\end{aligned}$$

Por lo tanto,

$$||F^2x - F^2y|| \leq \tfrac{(LT)^2}{2}||x-y||,$$

de donde $k(F^2) \leq \frac{(LT)^2}{2}$. Procediendo de manera análoga obtenemos que para toda $n \in \mathbb{N}$, $k(F^n) \leq \frac{(LT)^n}{n!}$. Como $\lim\limits_{n\to\infty} \frac{(LT)^n}{n!} = 0$, existe $n_0 \in \mathbb{N}$ tal que $\frac{(LT)^{n_0}}{n_0!} < 1$, es decir, F^{n_0} es una contracción. Por lo tanto, por el teorema de punto fijo de Banach, F^{n_0} tiene un único punto fijo, digamos x_0. Así, como $x_0 = F^{n_0}(x_0)$, aplicando F a la igualdad anterior tenemos que

$$F(x_0) = F(F^{n_0}(x_0)) = F^{n_0}(F(x_0)),$$

de donde $F(x_0)$ también es punto fijo de F^{n_0}. Por lo tanto, como F^{n_0} es una contracción, $F(x_0) = x_0$, de donde x_0 es un punto fijo de F. Además, si y_0 es un punto fijo de F, también lo es de F^{n_0}, y como probamos que F^{n_0} tiene un único punto fijo al cual llamamos x_0, entonces $x_0 = y_0$. Por lo tanto, F tiene un único punto fijo, es decir, (2.1) tiene una solución única. ■

Una vez que hemos probado la existencia de un punto fijo para contracciones definidas en espacios métricos completos, es natural preguntarnos qué es lo que ocurre con funciones de constante de Lipschitz $k = 1$. Para enfrentar este problema, en lugar de considerar espacios métricos, consideraremos espacios de Banach, esto es porque los espacios de Banach tienen estructura que los espacios métricos en general no tienen, pero aún así, los espacios de Banach siguen siendo suficientemente generales como para que los resultados obtenidos sean de interés.

La siguiente definición resulta de considerar funciones con constante de Lipschitz igual a 1 cuando la métrica en cuestion es inducida por la norma de un espacio de Banach.

Definición 29 Sean X un espacio de Banach, $K \subset X$ y $T : K \to K$. Diremos que T es una función no expansiva si para todo $x, y \in K$,

$$||Tx - Ty|| \leq ||x - y||.$$

Sean X un espacio de Banach, $K \subset X$ no vacío y $T : K \to$ una función no expansiva. Si K es convexo y compacto, veremos más adelante en el corolario 35 o en el teorema de Schauder (ver capítulo 4) que T tiene un punto fijo. Por otra parte, si T es una contracción y K es cerrado, por el teorema de contracción de Banach sabemos que T tiene un único punto fijo. Los ejemplos siguientes muestran que si K no es convexo o no es cerrado o no es acotado, entonces T puede no tener puntos fijos.

Ejemplo 30: Sea $f : (0,1] \to (0,1]$ definida como $f(x) = x/2$. Notemos que la función f es no expansiva, pero de tener un punto fijo, éste satisfaría la ecuación $x = x/2$, cuya única solución es $x = 0$, que no se encuentra en el dominio de f. Por lo tanto f no tiene puntos fijos.

Ejemplo 31: Consideremos la función $f : S^1 \to S^1$ dada por $f(x) = -x$. Claramente f es no expansiva, pero el único punto que satisface la ecuación $x = -x$ es el 0, que no se encuentra en S^1. Por lo tanto f no tiene puntos fijos.

Ejemplo 32: Sea $f : \mathbb{R} \to \mathbb{R}$ dada por $f(x) = x + 1$. Notemos que f es no expansiva, pero que no existe solución para la ecuación $x = x + 1$. Por lo tanto f no tiene puntos fijos.

Los ejemplos anteriores señalan que si el dominio de una función no expansiva f es no convexo, no cerrado o no acotado, no podemos asegurar la existencia de un punto fijo para f. Por lo anterior, restringiremos nuestra búsqueda de puntos fijos a funciones no expansivas definidas en subconjuntos no vacíos, convexos, cerrados y acotados de un espacio de Banach. En consecuencia damos la siguiente definición.

Definición 33 Sean X un espacio de Banach y $K \subset X$ no vacío, convexo, cerrado y acotado.

(a) Diremos que K tiene la propiedad del punto fijo (PPF) si toda función no expansiva de K en K tiene al menos un punto fijo.

(b) Diremos que X tiene la propiedad del punto fijo (PPF) si cada $K \subset X$ no vacío, cerrado, convexo y acotado tiene la PPF.

El siguiente lema nos habla de la existencia de puntos "casi fijos" para funciones no expansivas.

Lema 34 Sean X un espacio de Banach, $K \subset X$ no vacío, convexo, cerrado y acotado y $T : K \to K$ no expansivo. Entonces

$$\inf\{||x - Tx|| : x \in K\} = 0.$$

Demostración Sean $z \in K$ y $\varepsilon \in (0, 1)$. Definimos el mapeo $T_\varepsilon : K \to K$ por $T_\varepsilon(x) = \varepsilon z + (1 - \varepsilon)Tx$. Notemos que el mapeo anterior está bien definido por la convexidad de K. Ahora, dados $x, y \in K$ se tiene que

$$\begin{aligned} ||T_\varepsilon(x) - T_\varepsilon(y)|| &= ||\varepsilon z + (1 - \varepsilon)Tx - \varepsilon z - (1 - \varepsilon)Ty|| \\ &= (1 - \varepsilon)||Tx - Ty|| \\ &\leq (1 - \varepsilon)||x - y||. \end{aligned}$$

Así, T_ε es una contracción y existe x_ε punto fijo de T_ε. Entonces

$$\begin{aligned} ||x_\varepsilon - Tx_\varepsilon|| &= ||T_\varepsilon x_\varepsilon - Tx_\varepsilon|| \\ &= ||\varepsilon z + (1 - \varepsilon)Tx_\varepsilon - Tx_\varepsilon|| \\ &= ||\varepsilon z - \varepsilon Tx_\varepsilon|| \\ &= \varepsilon||z - Tx_\varepsilon|| \\ &\leq \varepsilon \; diamK. \end{aligned}$$

Como K es acotado su diámetro es finito, y puesto que $\varepsilon \in (0, 1)$ fue arbitrario obtenemos que $\inf\{||x - Tx|| : x \in K\} = 0$. ■

A continuación veremos algunas consecuencias del lema anterior.

Corolario 35 Sean X un espacio de Banach y $K \subset X$ no vacío y convexo. Si K es compacto entonces K tiene la PPF.

Demostración Sea K compacto. Por el lema anterior $\inf\{||x - Tx|| : x \in K\} = 0$. Como K es compacto este ínfimo se alcanza, digamos, en $y \in K$. Así, $||y - Ty|| = 0$, es decir, y es un punto fijo de T, por lo que K tiene la PPF. ■

Corolario 36 Sea X un espacio de Banach. Si X es de dimensión finita entonces X tiene la PPF.

Demostración Sea $K \subset X$ no vacío, convexo, cerrado y acotado. Como X es de dimensión finita, K es compacto, por lo que el corolario anterior nos asegura que K tiene la PPF. Por lo tanto, X tiene la PPF. ■

En el siguiente ejemplo veremos que si K es no vacío, convexo y cerrado, más no compacto, puede que no tenga la PPF.

Ejemplo 37: Consideremos el espacio de las sucesiones reales convergentes a 0, $(c_0, ||\cdot||_\infty)$, y definamos $T : B_{c_0} \to B_{c_0}$ como

$$T(x_1, x_2, ...) = (1, x_1, x_2, ...).$$

Notemos que B_{c_0} es no vacío, convexo, cerrado y acotado. Además, T es no expansiva puesto que si $x = (x_1, x_2, ...), y = (y_1, y_2, ...) \in B_{c_0}$, entonces

$$\begin{aligned}
||Tx - Ty||_\infty &= ||(1, x_1, x_2, ...) - (1, y_1, y_2, ...)||_\infty \\
&= ||(0, x_1 - y_1, x_2 - y_2, ...)||_\infty \\
&= \sup_{i \in \mathbb{N}}\{|x_i - y_i|\} \\
&= ||(x_1, x_2, ...) - (y_1, y_2, ...)||_\infty \\
&= ||x - y||_\infty.
\end{aligned}$$

Por otro lado, $Tx = x$ implica que $1 = x_1 = x_2 = ...$, pero $(1, 1, ...) \notin B_{c_0}$. Por lo tanto T no tiene puntos fijos y B_{c_0} no tiene la PPF.

Capítulo 3

El teorema de Kirk

En el capítulo anterior vimos que en un espacio de Banach X, los conjuntos no vacíos, convexos y compactos tienen la PPF. En general la compacidad en un espacio de Banach es una condición demasiado fuerte, por lo cual, surgió la necesidad de buscar condiciones más débiles que impliquen la PPF. Por ejemplo, apareció la pregunta siguiente: ¿los conjuntos no vacíos, convexos y ω-compactos tienen la PPF? Durante mucho tiempo no se tuvo una respuesta a esta pregunta hasta que en 1981 D. E. Alspach la respondió negativamente, mostrando un subconjunto no vacío convexo y ω-compacto de $L^1[0,1]$ sin la PPF.

El problema de determinar condiciones para un conjunto K (o para el espacio X que contiene a K) que siempre aseguren que K tiene la PPF para funciones no expansivas tiene su origen en 4 artículos, los cuales aparecieron en 1965. En el primero de ellos (Browder, 1965), F. Browder probó que un conjunto no vacío, convexo, cerrado y acotado $K \subset X$ tiene la PPF si X es un espacio de Hilbert. Casi inmediatamente, ambos Browder (1965) y Göhde (1965) probaron que lo mismo es cierto si X pertenece a la mucho más amplia clase de espacios "uniformemente convexos", de los cuales no hablaremos en este trabajo. Al mismo tiempo, Kirk (1965) observó que la presencia de una propiedad geométrica llamada "estructura normal" garantiza que $K \subset X$ no vacío, convexo, cerrado y acotado tiene la PPF si X es reflexivo. El concepto de estructura normal fue introducido en 1948 por Brodskii y Milman para estudiar puntos fijos de isometrías, y es una propiedad presente en todos los espacios uniformemente convexos. Los resultados de este capítulo pueden encontrarse en [3].

En esta sección, como su nombre lo indica, estudiaremos el teorema de Kirk, para lo cual comenzaremos con algunas definiciones.

Definición 38 Sean X un espacio normado, $D, H \subset X$ y $u \in X$. Definimos el radio de D respecto a u, el radio y el centro de Chebyshev de D respecto a H denotados por $r_u(D)$, $r_H(D)$ y $C_H(D)$ respectivamente, como

$$\begin{aligned} r_u(D) =& \sup\{||u-x|| : x \in D\}, \\ r_H(D) =& \inf\{r_u(D) : u \in H\}, \\ C_H(D) =& \{u \in H : r_u(D) = r_H(D)\}, \end{aligned}$$

y recordemos que el diámetro de D lo denotamos por $diam(D)$ y se define como

$$\begin{aligned} diam(D) &= \sup\{||u-v|| : u, v \in D\} \\ &= \sup\{r_u(D) : u \in D\}. \end{aligned}$$

Para entender mejor las definiciones anteriores notemos el significado geométrico que tienen. El número $r_u(D)$ representa el radio mínimo que puede tener una bola con centro en u y que contenga a D. El número $r_H(D)$ representa el radio mínimo que puede tener una bola que contiene a D con centro en un punto arbitrario de H y $C_H(D)$ es el conjunto de puntos en H tales que el radio de D respecto a ellos es menor o igual que el radio de D respecto a cualquier otro punto de H. Por último, como ya sabemos, $diam(D)$ es la distancia máxima a la que pueden estar dos elementos en D.

Por simplicidad se denotará por $r(D)$ a $r_D(D)$ y por $C(D)$ a $C_D(D)$. Notemos que

$$r(D) = \inf\{r_u(D) : u \in D\} \leq r_{u_0}(D) \leq \sup\{r_u(D) : u \in D\} = diam(D).$$

Definición 39 Sean X un espacio normado, $D \subset X$ y $u \in D$. Diremos que u es un punto diametral de D si $r_u(D) = diam(D)$. Si u es diametral para todo u en D (i.e. $r(D) = diam(D)$) diremos que D es un conjunto diametral.

Para comprender mejor el concepto de conjunto diametral veamos un par de ejemplos.

Ejemplo 40: La bola unitaria B_X de un espacio de Banach $X \neq \{0\}$ es un conjunto no diametral, ya que $r(B_X) = 1 < 2 = diam(B_X)$.

Ejemplo 41: La esfera unitaria S_X de un espacio de Banach X es claramente un conjunto diametral.

El ejemplo 40 induce a pensar que todo subconjunto no vacío, convexo, cerrado y acotado de un espacio de Banach es no diametral. Como muestra el siguiente ejemplo, esto no es cierto en general.

Ejemplo 42: En $C[0,1]$, el espacio de las funciones continuas de $[0,1]$ a $\mathbb{R}$, definimos

$$M = \{f \in C[0,1] : 0 = f(0) \leq f(t) \leq f(1) = 1, t \in [0,1]\}.$$

Notemos que todo elemento de M tiene norma 1, y por lo tanto M es acotado. Para ver que M es cerrado, sea $f \in C[0,1]$ un punto de acumulación de M. Así, para todo $\varepsilon > 0$ existe $g \in M$ tal que $g \neq f$ y

$$||f-g||_\infty = \sup\{|f(x) - g(x)| : x \in [0,1]\} < \varepsilon.$$

Ahora sea $\delta > 0$. Notemos que si $f(0) = \delta$, se tiene que para todo $h \in M$,

$$||f-h||_\infty \geq |f(0) - h(0)| = \delta.$$

Por lo tanto f no es punto de acumulación de M. En consecuencia $f(0) = 0$. Análogamente concluimos que $f(1) = 1$. Por otro lado, si $f(t_0) = 1 + \delta$ para algún $t_0 \in (0,1)$, entonces para toda $h \in M$ se tendría que

$$||f-h||_\infty \geq |f(t_0)-h(t_0)| \geq \delta.$$

Así, nuevamente f no sería punto de acumulación de M. Análogamente f no es punto de acumulación de M si $f(t_0) = -\delta$ para algún $t_0 \in (0,1)$. Por lo tanto, para todo $t \in (0,1)$, $0 \leq f(t) \leq 1$ y así $f \in M$, de donde M es cerrado. Veamos ahora que M es convexo. Sean $f, g \in M$ y $\lambda \in [0,1]$. Notemos que si $t \in [0,1]$,

$$\begin{aligned}(\lambda f + (1-\lambda)g)(t) &= \lambda f(t) + (1-\lambda)g(t)\\ &\leq \lambda f(1) + (1-\lambda)g(1)\\ &= 1\\ &= (\lambda f + (1-\lambda)g)(1)\end{aligned}$$

y además

$$\begin{aligned}(\lambda f + (1-\lambda)g)(t) &= \lambda f(t) + (1-\lambda)g(t)\\ &\geq \lambda f(0) + (1-\lambda)g(0)\\ &= 0\\ &= (\lambda f + (1-\lambda)g)(0).\end{aligned}$$

Así $\lambda f + (1-\lambda)g \in M$ y por lo tanto M es convexo. Finalmente veremos que M es diametral, es decir $r(M) = diam(M)$. Notemos que $diam(M) = 1$, ya que si definimos f_1, f_2 como

$$f_1(t) = \begin{cases} 3t, \text{ si } 0 \leq t \leq 1/3\\ 1, \text{ si } 1/3 \leq t \leq 1\end{cases}$$

y

$$f_2(t) = \begin{cases} 0, \text{ si } 0 \leq t \leq 2/3\\ 3(t-2/3), \text{ si } 2/3 \leq t \leq 1,\end{cases}$$

tenemos que

$$1 = |f_1(1/2) - f_2(1/2)| \leq ||f_1 - f_2||_\infty \leq |1-0| = 1,$$

de donde como $f_1, f_2 \in M$, $diam(M) \geq 1$. Por otro lado, dadas $f_1, f_2 \in M$,

$$||f_1 - f_2||_\infty = \sup\{|f_1(x) - f_2(x)| : x \in [0,1]\} \leq |1-0| = 1.$$

Así, $diam(M) = 1$. Resta probar que $r(M) = 1$. Procediendo por contradicción supongamos que $r(M) < 1$. Entonces existen $f \in M$ y $\varepsilon \in (0,1)$ tales que $r_f(M) = 1-\varepsilon$. Como f es continua, por el teorema del valor intermedio existe $t_0 \in (0,1)$ tal que $f(t_0) = \frac{\varepsilon}{2}$. Sea g dada por

$$g(t) = \begin{cases} \dfrac{t}{t_0}, \text{ si } 0 \leq t \leq t_0\\ 1, \text{ si } t_0 \leq t \leq 1.\end{cases}$$

Notemos que $g \in M$ y además

$$\begin{aligned}||f-g||_\infty &\geq |f(t_0) - g(t_0)|\\ &= 1 - \frac{\varepsilon}{2}.\end{aligned}$$

Así,

$$r_f(M) \geq 1 - \tfrac{\varepsilon}{2},$$

lo cual es una contradicción a la hipótesis de que $r_f(M) = 1 - \varepsilon$. Por lo tanto, el conjunto M es no vacío, convexo, cerrado, acotado y diametral, que es lo que buscábamos.

Para enunciar y probar el teorema de Kirk, introducimos a continuación el concepto de estructura normal.

Definición 43 (Estructura Normal) Sean X un espacio de Banach y $D \subset X$ convexo. Diremos que D tiene estructura normal (NS por sus siglas en inglés) si cada subconjunto acotado y convexo $S \subset D$ con $diamS > 0$ contiene a un punto no diametral, es decir $r(S) < diamS$.

Dicho de otra manera, si X es un espacio de Banach y $D \subset X$ tiene estructura normal, no pueden haber subconjuntos convexos, acotados y diametrales con diámetro positivo en D.

El siguiente concepto nos permite dar una caracterización de la estructura normal, la cual se debe nuevamente a Brodskii y Milman.

Definición 44 Sean X un espacio de Banach y (x_n) una sucesión acotada en X. Diremos que (x_n) es diametral si no es eventualmente constante y

$$\lim_{n\to\infty} d(x_{n+1}, conv(\{x_1, ..., x_n\})) = diam(\{x_n\}_{n=1}^{\infty}).$$

Lema 45 Sean X un espacio de Banach y (x_n) una sucesión diametral en X. Entonces toda subsucesión de (x_n) es diametral.

Demostración Sea (x_{n_k}) una subsucesión de (x_n). Puesto que (x_n) es diametral, se tiene que

$$\begin{aligned} diam\{x_n\} &= \lim_{k\to\infty} d(x_{n_{k+1}}, conv\{x_1, x_2, ...x_{n_{k+1}-1}\}) \\ &\leq \lim_{k\to\infty} d(x_{n_{k+1}}, conv\{x_{n_1}, x_{n_2}, ..., x_{n_k}\}) \\ &\leq diam\{x_{n_k}\} \\ &\leq diam\{x_n\}. \end{aligned} \tag{3.1}$$

Por (3.1) se tiene que (x_{n_k}) no es eventualmente constante y que

$$\lim_{k\to\infty} d(x_{n_{k+1}}, conv\{x_{n_1}, x_{n_2}, ..., x_{n_k}\}) = diam\{x_{n_k}\}.$$

Por lo tanto x_{n_k} es diametral. ■

Ahora veremos la caracterización antes mencionada de la estructura normal en términos de las sucesiones diametrales.

Lema 46 Sean X un espacio de Banach y $K \subset X$ convexo y acotado. Entonces K tiene estructura normal si y sólo si no contiene sucesiones diametrales.

Demostración Supongamos que K tiene estructura normal. Sea (x_n) una sucesión acotada y no eventualmente constante y definamos $S = conv(\{x_n\}_{n=1}^{\infty})$. Como K tiene estructura normal, existe $u \in S$ no diametral, es decir, $r_u(S) < diam(S)$. Como $u \in S$, existe m tal que $u \in conv(\{x_1, x_2..., x_m\})$, de donde para $n \geq m$ se tiene que

$$d(x_{n+1}, conv(\{x_1, x_2..., x_n\})) \leq d(x_{n+1}, u) \leq r_u(S).$$

Así,

$$\begin{aligned}\limsup_{n\to\infty} d(x_{n+1}, conv(\{x_1, x_2..., x_n\})) &\leq r_u(s) \\ &< diam(S).\end{aligned}$$

Por **(1)** del lema 6 tenemos que $diam(S) = diam(\{x_n\}_{n=1}^{\infty})$, por lo cual (x_n) no es una sucesión diametral. Probaremos ahora la implicación contraria por contrapositiva. Supongamos que $S \subset K$ es un conjunto diametral convexo con $d = diam(S) > 0$ y sean $\varepsilon \in (0, d)$ y $x_1 \in S$. Notemos que como S es convexo, dada (y_n) sucesión en S se tiene que para todo n, $\sum_{i=1}^{n} \frac{y_i}{n} \in S$. Dado que S es diametral, existe $x_2 \in S$ tal que $||x_1 - x_2|| > d - \varepsilon$, y nuevamente por convexidad y por ser S diametral, existe $x_3 \in S$ tal que $||\frac{x_1+x_2}{2} - x_3|| > d - \frac{\varepsilon}{2^2}$. Así, inductivamente podemos definir una sucesión (x_n) tal que

$$||y_n - x_{n+1}|| > d - \tfrac{\varepsilon}{n^2}, \text{ donde } y_n = \sum_{i=1}^{n} \frac{x_i}{n}.$$

Sea ahora $x \in conv(x_1, ..., x_n)$ con $x = \sum_{i=1}^{n} \alpha_i x_i$ con $\alpha_i \geq 0$ y $\sum_{i=1}^{n} \alpha_i = 1$. Definamos $\alpha = \max\{\alpha_1, ..., \alpha_n\}$ y notemos que $y_n = \frac{x}{n\alpha} + \sum_{i=1}^{n}(\frac{1}{n} - \frac{\alpha_i}{n\alpha})x_i$, $\frac{1}{n\alpha} + \sum_{i=1}^{n}(\frac{1}{n} - \frac{\alpha_i}{n\alpha}) = 1$ y $\frac{1}{n} - \frac{\alpha_i}{n\alpha} \geq 0$. Así, por construcción tenemos que

$$\begin{aligned}d - \frac{\varepsilon}{n^2} &< ||y_n - x_{n+1}|| \\ &= \Big|\Big|\frac{x}{n\alpha} + \sum_{i=1}^{n}\Big(\frac{1}{n} - \frac{\alpha_i}{n\alpha}\Big)x_i - \Big(\frac{1}{n\alpha} + \sum_{i=1}^{n}\Big(\frac{1}{n} - \frac{\alpha_i}{n\alpha}\Big)\Big)x_{n+1}\Big|\Big| \\ &\leq \frac{1}{n\alpha}||x - x_{n+1}|| + \sum_{i=1}^{n}\Big(\frac{1}{n} - \frac{\alpha_i}{n\alpha}\Big)||x_i - x_{n+1}|| \\ &\leq \frac{1}{n\alpha}||x - x_{n+1}|| + \Big(1 - \frac{1}{n\alpha}\Big)d.\end{aligned}$$

Así, despejando $||x - x_{n+1}||$ de la expresión anterior tenemos que

$$\begin{aligned}d &\geq ||x - x_{n+1}|| \\ &> n\alpha\Big(d - \frac{\varepsilon}{n^2} - \Big(1 - \frac{1}{n\alpha}\Big)d\Big) \\ &= d - \frac{\varepsilon\alpha}{n} \\ &\geq d - \frac{\varepsilon}{n}.\end{aligned}$$

Como $x \in conv(\{x_1, ..., x_n\})$ y $\varepsilon \in (0, d)$ fueron arbitrarios, entonces

$$\lim_{n\to\infty} d(x_{n+1}, conv(\{x_1, ..., x_n\})) = d,$$

y **(1)** del lema 6 se tiene que (x_n) es diametral. Por lo tanto, si K no posee sucesiones diametrales, entonces tiene estructura normal. ■

Corolario 47 Sean X un espacio de Banach y $K \subset X$ convexo y compacto. Entonces K tiene estructura normal.

Demostración Supongamos que K no tiene estructura normal. Entonces existe una sucesión diametral (x_n) en K. Como K es compacto existe una subsucesión convergente (x_{n_k}). Notemos que $d := diam(\{x_n\}_{n=1}^{\infty}) > 0$, ya que no es una sucesión constante. Además, como (x_{n_k}) es convergente también es de Cauchy, por lo que dado $\varepsilon > 0$ existe k_0 tal que si $k \geq k_0$ entonces

$$||x_{n_{k+1}} - x_{n_k}|| < \varepsilon.$$

Así, dado $\varepsilon > 0$, para k suficientemente grande se tiene que

$$\begin{aligned} d(x_{n_{k+1}}, conv(\{x_{n_1}, ..., x_{n_k}\})) &\leq d(x_{n_{k+1}}, \{x_{n_1}, ..., x_{n_k}\}) \\ &< \varepsilon, \end{aligned}$$

de donde

$$\lim_{k \to \infty} d(x_{n_{k+1}}, conv(\{x_{n_1}, ..., x_{n_k}\})) = 0 < d.$$

Por lo tanto, (x_{n_k}) no es diametral, y por el lema 45, (x_n) tampoco es diametral, lo cual es una contradicción. ■

Corolario 48 Sea X un espacio de Banach. Si X es de dimensión finita, entonces X tiene estructura normal.

Demostración Sea $K \subset X$ convexo, acotado y con $diam(K) > 0$. Como X es de dimensión finita, $\overline{K}$ es compacto, y así, por el corolario anterior, $\overline{K}$ tiene estructura normal. Luego, como $K \subset \overline{K}$, todo subconjunto acotado y convexo de diámetro positivo de K es también subconjunto de $\overline{K}$, y por la estructura normal de $\overline{K}$, dicho conjunto tiene un punto no diametral. Por lo tanto, K también tiene estructura normal. Así, X tiene estructura normal. ■

Ejemplo 49: Como vimos en el ejemplo 42, el espacio $C[0,1]$ tiene un subconjunto M, el cual es no vacío, acotado, convexo, cerrado y diametral, con $diam(M) > 0$. Por lo tanto, $C[0,1]$ no tiene estructura normal.

Las definiciones siguientes nos permitirán enunciar algunos lemas que nos servirán para probar el teorema de Kirk.

Definición 50 Sean K un conjunto, $T : K \to K$ una función y $D \subset K$. Diremos que D es $T-$invariante si $T(D) \subset D$.

Definición 51 Sean K un subconjunto de un espacio de Banach, $T : K \to K$ una función y $D \subset K$ no vacío, convexo, cerrado y $T-$invariante. Diremos que D es $T-$invariante minimal si no contiene subconjuntos propios no vacíos, convexos, cerrados y $T-$invariantes.

Como veremos en el siguiente lema, si K es no vacío, convexo y ω-compacto, entonces podemos garantizar la existencia de un conjunto $T-$invariante minimal.

Lema 52 Sean X un espacio de Banach y $K \subset X$ no vacío, ω-compacto y convexo. Entonces para todo mapeo $T : K \to K$ existe $D \subset K$ tal que D es $T-$invariante minimal.

Demostración Consideremos la familia Ω de todos los subconjuntos de K que son no vacíos, convexos, cerrados y $T-$invariantes. Sea $\leq$ el orden en Ω dado por el inverso de la inclusión, es decir, dados $K_1, K_2 \in \Omega$, $K_1 \leq K_2$ si $K_1 \supset K_2$. Como K es $\omega-$compacto por el teorema de Mazur, se tiene que todo elemento de Ω es ω-compacto, luego por el teorema 10, todo subconjunto totalmente ordenado $\{E_i\}_{i\in I}$ en Ω tiene intersección no vacía, la cual además es $T-$invariante. Así, al ser $E = \cap_{i\in I} E_i$ no vacío, convexo y cerrado, se sigue que $E \in \Omega$. Luego, por la definición de $(\Omega, \leq)$, lo anterior es equivalente a que cada subconjunto totalmente ordenado en Ω tenga una cota superior. Así, por el lema de Zorn, existe $D \in \Omega$ maximal. Dicho $D \in \Omega$ es $T-$invariante y al ser maximal en $(\Omega, \leq)$, no tiene subconjuntos propios $T-$invariantes. Por lo tanto, D es $T-$invariante minimal. ■

Antes de probar el teorema de Kirk es necesario demostrar los siguientes dos lemas.

Lema 53 Sean X un espacio de Banach y $D \subset X$ un subconjunto $T-$invariante minimal. Entonces $D = \overline{conv}T(D)$.

Demostración Primero notemos que $\overline{conv}T(D)$ es no vacío, convexo y cerrado. Ahora, como $T(D) \subset D$ y D es convexo y cerrado se sigue que

$$\overline{conv}T(D) \subset \overline{conv}(D) = D.$$

Entonces, $\overline{conv}T(D)$ es $T-$invariante ya que

$$T(\overline{conv}T(D)) \subset T(D) \subset \overline{conv}T(D).$$

Así, de la minimalidad de D se tiene que $D = \overline{conv}T(D)$. ■

Lema 54 Sean X un espacio de Banach y $K \subset X$ no vacío, convexo y cerrado. Entonces el centro de K, $C(K)$, es cerrado y convexo. Si además K es $\omega-$compacto, entonces $C(K) \neq \emptyset$.

Demostración Definamos $C_n(K)$ para $n \in \mathbb{N}$ como

$$C_n(K) = \{u \in K : r_u(K) \leq r(K) + \tfrac{1}{n}\},$$

y notemos que

$$C_n(K) = \{u \in K : \sup_{x\in K}\{||u - x||\} \leq r(K) + \frac{1}{n}\},$$

de donde los subconjuntos $C_n(K)$ son convexos y cerrados. Por lo tanto

$$C(K) = \{u \in K : \sup_{x \in K}\{||u - x||\} \leq r(K)\} = \bigcap_{n \in \mathbb{N}} C_n(K)$$

también es convexo y cerrado. Por último, si K es $\omega-$compacto, toda familia de subconjuntos $\omega-$cerrados en K con la propiedad de la intersección finita tiene intersección no vacía, en particular $\{C_n(K)\}_{n\in\mathbb{N}}$, de donde $C(K) = \cap_{n\in\mathbb{N}} C_n(K) \neq \emptyset$. ■

Finalmente tenemos lo necesario para demostrar el teorema de Kirk, el cual es uno de los teoremas más importantes en la teoría del punto fijo.

Teorema 55 (Kirk) Sean X un espacio de Banach y $K \subset X$ no vacío, convexo, $\omega-$compacto y con estructura normal. Entonces K tiene la PPF.

Demostración Sea $T : K \to K$ un mapeo no expansivo. Por el lema 52 existe $D \subset K$ invariante minimal y por el lema 53 $\overline{conv}(T(D)) = D$. Como $D \subset K$ es cerrado y convexo y K es $\omega-$compacto, D también es $\omega-$compacto. Así, por el lema 54, $C(D) \neq \emptyset$. Sea $u \in C(D)$, es decir, $r_u(D) = r(D)$. Como T es no expansivo

$$||Tu - Tv|| \leq ||u - v|| \leq r(D)$$

para todo $v \in D$, de donde

$$T(D) \subset \overline{B}(Tu, r(D)).$$

Así,

$$D = \overline{conv}(T(D)) \subset \overline{B}(Tu, r(D)),$$

por lo que $r_{Tu}(D) \leq r(D)$ y por lo tanto $r_{Tu}(D) = r(D)$. Como $u \in D$ fue arbitrario, esto prueba que $C(D)$ es $T-$invariante, y por la minimalidad de D se tiene que $C(D) = D$, por lo que D es diametral, lo cual sólo es posible si $D = \{u\}$, ya que K tiene estructura normal. Por lo tanto $Tu = u$, de donde u es un punto fijo de T. ■

Como un corolario del teorema de Kirk podemos ver una demostración alternativa del corolario 35.

Corolario 56 (Prueba alternativa de corolario 35) Sean X un espacio de Banach y $K \subset X$ no vacío y convexo. Si K es compacto entonces K tiene la PPF.

Demostración Por el corolario 47, si K es compacto entonces tiene estructura normal y además, también es ω-compacto, entonces, por el teorema de Kirk, K tiene la PPF. ■

Cabe señalar en este momento que para poder utilizar el teorema de Kirk es necesario conocer espacios o conjuntos con estructura normal. Sabemos hasta ahora que los subconjuntos convexos y compactos de un espacio de Banach tienen estructura normal, y en consecuencia que los espacios de Banach de dimensión finita también la tienen. Existen condiciones o criterios que garantizan la existencia de estructura normal en subconjuntos convexos de espacios de Banach. Para un estudio detallado al respecto (véase [5]). Sólo por mostrar un ejemplo, veremos que los espacios de Banach uniformemente convexos, entre ellos los espacios de Hilbert y los espacios L^p $(1 < p < \infty)$, tienen estructura normal.

Definición 57 Sea X un espacio de Banach.

(a) Diremos que X es estrictamente convexo si $||x||, ||y|| \leq 1$ y $||x - y|| > 0$ implican que $||\frac{x+y}{2}|| < 1$.

(b) Diremos que X es uniformemente convexo si para toda $\varepsilon > 0$, existe $\delta > 0$ tal que si $||x||, ||y|| \leq 1$ y $||x - y|| > \varepsilon$, entonces $||\frac{x+y}{2}|| < 1 - \delta$.

Proposición 58 Sea X un espacio de Banach. Si X es uniformemente convexo, entonces X tiene estructura normal.

Demostración Supongamos que X no tiene estructura normal. Entonces existe $S \subset X$ acotado, convexo, diametral y con diámetro positivo. Sean $x, y \in S$ con $x \neq y$ y definamos $u = \frac{x+y}{2}$. Tomemos $\varepsilon = \frac{||x-y||}{2diam(S)} > 0$. Como u es diametral, existe una sucesión (u_n) en S tal que

$$\lim_{n\to\infty} ||u_n - u|| = diam(S).$$

Para cada $n \in \mathbb{N}$ definamos $v_n = \frac{x-u_n}{diam(S)}$ y $w_n = \frac{y-u_n}{diam(S)}$. Así, $||v_n||, ||w_n|| \leq 1$ y $||v_n - w_n|| = ||\frac{x-y}{diam(S)}|| = 2\varepsilon > \varepsilon$. Por otra parte,

$$\lim_{n\to\infty} ||\frac{v_n+w_n}{2}|| = \lim_{n\to\infty} ||\frac{u-u_n}{diam(S)}|| = 1.$$

Por lo tanto X no es uniformemente convexo. $\blacksquare$.

Capítulo 4

Los Teoremas de Brouwer y de Schauder

Como ya mencionamos en el prólogo, en 1912 Brouwer probó su teorema de punto fijo para el cuadrado, para la bola, y posteriormente para sus contrapartes de n-dimensiones. Más tarde, en 1930 Schauder extendió dicho teorema a subconjuntos convexos y compactos de espacios de Banach.

En este capítulo enunciaremos y probaremos los teoremas de Brouwer y Schauder. Para ello nos introduciremos brevemente a la teoría del grado, la cual resultará de gran utilidad.

4.1 Teoría del grado en $\mathbb{R}^n$

En esta sección nos interesa estudiar ecuaciones de la forma $f(x) = y$, donde $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ es una función continua, $\Omega \subset \mathbb{R}^n$ es abierto y acotado y y es un punto en $\mathbb{R}^n$. Primero que nada queremos saber cuándo una tal ecuación tiene una solución, y si la tiene, queremos saber si es única o no. También nos interesa saber de qué manera están distribuidas las soluciones en Ω, si las hay. Una vez que tenemos información sobre una ecuación en particular, queremos también saber si las soluciones se mantienen o cambian drásticamente si cambiamos f y y de "cierta" forma.

Estos problemas, debido a su generalidad, son bastante complicados. Para tener más herramientas a la mano, en particular la compacidad en $\mathbb{R}^n$, supondremos también que f está definida en $\overline{\Omega}$. Por otra parte, esta nueva suposición abre la posibilidad de tener soluciones de la ecuación $f(x) = y$ en $\overline{\Omega} \backslash \Omega = \partial\Omega$ (la última igualdad de conjuntos se cumple pues Ω es abierto). Para evitar esto último y asegurar que, de existir, las soluciones de la ecuación $f(x) = y$ se encuentran en Ω, supondremos además que $y \notin f(\partial\Omega)$.

Para resolver los problemas anteriores nos gustaría contar con una función d que asigne a la terna (f, Ω, y) un entero que represente el número de soluciones en Ω de la ecuación $f(x) = y$. Supongamos por un momento que tal función d existe. Si $I : \overline{\Omega} \to \overline{\Omega}$ es la función identidad y $y \in \Omega$, entonces la función d debe de cumplir que

(d1) $d(I, \Omega, y) = 1$.

Respecto a la distribución de las soluciones, si Ω_1 y Ω_2 son subconjuntos abiertos y ajenos de Ω tales que $y \notin f(\overline{\Omega} \backslash (\Omega_1 \cup \Omega_2))$, entonces $\partial\Omega_j \subset \overline{\Omega} \backslash (\Omega_1 \cup \Omega_2)$, de donde $y \notin \mathbb{R}^n \backslash f(\partial\Omega_j)$, $j = 1, 2$. Así, la función d debe de cumplir que

(d2) $d(f, \Omega, y) = d(f, \Omega_1, y) + d(f, \Omega_2, y)$.

Finalmente, si la ecuación $f(x) = y$ tiene una solución y podemos deformar "continuamente" la función f a una función g de manera que en ningún paso de la deformación se obtengan soluciones en la frontera de Ω, nos gustaría que la función d cumpla la siguiente condición.

(d3) $d(f, \Omega, y) = d(g, \Omega, y)$.

Como veremos en las próximas secciones, existe una única función

$$d : \{(f, \Omega, y) : \Omega \subset \mathbb{R}^n \text{ abierto y acotado, } f : \overline{\Omega} \to \mathbb{R}^n \text{ continua,} \\ y \in \mathbb{R}^n \backslash f(\partial\Omega)\} \to \mathbb{Z},$$

llamada el grado topológico o el grado de Brouwer o simplemente el grado (en $\mathbb{R}^n$) que cumple las condiciones **(d1)**, **(d2)** y **(d3)**. Además, bajo ciertas condiciones, $d(f, \Omega, y)$ representa el número de soluciones en Ω de la ecuación $f(x) = y$.

Como comentamos anteriormente, el grado topológico nos será de bastante utilidad para demostrar el teorema de Brouwer, que es uno de los teoremas más importantes que veremos en este capítulo, y como ya se mencionó, es uno de los pilares de la teoría del punto fijo. La mayoría de los resultados de esta sección pueden ser encontrados en [1].

Una de las primeras personas en estudiar la teoría del grado en $\mathbb{R}^n$ de manera puramente analítica fue Mitio Nagumo en 1950 (véase [8]). En esta tesis estudiaremos la teoría del grado en $\mathbb{R}^n$ desde este mismo enfoque.

De aquí en adelante, a menos que se indique lo contrario, Ω será un subconjunto no vacío, abierto y acotado de $\mathbb{R}^n$.

La siguiente definición la usaremos a lo largo de este capítulo.

Definición 59 Sean $f \in C^1(\Omega)$ y $x \in \Omega$. Diremos que x es un punto crítico de f si $J_f(x) = 0$. Definimos también $S_f(\Omega) = \{x \in \Omega : J_f(x) = 0\}$, y lo denotaremos por S_f cuando Ω sea claro por el contexto. Además, diremos que un punto $y \in \mathbb{R}^n$ es un valor regular de $f : \Omega \to \mathbb{R}^n$ si $f^{-1}(y) \cap S_f(\Omega) = \emptyset$ y valor singular de otra manera.

De la definición anterior notemos que y es un valor regular de f si y sólo si $y \in f(S_f(\Omega))^c$.

4.1.1 Unicidad del grado en $\mathbb{R}^n$

En esta sección y en la próxima probaremos que existe una única función

$$d : \{(f, \Omega, y) : \Omega \subset \mathbb{R}^n \text{ abierto y acotado, } f : \overline{\Omega} \to \mathbb{R}^n \text{ continua,}\\ y \in \mathbb{R}^n \backslash f(\partial\Omega)\} \to \mathbb{Z},$$

que satisface las propiedades siguientes,

(d1) $d(I, \Omega, y) = 1$, para toda $y \in \Omega$, donde I es la función identidad en $\overline{\Omega}$.

(d2) $d(f, \Omega, y) = d(f, \Omega_1, y) + d(f, \Omega_2, y)$ para $\Omega_1, \Omega_2 \subset \Omega$ abiertos y disjuntos tales que $y \notin f(\overline{\Omega} \backslash (\Omega_1 \cup \Omega_2))$.

(d3) $d(h(t, \cdot), \Omega, y(t))$ es independiente de $t \in [0, 1]$, donde $h : [0, 1] \times \overline{\Omega} \to \mathbb{R}^n$ y $y : [0, 1] \to \mathbb{R}^n$ son continuas y $y(t) \notin h(t, \partial\Omega)$, para todo $t \in [0, 1]$.

Como comentamos anteriormente, la función que satisface estas propiedades es conocida como el grado topológico o el grado de Brouwer o simplemente el grado (en $\mathbb{R}^n$). Al valor $d(f, \Omega, y)$ le llamaremos el grado de f respecto a Ω y y.

En esta sección supondremos que el grado existe y nos dedicaremos a probar su unicidad, para lo cual procederemos por etapas. Primero se mostrará que para probar la unicidad del grado para funciones en $C(\overline{\Omega})$, es suficiente la unicidad del grado para funciones en $\overline{C}^\infty(\Omega)$. Posteriormente se mostrará que nos basta la unicidad del grado para funciones en $\overline{C}^\infty(\Omega)$ y valores regulares y finalmente que sólo requerimos la unicidad del grado para funciones lineales y valores regulares. Al final de la sección se probará que, efectivamente, el grado es único. En la próxima sección se probará la existencia del grado.

Comenzaremos con el siguiente lema.

Lema 60 Dada $\alpha > 0$ definamos $\varphi_\alpha : \mathbb{R}^n \to \mathbb{R}$ de la manera siguiente. Para $\alpha = 1$ definimos

$$\varphi_1(x) = \begin{cases} c \cdot \exp(-\frac{1}{1-||x||^2}) & \text{si } ||x|| < 1, \\ 0 & \text{si } ||x|| \geq 1, \end{cases}$$

donde $c \in \mathbb{R}$ es tal que $\int_{\mathbb{R}^n} \varphi_1(x)dx = 1$, y para $\alpha > 0$ definimos

$$\varphi_\alpha(x) = \alpha^{-n}\varphi_1(x/\alpha).$$

Entonces, para todo $\alpha > 0$, $\varphi_\alpha \in C^\infty(\mathbb{R}^n)$ y tiene soporte en $\overline{B}_\alpha$. Más aún,

$$\int_{\mathbb{R}^n} \varphi_\alpha(x)dx = 1.$$

Demostración Probaremos primero que $\varphi_\alpha \in C^\infty(\mathbb{R}^n)$, para lo cual basta probar que la función φ_1 es de clase C^∞. Notemos que $\varphi_1(x) = cf(1 - ||x||^2)$, donde $f : \mathbb{R} \to \mathbb{R}$ es la función definida por $f(t) = \exp(-1/t)$, si $t > 0$ y $f(t) = 0$, si $t \leq 0$. Así, para probar que φ_1 es clase C^∞ basta mostrar que f es clase C^∞. Notemos que para $x \neq 0$ las derivadas de cualquier orden de f existen en x, ya que f es constante en $(-\infty, 0)$. Por otro lado, como $-1/t$ es de clase C^∞ en $(0, \infty)$ y la función exponencial es de clase C^∞ en $\mathbb{R}$, por la regla de la cadena se tiene que f es de clase C^∞ en $(0, \infty)$. Por lo que resta probar que existen las derivadas de cualquier orden de f en 0. Para ello, primero probaremos por inducción que si $t \in (0, \infty)$, $f^{(k)}(t)$ está dada por

$$f^k(t) = p_k(1/t)\exp(-1/t),$$

para algún polinomio p_k. Para $k = 1$ se tiene que

$$f'(t) = p_1(1/t)\exp(-1/t),$$

donde $p_1(t) = t^2$. Supongamos ahora que $f^{(k)}(t) = p_k(1/t)\exp(-1/t)$ para algún polinomio p_k. Puesto que

$$f^{(k+1)}(t) = \tfrac{p_k(1/t)-p_k'(1/t)}{t^2}\exp(-1/t),$$

obtenemos que $f^{k+1}(t) = p_{k+1}(1/t)\exp(-1/t)$, donde

$$p_{k+1}(t) = (p_k(t) - p_k'(t))t^2,$$

que es lo que buscábamos. Notemos también que, como la función exponencial crece más rápido que cualquier polinomio, $\lim_{t\to 0^+} f^{(k)}(t) = 0$. Además es claro que $\lim_{t\to 0^-} f^{(k)}(t) = 0$. Usaremos estos hechos más adelante.

Nuevamente por medio de inducción probaremos que las derivadas de f de cualquier orden en 0 existen y son todas 0. Como caso base veremos que $f'(0)$ existe y $f'(0) = 0$. Notemos que

$$\begin{aligned}\lim_{h\to 0^+}\frac{f(h)-f(0)}{h} &= \lim_{h\to 0^+}\frac{f(h)}{h}\\ &= \lim_{h\to 0^+}\frac{\exp(-1/h)}{h}\\ &= 0\\ &= \lim_{h\to 0^-}\frac{f(h)-f(0)}{h},\end{aligned}$$

ya que la función exponencial crece más rápido que cualquier polinomio. Por lo tanto, $f'(0)$ existe y $f'(0) = 0$. Supongamos ahora que $f^{(k)}(0)$ existe y $f^{(k)}(0) = 0$. Calculemos nuevamente las derivadas laterales. Anteriormente vimos que $f^{(k)}(x) = 0$ si $x \in (-\infty, 0)$ y $f^{(k)}(x) \to 0$ cuando $x \to 0^+$. Así,

$$\begin{aligned}\lim_{h\to 0^+}\frac{f^{(k)}(h)-f^{(k)}(0)}{h} &= \lim_{h\to 0^+}\frac{f^{(k)}(h)}{h}\\ &= \lim_{h\to 0^+}\frac{p_k(1/h)\exp(-1/h)}{h}\\ &= \lim_{h\to 0^+}\frac{p_k(1/h)}{h}\exp(-1/h)\\ &= 0\\ &= \lim_{h\to 0^-}\frac{f^{(k)}(h)-f^{(k)}(0)}{h},\end{aligned}$$

ya que la función exponencial crece más rápido que cualquier polinomio. Por lo tanto, $f^{(k+1)}(0)$ existe y $f^{(k+1)}(0) = 0$, que es lo que queríamos demostrar. Por lo tanto, la función f es de clase C^∞.

A continuación veremos que $\int_{\mathbb{R}^n} \varphi_\alpha(x)dx = 1$ y $\overline{B_\alpha}$ es el soporte de φ_α para toda $\alpha > 0$. Para ver que $\overline{B_\alpha}$ es el soporte de φ_α recordemos que $\varphi_\alpha(x) = \alpha^{-n}\varphi_1(x/\alpha)$ para todo $x \in \mathbb{R}^n$ y $\alpha > 0$. Así,

$$\varphi_\alpha(x) = 0$$

si y sólo si

$$\alpha^{-n}\varphi_1(x/\alpha) = 0,$$

lo cual por la definición de φ_1 ocurre si y sólo si $||x/\alpha|| \geq 1$, lo cual ocurre si y sólo si $||x|| \geq \alpha$. Por lo tanto, el soporte de φ_α es $\overline{B_\alpha}$. Por otro lado, notemos que, como $\overline{B_\alpha}$ es el soporte de φ_α,

$$\begin{aligned}\int_{\mathbb{R}^n} \varphi_\alpha(x)dx &= \int_{B_\alpha} \varphi_\alpha(x)dx \\ &= \int_{B_\alpha} \alpha^{-n}\varphi_1(x/\alpha)dx.\end{aligned}$$

Por el teorema de cambio de variable con $g : \mathbb{R}^n \to \mathbb{R}^n$ definida como $g(x) = \alpha x$ se tiene que

$$\begin{aligned}\int_{B_\alpha} \alpha^{-n}\varphi_1(x/\alpha)dx &= \int_{B_1} \alpha^{-n}\varphi_1(x).\, \alpha^n dx \\ &= \int_{B_1} \varphi_1(x)dx \\ &= 1.\end{aligned}$$

Por lo tanto $\int_{\mathbb{R}^n} \varphi_\alpha(x)dx = 1$. ■

A las funciones del lema anterior se les comoce como funciones "suavizantes".

Proposición 61 Sean $K \subset \mathbb{R}^n$ compacto, $f \in C(K)$ y $\varepsilon > 0$. Entonces existe una función $g \in C^\infty(\mathbb{R}^n)$ tal que $||f(x) - g(x)|| < \varepsilon$ en K.

Demostración Por la proposición 8 existe una extensión continua $\widetilde{f}$ a $\mathbb{R}^n$ de la función f. Definamos ahora

$$f_\alpha(x) = \int_{\mathbb{R}^n} \widetilde{f}(\xi)\varphi_\alpha(x - \xi)d\xi,$$

para $x \in \mathbb{R}^n$ y $\alpha > 0$, donde las φ_α son las funciones suavizantes del lema anterior. Notemos que como las funciones suavizantes φ_α son de clase C^∞ para todo $\alpha > 0$, entonces f_α también es de clase C^∞ para toda $\alpha > 0$. A continuación veremos que $f_\alpha \to \widetilde{f}$ uniformemente en K cuando $\alpha \to 0$. Para ello, notemos que, si $\alpha > 0$,

$$\begin{aligned}f_\alpha(x) - \widetilde{f}(x) &= \int_{\mathbb{R}^n} \widetilde{f}(\xi)\varphi_\alpha(\xi - x) - \widetilde{f}(x)\varphi_\alpha(\xi - x)d\xi \\ &= \int_{B_\alpha(x)} (\widetilde{f}(\xi) - \widetilde{f}(x))\varphi_\alpha(\xi - x)d\xi.\end{aligned}$$

Luego, puesto que $\widetilde{f}$ es uniformemente continua en K, dado $\varepsilon > 0$ existe $\alpha_0 > 0$ tal que para todo $x \in K$, $0 < \alpha \leq \alpha_0$ se tiene que $||\widetilde{f}(x) - \widetilde{f}(y)|| < \varepsilon$ si $||x - y|| < \alpha$, de donde

$$\begin{aligned} ||f_\alpha(x) - \widetilde{f}(x)|| &\leq \int_{B_\alpha(x)} ||\widetilde{f}(\xi) - \widetilde{f}(x)||\varphi_\alpha(\xi - x)d\xi \\ &= \varepsilon. \end{aligned}$$

Así, concluimos que, como $\widetilde{f}$ es una extensión de f, $f_\alpha \to f$ uniformemente en K. Por lo tanto $g = f_\alpha$ satisface lo deseado para α suficientemente pequeña. ■

Antes de continuar veamos la siguiente definición.

Definición 62 Sean X y Y espacios normados y $f : X \to Y$ continua y acotada. Definimos la norma de f, $||f||_\infty$ como

$$||f||_\infty = \sup_{x \in X} ||f(x)||.$$

Proposición 63 La unicidad del grado para funciones en $\overline{C}^\infty(\Omega)$ implica la unicidad del grado para funciones en $C(\overline{\Omega})$.

Demostración Supongamos que el grado es único para funciones en $\overline{C}^\infty$. Sea d' una función que cumple **(d1)**, **(d2)** y **(d3)**. Consideremos una función $f \in C(\overline{\Omega})$, $y \notin f(\partial\Omega)$ y definamos $\alpha = d(y, f(\partial\Omega)) > 0$. Por la proposición anterior, existe $g \in \overline{C}^\infty(\Omega)$ tal que $||f - g||_\infty < \alpha$. Notemos que la función $h : [0,1] \times \overline{\Omega} \to \mathbb{R}^n$ definida como $h(t, x) = (1 - t)f(x) + tg(x)$ es continua. Si $x \in \partial\Omega$, se tiene que $||f(x) - y|| \geq \alpha$, de donde

$$\begin{aligned} ||h(t, x) - y|| &= ||(1 - t)f(x) + tg(x) - y|| \\ &\geq ||f(x) - y|| - t||f(x) - g(x)|| \\ &\geq \alpha - ||f - g||_\infty \\ &> 0. \end{aligned}$$

Así, por **(d3)** se tiene que $d(f, \Omega, y) = d(g, \Omega, y)$ y $d'(f, \Omega, y) = d'(g, \Omega, y)$. Entonces por la unicidad del grado en $\overline{C}^\infty(\Omega)$ se tiene que

$$d(f, \Omega, y) = d(g, \Omega, y) = d'(g, \Omega, y) = d'(f, \Omega, y).$$

De la arbitrariedad de f y y se concluye que $d = d'$. ■

Como mencionamos anteriormente, nos basta probar la unicidad del grado en $\overline{C}^\infty(\Omega)$ con valores regulares $y \notin f(\partial\Omega)$. Éste será nuestro próximo paso.

Proposición 64 Sean $f \in \overline{C}^1(\Omega)$ y $y \notin f(\partial\Omega)$ un valor regular de f. Entonces el conjunto de soluciones de la ecuación $y = f(x)$ es finito.

Demostración Supongamos que el conjunto de soluciones de la ecuación $y = f(x)$ es infinito. Sea $(x_n)_{n=1}^\infty$ una sucesión de soluciones distintas de la ecuación $y = f(x)$. Como $\overline{\Omega}$ es compacto, entonces $(x_n)_{n=1}^\infty$ tiene una subsucesión convergente $(x_{n_k})_{k=1}^\infty$, digamos a $x_0 \in \overline{\Omega}$. Puesto que f es continua, entonces x_0

también es solución de la ecuación $y = f(x)$. Notemos que $x_0 \notin \partial\Omega$, ya que $y \notin f(\partial\Omega)$. Entonces $x_0 \in \Omega$. Como y es un valor regular, por definición $f^{-1}(y) \cap S_f(\Omega) = \emptyset$, y así $x_0 \notin S_f(\Omega)$, por lo cual $J_f(x_0) \neq 0$. Por el teorema de la función inversa existe una vecindad U_0 de x_0 tal que f es un homeomorfismo entre U_0 y una vecindad de y, lo cual es una contradicción, pues x_0 es el límite de una sucesión de soluciones distintas de la ecuación $y = f(x)$. ■

La siguiente proposición implicará que si $f \in C^1(\Omega)$, entonces toda bola abierta de $\mathbb{R}^n$ contiene valores regulares de f.

Teorema 65 (de Sard) Sean $\Omega \subset \mathbb{R}^n$ abierto y $f \in C^1(\Omega)$. Entonces $\mu_n(f(S_f)) = 0$, donde μ_n denota la medida n-dimensional de Lebesgue.

Demostración Como $\Omega \subset \mathbb{R}^n$ es abierto, puede ser escrito como una unión numerable de cubos abiertos, digamos, $\Omega = \cup_{i \in I} Q_i$, los cuales además pueden ser elegidos de manera que $\overline{Q_i} \subset \Omega$. Así, es suficiente probar que $\mu_n(f(S_f(Q))) = 0$ para un cubo abierto Q tal que $\overline{Q} \subset \Omega$, ya que $f(S_f(\Omega)) = \cup_{i \in I} f(S_f(Q_i))$. Sean pues Q un cubo abierto contenido en Ω tal que su cerradura está contenida en Ω y ϱ la longitud lateral de Q. Por la continuidad uniforme de f' en $\overline{Q}$, dado $\varepsilon > 0$ existe $m \in \mathbb{N}$ tal que $||f'(x) - f'(\overline{x})|| < \varepsilon$ para todo $x, \overline{x} \in Q$ con $||x - \overline{x}|| \leq \sqrt{n}\varrho/m$. Además, podemos tomar $m \in \mathbb{N}$ tal que $\sqrt{n}\varrho/m < \varrho$. Sea $\delta = \sqrt{n}\varrho/m$. Notemos que

$$f(x) - f(\overline{x}) - f'(\overline{x})(x - \overline{x}) = \int_0^1 (f'(\overline{x} + t(x - \overline{x})) - f'(\overline{x}))(x - \overline{x})dt.$$

Así, usando la desigualdad del triángulo para integrales obtenemos que para toda $x, \overline{x} \in Q$ con $||x - \overline{x}|| < \delta$,

$$\begin{aligned}
||f(x) - f(\overline{x}) - f'(\overline{x})(x - \overline{x})|| &\leq \left|\left| \int_0^1 (f'(\overline{x} + t(x - \overline{x})) - f'(\overline{x}))(x - \overline{x})dt \right|\right| \\
&\leq \int_0^1 ||f'(\overline{x} + t(x - \overline{x})) - f'(\overline{x})|| \cdot ||x - \overline{x}||dt \\
&< \varepsilon ||x - \overline{x}||.
\end{aligned}$$

Puesto que ϱ es la longitud lateral de Q, podemos descomponer a Q en $Q^1, ..., Q^{m^n}$ cubos de longitud lateral ϱ/m. El diámetro de cada Q^k es entonces δ. Por la desigualdad anterior para $x, \overline{x} \in Q^k$ se tiene que

$$f(x) = f(\overline{x}) + f'(\overline{x})(x - \overline{x}) + R(x, \overline{x}), \tag{4.1}$$

donde $||R(x, \overline{x})|| < \varepsilon\delta$. Ahora notemos que si $Q^k \cap S_f = \emptyset$ para todo $k \in \{1, ..., m^n\}$, habremos acabado, por lo que supondremos que $Q^k \cap S_f \neq \emptyset$ para algún $k \in \{1, .., m^n\}$. Sea $\overline{x} \in Q^k \cap S_f$. Definamos $A = f'(\overline{x})$, $\widetilde{Q}^k = Q^k - \overline{x}$ y $g(y) = f(\overline{x} + y) - f(\overline{x})$ para $y \in \widetilde{Q}^k$. Entonces por (4.1) tenemos que

$$g(y) = Ay + \widetilde{R}(y) \tag{4.2}$$

con $\widetilde{R}(y) = R(\overline{x} + y, \overline{x})$ y $||\widetilde{R}(y)|| = ||R(\overline{x} + y, \overline{x})|| < \varepsilon\delta$ en $\widetilde{Q}^k$. Ahora, como $\overline{x} \in S_f$, se tiene que $\det A = 0$, de donde $A(\widetilde{Q}^k)$ está contenido en un subespacio $n-1$-dimensional de $\mathbb{R}^n$. Así, existe un vector $b^1 \in \mathbb{R}^n$ tal que $||b^1|| = 1$ y

$(z,b^1)=\sum_{i=1}^n z_i b_i^1=0$ para todo $z\in A(\widetilde{Q}^k)$. Sea $y\in\widetilde{Q}^k$. Extendiendo $\{b^1\}$ a una base ortonormal $\{b^1,...,b^n\}$ de $\mathbb{R}^n$, se tiene que $g(y)=\sum_{i=1}^n(g(y),b^i)b^i$, y por la ecuación (4.2) se tiene que

$$|(g(y),b^1)|=|(\widetilde{R}(y),b^1)|\le||\widetilde{R}(y)||\cdot||b^1||<\varepsilon\delta \tag{4.3}$$

y como $y,0\in\widetilde{Q}^k$ y $diam(\widetilde{Q}^k)=\delta$, se tiene que $||y||<\delta$, de donde

$$|(g(y),b^i)|\le||A||\cdot||y||+||\widetilde{R}(y)||\le||A||\delta+\varepsilon\delta, \tag{4.4}$$

para $i=2,...,n$, donde $||A||=||(a_{ij})||=(\sum_{i,j=1}^n a_{ij}^2)^{\frac{1}{2}}$. Entonces, por (4.3) y (4.4) se tiene que $f(Q^k)=f(\overline{x})+g(\widetilde{Q}^k)$ está contenido en una n-caja J_k alrededor de $f(\overline{x})$ de manera que

$$\mu_n(J_k)=(2(||A||\delta+\varepsilon\delta))^{n-1}\cdot 2\varepsilon\delta=2^n(||A||+\varepsilon)^{n-1}\varepsilon\delta^n. \tag{4.5}$$

Notemos que $\overline{Q}$ es un compacto, por lo que f' está acotada en $\overline{Q}$, por, digamos, $c>0$. En particular se tiene que $||A||<c$, ya que $A=f'(\overline{x})$ y $\overline{x}\in\overline{Q}$, por lo cual de (4.5) se tiene que

$$\mu_n(J_k)<2^n(c+\varepsilon)^{n-1}\varepsilon\delta^n.$$

Como

$$\begin{aligned} f(S_f(Q)) &\subset \cup_{k:Q_k\cap S_f(Q)\neq\emptyset} f(Q_k\cap S_f(Q)) \\ &\subset \cup_{k:Q_k\cap S_f(Q)\neq\emptyset} f(Q_k) \\ &\subset \cup_{k:Q_k\cap S_f(Q)\neq\emptyset} J_k \end{aligned}$$

y

$$\sum_{k:Q_k\cap S_f(Q)\neq\emptyset}\mu_n(J_k)\le m^n\cdot 2^n(c+\varepsilon)^{n-1}\varepsilon\delta^n=2^n(c+\varepsilon)^{n-1}(\sqrt{n}\varrho)^n\varepsilon,$$

por la arbitrariedad de $\varepsilon>0$, $f(S_f(Q))$ tiene medida cero. ■

Corolario 66 Si $f\in C^1(\Omega)$, entonces toda bola abierta en $\mathbb{R}^n$ contiene valores regulares de f.

Demostración Sea B una bola abierta en $\mathbb{R}^n$. Como $\mu_n(B)>0$ y $\mu_n f(S_f)=0$, existe $y\in B\cap f(S_f)^c$. Así $y\in B$ y y es un valor regular de f (recordemos que $f(S_f)^c$ es el conjunto de valores regulares de f). ■

A continuación veremos que dada $f\in C(\overline{\Omega})$ y $y\notin f(\partial\Omega)$, entonces existe una vecindad de y en la cual $d(f,\Omega,\cdot)$ es constante.

Proposición 67 Sean $f\in C(\overline{\Omega})$ y $y\notin f(\partial\Omega)$. Entonces existe $\alpha>0$ tal que $d(f,\Omega,y_1)=d(f,\Omega,y)$ para todo $y_1\in B_\alpha(y)$.

Demostración Definamos $\alpha=d(y,f(\partial\Omega))$. Es claro que $B_\alpha(y)\cap f(\partial\Omega)=\emptyset$, por lo cual, si $y_1\in B_\alpha(y)$, por la propiedad **(d3)** utilizando $h(t,x)=f(x)$ y $y(t)=ty+(1-t)y_1$ obtenemos que

$$d(f, \Omega, y_1) = d(f, \Omega, y)$$

para todo $y_1 \in B_\alpha(y)$. ■

Del corolario 66 y de la proposición 67 obtenemos el siguiente corolario.

Corolario 68 Sean $f \in C^1(\Omega)$ y $y \notin f(\partial\Omega)$. Entonces existe y' valor regular de f tal que

$$d(f, \Omega, y) = d(f, \Omega, y').$$

Demostración Por la proposición 67, existe $\alpha > 0$ tal que

$$d(f, \Omega, y_1) = d(f, \Omega, y) \text{ para todo } y_1 \in B_\alpha(y).$$

Por otra parte, por el corolario 66 sabemos que $B_\alpha(y)$ contiene valores regulares de f. ■.

De la proposición 63 y del corolario 68 resulta que para probar la unicidad del grado es suficiente probar la unicidad del grado para funciones en $\overline{C}^\infty(\Omega)$ y valores regulares, por lo que de ahora en adelante, salvo que se indique lo contrario, supondremos que $f \in \overline{C}^\infty(\Omega)$ y $y \notin f(\partial\Omega) \cup f(S_f) = f(\partial\Omega \cup S_f)$.

Proposición 69 Sean $f \in \overline{C}^\infty(\Omega)$ y $y \notin f(\partial\Omega \cup S_f)$. Entonces podemos escribir el grado de f en y como

$$d(f, \Omega, y) = \sum_{i=1}^{k} d(f, U_i, y),$$

donde las U_i son vecindades disjuntas de las preimágenes de y, y $\sum_{\emptyset} = 0$.

Demostración Usando **(d2)** con $\Omega_1 = \Omega$ y $\Omega_2 = \emptyset$ se tiene que

$$\begin{aligned} d(f, \Omega, y) &= d(f, \Omega_1, y) + d(f, \Omega_2, y) \\ &= d(f, \Omega, y) + d(f, \emptyset, y), \end{aligned}$$

de donde $d(f, \emptyset, y) = 0$. Por otro lado, si Ω_1 es un subconjunto abierto de Ω tal que $y \notin f(\overline{\Omega} \backslash \Omega_1)$ y $\Omega_2 = \emptyset$, entonces nuevamente por **(d2)** se tiene que

$$\begin{aligned} d(f, \Omega, y) &= d(f, \Omega_1, y) + d(f, \Omega_2, y) \\ &= d(f, \Omega_1, y) + 0. \end{aligned} \tag{4.6}$$

Recordemos que como y es un valor regular, entonces por la proposición 64 $f^{-1}(y)$ es finito. Si $f^{-1}(y) = \emptyset$, al tomar $\Omega_1 = \emptyset$ en (4.6), se tiene que $d(f, \Omega, y) = d(f, \emptyset, y) = 0$. En caso de que $f^{-1}(y) = \{x^1, x^2, ..., x^k\}$, podemos escoger vecindades disjuntas U_i de x^i y, nuevamente de **(d2)**, obtenemos

$$d(f, \Omega, y) = \sum_{i=1}^{k} d(f, U_i, y). \blacksquare$$

A continuación probaremos que podemos expresar cada $d(f, U_i, y)$ de la proposición anterior como

$$d(f, U_i, y) = d(f'(x^i), B_r, 0),$$

donde $r > 0$ es arbitrario, y en consecuencia, para probar la unicidad del grado será suficiente probar que $d(A, B_r, 0)$ se obtiene de manera única si A es una función lineal con detrerminante no nulo.

Antes de continuar introduciremos la siguiente notación que resulta de utilidad cuando se comparan funciones al momento de tomar ciertos límites.

Notación 70 (o de Landau) Sean $f, g : \mathbb{R}^n \to \mathbb{R}$. Decimos que $g = o(f)$ cuando $x \to p$ si existe $r > 0$ tal que f no se anula en $B_r(p) \backslash \{p\}$ y

$$\lim_{x \to p} \frac{g(x)}{f(x)} = 0.$$

Proposición 71 Sean f, U_i y x^i como en la proposición anterior. Entonces

$$d(f, U_i, y) = d(f'(x^i), B_r, 0),$$

donde $r > 0$ es arbitrario.

Demostración Sea $A = f'(x^i)$ y notemos que, por el teorema de Taylor,

$$f(x) = y + A(x - x^i) + o(||x - x^i||), \text{ cuando } ||x - x^i|| \to 0. \tag{4.7}$$

Como y es un valor regular de f, se tiene que $\det A \neq 0$, por lo cual A^{-1} existe y obtenemos que, si $z \in \mathbb{R}^n$,

$$\begin{aligned} ||z|| &= ||A^{-1}Az|| \\ &\leq ||A^{-1}|| \cdot ||Az||, \end{aligned}$$

es decir, existe un $c > 0$ tal que $||Az|| \geq c||z||$, para todo $z \in \mathbb{R}^n$. Por lo cual si $\widetilde{y}(t) = ty$ y $h(t, x) = tf(x) + (1 - t)A(x - x^i)$, por (4.7) se tiene que

$$\begin{aligned} ||h(t, x) - \widetilde{y}(t)|| &= ||A(x - x^i) + t \cdot o(||x - x^i||)|| \\ &\geq c||x - x^i|| - o(||x - x^i||) \\ &> 0, \end{aligned}$$

para todo $t \in [0, 1]$ y $||x - x^i|| \leq \delta$ con $\delta > 0$ suficientemente pequeño. Entonces $\widetilde{y}(t) \notin h(t, \partial\Omega)$, y en consecuencia, por **(d3)** se tiene que

$$d(f, B_\delta(x^i), y) = d(A - Ax^i, B_\delta(x^i), 0). \tag{4.8}$$

Como x^i es la única solución a $Ax - Ax^i = 0$, si $B_\delta(x^i) \subset B_r$ para algún $r > 0$, entonces $0 \notin (A(\cdot) - Ax^i)(B_r \backslash B_\delta(x^i))$. Por lo cual, por **(d2)** resulta que

$$d(A - Ax^i, B_\delta(x^i), 0) = d(A - Ax^i, B_r, 0). \tag{4.9}$$

Por otra parte, podemos suponer que $\delta > 0$ también cumple que $B_\delta(x^i) \subset U_i$, y puesto que x^i es la unica solución en U_i de la ecuación $f(x) = y$, entonces por **(d2)** obtenemos que

$$d(f, B_\delta(x^i), y) = d(f, U_i, y). \tag{4.10}$$

Luego, si definimos $h(t,x) = Ax - tAx^i$, se tiene que $h(t,x) \neq 0$ en $[0,1] \times \partial B_r$, de donde **(d3)** y las ecuaciones (4.10), (4.8) y (4.9) implican que

$$d(f, U_i, y) = d(f'(x^i), B_r, 0).$$

Finalmente, por **(d2)**, $r > 0$ puede ser arbitrario. ■

Como comentamos anteriormente, para probar la unicidad del grado, lo único que resta probar es que el grado es único para funciones lineales con determinante no nulo, lo que demostraremos viendo que, si A es una transformación lineal con $\det A \neq 0$, entonces

$$d(A, B_r, 0) = Sgn(\det A).$$

Para probar lo mencionado, utilizaremos un par de resultados de álgebra lineal. La siguiente definición y proposción pueden ser consultados en [10]

Definición 72 Si A es una matriz de $n \times n$ con entradas en un campo $\mathbb{K}$, su polinomio característico es

$$P_A(x) = \det(xI - A),$$

donde I es la matriz identidad de $n \times n$.

Proposición 73 Sean $A = [a_{ij}]$ una matriz de $n \times n$ con entradas en un campo $\mathbb{K}$, y $P_A(x) = \prod_{i=1}^{n}(x - \alpha_i)$ su polinomio característico. Entonces

$$tr(A) = \sum_{i=1}^{n} \alpha_i \text{ y } \det(A) = \prod_{i=1}^{n} \alpha_i.$$

Proposición 74 Sean A una matriz real con $\det A \neq 0$, $\lambda_1, ..., \lambda_m$ los valores propios negativos de A y $\alpha_1, ..., \alpha_m$ sus multiplicidades como ceros de $\det(A - \lambda I)$, suponiendo que tiene dichos valores propios. Entonces $\mathbb{R}^n$ es la suma directa de dos subespacios, N y M, tales que,

(a) N y M son invariantes bajo A.
(b) $A\big|_N$ tiene sólo los valores propios $\lambda_1, ..., \lambda_m$ y $A\big|_M$ no tiene valores propios negativos.
(c) $\dim(N) = \sum_{k=1}^{m} \alpha_k$.

Antes de continuar enunciaremos la definición de homotopía, que nos será de utilidad en la proposición siguiente.

Definición 75 Sean X, Y espacios topológicos y $f, g : X \to Y$ funciones continuas. Decimos que f es homotópica a g si existe una función continua $F : X \times [0,1] \to Y$ tal que $F(x,0) = f(x)$ y $F(x,1) = g(x)$. A una función F que satisface lo anterior se le llama homotopía.

Proposición 76 Si $A : \mathbb{R}^n \to \mathbb{R}^n$ es una transformación lineal invertible y $r > 0$, entonces

$$d(A, B_r, 0) = Sgn(\det(A)).$$

Demostración Sean $\lambda_1, ...\lambda_m$ los valores propios negativos de A y $\alpha_1, ..., \alpha_m$ sus multiplicidades como ceros de $\det(A - \lambda I)$, suponiendo que tiene dichos valores propios. Sean N y M como en la proposición anterior. Expresemos

$$\det(A - \lambda I) = (-1)^n \prod_{k=1}^{m} (\lambda - \lambda_k)^{\alpha_k} \prod_{j=m+1}^{n} (\lambda - \mu_j)^{\beta_j}.$$

Entonces si $\alpha = \sum_{k=1}^{m} \alpha_k = \dim(N)$,

$$\det A = (-1)^\alpha \prod_{k=1}^{m} |\lambda_k|^{\alpha_k} \prod_{j=m+1}^{n} \mu_j^{\beta_j}.$$

Notemos que como A es una transformación lineal real, si $\mu = a + ib$ es un valor propio de A con $b \neq 0$, entonces $\overline{\mu} = a - ib$ también es valor propio de A, y $\mu\overline{\mu} = a^2 + b^2 > 0$. Por lo tanto

$$\prod_{j=m+1}^{n} \mu_j^{\beta_j} > 0.$$

Así, como $|\lambda_k| > 0$ para $k \in \{1, ..., m\}$, $Sgn(\det A) = (-1)^\alpha$. Puesto que $\det(A - \lambda I) = 0$ si y sólo si λ es un valor propio de A, entonces se tiene que

$$\det(tA + (1-t)I) = 0$$

si y sólo si $\frac{-(1-t)}{t} < 0$ es un valor propio de A. Por lo tanto, si A no tiene valores propios negativos se tiene que $\det(tA + (1-t)I) \neq 0$ en $[0, 1]$. Entonces si A no tiene valores propios negativos, usando $h(t, x) = tAx + (1-t)x$, obtenemos que $0 \notin h(t, \partial B_r)$ y en consecuencia de **(d3)** y de **(d1)** obtenemos que

$$d(A, B_r, 0) = d(I, B_r, 0) = 1 = Sgn(\det A).$$

Supongamos ahora que A tiene valores propios negativos, es decir, consideraremos el caso en que $N \neq \{0\}$. Denotemos por Ω a B_r. Examinaremos este caso a través de tres pasos.

Paso 1 Supongamos que $\alpha = \dim N$ es par, digamos que $\alpha = 2p$. Como $\mathbb{R}^n = N \oplus M$, todo $x \in \mathbb{R}^n$ tiene una única expresión como $x = P_1x + P_2x$, con $P_1x \in N$ y $P_2x \in M$, donde $P_1 : \mathbb{R}^n \to N$ $P_2 = I - P_1 : \mathbb{R}^n \to M$ son proyecciones lineales. Por la proposición 74, podemos descomponer la matriz A como $A = AP_1 + AP_2$, ya que $A(N) \subset N$ y $A(M) \subset M$. Probaremos que $h(t, x) = tAx + (1-t)(-P_1x + P_2x) \neq 0$ en $[0, 1] \times \partial\Omega$. Tomemos $(t, x) \in [0, 1] \times \partial\Omega$. Notemos que $h(0, x) = 0$ implica que $P_1x = P_2x$, de donde $P_1x = P_2x \in N \cap M = \{0\}$, por lo cual $x = 0$. Por otro lado, si $h(t, x) = 0$ con $t \neq 0$, entonces

$$AP_1x = t^{-1}(1-t)P_1x \in N \text{ y } AP_2x = -t^{-1}(1-t)P_2x \in M.$$

Por (b) de la proposición anterior se tiene que AP_1 sólo tiene valores propios negativos, de donde AP_1x no puede ser un múltiplo positivo de P_1x. Por otro lado, AP_2 sólo tiene valores propios positivos, de donde AP_2x no puede ser un múltiplo negativo de P_2x, en consecuencia $P_1x = P_2x = 0$, lo cual implica que $x = 0$. Por lo tanto, $h(t,x) \neq 0$ en $[0,1] \times \partial\Omega$, de donde **(d3)** implica

$$d(A,\Omega,0) = d(-P_1 + P_2, \Omega, 0). \tag{4.11}$$

Consideremos la matriz

$$\widetilde{B} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

y notemos que $\widetilde{B}^2 = -I$, donde I es la matriz identidad. Definamos a B como la matriz de $\alpha \times \alpha$ que tiene p veces la matriz $\widetilde{B}$ en su diagonal y 0 en los demás registros. Entonces $B^2 = -I|_N$. Definamos $h_1(t,x)$ homotopía de $-P_1 + P_2$ a $BP_1 + P_2$ como

$$h_1(t,x) = t(BP_1x + P_2x) + (1-t)(-P_1x + P_2x).$$

Veremos que $0 \notin h_1([0,1], \partial\Omega)$. Sea $(t,x) \in [0,1] \times \partial\Omega$. Notemos que, en términos de sus componentes, $h_1(t,x) = 0$ si y sólo si

$$t(BP_1x) + (1-t)(-P_1x) = 0 \text{ y } P_2x = 0, \tag{4.12}$$

ya que P_1 y P_2 son proyecciones en subespacios cuya intersección es $\{0\}$. Supongamos que $h(t,x) = 0$ con $t \neq 0$. De la primera igualdad de (4.12) resulta que

$$BP_1x = (\tfrac{1-t}{t})P_1x,$$

lo cual no es posible, pues B sólo tiene valores propios complejos no reales. Por otro lado notemos que

$$h(0,x) = -P_1x + P_2x = 0$$

si y sólo si $P_1x = 0$ y $P_2x = 0$, lo cual implica que $x = 0$, lo cual es una contradicción. Así, $0 \notin h_1([0,1], \partial\Omega)$. Definamos ahora $h_2(t,x)$ homotopía de $BP_1 + P_2$ a $I = P_1 + P_2$ como

$$\begin{aligned} h_2(t,x) &= tIx + (1-t)(BP_1x + P_2x) \\ &= t(P_1 + P_1)x + (1-t)(BP_1x + P_2x). \end{aligned}$$

Similarmente a h_1, se tiene que $0 \notin h_2([0,1], \partial\Omega)$. Así, por **(d3)** usando h_1 y h_2 y por (4.11),

$$\begin{aligned} d(A,\Omega,0) &= d(-P_1 + P_2, \Omega, 0) \\ &= d(I, \Omega, 0) = 1 \\ &= (-1)^{2p} = (-1)^{\alpha} = Sgn(\det A). \end{aligned}$$

Paso 2 Supongamos ahora que $\alpha = \dim N = 2p + 1$ para algún $p \geq 0$. Entonces descompongamos $N = N_1 \oplus N_2$, con $\dim N_1 = 1$ y $\dim N_2 = 2p$, con sus respectivas proyecciones $\widetilde{Q}_1 : N \to N_1$ y $\widetilde{Q}_2 = I\big|_N - \widetilde{Q}_1 : N \to N_2$. Entonces $P_1 = \widetilde{Q}_1 P_1 + \widetilde{Q}_2 P_1$, y como en el primer paso, encontramos homotopías indicadas por "$\to$" tales que

$$A \to -P_1 + P_2 \to -\widetilde{Q}_1 P_1 + B\widetilde{Q}_2 P_1 + P_2 \to -\widetilde{Q}_1 P_1 + \widetilde{Q}_2 P_1 + P_2.$$

Entonces, si definimos $Q_1 = \widetilde{Q}_1 P_1$ y $Q_2 = \widetilde{Q}_2 P_1 + P_2$, se tiene que

$$d(A, \Omega, 0) = d(-Q_1 + Q_2, \Omega, 0).$$

Notemos que Q_1 y $Q_2 = I - Q_1$ son las proyecciones de la descomposición $\mathbb{R}^n = N_1 \oplus (N_2 \oplus M)$. Al igual que en el Paso 1, $x = 0$ es el único cero de $-Q_1 + Q_2$, esto porque $-Q_1 x + Q_2 x = 0$ si y sólo si $Q_1 x = Q_2 x \in N_1 \cap (N_2 \oplus M) = \{0\}$ por ser $\mathbb{R}^n$ suma directa de N_1 y $(N_2 \oplus M)$. Por lo tanto, por **(d2)**, podemos reemplazar a $\Omega = B_r$ en $d(-Q_1 + Q_2, \Omega, 0)$ por cualquier conjunto abierto y acotado que contenga a $x = 0$ sin cambiar d, por ejemplo, por el cilindro $(B_r \cap N_1) + \widetilde{B}_r$, donde $\widetilde{B}_r = B_r \cap (N_2 \oplus M)$. Veremos ahora que el problema ha sido reducido esencialmente a una dimensión. Dados $\Omega_{N_1} \subset N_1$ abierto y acotado y $g : \overline{\Omega_{N_1}} \to N_1$ continua con $0 \notin g(\partial \Omega_{N_1})$, definamos $\widetilde{d}(g, \Omega_{N_1}, 0) = d((g \circ Q_1) + Q_2, \Omega_{N_1} + \widetilde{B}_r, 0)$. Así, **(d1)**, **(d2)** y **(d3)** implican

$\widetilde{d}$**1)** $\widetilde{d}(I, \Omega_{N_1}, 0) = 1$, cuando $0 \in \Omega_{N_1}$, donde I es la función identidad en Ω_{N_1},

$\widetilde{d}$**2)** $\widetilde{d}(g, \Omega_{N_1}, 0) = \widetilde{d}(g, \Omega_1, 0) + \widetilde{d}(g, \Omega_2, 0)$ para $\Omega_1, \Omega_2 \subset \Omega_{N_1}$ abiertos y disjuntos tales que $0 \notin g(\overline{\Omega_{N_1}} \backslash (\Omega_1 \cup \Omega_2))$,

y

$\widetilde{d}$**3)** $\widetilde{d}(h(t, \cdot), \Omega_{N_1}, 0)$ es constante en $[0, 1]$, si $h : [0, 1] \times \overline{\Omega_{N_1}} \to N_1$, es continua y $0 \notin h([0, 1] \times \partial \Omega_{N_1})$,

pues, si $0 \in \Omega_{N_1}$, por **(d1)** se tiene que

$$\begin{aligned} \widetilde{d}(I, \Omega_{N_1}, 0) &= d((I \circ Q_1) + Q_2, \Omega_{N_1} + \widetilde{B}_r, 0) \\ &= d(Q_1 + Q_2, \Omega_{N_1} + \widetilde{B}_r, 0) \\ &= d(I, \Omega_{N_1} + \widetilde{B}_r, 0) = 1. \end{aligned}$$

Por otra parte, por **(d2)**,

$$\begin{aligned} \widetilde{d}(g, \Omega_{N_1}, 0) &= d((g \circ Q_1) + Q_2, \Omega_{N_1} + \widetilde{B}_r, 0) \\ &= d((g \circ Q_1) + Q_2, \Omega_1 + \widetilde{B}_r, 0) + d((g \circ Q_1) + Q_2, \Omega_2 + \widetilde{B}_r, 0) \text{ (por } \mathbf{(d2)}) \\ &= \widetilde{d}(g, \Omega_1, 0) + \widetilde{d}(g, \Omega_2, 0), \end{aligned}$$

y por **(d3)**,

$$\widetilde{d}(h(t, \cdot), \Omega_{N_1}, 0) = d((h(t, \cdot) \circ Q_1) + Q_2, \Omega_{N_1} + \widetilde{B}_r, 0),$$

que es constante en $[0, 1]$. Sea $\Omega_{N_1} \subset N_1$ un subconjunto abierto y acotado con $0 \in \Omega_{N_1}$. Como

$$d(-Q_1 + Q_2, \Omega, 0) = d(-Q_1 + Q_2, \Omega_{N_1} + \tilde{B}_r, 0)$$

y

$$d(-Q_1 + Q_2, \Omega_{N_1} + \tilde{B}_r, 0) = \tilde{d}(-I|_{N_1}, \Omega_{N_1}, 0),$$

tenemos que probar que

$$\tilde{d}(-I|_{N_1}, \Omega_{N_1}, 0) = -1 = (-1)^{2p+1} = Sgn(\det A).$$

Como **($\widetilde{d}$1)** es el único cálculo concreto que tenemos, es natural buscar una función g y conjuntos abiertos y disjuntos $\Omega_1, \Omega_2 \subset \Omega_{N_1}$, con $0 \notin g(\overline{\Omega_{N_1}} \setminus (\Omega_1 \cup \Omega_2))$, tales que, $\widetilde{d}(g, \Omega_{N_1}, 0) = 0$, $g\big|_{\Omega_1}$ es homotópico a $-I\big|_{\Omega_1}$ y $g\big|_{\Omega_2}$ es homotópico a $I\big|_{\Omega_2}$, ya que de esta manera, de **(d2)** y **(d3)** se obtendrá que

$$\widetilde{d}(-I\big|_{N_1}, \Omega_{N_1}, 0) = -\widetilde{d}(I\big|_{N_1}, \Omega_{N_1}, 0) = -1.$$

Efectivamente, de **($\widetilde{d}$2)** tendremos que

$$0 = \widetilde{d}(g, \Omega_1, 0) + \widetilde{d}(g, \Omega_2, 0)$$

y de **($\widetilde{d}$3)** tendremos que

$$\widetilde{d}(g, \Omega_1, 0) = \widetilde{d}(-I\big|_{N_1}, \Omega_1, 0)$$

y

$$\widetilde{d}(g, \Omega_2, 0) = \widetilde{d}(I\big|_{N_1}, \Omega_1, 0) = 1,$$

concluyendo que

$$\widetilde{d}(-I\big|_{N_1}, \Omega_{N_1}, 0) = -\widetilde{d}(I\big|_{N_1}, \Omega_{N_1}, 0) = -1.$$

En el siguiente paso encontraremos una función g y conjuntos $\Omega_1, \Omega_2 \subset \Omega_{N_1}$ para realizar el cálculo anterior.

Paso final Como $\dim N_1 = 1$, existe un $e \in \mathbb{R}^n$ con $||e|| = 1$ tal que $N_1 = \{\lambda e : \lambda \in \mathbb{R}\}$. Consideremos los subconjuntos

$$\Omega_{N_1} = \{\lambda e : \lambda \in (-2, 2)\},\ \Omega_1 = \{\lambda e : \lambda \in (-2, 0)\} \text{ y } \Omega_2 = \{\lambda e : \lambda \in (0, 2)\},$$

y las funciones $g : N_1 \to N_1$ y $h : [0,1] \times \overline{\Omega_{N_1}} \to N_1$ definidas como

$$g(\lambda e) = (|\lambda| - 1)e \text{ y } h(t, \lambda e) = t(|\lambda| - 2)e + e.$$

Notemos que $h(0, \lambda e) = e$ y $h(1, \lambda e) = g(\lambda e)$. Veremos que estos son los conjuntos y la función que buscábamos en el paso anterior.

Notemos que $\widetilde{d}(e, \Omega_{N_1}, 0) = 0$, ya que la preimagen de 0 bajo la función constante e es vacía. Además, si $\lambda e \in \partial\Omega_{N_1}$, entonces $|\lambda| = 2$, por lo que, en este caso, $h(t, x) = e \neq 0$ en $[0,1] \times \partial\Omega_{N_1}$. Notemos también que $g(0) = -e \neq 0$. Teniendo en cuenta estas observaciones, por ($\widetilde{d}$**3**) aplicada a h y por **($\widetilde{d}$2)** respectivamente se tiene que

$$\begin{aligned} 0 &= \widetilde{d}(e, \Omega_{N_1}, 0) \\ &= \widetilde{d}(g, \Omega_{N_1}, 0) \\ &= \widetilde{d}(g, \Omega_1, 0) + \widetilde{d}(g, \Omega_2, 0). \end{aligned}$$

Ahora notemos que $g\big|_{\Omega_1}(\lambda e) = -(\lambda+1)e$ tiene como único cero a $-e \in \Omega_1 \subset \Omega_{N_1}$. Definamos $h_1 : [0,1] \times \overline{\Omega_{N_1}} \to N_1$ como $h_1(t,\lambda e) = -\lambda e - te$. Como $g = -I\big|_{N_1} - e$ en Ω_1 y $-I\big|_{N_1} - e$ no se anula en $\overline{\Omega_{N_1}} \backslash \Omega_1$, usando **($\widetilde{d}$2)** y **($\widetilde{d}$3)** con h_1 se tiene que

$$\begin{aligned}
\widetilde{d}(g,\Omega_1,0) &= \widetilde{d}(-I\big|_{N_1} - e, \Omega_1, 0) \\
&= \widetilde{d}(-I\big|_{N_1} - e, \Omega_{N_1}, 0) \\
&= \widetilde{d}(-I\big|_{N_1}, \Omega_{N_1}, 0).
\end{aligned}$$

Por un argumento análogo, definiendo $h_2 : [0,1] \times \overline{\Omega_{N_1}} \to N_1$ como $h_1(t,\lambda e) = \lambda e - te$, usando **($\widetilde{d}$3)** con h_2 obtenemos que $\widetilde{d}(g,\Omega_2,0) = \widetilde{d}(I\big|_{N_1}, \Omega_{N_1}, 0) = 1$, de donde $\widetilde{d}(-I\big|_{N_1}, \Omega_{N_1}, 0) = -1$, que es lo que queríamos probar. ■

Hemos probado entonces que $d(A,\Omega,0) = Sgn(\det(A))$ para transformaciones lineales invertibles, lo que, como se comentó anteriormente, era lo que faltaba para probar la unicidad del grado, es decir, demostramos el siguiente teorema.

Teorema 77 Sea

$$M = \{(f,\Omega,y) : \Omega \subset \mathbb{R}^n \text{ abierto y acotado, } f \in C(\overline{\Omega}) \text{ y } y \in \mathbb{R}^n \backslash f(\partial\Omega)\}.$$

Entonces existe a lo más una función $d : M \to \mathbb{Z}$ con las propiedades **(d1)**, **(d2)** y **(d3)**. Más aún, estas propiedades implican que $d(A,\Omega,0) = Sgn(\det A)$ para transformaciones lineales A con $\det A \neq 0$ y $0 \in \Omega$.

Ya probamos la unicidad del grado en $\mathbb{R}^n$, pero eso de poco nos sirve sin saber que dicha función existe. Eso es lo que probaremos en la siguiente sección.

4.1.2 Existencia del grado en $\mathbb{R}^n$

En la sección anterior vimos que a lo más existe una función d que cumple **(d1)**, **(d2)** y **(d3)**, restringiéndonos al caso más simple, siendo este el grado para funciones lineales. Además, vimos cómo calcular el grado para funciones lineales invertibles. Comencemos ahora con la siguiente definición, la cual está motivada por los resultados obtenidos en la sección anterior, en particular por las proposiciones 69, 71 y 76.

Definición 78 Sean $\Omega \subset \mathbb{R}^n$ abierto y acotado, $f \in \overline{C}^1(\Omega)$ y $y \in \mathbb{R}^n \backslash f(\partial\Omega \cup S_f)$. Entonces definimos

$$d(f,\Omega,y) = \sum_{x \in f^{-1}(y)} Sgn J_f(x), \tag{4.13}$$

donde $\sum_{\emptyset} = 0$.

Algo que podría parecer problemático sobre la definición anterior es el requerir que $y \notin f(S_f)$, es decir que y sea un valor regular de f. En la sección anterior probamos que el conjunto $f(S_f)$ tiene medida cero, y en consecuencia es irrelevante para fines de integración, por lo que, para eliminar la hipótesis de que y es un valor regular de f, reemplazaremos $\sum SgnJ_f(x)$ por una integral más manejable.

Proposición 79 Sean f, Ω y y como en la definición anterior y $(\varphi_\varepsilon)_{\varepsilon>0}$ las funciones suavizantes introducidas en la sección anterior. Entonces existe $\varepsilon_0 = \varepsilon_0(y, f) > 0$ tal que

$$d(f, \Omega, y) = \int_\Omega \varphi_\varepsilon(f(x) - y)J_f(x)dx, \tag{4.14}$$

para $0 < \varepsilon \leq \varepsilon_0$.

Demostración Por la proposición 64 sabemos que $f^{-1}(y)$ es un conjunto finito. Notemos que si $f^{-1}(y) = \emptyset$, entonces para $x \in \Omega$ y $\varepsilon > 0$ suficientemente pequeño se tiene que $\left|\left|\frac{f(x)-y}{\varepsilon}\right|\right| > 1$, de donde $\varphi_\varepsilon(f(x)-y) = 0$, y en consecuencia se cumple (4.14). Veamos entonces el caso en que $f^{-1}(y) = \{x^1, ..., x^m\}$. Puesto que y es como en la definición anterior, por el teorema de la función inversa, existen bolas disjuntas $B_\varrho(x^i)$ tales que $f\big|_{B_\varrho(x^i)}$ es un homeomorfismo sobre una vecindad V_i de y. Por otra parte, como $f \in C^1(\Omega)$, entonces J_f es continuo en Ω, por lo cual podemos suponer también que $SgnJ_f(x) = SgnJ_f(x^i)$ en $B_\varrho(x^i)$. Sean $B_r(y) \subset \cap_{i=1}^m V_i$ y $U_i = B_\varrho(x^i) \cap f^{-1}(B_r(y))$. Entonces $||f(x)-y|| > r$ en $\overline{\Omega}\backslash\cup_{i=1}^m U_i$, de donde $\varepsilon < r$ y $x \notin \cup_{i=1}^\infty U_i$ implica $\left|\left|\frac{f(x)-y}{\varepsilon}\right|\right| > 1$. Luego, si $x \notin \cup_{i=1}^m U_i$ y $\varepsilon < r$, se tiene que $\varphi_\varepsilon(f(x) - y) = 0$. En consecuencia, si $\varepsilon < r$ se tiene que

$$\begin{aligned}\int_\Omega \varphi_\varepsilon(f(x) - y)J_f(x)dx &= \sum_{i=1}^m \int_{U_i} \varphi_\varepsilon(f(x) - y)J_f(x)dx \\ &= \sum_{i=1}^m SgnJ_f(x^i) \int_{U_i} \varphi_\varepsilon(f(x) - y)|J_f(x)|dx.\end{aligned}$$

Por otra parte, puesto que

$$U_i = (f|_{B_\varrho(x^i)})^{-1}(B_r(y))$$

y $f\big|_{B_\varrho(x^i)}$ homeomorfismo sobre V_i, entonces $f(U_i) = B_r(y)$. Así, $f(U_i)-y = B_r$ y además $J_f(x) = J_{f-y}(x)$, por lo que el teorema de cambio de variable implica que

$$\int_{U_i} \varphi_\varepsilon(f(x) - y)|J_{f-y}(x)|dx = \int_{B_r} \varphi_\varepsilon(x)dx = 1,$$

para $\varepsilon < r$. (Notemos que como $\varepsilon < r$, entonces $sop(\varphi_\epsilon) \subset \overline{B_\epsilon} \subset \overline{B_r}$). De lo anterior se sigue el resultado con $\epsilon_0 = \frac{r}{2}$. ■

El siguiente paso es eliminar en la definición 78 la hipótesis de que $y \notin f(S_f)$. Supongamos que $f \in \overline{C}^2(\Omega)$ y $y \notin f(\partial\Omega)$, no necesariamente un valor regular de f. Sea $\alpha = d(y, f(\partial\Omega))$ y supongamos que $y^1, y^2 \in B_\alpha(y)$ son dos valores regulares de f. Tales valores regulares de f existen por el corolario 66. Además sea $\delta = \alpha - \max\{||y^i - y|| : i = 1, 2\}$. Por la proposición anterior podemos encontrar $\varepsilon < \delta$ tal que

$$d(f, \Omega, y^i) = \int_\Omega \varphi_\varepsilon(f(x) - y^i)J_f(x)dx, \tag{4.15}$$

para $i = 1, 2$. Más adelante probaremos que estas integrales son iguales, lo cual justifica la siguiente definición.

Definición 80 Sean $\Omega \subset \mathbb{R}^n$ abierto y acotado, $f \in \overline{C}^2(\Omega)$ y $y \notin f(\partial\Omega)$. Definimos

$$d(f, \Omega, y) = d(f, \Omega, y^1),$$

donde y^1 es cualquier valor regular de f tal que $||y^1 - y|| < d(y, f(\partial\Omega))$ y $d(f, \Omega, y^1)$ está dado por la definición 78.

Tenemos todo listo para probar que las integrales de (4.15) son iguales y en consecuencia para probar el siguiente lema.

Lema 81 Sean $f \in \overline{C}^2(\Omega)$ y $y \notin f(\partial\Omega)$. Sea $\alpha = d(y, f(\partial\Omega))$. Entonces

$$d(f, \Omega, y^1) = d(f, \Omega, y^2)$$

para todo par y^1, y^2 de valores regulares de f en $B_\alpha(y)$.

Demostración Para $x \in \mathbb{R}^n$ definamos

$$w(x) = \left[\int_0^1 \varphi_\varepsilon(x - y^1 + t(y^1 - y^2))dt \right] (y^1 - y^2).$$

Notemos que si $\sigma(t) = x - y^1 + t(y^1 - y^2)$, entonces

$$\tfrac{d}{dt}\varphi_\varepsilon(x - y^1 + t(y^1 - y^2)) = \sum_{i=1}^{n} \partial_i \varphi_\epsilon(\sigma(t))(y^1 - y^2)_i.$$

Por lo tanto,

$$\begin{aligned} div(w(x)) &= \sum_{i=1}^{n} \partial_i w_i(x) \\ &= \sum_{i=1}^{n} \partial_i \Big[\int_0^1 \varphi_\epsilon(\sigma(t))(y^1 - y^2)_i dt \Big] \\ &= \sum_{i=1}^{n} \int_0^1 \partial_i \varphi_\epsilon(\sigma(t))(y^1 - y^2)_i dt \\ &= \int_0^1 \sum_{i=1}^{n} \partial_i \varphi_\epsilon(\sigma(t))(y^1 - y^2)_i dt \\ &= \int_0^1 \frac{d}{dt} \varphi_\epsilon(\sigma(t))dt, \end{aligned}$$

de donde por el teorema fundamental del cálculo,

$$div(w(x)) = \varphi_\varepsilon(x - y^2) - \varphi_\varepsilon(x - y^1). \tag{4.16}$$

Recordemos que $sop\varphi_\varepsilon = \overline{B}_\varepsilon$, por lo cual, si $||x - (y^1 - t(y^1 - y^2))|| > \varepsilon$, entonces $w(x) = 0$, de donde $sop(w) \subset \overline{B}_r(y)$, donde $r = \alpha - (\delta - \varepsilon)$. Esto en particular implica que $f(\partial\Omega) \cap sop(w) = \emptyset$. Ahora probaremos que esto último implica la existencia de una función $v \in C^1(\mathbb{R}^n)$ tal que $sop(v) \subset \Omega$ y

$$(\varphi_\varepsilon(f(x) - y^2) - \varphi_\varepsilon(f(x) - y^1))J_f(x) = div(v(x)) \text{ en } \Omega, \tag{4.17}$$

de lo cual se seguirá el resultado, pues si $Q = [-a, a]^n$ es un hipercubo tal que $\Omega \subset Q$, se tendrá que

$$\begin{aligned} d(f, \Omega, y^2) - d(f, \Omega, y^2) &= \int_\Omega divv(x)dx \\ &= \int_Q divv(x)dx \\ &= \sum_{i=1}^n \int_{-a}^a \dots \int_{-a}^a \left(\int_{-a}^a \frac{\partial v_i}{\partial x_i} dx_i \right) dx_1 \dots dx_{i-1} dx_{i+1} \dots dx_n \\ &= 0, \end{aligned}$$

pues $\Omega \cap \partial Q = \emptyset$ y $sop(v) \subset \Omega$. Para mostrar la existencia de la función v, definamos su i-ésima componente como $v_i(x) = \sum_{j=1}^n w_j(f(x))d_{ij}(x)$ en $\overline{\Omega}$ (donde w_j es la j-ésima componente de w y d_{ij} es como en la proposición 19) y $v_i(x) = 0$ en $\mathbb{R}^n \backslash \overline{\Omega}$, para $i = 1, ..., n$. Como $sop(w) \subset \overline{B}_r(y) \subset B_\alpha(y)$, entonces $v(x) = 0$, si $x \in \partial\Omega$. Así $sop(v) \subset \Omega$. Además, por regla de la cadena tenemos que

$$\begin{aligned} \partial_i v_i(x) &= \sum_{j=1}^n w_j(f(x))\partial_i d_{ij}(x) + \sum_{j=1}^n d_{ij}(x)\partial_i(w_j(f(x))) \\ &= \sum_{j=1}^n w_j(f(x))\partial_i d_{ij}(x) + \sum_{j,k=1}^n d_{ij}(x)\partial_k w_j(f(x))\partial_i f_k(x). \end{aligned} \tag{4.18}$$

Notemos que $\sum_{i=1}^n d_{ij}(x)\partial_i f_k(x) = J_f(x)$, cuando $j = k$, ya que es precísamente la definición del determinante por cofactores de la matriz jacobiana. Por otro lado, si $j \neq k$, $\sum_{i=1}^n d_{ij}(x)\partial_i f_k(x)$ resulta ser el determinante de la matriz que se obtiene de $J_f(x)$ al sustituir el j-ésimo renglón por el k-ésimo renglón. Como los renglones j y k tal matriz son iguales, entonces su determinante es cero. Por lo tanto,

$$\sum_{i=1}^n d_{ij}(x)\partial_i f_k(x) = \delta_{jk} J_f(x), \tag{4.19}$$

donde δ_{jk} es la delta de Kronecker. Así, por (4.18), por (4.19) y por la proposición

19 se tiene que

$$\begin{aligned}
divv(x) &= \sum_{i=1}^{n} \partial_i v_i(x) \\
&= \sum_{i=1}^{n} \sum_{j,k=1}^{n} d_{ij}(x)\partial_k w_j(f(x))\partial_i f_k(x) + \sum_{i=1}^{n}\sum_{j=1}^{n} w_j(f(x))\partial_i d_{ij}(x) \\
&= \sum_{k,j=1}^{n} \partial_k w_j(f(x))\delta_{jk} J_f(x) + \sum_{j=1}^{n}\sum_{i=1}^{n} w_j(f(x))\partial_i d_{ij}(x) \\
&= \Big[\sum_{k=1}^{n} \partial_k w_k(f(x))\Big] J_f(x) + \sum_{j=1}^{n} w_j(f(x)) \sum_{i=1}^{n} \partial_i d_{ij}(x) \\
&= \Big[\sum_{k=1}^{n} \partial_k w_k(f(x))\Big] J_f(x) \\
&= [divw(f(x))]J_f(x),
\end{aligned}$$

lo cual por (4.16) es precisamente la igualdad (4.17). ■

Como último paso resta probar que el grado de la definición 80 es el mismo para toda función de clase $\overline{C}^2(\Omega)$ suficientemente cercana a una función continua dada. Para este fin, usaremos un caso especial del teorema de la función implícita.

Teorema 82 Sea $h : \mathbb{R} \times \Omega \to \mathbb{R}^n$ continuamente diferenciable tal que $h(t_0, x_0) = 0$ y $J_{h(t_0,\cdot)}(x_0) \neq 0$ para algún $(t_0, x_0) \in \mathbb{R} \times \Omega$. Entonces existen un intervalo $(t_0 - r, t_0 + r)$, una bola $B_\delta(x_0) \subset \Omega$ y una función continua $x : (t_0 - r, t_0 + r) \to B_\delta(x_0)$ tal que $x(t_0) = x_0$ y $x(t)$ es la única solución en $B_\delta(x_0)$ de $h(t, x) = 0$.

Probaremos entonces la siguiente proposición.

Proposición 83 Sean $f \in \overline{C}^2(\Omega)$ y $y \notin f(\partial\Omega)$. Entonces, para $g \in \overline{C}^2(\Omega)$ existe un $\delta = \delta(f, y, g) > 0$ tal que

$$d(f + tg, \Omega, y) = d(f, \Omega, y) \tag{4.20}$$

para $|t| < \delta$.

Demostración Sea $g \in \overline{C}^2(\Omega)$. Supongamos primero que y es un valor regular de f. El caso en que $f^{-1}(y) = \emptyset$ es claro, ya que dada $x \in \overline{\Omega}$, puesto que $\lim_{t\to 0} f(x) + tg(x) = f(x) \neq y$, se tiene que $f(x) + tg(x) \neq y$ en $\overline{\Omega}$ para $|t|$ suficientemente pequeño, y en consecuencia ambos miembros de (4.20) son cero. Supongamos ahora que $f^{-1}(y) = \{x^1, ..., x^m\}$ y $J_f(x^i) \neq 0$ para $i = 1, ..., m$. Sean $f_t = f + tg$ y $h(t, x) = f_t(x) - y$. Tenemos que $h(0, x^i) = 0$ y $J_{h(0,\cdot)}(x^i) = J_f(x^i) \neq 0$. Por el teorema anterior, podemos encontrar un intervalo $(-r_i, r_i)$, bolas disjuntas $\overline{B}_{\varrho_i}(x^i)$ en Ω y funciones continuas $z^i : (-r_i, r_i) \to B_{\varrho_i}(x^i)$ tales que $z^i(0) = x^i$ y $z^i(t)$ es la única solución de $h(t, x)$ en $B_{\varrho_i}(x^i)$. Sean $r = \min\{r_1, ..., r_m\}$ y $\varrho = \min\{\varrho_1, ..., \varrho_m\}$. En consecuencia, si $t \in (-r, r)$ y definimos $V = \cup_{i=1}^m B_\varrho(x_i)$ se tiene que

$$\begin{aligned} f_t^{-1}(y) \cap V &= \{x \in \overline{\Omega} : f_t(x) = y\} \cap V \\ &= \{x \in \overline{\Omega} : f(x) + tg(x) = y\} \cap V \\ &= \{x \in \overline{\Omega} : h(t,x) = 0\} \cap V \\ &= \{z^1(t), ..., z^m(t)\}. \end{aligned}$$

Podemos además escoger ϱ de manera que

$$SgnJ_f(x) = SgnJ_f(x^i) \text{ en } \overline{B}_\varrho(x^i). \tag{4.21}$$

Como $||f(x) - y|| > \beta$ en $\overline{\Omega}\backslash V$ para algún $\beta > 0$, entonces, si $x \in \overline{\Omega}\backslash V$ y $|t| < \delta_0 = \min\{r, \beta||g||_\infty^{-1}\}$, al notar que

$$\begin{aligned} ||f_t - f||_\infty &= ||tg||_\infty \\ &\leq \beta||g||_\infty^{-1}||g||_\infty \\ &= \beta \end{aligned}$$

obtenemos que $||f(x) - y|| - ||f_t - f||_\infty > 0$, de donde

$$\begin{aligned} ||f_t(x) - y|| &= ||f_t(x) - f(x) + f(x) - y|| \\ &= ||(f(x) - y) - (f(x) - f_t(x))|| \\ &\geq ||f(x) - y|| - ||f(x) - f_t(x)|| \\ &> 0. \end{aligned}$$

Por lo tanto, si $|t| < \delta_0 = \min\{r, \beta||g||_\infty^{-1}\}$ obtenemos que $f_t^{-1}(y) \cap \overline{\Omega}\backslash V = \emptyset$ y en consecuencia que

$$f_t^{-1}(y) = \{z^1(t), ..., z^m(t)\}. \tag{4.22}$$

Finalmente, como $J_{f_t}(x)$ es continua con respecto a t y x, existe $\delta \leq \delta_0$ tal que

$$|J_{f_t}(x) - J_f(x)| < \min\{|J_f(z)| : z \in \overline{V}\} \tag{4.23}$$

para $|t| < \delta$ y $x \in \overline{V}$ (notemos que como el mínimo anterior se alcanza en algún punto de $\overline{V}$, de (4.21) se tiene que tal mínimo no es cero). Entonces, si $|t| < \delta$, de (4.23) tenemos que

$$SgnJ_{f_t}(z^i(t)) = SgnJ_f(z^i(t))$$

y de (4.21) resulta que

$$SgnJ_{f_t}(z^i(t)) = SgnJ_f(z^i(t)) = SgnJ_f(x^i),$$

de donde, usando la definición 78 y (4.22) obtenemos que $d(f_t, \Omega, y) = d(f, \Omega, y)$ para $|t| < \delta$.

Antes de continuar, notemos que por (4.22) y (4.21) se tiene que $y \notin f_t(S_{f_t})$, si $|t| < \delta_0$. Así, se ha obtenido que si y es un valor regular de f, entonces y es un valor regular de f_t si $|t|$ es suficientemente pequeño.

Por último, supongamos que y no es un valor regular de f. Escojamos entonces un valor regular de f, digamos $y_0 \in B_{\alpha/3}(y)$, donde $\alpha = d(y, f(\partial\Omega))$. Por el paso anterior podemos tomar $\delta_0 > 0$ tal que

$$d(f_t, \Omega, y_0) = d(f, \Omega, y_0) = d(f, \Omega, y) \tag{4.24}$$

para $|t| < \delta_0$ (recordemos que la última igualdad se da por el lema 81). Sea $\delta = \min\{\delta_0, \frac{1}{3}||g||_\infty^{-1}\alpha\}$. Notemos que si $x \in \partial\Omega$,

$$\begin{aligned} ||y_0 - f(x)|| &= ||y_0 - y + y - f(x)|| \\ &\geq ||y - f(x)|| - ||y - y_0|| \\ &> \alpha - \frac{\alpha}{3} \\ &\geq \frac{2\alpha}{3}. \end{aligned}$$

Por otro lado, si $|t| < \delta$,

$$\begin{aligned} ||f_t - f||_\infty &= ||tg||_\infty \\ &\leq \frac{1}{3}||g||_\infty^{-1}\alpha||g||_\infty \\ &= \frac{\alpha}{3}. \end{aligned}$$

Así, si $x \in \partial\Omega$ y $|t| < \delta$,

$$\begin{aligned} ||y_0 - f_t(x)|| &= ||y_0 - f(x) + f(x) - f_t(x)|| \\ &\geq ||y_0 - f(x)|| - ||f_t(x) - f(x)|| \\ &> \frac{\alpha}{3}. \end{aligned}$$

y entonces $0 < ||y_0 - y|| < \alpha/3 \leq d(y_0, f_t(\partial\Omega))$. Así, $y_0 \notin f_t(\partial\Omega)$, y como y_0 es un valor regular de f_t para $|t| < \delta$, entonces por la definición 80 y por (4.24) se tiene que

$$\begin{aligned} d(f_t, \Omega, y) &= d(f_t, \Omega, y_0) \\ &= d(f, \Omega, y) \end{aligned}$$

para $|t| < \delta$. $\blacksquare$

La proposición anterior nos dice entonces que el grado es una función constante en todos los mapeos de clase $\overline{C}^2(\Omega)$ suficientemente "cercanos" a una función continua dada. En efecto, sean $f \in C(\overline{\Omega})$, $y \notin f(\partial\Omega)$ y $\alpha = d(y, f(\partial\Omega))$. Consideremos dos funciones $g, \tilde{g} \in \overline{C}^2(\Omega)$ tales que $||g - f||_\infty < \alpha$ y $||\tilde{g} - f||_\infty < \alpha$. Sean $h(t, x) = g(x) + t(\tilde{g}(x) - g(x))$ y $\varphi(t) = d(h(t, \cdot), \Omega, y)$ para $t \in \mathbb{R}$. Sea t_0 un punto arbitrario de $[0, 1]$. Como $h(t, x) = h(t_0, x) + (t - t_0)(\tilde{g}(x) - g(x))$, la proposición anterior nos dice que $\varphi(t)$ es constante en una vecindad de t_0. Como $[0, 1]$ es compacto y conexo, entonces existen $t_1, ..., t_m \in [0, 1]$ y vecindades $V_1, ..., V_m$ de $t_1, ..., t_m$, respectivamente tales que $\{V_i\}_{i=1}^m$ es una cubierta de $[0, 1]$, $V_i \cap V_{i+1} \neq \emptyset$ para $i = 1, ..., m - 1$, y $\varphi(t)$ es constante en cada V_i. Así

φ es constante en $[0,1]$, y en consecuencia, $d(g,\Omega,y) = d(\tilde{g},\Omega,y)$. Así, podemos dar la siguiente definición.

Definición 84 (El grado en $\mathbb{R}^n$) Sean $f \in C(\overline{\Omega})$ y $y \in \mathbb{R}^n \backslash f(\partial\Omega)$. Definimos $d(f,\Omega,y) = d(g,\Omega,y)$, donde $g \in \overline{C}^2(\Omega)$ es cualquier función tal que $||g-f||_\infty < d(y, f(\partial\Omega))$ y $d(g,\Omega,y)$ está dada por la definición 80.

Ahora veremos que, efectivamente, el grado construido en esta sección cumple las propiedades **(d1)** a **(d3)** señaladas al inicio de esta sección.

Proposición 85 El grado en $\mathbb{R}^n$ cumple con las siguientes propiedades.

(d1) $d(I,\Omega,y) = 1$, para toda $y \in \Omega$, donde I es la función identidad en Ω.
(d2) $d(f,\Omega,y) = d(f,\Omega_1,y) + d(f,\Omega_2,y)$ para $\Omega_1,\Omega_2 \subset \Omega$ abiertos y disjuntos tales que $y \notin f(\overline{\Omega}\backslash(\Omega_1\cup\Omega_2))$.
(d3) $d(h(t,\cdot),\Omega,y(t))$ es independiente de $t \in [0,1]$, donde $h : [0,1]\times\overline{\Omega} \to \mathbb{R}^n$ y $y : [0,1] \to \mathbb{R}^n$ son continuas y $y(t) \notin h(t,\partial\Omega)$, para todo $t \in [0,1]$.

Demostración (d1) Por la definición 78 tenemos que

$$\begin{aligned} d(I,\Omega,y) &= Sgn(\det(I)) \\ &= 1. \end{aligned}$$

(d2) Sean (f,Ω,y) en el domino de d y tomemos Ω_1 y Ω_2 subconjuntos abiertos y disjuntos de Ω tales que $y \notin f(\overline{\Omega}\backslash(\Omega_1\cup\Omega_2))$. Tomemos $N > 2$ tal que

$$\tfrac{1}{N}\min\{d(y,f(\partial\Omega_1)), d(y,f(\partial\Omega_2))\} < d(y,f(\partial\Omega)).$$

Como $\overline{\Omega}\backslash(\Omega_1\cup\Omega_2)$ es compacto, entonces $f(\overline{\Omega}\backslash(\Omega_1\cup\Omega_2))$ es compacto. Puesto que $y \notin f(\overline{\Omega}\backslash(\Omega_1\cup\Omega_2))$, entonces $\lambda := d(y, f(\overline{\Omega}\backslash(\Omega_1\cup\Omega_2))) > 0$. Por la proposición 61 existe $g \in C^\infty(\mathbb{R}^n)$ tal que

$$||f(x)-g(x)|| < \tfrac{1}{N}\min\{d(y,f(\partial\Omega_1)), d(y,f(\partial\Omega_2)), \lambda\},$$

para $x \in \Omega$, así por la definición 84

$$d(f,\Omega,y) = d(g,\Omega,y).$$

Existe $M \in \mathbb{N}$ tal que

$$\tfrac{1}{M}\min d(y,g(\partial\Omega_1)), d(y,g(\partial\Omega_2)) < d(y,g(\partial\Omega)).$$

Por el corolario 66 existe y^1 valor regular de g tal que

$$||y-y^1|| < \tfrac{1}{M}\min\{d(y,g(\partial\Omega_1)), d(y,g(\partial\Omega_2)), \lambda\}.$$

Luego por la definición 80,

$$d(g,\Omega,y) = d(g,\Omega,y^1).$$

Sabemos que dada $x \in \Omega$ se tiene que $||f(x)-g(x)|| < \frac{\lambda}{2}$. Luego como $||y-y^1|| < \frac{\lambda}{2}$, resulta que

$$g^{-1}(y^1)\cap(\Omega\backslash(\Omega_1\cup\Omega_2)) = \emptyset. \tag{4.25}$$

Por la definición 78, teniendo en cuenta (4.25), resulta que

$$\begin{aligned}
d(g,\Omega,y^1) &= \sum_{x\in g^{-1}(y^1)} SgnJ_f(x)\\
&= \sum_{x\in g^{-1}(y^1)\cap\Omega_1} SgnJ_f(x) + \sum_{x\in g^{-1}(y^1)\cap\Omega_2} SgnJ_f(x)\\
&+ \sum_{x\in g^{-1}(y^1)\cap(\Omega\setminus(\Omega_1\cup\Omega_2))} SgnJ_f(x)\\
&= \sum_{x\in g^{-1}(y^1)\cap\Omega_1} SgnJ_f(x) + \sum_{x\in g^{-1}(y^1)\cap\Omega_2} SgnJ_f(x)\\
&= d(g,\Omega_1,y^1) + d(g,\Omega_2,y^1).
\end{aligned}$$

Si $x \in \Omega_1 \subset \Omega$, entonces

$$||f(x) - g(x)|| < \tfrac{d(y,f(\partial\Omega_1))}{N} < d(y, f(\partial\Omega_1)).$$

Luego, por la definición 84,

$$d(f,\Omega_1,y) = d(g,\Omega_1,y).$$

Por otra parte, como $||y - y^1|| < \frac{d(y,g(\partial\Omega_1))}{M} < d(y, g(\partial(\Omega_1)))$, entonces por la definición 80,

$$d(g,\Omega_1,y) = d(g,\Omega_1,y^1).$$

Así,

$$d(f,\Omega_1,y) = d(g,\Omega_1,y) = d(g,\Omega_1,y^1).$$

Análogamente obtenemos que

$$d(f,\Omega_2,y) = d(g,\Omega_2,y) = d(g,\Omega_2,y^1).$$

Así,

$$d(f,\Omega,y) = d(f,\Omega_1,y) + d(f,\Omega_2,y).$$

(d3) Probaremos primero lo siguiente.

(a) Sean $f : \overline{\Omega} \to \mathbb{R}^n$ y $y : [0,1] \to \mathbb{R}^n$ funciones continuas tales que $y(t) \notin f(\partial\Omega)$ para todo $t \in [0,1]$. Entonces $d(f,\Omega,y(t))$ es independiente de $t \in [0,1]$.
(b) Sea $h : [0,1] \times \overline{\Omega} \to \mathbb{R}^n$ una función continua tal que $y \notin h(t,\partial\Omega)$ para todo $t \in [0,1]$. Entonces $d(h(t,\cdot),\Omega,y)$ es independiente de $t \in [0,1]$.

Probaremos primero **(a)**. Dada $t \in [0,1]$, definamos

$$u(t) = d(y(t), f(\partial\Omega)).$$

Como y es continua, entonces u es continua. Puesto que $[0,1]$ es compacto, existe $p \in [0,1]$ tal que $u(p) = \min\{u(t) : t \in [0,1]\}$. Puesto que $y(p) \notin f(\partial\Omega)$, entonces $u(p) > 0$, y así,

$$\eta := \min\{d(y(t), f(\partial\Omega)) : t \in [0,1]\} = u(p) > 0.$$

Por la proposición 61, existe $g \in C^\infty(\mathbb{R}^n)$ tal que $||f(x) - g(x)|| < \eta$ en Ω. Luego, por la definición 84,

$$d(f, \Omega, y(t)) = d(g, \Omega, y(t)), t \in [0, 1].$$

Dada $t \in [0, 1]$, definamos

$$v(t) = d(y(t), g(\partial\Omega)).$$

Como en el caso de la función u, se tiene que

$$\gamma := \min\{d(y(t), g(\partial\Omega)) : t \in [0, 1]\} = \min\{v(t) : t \in [0, 1]\} > 0.$$

Notemos que como $[0, 1]$ es compacto y conexo, $y([0, 1])$ también lo es, de donde existen $m \in \mathbb{N}$ y $t_1, ..., t_m \in [0, 1]$ tales que

$$y([0, 1]) \subset \bigcup_{i=1}^{m} B(y(t_i), \gamma) \tag{4.26}$$

y

$$B(y(t_i), \gamma) \cap B(y(t_{i+1}), \gamma) \neq \emptyset,$$

$i = 1, ..., m - 1$. Por el corolario 66, existen valores regulares de g, digamos $y^{t_i} \in B(y(t_i), \gamma) \cap B(y(t_{i+1}), \gamma)$, para $i = 1, ..., m - 1$. Así, por la definición 80,

$$d(g, \Omega, y^{t_i}) = d(g, \Omega, y(t_{i+1})) = d(g, \Omega, y^{t_{i+1}}), \tag{4.27}$$

para $i = 1, ..., m - 1$. Por otra parte, de (4.26) resulta que dado $t \in [0, 1]$, existe $j \in \{1, 2, ..., m-1\}$ tal que $y(t) \in B(y(t_j), \gamma)$, y en consecuencia por la definición 80,

$$d(g, \Omega, y(t)) = d(g, \Omega, y^{t_j}). \tag{4.28}$$

De (4.27) y de (4.28) obtenemos que $d(g, \Omega, y(t))$ es independiente de $t \in [0, 1]$, y en consecuencia $d(f, \Omega, y(t))$ también.

Probemos ahora **(b)**. Como $[0, 1] \times \partial\Omega$ es compacto y h es continua, entonces $h([0, 1] \times \partial\Omega)$ es compacto. Puesto que $y \notin h(t, \partial\Omega)$ para todo $t \in [0, 1]$, entonces $y \notin h([0, 1] \times \partial\Omega)$. Así,

$$\eta := d(y, h([0, 1] \times \partial\Omega)) > 0.$$

Puesto que h es continua y su dominio es compacto, entonces h es uniformemente continua. Así, existe $\delta > 0$ tal que si (t, x), $(r, y) \in [0, 1] \times \partial\Omega$ y $||(t, x) - (r, y)|| < \delta$, entonces

$$||h(t, x) - h(r, y)|| < \tfrac{\eta}{2}.$$

Tomemos una partición $\{t_0, ..., t_m\}$ de $[0, 1]$ tal que $|t_i - t_{i+1}| < \delta$, $i = 1, ...m-1$. Por la proposición 61, existe $g_i \in C^\infty(\mathbb{R}^n)$ tal que $||h(t_i, x) - g_i(x)|| < \frac{\eta}{2}$ para $x \in \Omega$. Para cada $i \in 0, 1, 2, ..., m - 1$ tomemos un punto arbitrario $p^i \in [t_i, t_{i+1}]$. Como

$$\begin{aligned} ||h(p^i, x) - g_i(x)|| &\leq ||h(p^i, x) - h(t_i, x)|| + ||h(t_i, x) - g_i(x)|| \\ &\leq \frac{\eta}{2} + \frac{\eta}{2} \\ &= \eta \\ &\leq d(y, h(t, \partial\Omega)), \end{aligned}$$

y

$$\begin{aligned} ||h(p^i,x)-g_{i+1}(x)|| &\leq ||h(p^i,x)-h(t_{i+1},x)|| + ||h(t_{i+1},x)-g_{i+1}(x)|| \\ &\leq \frac{\eta}{2}+\frac{\eta}{2} \\ &= \eta \\ &\leq d(y,h(t,\partial\Omega)), \end{aligned}$$

de la definición 84 obtenemos que

$$d(h(p^i,\cdot),\Omega,y) = d(g_i,\Omega,y) \tag{4.29}$$

y

$$d(h(p^i,\cdot),\Omega,y) = d(g_{i+1},\Omega,y). \tag{4.30}$$

De (4.29) y de (4.30) resulta que

$$d(g_i,\Omega,y) = d(g_{i+1},\Omega,y), i=0,1,...,m-1. \tag{4.31}$$

De la arbitrariedad de $p^i \in [t_i,t_{i+1}]$ para $i=0,1,2,...,m-1$, de (4.29) y de (4.31) obtenemos que $d(h(t,\cdot),\Omega,y)$ es independiente de $t\in[0,1]$.

Finalmente, de **(a)** y de **(b)** obtenemos que para $t\in[0,1]$,

$$d(h(t,\cdot),\Omega,y(t)) = d(h(0,\cdot),\Omega,y(t)) = d(h(0,\cdot),\Omega,y(0)),$$

de lo cual obtenemos **(d3)**. ■

Antes de continuar, recordemos que si $f\in\overline{C}^1(\Omega)$ y $y\notin f(\partial\Omega)$ es un valor regular de f, entonces $d(f,\Omega,y)$ está dado por la fórmula (4.21). En consecuencia, en este caso, $d(f,\Omega,y)$ es la suma de las soluciones de la ecuación $y=f(x)$ con jacobiano positivo menos la suma de las soluciones de la ecuación $y=f(x)$ con jacobiano negativo. En particular si $J_f(x)$ tiene el mismo signo para toda $x\in f^{-1}(y)$, entonces $|d(f\Omega,y)|$ es el número de ceros de la ecuación $y=f(x)$. Esta es la razón por la cual, en ocaciones especiales, al grado topológico se le llama el número de ceros de f.

Previo a utilizar la teoría del grado para probar el teorema de Brouwer, veamos una aplicación de la teoría del grado en $\mathbb{R}^n$. Podemos comprender el siguiente teorema de la siguiente manera: si tomamos un erizo, no importa cómo lo peinemos, siempre habrá alguna púa ortogonal al erizo.

Teorema 86 (del erizo) Sean $\Omega\subset\mathbb{R}^n$ abierto y acotado con $0\in\Omega$ y $f:\partial\Omega\to\mathbb{R}^n\backslash\{0\}$ continua. Supongamos también que n es impar. Entonces existen $x\in\partial\Omega$ y $\lambda\neq 0$ tales que $f(x)=\lambda x$.

Demostración Por la proposición 8 podemos suponer que $f\in C(\overline{\Omega})$. Como n es impar, se sigue de la definición 84, más específicamente de la fórmula (4.13), que

$$d(-I,\Omega,0) = SgnJ_{-I}(0) = (-1)^n = -1.$$

Consideremos los casos, $d(f,\Omega,0)=-1$ y $d(f,\Omega,0)\neq -1$.

Supongamos que $d(f,\Omega,0)=-1$. Como $d(f,\Omega,0)\neq d(I,\Omega,0)$, **(d3)** implica que la homotopía $h(t,x)=(1-t)f(x)+tx$ debe tener un cero $(t_0,x_0)\in(0,1)\times\partial\Omega$ (pues $h(0,\cdot)$ y $h(1,\cdot)$ no son cero en $\partial\Omega$). Entonces, despejando f obtenemos que

$$f(x_0)=-t_0(1-t_0)^{-1}x_0,$$

que es lo que buscábamos.

Supongamos ahora que $d(f,\Omega,0)\neq -1$. Al realizar un proceso análogo al anterior con $h(t,x)=(1-t)f(x)-tx$, concluimos que existe $(t_0,x_0)\in(0,1)\times\partial\Omega$ tal que

$$f(x_0)=t_0(1-t_0)^{-1}x_0,$$

que es lo que queríamos demostrar.

4.2 Los teoremas de Brouwer y de Schauder

El teorema de Brouwer nos dice que si $D\subset\mathbb{R}^n$ es un conjunto no vacío, compacto y convexo y $f:D\to D$ es continua entonces f tiene un punto fijo. El teorema de Brouwer es uno de los pilares de la teoria de punto fijo y ha motivado el estudio de diversas áreas de las mátemáticas. Al igual que otros teoremas de punto fijo, el teorema de Brouwer juega un papel importante en las matemáticas y en otras áreas del conocimiento debido a sus aplicaciones.

En esta sección utilizaremos la teoría del grado en $\mathbb{R}^n$ para poder demostrar el teorema de Brouwer. También enunciaremos y probaremos el teorema de Schauder, el cual es una generalización del teorema de Brouwer. Los resultados de esta sección pueden ser encontrados en [1], salvo el teorema de Schauder, el cual se encuentra en [2].

Comencemos recordando las propiedades que definen al grado en $\mathbb{R}^n$.

Propiedades del grado en $\mathbb{R}^n$

Recordemos que en las secciones anteriores se probó la existencia y unicidad de una función

$$d:\{(f,\Omega,y):\Omega\subset\mathbb{R}^n \text{ abierto y acotado, } f:\overline{\Omega}\to\mathbb{R}^n \text{ continua,}\\ y\in\mathbb{R}^n\backslash f(\partial\Omega)\}\to\mathbb{Z},$$

llamada el grado en $\mathbb{R}^n$, que cumple las propiedades **(d1)**, **(d2)** y **(d3)** siguientes.

(d1) $d(I,\Omega,y)=1$, para toda $y\in\Omega$, donde I es la función identidad en Ω.
(d2) $d(f,\Omega,y)=d(f,\Omega_1,y)+d(f,\Omega_2,y)$ para $\Omega_1,\Omega_2\subset\Omega$ abiertos y disjuntos tales que $y\notin f(\overline{\Omega}\backslash(\Omega_1\cup\Omega_2))$.

(d3) $d(h(t,\cdot),\Omega,y(t))$ es independiente de $t\in[0,1]$, donde $h:[0,1]\times\overline{\Omega}\to\mathbb{R}^n$ y $y:[0,1]\to\mathbb{R}^n$ son continuas y $y(t)\notin h(t,\partial\Omega)$, para todo $t\in[0,1]$.

Las propiedades **(d1)**−**(d3)** son las propiedades básicas del grado d. En la proposición siguiente veremos propiedades del grado resultantes de estas.

Proposición 87 (Otras propiedades del grado en $\mathbb{R}^n$)

(d4) $d(f,\Omega,y)=0$ si $f^{-1}(y)=\emptyset$.
(d5) $d(\cdot,\Omega,y)$ es constante en $\{g\in C(\overline{\Omega}):||f-g||_\infty<r\}$, donde $r=\varrho(y,f(\partial\Omega))$.
(d5.1) $d(f,\Omega,\cdot)$ es constante en $B_r(y)\subset\mathbb{R}^n$ donde $r=\varrho(y,f(\partial\Omega))$. Más aún, $d(f,\Omega,\cdot)$ es constante en cada componente conexa de $\mathbb{R}^n\backslash f(\partial\Omega)$.
(d6) $d(f,\Omega,y)=d(g,\Omega,y)$ si $f\big|_{\partial\Omega}=g\big|_{\partial\Omega}$.
(d7) $d(f,\Omega,y)=d(f,\Omega_1,y)$ para todo $\Omega_1\subset\Omega$ abierto tal que $y\notin f(\overline{\Omega}\backslash\Omega_1)$.

Demostración (d4) La prueba de esta afirmación se encuentra en la demostración de la proposición 69.

(d5) Esto se probó en la demostración de la proposición 63.

(d5.1) Sea (f,Ω,y_1) en el dominio de d. Notemos que si $y_2\in B_r(y_1)$, donde $r=\varrho(y_1,f(\partial\Omega))$, usando **(d3)** con $y(t)=ty_1+(1-t)y_2$, se tiene que $d(f,\Omega,y_1)=d(f,\Omega,y_2)$. Para la segunda parte recordemos que Ω es abierto y acotado, por lo que $\overline{\Omega}$ es compacto. Además, como $\partial\Omega\subset\overline{\Omega}$ es cerrado, también es compacto y por lo tanto $f(\partial\Omega)$ es cerrado, de donde $\mathbb{R}^n\backslash f(\partial\Omega)$ es abierto. Así, las componentes conexas de $\mathbb{R}^n\backslash f(\partial\Omega)$ son abiertas, y por lo tanto arcoconexas. Luego, dada una componente conexa C de $\mathbb{R}^n\backslash f(\partial\Omega)$, y $y_1,y_2\in C$, existe una curva continua α en C tal que $\alpha(0)=y_1$ y $\alpha(1)=y_2$, de donde usando **(d3)** se tiene lo deseado.

(d6) Sean (f,Ω,y) y (g,Ω,y) en el dominio de d y supongamos que $f|_{\partial\Omega}=g|_{\partial\Omega}$. Consideremos la homotopía h dada por

$$h(t,x)=tf(x)+(1-t)g(x).$$

Como $y\notin f(\partial\Omega)=g(\partial\Omega)$, entonces $y\notin h(t,\partial\Omega))$ para todo $t\in[0,1]$, pues si $y=h(t,p)$ para algunos $t\in[0,1]$ y $p\in\partial\Omega$, entonces

$$y=h(t,p)=tf(p)+(1-t)g(p)=tf(p)+(1-t)f(p)=f(p)\in f(\partial\Omega),$$

lo cual no es posible. Por **(d3)** aplicada a h obtenemos lo afirmado.

(d7) Sean (f,Ω,y) y (f,Ω_1,y) en el dominio de d con $\Omega_1\subset\Omega$ abierto tal que $y\notin f(\overline{\Omega}\backslash\Omega_1)$ y definamos $\Omega_2=\Omega\backslash\overline{\Omega_1}$. Como $y\notin f(\partial\Omega)$ y $y\notin f(\partial\Omega_1)$, entonces $y\notin f(\overline{\Omega}\backslash(\Omega_1\cup\Omega_2))$. Por hipótesis $y\notin f(\overline{\Omega}\backslash\Omega_1)$, de donde $y\notin f(\Omega_2)$ y por lo tanto por **(d4)** $d(f,\Omega_2,y)=0$. Así, de **(d2)** se tiene que

$$\begin{aligned} d(f,\Omega,y) &= d(f,\Omega_1,y)+d(f,\Omega_2,y)\\ &= d(f,\Omega_1,y), \end{aligned}$$

que es lo que queríamos demostrar. ■

Finalmente estamos preparados para probar el teorema de Brouwer.

Teorema 88 (Brouwer) Sean $D \subset \mathbb{R}^n$ un subconjunto no vacío, compacto, convexo y $f : D \to D$ continua. Entonces f tiene un punto fijo. Lo anterior es cierto si D es homeomorfo a un subconjunto convexo y compacto de $\mathbb{R}^n$.

Demostración Comenzaremos suponiendo que $D = \overline{B}_r(0)$ para algún $r > 0$. Si f tiene algún punto fijo en ∂D habremos terminado, por lo que supongamos que no es así. Definamos $h : [0,1] \times D \to \mathbb{R}^n$ dada por $h(t,x) = x - tf(x)$. Notemos que h es una homotopía entre I e $I - f$, donde I es la función identidad en D. Así, si $(t,x) \in [0,1) \times \partial D$,

$$||h(t,x)|| \geq ||x|| - t||f(x)|| = r - t||f(x)|| \geq r(1-t) > 0.$$

Además, por hipótesis, $f(x) \neq x$ en ∂D. Por lo tanto, $0 \notin h([0,1] \times \partial D)$. Así, por **(d3)** se tiene que

$$d(I - f, B_r(0), 0) = d(I, B_r(0), 0),$$

y por **(d1)**,

$$d(I - f, B_r(0), 0) = d(I, B_r(0), 0) = 1,$$

lo cual por **(d4)** nos dice que existe $x \in B_r(0)$ tal que $(I - f)(x) = 0$, es decir, f tiene un punto fijo en $B_r(0)$.

Supongamos ahora que D es un subconjunto no vacío, compacto y convexo de $\mathbb{R}^n$. Por la proposición 8 existe una extensión continua $\widetilde{f} : \mathbb{R}^n \to \mathbb{R}^n$ de f dada por,

$$\widetilde{f}(x) = \begin{cases} f(x) & \text{si } x \in D \\ \left(\sum\limits_{i=1}^{\infty} 2^{-i}\varphi_i(x)\right)^{-1} \sum\limits_{i=1}^{\infty} 2^{-i}\varphi_i(x) f(a^i) & \text{si } x \notin D \end{cases},$$

donde $\{a^1, a^2, ...\}$ es un conjunto denso y numerable en D y

$$\varphi_i(x) = \max\{2 - \tfrac{x - a^i}{d(x,D)}, 0\} \text{ para } x \notin D.$$

Sea $m \in \mathbb{N}$ y definamos $\widetilde{f}_m : \mathbb{R}^n \to \mathbb{R}^n$ como

$$\widetilde{f}_m(x) = \begin{cases} f(x) & \text{si } x \in D \\ 0 & \text{si } x \notin D \text{ y } A(m,x) = 0 \\ A(m,x)^{-1} \sum\limits_{i=1}^{m} 2^{-i}\varphi_i(x) f(a^i) & \text{si } x \notin D \text{ y } A(m,x) \neq 0, \end{cases}$$

donde

$$A(m,x) = \sum_{i=1}^{m} 2^{-i}\varphi_i(x).$$

Sea $x \in D^c$. Puesto que $\{a^1, a^2, ...\}$ es denso en D, existe $i_0 \in \mathbb{N}$ tal que $\varphi_{i_0}(x) > 0$. En consecuencia, si $m > i_0$, se tiene que $\widetilde{f}_m(x)$ es una combinación convexa de los vectores $f(a^1), ..., f(a^m)$, por lo cual $\widetilde{f}_m(x) \in conv(D) = D$. Además, haciendo $m \to \infty$ se tiene que $\widetilde{f}_m \to \widetilde{f}$, y como D es cerrado y convexo, $\widetilde{f}(\mathbb{R}^n) \subset \overline{conv}(D) = D$. Sea $r > 0$ tal que $D \subset B_r$. Como $\widetilde{f}(B_r) \subset D$, la restricción $\widetilde{f}\Big|_{B_r} : B_r \to B_r$ tiene un punto fijo, el cual se encuentra en D. Por lo tanto, f tiene un punto fijo en D.

Por último supongamos que $D = h(D_0)$, con D_0 convexo y compacto y h homeomorfismo. Entonces la función $h^{-1}fh : D_0 \to D_0$ tiene un punto fijo x_0 por lo que acabamos de probar, y así, f tiene a $h(x_0)$ como punto fijo, que es lo que se quería demostrar. ■

Ahora veremos un corolario del teorema de Brouwer, el cual nos ayudará a demostrar el teorema de Schauder.

Corolario 89 Sean W un espacio normado real de dimensión n, $D \subset W$ convexo, compacto y no vacío y $f : D \to D$ continua. Entonces f tiene un punto fijo.

Demostración Notemos que W es isomorfo a $\mathbb{R}^n$, por lo que existe una transformación lineal $T : W \to \mathbb{R}^n$ biyectiva. Definamos $\widetilde{D} = T(D)$ y $\widetilde{f} = TfT^{-1}$. Así, $\widetilde{f} : \widetilde{D} \to \widetilde{D}$ es continua por ser una composición de funciones continuas y $\widetilde{D}$ es convexo y compacto, ya que es imagen de un conjunto convexo y compacto bajo una transformación lineal. Entonces, por el teorema de Brouwer $\widetilde{f}$ tiene un punto fijo en $\widetilde{D}$, digamos x_0. Así

$$x_0 = \widetilde{f}(x_0) = Tf(T^{-1}x_0),$$

luego

$$T^{-1}x_0 = f(T^{-1}x_0),$$

de donde $T^{-1}x_0 \in D$ es un punto fijo de f. ■

El teorema de Brouwer fue extendido a espacios de Banach por Schauder en 1930. A continuación enunciaremos dicha extensión.

Teorema 90 (Schauder) Sean X un subconjunto no vacío y convexo de un espacio de Banach, $Y \subset X$ compacto y no vacío y $f : X \to Y$ continua. Entonces f tiene un punto fijo.

Demostración Sea $\varepsilon > 0$. Consideremos la cubierta abierta de Y dada por $\{B_\varepsilon(y) : y \in Y\}$. Como Y es compacto existen $\{y_1, ..., y_n\}$ tales que $\{B_\varepsilon(y_i) : i = 1, ..., n\}$ es también cubierta de Y. Definamos

$$X_\varepsilon = conv\{y_1, ..., y_n\},$$

el cual es un conjunto convexo. Como $Y \subset X$ y X es convexo, entonces $X_\varepsilon \subset X$. Ahora mapearemos Y en X_ε mediante una función continua $P_\varepsilon : Y \subset X \to X_\varepsilon$ tal que $P_\varepsilon(y)$ aproxima a y, es decir, tal que

$$||P_\varepsilon(y) - y|| < \varepsilon \text{ para todo } y \in Y.$$

Para construir P_ε construiremos n funciones continuas θ_i tales que

$$\theta_i(y) \geq 0 \text{ y } \sum_{i=1}^{n} \theta_i(y) = 1 \text{ para todo } y \in Y.$$

Comencemos definiendo para $i = 1, ..., n$,

$$\varphi_i(y) = \begin{cases} 0 & \text{si } ||y_i - y|| \geq \varepsilon, \\ \varepsilon - ||y_i - y|| & \text{si } ||y_i - y|| < \varepsilon. \end{cases}$$

Notemos que cada φ_i es continua y por construcción de las y_i, dada $y \in Y$ existe i tal que $||y_i - y|| < \varepsilon$, de donde $\varphi_i(y) > 0$ para dicha $y \in Y$. Ahora definamos $s(y) = \sum_{i=1}^{n} \varphi_i(y)$ y notemos que por lo anterior $s(y) > 0$ para toda $y \in Y$. Luego definamos $\theta_i(y) = \frac{\varphi_i(y)}{s(y)}$, para $i = 1, ..., n$ y para todo $y \in Y$. Notemos que para todo $i = 1, ..., n$, $y \in Y$, $\theta_i(y) \geq 0$ y $\sum_{i=1}^{n} \theta_i(y) = 1$. Finalmente, definimos P_ε como

$$P_\varepsilon(y) = \theta_1(y)y_1 + \theta_2(y)y_2 + ... + \theta_n(y)y_n, \text{ para toda } y \in Y.$$

Así, $P_\varepsilon : Y \to X_\varepsilon$ es una función continua. Además, por construcción, si $||y - y_i|| \geq \varepsilon$, entonces $\theta_i(y) = 0$, de donde

$$\begin{aligned} ||P_\varepsilon(y) - y|| &= \left|\left| \sum_{i=1}^{n} \theta_y (y_i - y) \right|\right| \\ &< \left| \sum_{i=1}^{n} \theta_i \right| \varepsilon \\ &= \varepsilon. \end{aligned}$$

Ahora consideremos la función $f_\varepsilon : X_\varepsilon \to X_\varepsilon$ definida como $f_\varepsilon(x) = P_\varepsilon(f(x))$. Por el corolario 89, f_ε tiene un punto fijo $x_\varepsilon \in X_\varepsilon$. Consideremos una sucesión decreciente $(\varepsilon_m)_{m=1}^{\infty}$ con $\varepsilon_m > 0$ para todo $m \in \mathbb{N}$ y con $\lim_{m\to\infty} \varepsilon_m = 0$. Como Y es compacto y $f : X \to Y$, entonces (ε_m) tiene una subsucesión ε_{m_k} tal que la sucesión $(f(x_{\varepsilon_{m_k}}))$ converge, digamos, a $y^* \in Y$. Por simplicidad supongamos que (ε_m) es dicha subsucesión. Como

$$x_{\varepsilon_m} = f_{\varepsilon_m}(x_{\varepsilon_m}) = P_{\varepsilon_m}(f(x_{\varepsilon_m})),$$

si definimos $y_{\varepsilon_m} = f(x_{\varepsilon_m})$, entonces podemos expresar

$$x_{\varepsilon_m} = y_{\varepsilon_m} + (P_{\varepsilon_m}(y_{\varepsilon_m}) - y_{\varepsilon_m}).$$

Puesto que $||P_{\varepsilon_m}(y) - y|| < \varepsilon_m$ para todo $y \in Y$, obtenemos que $x_{\varepsilon_m} \to y^*$, y como f es continua, $y^* = \lim_{n\to\infty} f(x_{\varepsilon_n}) = f(y^*)$, que es lo que queríamos demostrar. $\blacksquare$

Ahora veremos una aplicación del teorema de Brouwer en álgebra lineal.

Teorema 91 (Perron-Frobenius) Sea $A = (a_{ij})$ una matriz de $n \times n$ tal que $a_{ij} \geq 0$ para todo i, j. Entonces existen $\lambda \geq 0$ y $x = (x_1, ..., x_n) \neq 0$ tales que $x_i \geq 0$ para cada i y $Ax = \lambda x$. En otras palabras, A tiene un eigenvector no negativo asociado a un eigenvalor no negativo.

Demostración Para probar este resultado sea

$$D = \{x \in \mathbb{R}^n; x_i \geq 0 \text{ para todo } i \text{ y } \sum_{i=1}^{n} x_i = 1\}.$$

Si $Ax = 0$ para algún $x \in D$ entonces ya acabamos con $\lambda = 0$. Si $Ax \neq 0$ para todo $x \in D$, entonces $\sum_{i=1}^{n}(Ax)_i > 0$, para todo $x \in D$. Definamos $f : D \to \mathbb{R}^n$ como

$$f(x) = \tfrac{Ax}{\sum_{i=1}^{n}(Ax)_i},$$

la cual es una función continua en D. Además $f(D) \subset D$, ya que $a_{ij} \geq 0$ para todo i, j, y por hipótesis Ax no se anula en D. Por el teorema de Brouwer, f tiene un punto fijo, es decir, un $x_0 \in D$ tal que $Ax_0 = \lambda x_0$, con $\lambda = \sum_{i=1}^{n}(Ax_0)_i$.$\blacksquare$

Para terminar esta sección veremos un ejemplo (véase [2]) del uso del teorema de Schauder en la solución de ecuaciones integrales. A diferencia del problema de valor inicial de Cauchy, el cual se resolvió mostrando que cierto operador es una contraccíon, en este ejemplo lo importante será mostrar que cierto operador continuo manda la bola unitaria B del espacio $C[0, 1]$ en un subconjunto compacto de B.

Ejemplo 92 Probaremos que la siguiente ecuación integral

$$x(t) = \int_0^1 e^{-st}\cos(7x(s))ds, \tag{4.32}$$

para $t \in [0, 1]$, tiene solución en el espacio $C[0, 1]$.

Sea $B = \{x \in C[0, 1] : ||x||_\infty \leq 1\}$ la bola unitaria de $C[0, 1]$. Si $x \in C[0, 1]$ y $t \in [0, 1]$ definimos

$$y(t) = \int_0^1 e^{-st}\cos(7x(s))ds,$$

entonces $y \in C[0, 1]$ y

$$\begin{aligned} |y(t)| &= \Big| \int_0^1 e^{-st}\cos(7x(s))ds \Big| \\ &\leq \int_0^1 e^{-st}ds \\ &\leq 1. \end{aligned}$$

Así, $y \in B$. Definamos $f : B \to B$ como

$$f(x)(t) = \int_0^1 e^{-st}\cos(7x(s))ds.$$

Probaremos que f es continua. Tomemos entonces (x_n) una sucesión en B convergente en $C[0, 1]$, digamos a $x \in B$. Como $x_n \to x$ uniformemente en $[0, 1]$, entonces $\cos(7x_n) \to \cos(7x)$ uniformemente en $[0, 1]$, luego, dado $\varepsilon > 0$ existe $N \in \mathbb{N}$ tal que si $n \geq N$, $||\cos(7x_n) - \cos(7x)||_\infty < \varepsilon$. Entonces, si $n \geq N$

y $t \in [0,1]$,

$$\begin{aligned}
|f(x_n)(t) - f(x)(t)| &= \Big| \int_0^1 e^{-st} \cos(7x_n(s))ds - \int_0^1 e^{-st} cos(7x(s))ds \Big| \\
&\leq \int_0^1 e^{-st} |\cos(7x_n(s)) - \cos(7x(s))| ds \\
&< \int_0^1 e^{-st} \varepsilon ds \\
&\leq \varepsilon.
\end{aligned}$$

Como la desigualdad anterior no depende de $t \in [0,1]$, entonces $||f(x_n) - f(x)||_\infty \to 0$, de donde f es continua. Para poder usar el teorema de Schauder resta probar que $f(B)$ es compacto. Para esto veremos que dada una sucesión (y_n) en $f(B)$, esta tiene una subsucesión convergente. Para ello utilizaremos el teorema de Arzelá-Ascoli. Primero notemos que la sucesión (y_n) es uniformemente acotada, ya que $|y_n(t)| \leq 1$ para todo $t \in [0,1]$. Para ver que la sucesíon (y_n) es equicontinua, sean $r, t, s \in [0,1]$ y sin pérdida de generalidad supongamos que $r < t$. Por el teorema de valor medio existe $c \in (r,t)$ tal que

$$\begin{aligned}
|e^{-sr} - e^{-st}| &= |se^{-sc}| \cdot |t - r| \\
&\leq s|t - r|.
\end{aligned}$$

Así,

$$\begin{aligned}
|y_n(t) - y_n(r)| &\leq \int_0^1 \Big| e^{-st} - e^{-sr} \Big| ds \\
&\leq \int_0^1 s|t - r| ds \\
&\leq \frac{1}{2}|t - r|.
\end{aligned}$$

Como el lado derecho de la última desigualdad tiende a 0 cuando $t \to r$ y es independiente de n, la sucesión (y_n) es equicontinua. Por lo tanto, el teorema de Arzelá-Ascoli dice que (y_n) tiene una subsucesión que converge uniformemente. Como (y_n) fue arbitraria, $f(B)$ es compacto. Por lo tanto, por el teorema de Schauder, f tiene un punto fijo, es decir, la ecuación (4.32) tiene una solución.

Bibliografía

[1] K. Deimling, *Nonlinear Functional Analysis*, Gesamthochschule Paderborn, Springer-Verlag, 1980.

[2] J. N. Franklin, *Methods of Mathematical Economics: Linear and Nonlinear Programming, Fixed-Point Theorems*, California Institute of Technology Pasadena, California, 2002.

[3] K. Goebel & W. A. Kirk, *Topics in metric fixed point theory*, Cambridge University Press, Cambridge, 1990.

[4] J. L. Hernández Barradas, *La propiedad de punto fijo en sumas directas de espacios de Banach*, tesis de licenciatura, Universidad de Guanajuato, 2015.

[5] W. A. Kirk & B. Sims, *Handbook of metric fixed point theory*, KLUWER ACADEMIC PUBLISHERS, 2001

[6] A. McLennan, *Advanced Fixed Point Theory for Economics*, Springer, 2014.

[7] R. E. Megginson, *An introduction to Banach Space Theory*, Springer, 1998.

[8] Mitio Nagumo, *A Theory of Degree of Mapping Based on Infinitesimal Analysis* American Journal of Mathematics Vol. 73, No. 3, 1951

[9] J. A. Munkres, *TOPOLOGY A First Course*, Massachusetts Institute of Technology.

[10] JOSEPH J. ROTMAN, *A FIRST COURSE IN ABSTRACT ALGEBRA*. University of Illinois at Urbana-Champaign, Third Edition, PRENTICE HALL, Upper Saddle River, New Jersey 07458, 2005

[11] W. Rudin, *Principios de análisis matemático*, 1980.