FACULTAD DE MATEMÁTICAS

DEPARTAMENTO: ECUACIONES DIFERENCIALES Y ANÁLISIS NUMÉRICO

Trabajo Fin de Grado:

Teoremas de punto fijo y aplicaciones.

Autor:

Mari Carmen Villarrubia Bizcocho

Dirigido por:

Inmaculada Gayte Delgado

2022 - 2023

Abstract

This paper deals with the study of Banach's, Brouwer's and Schauder's fixed point theorems and their applications. We will see how to apply these results to linear and nonlinear equations, to systems of equations, to initial value problems and to integral equations. In particular, we will look at some of the better known applications such as the Implicit Function Theorem, the Picard-Lindelof Theorem for the case of the Banach Point Theorem, the Fundamental of Algebra Theorem for the case of the Brouwer Fixed Point Theorem and Peano's Theorem for the case of the Schauder Fixed Point Theorem.

Resumen

Este trabajo trata sobre el estudio de los Teoremas de punto fijo de Banach, de Brouwer y de Schauder y sus aplicaciones. Veremos cómo aplicar estos resultados a ecuaciones lineales y no lineales, a sistemas de ecuaciones, a problemas de valor inicial y a ecuaciones integrales. En particular presentaremos unas aplicaciones clásicas, el Teorema de la Función Implícita, el Teorema de Picard-Lindelof para el caso del Teorema del punto de Banach, el Teorema Fundamental del Álgebra para el caso del Teorema del punto fijo de Brouwer y el Teorema de Peano para el caso del Teorema del punto fijo de Schauder.

Índice general

Introducción

Este trabajo estudia los principios fundamentales de la teoría de punto fijo. Para ello seguiremos los Capítulos 1 y 2 del libro de Zeidler [1].
Comenzaremos considerando una aplicación, $T : A \to \; B \subseteq A$, y una ecuación

$$y = T(y), \tag{1}$$

tal que toda solución x de la ecuación se llama punto fijo de T. Los teoremas de punto fijo garantizan la existencia de solución de una ecuación escrita como (1).
Cada ecuación de la forma $F(x) = 0$ donde F es una función real, o más generalmente una función en un espacio de Banach, puede escribirse como la ecuación de punto fijo:

$$x = x + F(x),$$

o de otras muchas formas.
Por lo tanto, los teoremas de punto fijo constituyen una herramienta importante para demostrar la existencia de soluciones de ecuaciones.
Los tres teoremas de punto fijo siguientes, de importancia fundamental, son los que presentamos en este trabajo:

(a) Teorema del punto fijo de Banach

(b) Teorema del punto fijo de Brouwer

(c) Teorema del punto fijo de Schauder

Los teoremas (a) y (b) son unas generalizaciones de dos conocidos métodos de demostración para el problema de valor inicial

$$\begin{cases} x'(t) = f(t, x(t)), \\ x(t_0) = y_0 \end{cases} \tag{2}$$

Si f es continua, (2) es equivalente a la ecuación integral

$$x(t) = y_0 + \int_{t_0}^{t} f(s, x(s))\, ds \tag{3}$$

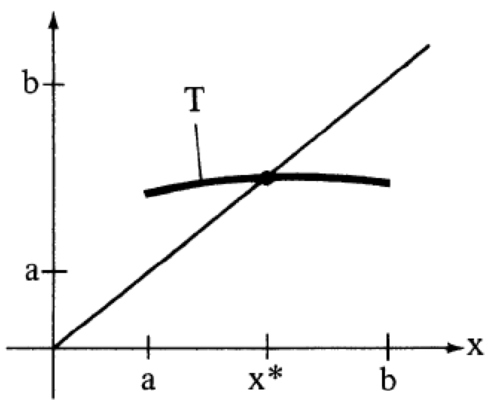
El teorema de Picard-Lindelöf garantiza la existencia y la unicidad de solución si $f(t,y)$ es continua y localmente Lipschitziana respecto a la segunda variable. Si es f solamente continua, entonces el teorema de Peano garantiza la existencia de solución de (2), pero no podemos decir nada sobre la unicidad de la solución. La construcción de soluciones aproximadas son una aplicación del Teorema del punto fijo de Banach.
Para demostrar estos dos teoremas de una manera funcional-analítica, escribimos (3) como una ecuación de operador no lineal,

$$x = Tx, \quad x \in M \subseteq X, \tag{4}$$

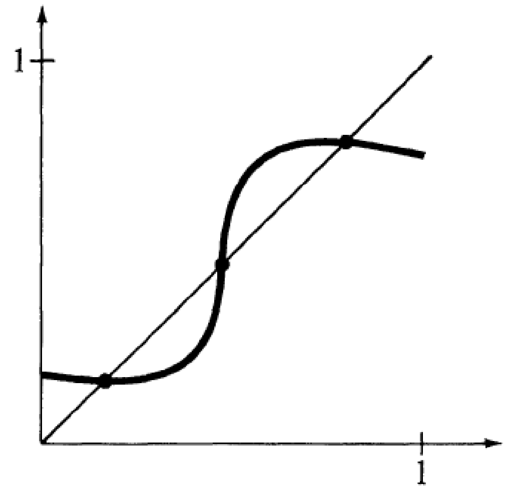
en un espacio de funciones X adecuado. Buscamos una solución de (4), es decir, un punto fijo de T en M. Como veremos en los Capítulos 1 y 2, se cumplen las siguientes relaciones:

Teorema del punto fijo de Banach $\Longrightarrow$ *Teorema del punto fijo de Picard-Lindelöf*
Teorema del punto fijo de Schauder $\Longrightarrow$ *Teorema de Peano*

El significado intuitivo de estos dos teoremas de punto fijo esta contenido en estas dos figuras.



(a) Teorema del punto fijo de Banach



(b) Teorema del punto fijo de Shauder

Para que los teoremas de punto fijo sean aplicables, es crucial una estructura que garantice que una determinada bola alrededor del origen se aplica en sí misma.

El Teorema del punto fijo de Banach puede aplicarse a aplicaciones T que sean diferenciables o, más generalmente, continuamente Lipschitzianas, con tal de que sean contractivas. Proporciona tanto la conclusión de unicidad como un método constructivo para determinar la solución, llamado método de aproximaciones sucesivas.

El Teorema del punto fijo de Schauder requiere la continuidad, y en los casos de espacios de dimension infinita, la compacidad de la aplicación T.
No proporciona ninguna conclusión de unicidad en general, ni un método para construir una solución.

Siempre que apliquemos teoremas de punto fijo a ecuaciones de la forma $x = Tx$, debemos tener cuidado de que no haya soluciones triviales, digamos $x = 0$, a mano. De lo contrario, la conclusión de existencia del teorema del punto fijo no aporta ninguna información nueva. En el caso de una solución trivial, necesitamos conclusiones sobre la existencia de varios puntos fijos.

En el Capítulo 1 estudiamos el Teorema del punto de Banach y sus aplicaciones.
Este Teorema dice que:
Supongamos que:

(i) *T es un operador $T : M \subseteq X \to M$, es decir, de M en sí mismo;*

(ii) *M es un conjunto cerrado no vacío en un espacio métrico completo (X, d);*

(iii) *T es k-contractiva, es decir,*

$$d(Tx, Ty) \leqslant kd(x, y)$$

para todo $x, y \in M$ y para k fijo, $0 \leqslant k < 1$.

Entonces se tiene:

(a) *Existencia y unicidad: La ecuación $x = Tx$ con $x \in M$, tiene exactamente una solución, es decir, T tiene exactamente un punto fijo en M;*

(b) *Convergencia de la iteración: La sucesión $\{x_n\}$ de aproximaciones sucesivas dada por $x_{n+1} = Tx_n$, con $x_0 \in M$ y $n = 0, 1, 2, \ldots$ converge a la solución, x, para una elección arbitraria del punto inicial x_0 en M;*

(c) *Estimaciones del error: Para todo $n = 0, 1, 2, \ldots$ tenemos la estimación del error a priori*

$$d(x_n, x) \leqslant k^n (1-k)^{-1} d(x_0, x_1),$$

y la estimación del error a posteriori

$$d(x_{n+1}, x) \leqslant k(1-k)^{-1} d(x_n, x_{n+1});$$

(d) *Tasa de convergencia: Para todo* $n = 0, 1, 2, \ldots$ *tenemos que*

$$d(x_{n+1}, x) \leqslant k d(x_n, x),$$

y, haciendo el proceso n-veces,

$$d(x_n, x) \leqslant k^n d(x_0, x).$$

Este teorema tiene un gran importancia ya que estudia la existencia y la unicidad de la solución de la ecuación, la estabilidad de la solución, la existencia de un método iterativo, las estimaciones del error y el orden de convergencia. Además, tiene importantes aplicaciones dentro de las matemáticas, nosotros vamos a estudiar en este trabajo cómo se puede aplicar el Teorema del punto fijo de Banach a ecuaciones lineales y no lineales, a problemas de valor inicial, en este caso estudiaremos en profundidad la aplicación el Teorema de Picard-Lidenof y sus corolarios y al Teorema de la Función Implícita. Veremos también cómo aplicarlo a ecuaciones con operadores lineales.

En el Capítulo 2 estudiamos los Teoremas de punto fijo de Brouwer y de Schauder.
En primer lugar, el Teorema del punto fijo de Brouwer nos dice que:
Supongamos que M *es un subconjunto no vacío, convexo y compacto de* $\mathbb{R}^N$, *donde* $N \geqslant 1$, *y que* $f : M \to M$ *es una aplicación continua. Entonces* f *tiene un punto fijo.*
También vamos a dar unos resultados equivalentes de este teorema. Las aplicaciones que vamos a ver serán la existencia de solución para sistemas de ecuaciones, y el Teorema Fundamental del Álgebra.

En segundo lugar, generalizamos el Teorema del punto fijo de Brouwer para espacios de Banach de dimension infinita. Esto se consigue con el Teorema del punto fijo de Schauder:
Sea M *un subconjunto no vacío, cerrado, acotado y convexo de un B-espacio* X, *y supongamos que* $T : M \to M$ *es un operador compacto. Entonces* T *tiene un punto fijo.*
Este teorema lo aplicaremos a problemas de valor inicial, y obtendremos el Teorema de Peano, a existencia de solución de ecuaciones integrales y sistemas de ecuaciones integrales, a ecuaciones diferenciales y sistemas de inecuaciones.

Capítulo 1

Teorema del punto fijo de Banach y métodos iterativos

Este capítulo sirve de introducción a los métodos iterativos de gran importancia para la matemática numérica moderna. Hemos seguido la referencia [1]. El teorema del punto fijo de Banach es el principal instrumento teórico para crear una visión unificada de la abundancia de métodos prácticos de iteración. La importancia esencial de este teorema de punto fijo se puede resumir como sigue:

(i) La convergencia de un método iterativo viene determinada por la k-contractividad de los operadores;

(ii) La calidad de la estimación del error depende del tamaño de la constante $k < 1$ de contractividad;

(iii) Para ecuaciones con operadores lineales k-contractividad significa que la norma del operador, más precisamente, el radio espectral es menor que 1.

Nuestra aplicación inmediata del teorema del punto fijo de Banach en este capítulo y en el siguiente será a sistemas de ecuaciones lineales y no lineales, a ecuaciones integrales lineales y no lineales, y a ecuaciones diferenciales ordinarias.
Un ejemplo de la importancia de los métodos iterativos la tenemos en los métodos en diferencias, que constituyen un esquema de aplicación universal, y en la práctica, muy significativo para determinar soluciones aproximadas de ecuaciones diferenciales ordinarias y parciales. En la teoría de reactores nucleares, por ejemplo , surgen problemas de valor límite para ecuaciones diferenciales elípticas que conducen a sistemas de ecuaciones con 10^5 o más incógnitas. Estos grandes sistemas sólo se pueden resolver de forma iterativa.

1.1. Teorema del punto fijo de Banach

Examinamos las condiciones en las que la ecuación no lineal

$$x = Tx, \quad x \in M, \tag{8}$$

puede resolverse mediante aproximaciones sucesivas

$$x_{n+1} = Tx_n, \quad x_0 \in M, \quad n = 0, 1, 2, \ldots. \tag{9}$$

Recordamos las definiciones siguientes:

Definición 1.1 Sea X un espacio métrico. Se dice que una sucesión $\{x_n\}$ es *de Cauchy* si para todo $\varepsilon > 0$ existe un $n_0 \in \mathbb{N}$ tal que para todos los números naturales $n, m \geq n_0$ se verifica que $d(x_n, x_m) < \varepsilon$.

Definición 1.2 Un espacio métrico X es *completo* si toda sucesión de Cauchy es convergente. Es decir, las sucesiones de Cauchy son siempre convergentes.

Definición 1.3 Un *espacio de Banach* es un espacio vectorial normado y completo respecto a la métrica derivada de su norma. Lo denotaremos por B-espacio.

Recordadas estas definiciones podemos enunciar el siguiente teorema fundamental:

Teorema 1.4 (Teorema del punto fijo de Banach (1922)). *Supongamos que:*

(i) *T es un operador $T : M \subseteq X \to M$,es decir, de M en sí mismo;*

(ii) *M es un conjunto cerrado no vacío en un espacio métrico completo (X, d);*

(iii) *T es k-contractiva, es decir,*

$$d(Tx, Ty) \leqslant kd(x, y) \tag{10}$$

para todo $x, y \in M$ y para k fijo, $0 \leqslant k < 1$.

Entonces se tiene:

(a) Existencia y unicidad*: La ecuación (8) tiene exactamente una solución, es decir, T tiene exactamente un punto fijo en M;*

(b) Convergencia de la iteración*: La sucesión $\{x_n\}$ de aproximaciones sucesivas dada por (9) converge a la solución, x, para una elección arbitraria del punto inicial x_0 en M;*

(c) Estimaciones del error*: Para todo $n = 0, 1, 2, \ldots$ tenemos la estimación del error a priori*

$$d(x_n, x) \leqslant k^n (1-k)^{-1} d(x_0, x_1), \tag{11}$$

y la estimación del error a posteriori

$$d(x_{n+1},x) \leqslant k(1-k)^{-1}d(x_n,x_{n+1}); \tag{12}$$

(d) Tasa de convergencia*: Para todo $n = 0,1,2,...$ tenemos que*

$$d(x_{n+1},x) \leqslant kd(x_n,x). \tag{13}$$

Y también obtendríamos esta desigualdad haciendo el proceso n-veces:

$$d(x_n,x) \leqslant k^n d(x_0,x). \tag{14}$$

En las demostraciones haremos uso de la siguiente definición:

Definición 1.5 Un operador $T: M \subseteq X \to X$ sobre un espacio métrico (X,d) se llama *k-contractivo* si (10) se cumple para todo $x,y \in M$ con k fijo, $0 \leqslant k < 1$. Si esto se cumple para $k = 1$, T se llama *no expansivo* ; y si esto se cumple para k fijo, $0 \leqslant k < \infty$, T se llama *continuamente Lipschitziano.*
Si $d(Tx,Ty) < d(x,y)$ para todo $x,y \in M$ con $x \neq y$, T se llama *contractivo.*

Para T, obviamente tenemos las siguientes implicaciones:

k-contractiva $\Rightarrow$ *contractiva* $\Rightarrow$ *no expansiva* $\Rightarrow$ *continuamente Lipschitziana.*

Todo B-espacio $(X,\|\cdot\|)$ es también un espacio métrico completo (X,d) bajo la métrica $d(x,y) = \|x-y\|$. En un B-espacio, (10) se convierte por tanto en $\|Tx-Ty\| \leqslant k\|x-y\|$. Para un operador lineal continuo $T: X \to X$ en un B-espacio X, se tiene que:

$$\|Tx-Ty\| \leqslant \|T\| \cdot \|x-y\|, \quad \text{para todo } x,y \in X \tag{15}$$

Por tanto, T es continuamente Lipschitziana. Si $\|T\| \leqslant 1$, T es no expansivo, y si $\|T\| < 1$, T es k-contractivo con $k = \|T\|$.

Para ver que una función real definida en un intervalo cerrado $[a,b]$ de $\mathbb{R}$ es contractiva, tenemos que comprobar que tiene la derivada acotada por un valor menor que 1 en todo punto del intervalo, es decir, toda función derivable $f:[a,b] \to \mathbb{R}$ tal que $|f'(x)| \leq k < 1$ para todo $x \in [a,b]$ es contractiva con constante de contractividad k.

Demostración del Teorema 1.4:

(I) Veamos que $\{x_n\}$ es una sucesión de Cauchy. Usando (10) se tiene que :

$$\begin{aligned} d(x_n, x_{n+1}) &= d(Tx_{n-1}, Tx_n) \leqslant \\ &\leqslant kd(x_{n-1}, x_n) \leqslant \\ &\leqslant k^2 d(x_{n-2}, x_{n-1}) \leqslant \ldots \\ \ldots &\leqslant k^n d(x_0, x_1), \text{ usando reiteradamente la desigualdad.} \end{aligned}$$

Y aplicando la desigualdad triangular reiteradamente y la fórmula de la suma de una serie geométrica se tiene

$$\begin{aligned} d(x_n, x_{n+m}) &\leqslant d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \ldots + d(x_{n+m-1}, x_{n+m}) \leqslant \\ &\leqslant (k^n + k^{n+1} + \ldots + k^{n+m-1}) d(x_0, x_1) = \\ &= \sum_{i=n}^{m-1} k^i d(x_0, x_1) = \frac{k^n - k^{n+m}}{1-k} d(x_0, x_1) \leqslant \\ &\leqslant k^n (1-k)^{-1} d(x_0, x_1), \text{ (tomando } m \to \infty) \end{aligned}$$

Obtenemos la estimación:

$$d(x_n, x_{n+m}) \leqslant \frac{k^n}{1-k} d(x_0, x_1) \tag{16}$$

El miembro de la derecha en la desigualdad (16) tiende a 0 cuando $n \to \infty$, luego la sucesión $\{x_n\}$ es de Cauchy.
Como X es completo, la sucesión $\{x_n\} \subset M$ converge a un elemento $x \in M$, por ser M cerrado.
Si tomamos límite cuando $m \to \infty$ en la desigualdad (16) tenemos la estimación de error (11).

(II) La estimación del error (12) se obtiene aplicando (10) y dejando $m \to \infty$ en

$$\begin{aligned} d(x_{n+1}, x_{n+m+1}) &\leqslant d(x_{n+1}, x_{n+2}) + \ldots + d(x_{n+m}, x_{n+m+1}) < \\ &< (k + k^2 + \ldots + k^m) d(x_n, x_{n+1}) \leqslant \\ &\leqslant k(1-k)^{-1} d(x_n, x_{n+1}). \end{aligned}$$

(III) Veamos que el punto x es una solución de (8). Como T es continua por (10), y puesto que $T(M) \subseteq M$ y $x_0 \in M$, tenemos que $x_n \in M$ también, para todo n. Como M es cerrado y $x_n \to x$ cuando $n \to \infty$, obtenemos que $x \in M$. La ecuación (9) implica que $Tx = x$ para $n \to \infty$.

(IV) La ecuación (13) se deduce de $d(x_{n+1}, x) = d(Tx_n, Tx) \leqslant kd(x_n, x)$. Aplicando reiteradamente esta desigualdad se tiene (14).

(V) Unicidad de la solución. Supongamos que $x = Tx$ e $y = Ty$; entonces $d(x,y) = d(Tx,Ty) \leqslant kd(x,y) < d(x,y)$, lo que obliga a $d(x,y) = 0$, es decir, $x = y$.

□

Observación 1.6 Podemos comprobar fácilmente que todas las hipótesis del Teorema 1.4 son esenciales.

1.2. Dependencia continua respecto de un parámetro

En muchas aplicaciones, T depende de un parámetro adicional p. Entonces, (8) se sustituye por la ecuación

$$x_p = T_p x_p, \quad x_p \in M, \tag{17}$$

donde $p \in P$.

El siguiente resultado es una consecuencia del Teorema 1.4.

Proposición 1.7 *Sean las siguientes condiciones:*

(i) P es un espacio métrico, llamado espacio paramétrico.

(ii) Para cada p, el operador T_p satisface las hipótesis del Teorema 1.4, pero con k en (10) independiente de p.

(iii) Para un $p_0 \in P$ fijo, y para todo $x \in M$, $\lim_{p\to p_0}(T_p x) = T_{p_0}x$.

Entonces, para cada $p \in P$, (17) tiene exactamente una solución $x_p \in M$, y $\lim_{p\to p_0}(x_p) = x_{p_0}$.

Demostración:

Sea x_p una solución de (17) dada por el Teorema 1.4; entonces, por la desigualdad triangular

$$\begin{aligned} d(x_p, x_{p_0}) &= d(T_p x_p, T_{p_0} x_{p_0}) \leqslant \\ &\leqslant d(T_p x_p, T_p x_{p_0}) + d(T_p x_{p_0}, T_{p_0} x_{p_0}) \leqslant \\ &\leqslant kd(x_p, x_{p_0}) + d(T_p x_{p_0}, T_{p_0} x_{p_0}), \end{aligned}$$

y por lo tanto, despejando $d(x_p, x_{p_0})$, se tiene que:

$$d(x_p, x_{p_0}) \leqslant (1-k)^{-1} d(T_p x_{p_0}, T_{p_0} x_{p_0}) \to 0 \quad \textit{cuando} \quad p \to p_0,$$

por (iii). □

1.3. La importancia del Teorema del punto fijo de Banach

La gran importancia del Teorema 1.4 y la Proposición 1.7 deriva del hecho de que estos teoremas contienen conjuntamente ocho elementos de importancia fundamental para el tratamiento teórico-práctico de las ecuaciones matemáticas, estos son:

(A) Existencia de una solución,

(B) Unicidad de la solución,

(C) Estabilidad de la solución bajo pequeñas perturbaciones de la ecuación (visto en la Proposición 1.7),

(D) Existencia de un método de aproximaciones convergentes,

(E) Estimaciones de error a priori,

(F) Estimaciones de error a posteriori,

(G) Estimaciones sobre el orden de convergencia,

(H) Estabilidad del método de aproximación.

Con (H) queremos decir que un cambio en el elemento inicial x_0 no cambia el valor límite de la iteración, mientras que x_0 pertenezca a M. No todo punto fijo tiene la propiedad de que un proceso iterativo iniciado en sus proximidades converja a este punto fijo. Más bien, algunos puntos fijos atraen y otros repelen. Para las matemáticas numéricas, la importancia de la estabilidad para un método iterativo reside en el hecho de que el elemento inicial sólo puede introducirse en el ordenador con una precisión limitada. Por tanto, los métodos iterativos que conducen a valores límites diferentes cuando se produce un ligero cambio en el valor inicial son numéricamente inestables y, por tanto, inútiles.

Las estimaciones del error se dividen en estimación a priori y a posteriori. Una estimación *a priori* proporciona una estimación del error antes de realizar cualquier cálculo. Una estimación *a posteriori* proporciona una estimación del error mientras se está realizando el cálculo. La estimación del error a priori $d(x_n, x) \leqslant k^n(1-k)^{-1}d(x_0, Tx_0)$ en el Teorema 1.4 permite utilizar el conocimiento del valor inicial para determinar el numero de pasos de iteración necesarios para alcanzar el nivel de precisión deseado. La estimación del error a posteriori $d(x_{n+1}, x) \leqslant k(1-k)^{-1}d(x_n, x_{n+1})$ en el Teorema 1.4 permite utilizar los valores calculados x_n y x_{n+1} para determinar la precisión de la aproximación x_{n+1}. La experiencia enseña que las estimaciones del error a posteriori son más precisas que las estimaciones del error a priori.

Una ventaja importante de los métodos iterativos es que, incluso en ausencia de estimaciones teóricas del error, se pueden extraer conclusiones en principio sobre la precisión de los valores calculados aprovechando la observación práctica de que un decimal se ha determinado con precisión una vez que ya no cambia en el transcurso del cálculo posterior.

En cuanto a la estabilidad de los métodos iterativos, otra ventaja es que es una prueba de errores: un valor calculado incorrectamente puede aumentar el número de iteraciones necesarias, pero no afecta a la convergencia del método. Supongamos que el ordenador ha calculado incorrectamente x_{n_0}. Entonces consideramos x_{n_0} como un nuevo valor inicial. Dado que el método converge para cada valor inicial en M, por el Teorema 1.4, el hecho de que x_{n_0} sea incorrecto no tendrá ninguna influencia, siempre y cuando x_{n_0} se encuentre en M. Esto supone, por supuesto, que los valores adicionales x_n, con $n > n_0$, se calculan correctamente.

El Teorema 1.4 estima la tasa de convergencia mediante

$$d(x_{n+1}, x) \leqslant kd(x_n, x)^s, \quad k < 1 \tag{18}$$

donde $s = 1$. Esto significa que, en el caso menos favorable, la precisión aumenta sólo en un factor de k con cada iteración (convergencia lineal). En la práctica, así como en ciertas cuestiones teóricas importantes relacionadas con el Teorema de la Función Implícita, uno está interesado en tasas de convergencias más rápidas, con $s > 1$. En el Ejemplo 1.10 describimos un método para acelerar la tasa de convergencia y lo aplicamos al método de Newton.

Al construir un método iterativo, a menudo resulta ventajoso elegir entre formulaciones equivalentes del problema. Lo ilustraremos con varios ejemplos importantes; para empezar, consideremos la ecuación

$$g(x) = 0, \quad g \text{ una función real.} \tag{19}$$

Al transformar (19) en una ecuación equivalente de punto fijo de la forma

$$x = Tx,$$

tenemos al menos las siguientes posibilidades:

$$Tx = x - g(x) \quad \text{(versión más sencilla)}, \tag{20}$$

$$Tx = x - \omega g(x) \quad \text{(relajación lineal)}, \tag{21}$$

$$Tx = x - \omega F(g(x)) \quad \text{(relajación no lineal)}, \tag{22}$$

$$Tx = x - g(x)/g'(x) \quad \text{(Método de Newton)}, \tag{23}$$

$$Tx = h^{-1}(g(x) - f(x)), \quad \text{donde } g(x) = h(x) + f(x) \quad \text{(método de división)}. \tag{24}$$

Aquí ω denota un parámetro real distinto de cero, mientras que F, f, y h son funciones adecuadas con h^{-1} denota la función inversa. Los métodos iterativos correspondientes derivan de $x_{n+1} = Tx_n$.

Por ejemplo, (21) conduce a

$$x_{n+1} = x_n - \omega g(x_n), \quad n = 0, 1, 2\ldots . \tag{25}$$

Para mejorar la convergencia, con frecuencia se altera el parámetro de alguna manera adecuada en cada iteración. Entonces

$$x_{n+1} = x_n - \omega_n g(x_n), \quad n = 0, 1, 2\ldots . \tag{26}$$

Veamos una aplicación sencilla del Teorema 1.4. Consideramos la ecuación no lineal

$$x = Tx, \quad x \in [a,b], \tag{27}$$

junto con el método iterativo

$$x_{n+1} = T(x_n), \quad x_0 \in [a,b], \quad n = 0, 1, 2, \ldots . \tag{28}$$

Las soluciones x^* de (27) son las intersecciones x^* de la gráfica de T con la recta $y = x$ como representamos en la Figura 1.1.

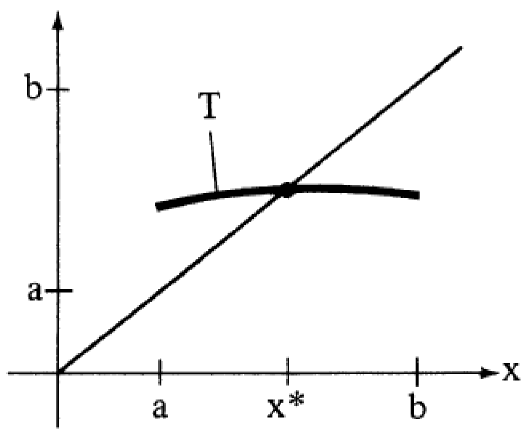


Figura 1.1: Intersección

Proposición 1.8 (Solución de (27)). *Supongamos que:*

(i) $T : [a,b] \to [a,b]$ *es una función real, donde* $-\infty < a < b < \infty$, *y*

(ii) $|T(x) - T(y)| \leqslant k|x-y|$ *para todo* $x, y \in [a,b]$ *y* $k \in [0,1)$.

Entonces (27) tiene exactamente una solución x, *y para* $n = 0,1,2,\ldots$ *es cierto que*

$$|x_n - x| \leqslant k^n (1-k)^{-1} |x_1 - x_0| \quad \text{(estimación del error a priori);} \tag{29}$$

$$|x_{n+1} - x| \leqslant k(1-k)^{-1} |x_{n+1} - x_n| \quad \text{(estimación del error a posteriori);} \tag{30}$$

$$|x_{n+1} - x| \leqslant k|x_n - x| \quad \text{(convergencia lineal).} \tag{31}$$

Por lo tanto $x_n \to x$ *cuando* $n \to \infty$ *para un* $x_0 \in [a,b]$

Observación 1.9 La condición (ii) significa geométricamente que la pendiente de la secante en la Figura 1.1 es, en valor absoluto, menor o igual que k, para $0 \leqslant k < 1$.
En particular, (ii) se cumple siempre que T sea continua en $[a,b]$ y diferenciable en (a,b), con $|T'(x)| \leqslant k < 1$ para todo $x \in (a,b)$, pues en ese caso, el Teorema del valor medio implica que:

$$|T(x) - T(y)| = |T'(\xi)(x-y)| \leqslant k|x-y|, \quad \text{para todo } x, y \in [a,b].$$

Veamos otra aplicación:
Queremos tratar ahora cómo diseñar un método iterativo donde la convergencia sea más rápida que meramente lineal, es decir, el problema de acelerar la convergencia del método. Sea $x \in (a,b)$ una solución de la ecuación real

$$x = F(x),$$

y supongamos la sucesión $\{x_n\}$, donde

$$x_{n+1} = F(x_n),$$

y $x_n \in (a,b)$ para todo n, converge a x cuando $n \to \infty$. Supongamos además que F es m-veces diferenciable en $[a,b]$ con

$$F'(x) = F''(x) = \cdots = F^{(m-1)}(x) = 0.$$

En este caso, el Teorema de Taylor dice que

$$|F(x_n) - F(x)| = |F^{(m)}(\xi_n)||x_n - x|^m/m!, \quad \xi^n \in (a,b).$$

Como $x_{n+1} = F(x)$ y $x = F(x)$, tenemos

$$|x_{n+1} - x| \leqslant \sup_{a<\xi<b} |F^{(m)}(\xi)||x_n - x|^m/m!. \tag{32}$$

Si el supremo en (32) es finito, obtenemos convergencia de orden m, frente a la convergencia ($m = 1$) de (31).

Ejemplo 1.10 El *truco del método de Newton* consiste en reescribir la ecuación $f(x) = 0$ en la forma equivalente

$$x = F(x), \quad \text{donde} \quad F(x) = x - \frac{f(x)}{f'(x)}. \tag{33}$$

Entonces el método iterativo se convierte en

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}. \tag{34}$$

Supongamos que $f'(x_n) \neq 0$ para todo n. Entonces, $F'(x) = f(x)f''(x)/(f'(x))^2$, de modo que si x es una solución de $f(x) = 0$ con $f'(x) \neq 0$, entonces $F'(x) = 0$. Así tenemos un método con $m = 2$ en (32), es decir, tenemos convergencia cuadrática.

La interpretación geométrica del método de Newton se muestra en la Figura 1.2(a). Para encontrar un cero, x, se toma el valor inicial, x_0, y se determina el valor funcional correspondiente, $f(x_0)$. El siguiente valor iterativo , x_1, es la intersección de la recta tangente en $(x_0, f(x_0))$ y el eje X, y así sucesivamente.
Es típico del método de Newton que converja muy rápidamente si el valor inicial x_0 se encuentra ya en las proximidades del cero. Pero la Figura 1.2(b) muestra una situación en la que el método oscila permanentemente y nunca converge al cero. La razón es la desafortunada elección del valor inicial.

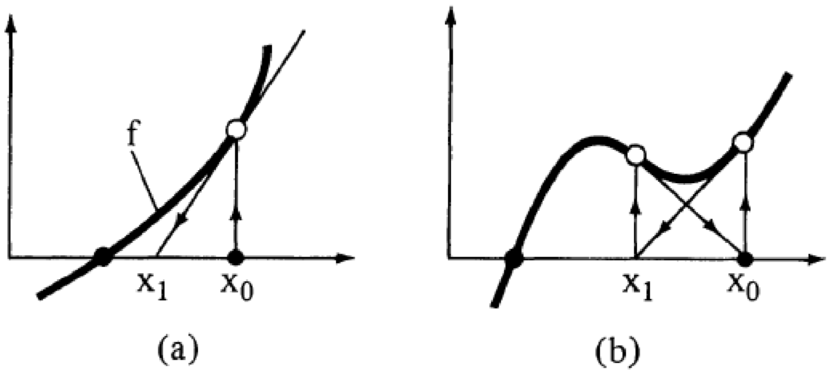


Figura 1.2: (a) Método de Newton
(b) Oscilación del método

Hemos visto el caso sencillo de una ecuación real a fin de ilustrar mejor las ideas subyacentes en los métodos iterativos.

Veamos ahora cómo aplicar esta teoría a un problema de valor inicial con una ecuación diferencial ordinaria de primer orden.

1.4. Teorema de Picard-Lindelöf

Como otra aplicación básica del Teorema 1.4 y de la Proposición 1.7, consideramos el problema del valor inicial

$$\begin{cases} x'(t) = f(t, x(t)), \\ x(t_0) = p \end{cases} \tag{35}$$

para una ecuación diferencial ordinaria en $[t_0 - c, t_0 + c]$. Geométricamente, (35) significa que buscamos una curva que satisfaga la ecuación diferencial y pase por (t_0, p), como en la Figura 1.3, con $p = p_0$. En los puntos extremos $t = t_0 \pm c$, $x'(t)$ debe interpretarse como la derivada correspondiente.

La ecuación (35), es equivalente a la ecuación integral siguiente

$$x(t) = p + \int_{t_0}^{t} f(s, x(s))ds, \quad \text{para todo } t \in [t_0 - c, t_0 + c], \tag{36}$$

si f es continua en un entorno adecuado de (t_0, p).

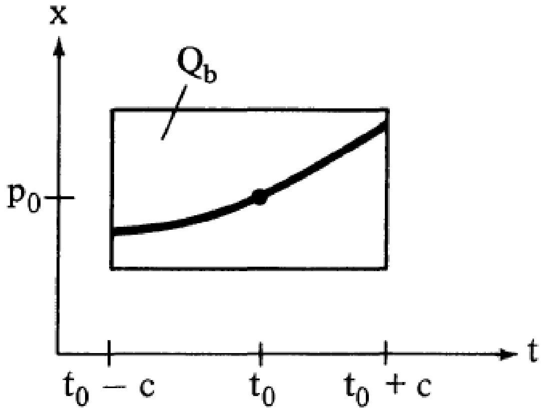


Figura 1.3: Representación geométrica Teorema de Picard-Lindelöf

Teorema 1.11 (Picard(1890), Lindelöf(1894)). *Sean los números reales t_0, p_0 y el rectángulo*

$$Q_b = \{(t, x) \in \mathbb{R}^2 : |t - t_0| \leqslant a, |x - p_0| \leqslant b\}$$

para $a, b > 0$ fijos. Supongamos que $f : Q_b \to \mathbb{R}$ es continua y que

$$|f(t, x) - f(t, y)| \leqslant L|x - y| \quad \textit{para todo } (t, x), (t, y) \in Q_b, \tag{37}$$

$$|f(t, x)| < K, \quad \textit{para todo } (t, x) \in Q_b, \tag{38}$$

donde $L \geqslant 0$ y $K \geqslant 0$ son fijos.

Entonces:

(a) Existencia y unicidad. *Si fijamos $c = min(a, b/K)$ y $p = p_0$, entonces la ecuación integral (36) tiene exactamente una solución continua $x(\cdot)$ en el intervalo $[t_0 - c, t_0 + c]$. Esta función también es la única solución del problema de valor inicial (35) en $[t_0 - c, t_0 + c]$.*

(b) Aproximación sucesiva. *Las aproximaciones sucesivas*

$$x_{n+1}(t) = p_0 + \int_{t_0}^{t} f(s, x_n(s))ds, \quad x_0(t) = p_0, \quad n = 0, 1, 2, \ldots$$

convergen uniformemente en $[t_0 - c, t_0 + c]$ a la solución $x(\cdot)$.

(c) Dependencia continua de la solución respecto de los valores iniciales. *Sea $[t_0 - d, t_0 + d]$ con d fijo, $0 < d < c$, elegido. Entonces la ecuación integral (36) tiene exactamente una solución continua $x_p(\cdot)$ en el intervalo $[t_0 - d, t_0 + d]$ para cada p en un entorno suficientemente*

pequeño de p_0.

Además, es cierto lo siguiente:

$$\textit{Si } p \to p_0, \textit{ entonces } x_p(t) \to x_{p_0}(t) \textit{ uniformemente en } [t_0 - d, t_0 + d]. \tag{39}$$

Corolario 1.12 (Estimación del error). *Sea $||x||_1 = max_{t_0 - c \leqslant t \leqslant t_0 + c}|x(t)|exp(-L|t - t_0|)$. Para $k = 1 - exp(-Lc)$ y para todo $n = 0, 1, 2, ...,$ tenemos*

$$||x_n - x||_1 \leqslant k^n (1-k)^{-1} ||x_1 - x_0||_1$$

$$||x_{n+1} - x||_1 \leqslant k(1-k)^{-1} ||x_{n+1} - x_n||_1$$

$$||x_{n+1} - x||_1 \leqslant k||x_n - x||_1$$

En el siguiente corolario, consideramos una situación en la que, a diferencia del Teorema 1.11, se puede garantizar la existencia de un intervalo mayor para la solución.

Corolario 1.13 *Dados $a > 0$ y $L \geqslant 0$ fijo, y sea $Q = \{(t,x) \in \mathbb{R}^2 : |t - t_0| \leqslant a\}$. Si $f : Q \to \mathbb{R}$ es continuo con*

$$|f(t,x) - f(t,y)| \leqslant L|x - y| \quad \textit{para todo } (t,x), (t,y) \in Q,$$

entonces el problema de valor inicial (35) tiene exactamente una solución para cada p real, en $C^1([t_0 - a, t_0 + a])$.

Demostración:

Podemos reescribir la ecuación integral (36) como una ecuación de operador

$$x = T_p x, \quad x \in M \subseteq X, \quad X = C([t_0 - c, t_0 + c]), \tag{40}$$

donde $C([t_0 - c, t_0 + c])$ es el B-espacio de las funciones reales continuas en $[t_0 - c, t_0 + c]$ con la norma uniforme

$$||x|| = max_{t \in [t_0 - c, t_0 + c]} |x(t)|;$$

y sea $M = \{x \in X : ||x - p_0|| \leqslant b\}$. Con esta elección, la solución existiría en el intervalo $[t_0 - c, t_0 + c]$, donde $c < min(a, b/K, 1/L)$. Se trata de un intervalo más pequeño que el reclamado por el teorema. Para obtener este último, proporcionamos X con una norma nueva, $|| \cdot ||_1$. Dado que

$$||x|| exp(-Lc) \leqslant ||x||_1 \leqslant ||x||, \quad \text{para todo } x \in X, \tag{41}$$

las normas son equivalentes y $(X, || \cdot ||_1)$ y $(X, || \cdot ||)$ son ambos B-espacios.

Demostración del Teorema 1.11

Probamos (a):

(I) Definimos un operador T_p por $T_p x = y$, donde

$$y(t) = p + \int_{t_0}^{t} f(s, x(s))ds, \quad \text{para todo } t \in [t_0 - c, t_0 + c]$$

y comprobamos que se cumplen las condiciones del Teorema 1.4:

(I-1) M es cerrado en $(X, \|\cdot\|_1)$. Supongamos que $x_n \in M$ para todo n, es decir, $\|x_n - p_0\| \leqslant b$ para todo n, y sea $\|x - x_n\|_1 \to 0$ cuando $n \to \infty$. Entonces (41) implica que $\|x - x_n\| \to 0$ cuando $n \to \infty$. Por lo tanto $\|x - p_0\| \leqslant b$, es decir, $x \in M$.

(I-2) T_{p_0} asigna M en M. Porque si $x \in M$, entonces $\|x - p_0\| \leqslant b$, y por lo tanto

$$|x(t) - p_0| \leqslant b \text{ para todo } t \in [t_0 - c, t_0 + c].$$

Por (38),

$$\begin{aligned}
\|T_{p_0} x - p_0\| = & \max_{t \in [t_0 - c, t_0 + c]} \left| \int_{t0}^{t} f(s, x(s))ds \right| \leqslant \\
\leqslant & \max_{t \in [t_0 - c, t_0 + c]} \int_{t0}^{t} |f(s, x(s))| ds \leqslant \\
\leqslant & \max_{t \in [t_0 - c, t_0 + c]} \int_{t0}^{t} K ds = cK \leqslant b.
\end{aligned}$$

Luego, $T_{p_0} x \in M$.

(I-3) T_{p_0} es k-contractivo en M. Supongamos $x, y \in M$, es decir, $\|x - p_0\| \leqslant b$, $\|y - p_0\| \leqslant b$. Entonces, usando (37),

$$\begin{aligned}
\|T_{p_0} x - T_{p_0} y\|_1 = & max_{t \in [t_0 - c, t_0 + c]} \left| \int_{t_0}^{t} f(s, x(s)) - f(s, y(s))ds \right| e^{-L|t - t_0|} \leqslant \\
\leqslant & max_t \left| \int_{t_0}^{t} e^{L|s - t_0| - L|t - t_0|} ds \right| L \|x - y\|_1 \leqslant \\
\leqslant & k\|x - y\|_1, \quad \text{donde } k = 1 - e^{-Lc} < 1.
\end{aligned}$$

La integral se calcula por separado para $t \geqslant t_0$ y $t \leqslant t_0$.

(I-4) El Teorema 1.4 proporciona la existencia de exactamente una solución de $x = T_{p_0} x$, con $x \in M$. Esta solución satisface (36) con $p = p_0$, y por lo tanto también satisface (35).

(II) Demostramos además que toda solución $y(\cdot)$ de (36) que sea continua en $[t_0 - c, t_0 + c]$ coincide con $x(\cdot)$ en (I-4). De lo contrario, $y(\cdot)$ tendríamos que difiere de $x(\cdot)$ en algún punto (t_1, p_1) (Figura 1.4). Entonces, podríamos fijar el problema de valor inicial en

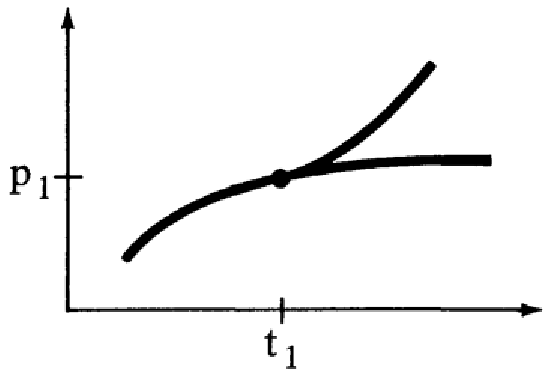


Figura 1.4: $y(\cdot)$ difiere de $x(\cdot)$

este punto, y los resultados anteriores indicarían la unicidad local de la solución. Por lo tanto, no puede producirse tal desviación.

Demostración de (b) y del corolario 1.12:
Por el Teorema 1.4, observemos que $\|x - x_n\|_1 \to 0$, lo que significa la convergencia uniforme en $[t_0 - c, t_0 + c]$.

Demostración de (c):
Comprobamos las hipótesis de la Proposición 1.7. Puesto que queremos dejar que p varíe cerca de p_0, tenemos que reducir el rectángulo Q_b de la Figura 1.3. Así, sustituimos el intervalo $[t_0 - c, t_0 + c]$ en la definición de X y M por $[t_0 - d, t_0 + d]$, y hacemos b un poco más pequeño. Los mismos argumentos que en la demostración de la afirmación (a) demuestran que

$$x_p = T_p x_p$$

tiene exactamente una solución $x_p \in M$ para todo p en un entorno pequeño de p_0, y $x_p(\cdot)$ es la única solución continua de (36) en $[t_0 - d, t_0 + d]$. Como $p_n \to p_0$ y $n \to \infty$, tenemos

$$\|T_{p_n} x - T_{p_0} x\|_1 = |p_n - p_0| \to 0, \quad n \to \infty, \quad \text{para todo } x \in M.$$

Por la Proposición 1.7 entonces, $\|x_{p_n} - x_{p_0}\|_1 \to 0$ cuando $n \to \infty$.

Demostración del corolario 1.13:
En la demostración de (a), fijemos $M = X = C[t_0 - a, t_0 + a]$. La solución $x \in M$ de la ecuación integral (36) es continuamente diferenciable, es decir, $x \in C^1([t_0 - a, t_0 + a])$. $\square$

1.5. Teorema de la Función Implícita

En esta sección mostraremos el Teorema de la Función Implícita siguiendo la referencia [3] de la bibliografía. Este teorema clásico en Análisis Funcional es otra aplicación del Teorema 1.4.

Teorema 1.14 (Teorema de la Función Implícita). *Sea N un entorno de un punto $(a,b) \in \mathbb{R}^2$. Supongamos que f es una función continua de N en N verificando que:*

1. *$\frac{\partial f}{\partial y}(x,y)$ existe en N y es continua en (a,b).*
2. *$\frac{\partial f}{\partial y}(a,b) \neq 0$.*
3. *$f(a,b) = 0$.*

Entonces, existe una única función y_0 definida en cierto entorno suficientemente pequeño de a tal que $f(x,y_0(x)) = 0$ en dicho entorno.

Demostración:
A partir de ahora, usaremos la notación $D_f = \frac{\partial f}{\partial y}(a,b)$. Denotaremos por M al conjunto de funciones que a lo largo de la demostración les impondremos condiciones para poder probar la existencia y unicidad de punto fijo de la función T de M en M, que para cada x, viene dada por

$$[Tz](x) = z(x) - \frac{f(x,z(x))}{D_f}.$$

Está claro que, si probamos que dicha función tiene un único punto fijo y, dicho punto fijo será la función que buscamos.
Consideramos el rectángulo $R = [a-\varepsilon, a+\varepsilon] \times [b-\delta, b+\delta]$, donde ε y δ son números positivos suficientemente pequeños para los cuales

$$\left| \frac{1}{D_f} \frac{\partial f}{\partial y}(x,y) - 1 \right| < \frac{1}{2} \text{ para } (x,y) \in R,$$

$$\left| \frac{1}{D_f} f(x,y) \right| < \frac{1}{2}\delta \text{ para } |x| < \varepsilon.$$

Ahora, sea $C = C([a-\varepsilon, a+\varepsilon], \mathbb{R})$ y tomamos

$$M = \{y \in C : y(a) = b,\ \|y - \beta\| \leqslant \delta\},$$

donde, en este caso, β denota a la función que vale b en todo punto y la norma es la norma infinito o de la convergencia uniforme. Claramente T es una aplicación que lleva elementos de M en C. Tenemos que

$$\|T\beta - \beta\| = \left\| \frac{1}{D_f} f(x,b) \right\| < \frac{1}{2}\delta.$$

Para $(x,y) \in R$, tenemos que

$$\left|\frac{\partial}{\partial y}\left(y - \frac{1}{D_f} f(x,y)\right)\right| = \left|1 - \frac{1}{D_f}\frac{\partial}{\partial y} f(x,y)\right| < \frac{1}{2},$$

de lo que se sigue que, para $y, z \in M$,

$$|[Ty](x) - [Tz](x)| \leqslant \frac{1}{2}|y(x) - z(x)|, \ x \in [a-\varepsilon, a+\varepsilon].$$

Por tanto, $\|Ty - Tz\| \leqslant \frac{1}{2}\|y - z\|$, es decir, T es una aplicación contractiva con constante de contractividad $k = 1/2$.

Además, T aplica M en M ya que:

$$\begin{aligned} \|Ty - \beta\| &\leqslant \|Ty - T\beta\| + \|T\beta - \beta\| \leqslant \\ &\leqslant \frac{1}{2}\|y - z\| + \|T\beta - \beta\| < \\ &< \frac{1}{2}\delta + \frac{1}{2}\delta = \delta. \end{aligned}$$

Además, por ser M subespacio cerrado de un espacio completo, M es completo. Por el Teorema 1.4, T tiene un único punto fijo, lo que es equivalente a que existe una única función continua y definida en un entorno $[a-\varepsilon, a+\varepsilon]$ de a tal que $f(x, y(x)) = 0$, para todo x en dicho entorno. □

Este teorema puede ser generalizado para los casos m-dimensional e infinito dimensional. Introduciremos la derivada de Fréchet para ello:

Sean X un espacio normado, U un abierto no vacío de X y $F : U \subseteq X \to \mathbb{R}$ una función. Se dice F es *diferenciable* en $u_0 \in U$ en el sentido de Fréchet si existe una aplicación lineal y continua $L : X \to \mathbb{R}$ tal que

$$\lim_{h \to 0} \frac{F(u_0 + h) - F(u_0 - L(h)}{\|h\|} = 0.$$

La aplicación L se denomina diferencial en el sentido de Fréchet de F en u_0.

Sea f una aplicación de $B \times C$ en C, donde B y C son B-espacios, D_f es la diferencial en el sentido de Fréchet de f en (a,b). Reemplazamos la segunda condición en el Teorema 1.14 por la existencia $(D_f)^{-1}$, en particular, si f está definida en $\mathbb{R}^m \times \mathbb{R}^n$ con valores en $\mathbb{R}^n$, interpretaremos D_f como la matriz de dimension $n \times n$ de derivadas parciales con respecto a las n variables de $\mathbb{R}^n$.

1.6. Teorema principal de los métodos iterativos para ecuaciones con operadores lineales

A partir de esta sección volvemos a seguir [1].

Veamos ahora importantes consecuencias del Teorema del punto fijo de Banach cuando nuestra ecuación de punto fijo viene dada por un operador lineal. Esto tiene importantes aplicaciones, por ejemplo en los métodos iterativos para resolver sistemas de ecuaciones lineales. Queremos estudiar la ecuación de punto fijo

$$x = Ax + b, \quad x \in X, \tag{42}$$

y el método iterativo correspondiente

$$x_n = Ax_{n-1} + b, \quad x_0 \in X, \quad n = 1, 2, \dots, \tag{43}$$

dados el operador lineal continuo $A : X \to X$ en un B-espacio X, y el elemento $b \in X$. Las conclusiones que obtenemos serán de gran importancia para los problemas lineales y también servirán de base para las consideraciones de estabilidad de los métodos iterativos para ecuaciones no lineales.
Supondremos que X es un B-espacio complejo. Esto no es una restricción, porque siempre podemos extender el B-espacio real X al B-espacio complejo $X_{\mathbb{C}}$. El operador lineal continuo $A : X \to X$ está extendido entonces a un operador lineal continuo sobre $X_{\mathbb{C}}$.
Nuestro principal resultado en un B-espacio complejo será que $\{x_n\}$ siempre converge a una solución única x de (42) si el radio espectral $r(A)$ es menor estrictamente que 1. Recordemos que $r(A) = \lim_{n\to\infty}(\|A^n\|)^{1/n}$ y $r(A) \leqslant \|A\|$, y $\|A\|$ es la norma de un operador lineal continuo, es decir, si $A : X \to X$ lineal y continuo,

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \max_{\|x\|=1} \|Ax\| = \max_{\|x\| \leqslant 1} \|Ax\|.$$

Si A es una matriz, $\|A\|$ puede ser cualquier norma subordinada en una norma vectorial, por ejemplo la norma fila, la norma columna o la norma espectral.

Un sistema de ecuaciones lineales

$$\xi_i = \sum_{j=1}^{N} a_{ij} \xi_j + b_i, \quad i = 1, \dots, N \tag{44}$$

con ξ_i, $a_{ij} \in \mathbb{C}$, $i, j = 1, \dots, N$, puede escribirse en la forma de (42) siendo $x = (\xi_1, \dots, \xi_N)$, $b = (b_1, \dots, b_N)$ y $X = \mathbb{C}^N$. El radio espectral $r(A)$ es entonces el autovalor de módulo máximo de la matriz compleja (a_{ij}).

Definición 1.15 La ecuación (42) tiene *un método iterativo estable* si y solo si (42) tiene exactamente una solución x para cada $b \in X$ y la sucesión $\{x_n\}$ converge a x a partir de un elemento inicial arbitrario $x_0 \in X$.
El número $R_n = log_{10}(\|A^n\|^{-1/n})$ se denomina *tasa media de convergencia* para n iteraciones. El número $R_\infty = lim_{n\to\infty} R_n$ se denomina *tasa asintótica de convergencia.*

Veremos que R_∞ siempre existe, y $R_\infty = log_{10} r(A)^{-1}$ e ilustraremos la importancia de ellos más adelante.

Teorema 1.16 (Teorema principal para ecuaciones de operadores lineales). *Supongamos que $A : X \to X$ es un operador lineal continuo sobre un B-espacio real o complejo X.*

(a) Criterio de la norma. *Si $\|A\| < 1$, entonces (42) tiene un método iterativo estable. El inverso $(I - A)^{-1}$ existe como operador lineal continuo sobre X y*

$$(I - A)^{-1} = \sum_{k=0}^{\infty} A^k. \tag{45}$$

Esta serie geométrica generalizada, o serie de Neumann, converge en la norma del operador. La única solución de (42) es $x = (I - A)^{-1} b$. Para $n = 0, 1, 2, \ldots$ tenemos

$$\|x - x_n\| \leqslant \|A\|^n \|x_1 - x_0\| / (I - \|A\|) \quad \text{(estimación del error a priori)}; \tag{46}$$

$$\|x - x_{n+1}\| \leqslant \|A\| \, \|x_{n+1} - x_n\| / (I - \|A\|) \quad \text{(estimación del error a posteriori)}; \tag{47}$$

$$\|x - x_{n+1}\| \leqslant \|A\| \, \|x - x_n\| \quad \text{(convergecia lineal)}. \tag{48}$$

(b) Criterio del radio espectral. *Supongamos que X es un B-espacio complejo y $r(A)$ es el radio espectral de A. Entonces la ecuación (42) tiene un método iterativo estable para $r(A) < 1$, y no tiene método iterativo estable para $r(A) > 1$. Para $r(A) < 1$ y x la única solución de (42), existe la estimación del error*

$$\|x - x_n\| \leqslant \|A^n\| \, \|x - x_0\| = 10^{-nR_n} \|x - x_0\|, \quad n = 1, 2, \ldots \tag{49}$$

Veamos una interpretación de R_∞. Por la Definición 1.15, $R_n / R_\infty \to 1$ cuando $n \to \infty$. Sea $a(n) = 1 - (R_n / R_\infty)$. Entonces (49) implica

$$\|x - x_n\| \leqslant (10^{1 - a(n)})^{-nR_\infty} \|x - x_0\|, \quad n = 1, 2, \ldots, \tag{50}$$

y $a(n) \to 0$ cuando $n \to \infty$. Así, para n grande, R_∞ es una media del número de potencias de 10 en que se reduce el error en cada paso de la iteración.

El Teorema 1.16 tiene multitud de aplicaciones a sistemas lineales de ecuaciones y ecuaciones integrales lineales. La solución iterativa de sistemas lineales de ecuaciones tiene una importancia fundamental para el tratamiento práctico de los diferentes métodos para ecuaciones diferenciales ordinarias y parciales.

Podemos mejorar los resultados del Teorema 1.16. Para ello necesitamos introducir algunas definiciones topológicas.

Definición 1.17 Sea X un espacio topológico. Un conjunto $A \subset X$ se llama *diseminado* si su clausura $\overline{A}$ tiene interior vacío, es decir, $int(\overline{A}) = \emptyset$.

Definición 1.18 Los conjuntos de *primera categoría* en X son aquellos expresables como unión numerable de conjuntos diseminados. Cualquier subconjunto de X que no sea de primera categoría se dice que es de *segunda categoría* en X.

Corolario 1.19 (Versión mejorada del Teorema 1.16). *Sea $A : X \to X$ un operador lineal continuo sobre un B-espacio complejo X.*

(a) Si $r(A) < 1$, entonces $\{x_n\}$ en (43) converge, para cada $b \in X$ y para un valor inicial arbitrario $x_0 \in X$, a una única solución x de (42). Para cada $\varepsilon > 0$, existe un número positivo $C(\varepsilon)$ tal que

$$\|x - x_n\| \leqslant C(\varepsilon)(r(A) + \varepsilon)^n \|x_1 - x_0\|, \quad n = 1, 2, \dots . \tag{51}$$

(b) Si $r(A) > 1$, entonces existe un punto $b \in X$ tal que $\{x_n\}$ diverge para $x_0 = 0$. Si denotamos K_0 el conjunto de todos los $b \in X$ para los que $\{x_n\}$ converge cuando $x_0 = 0$, entonces K_0 es de primera categoría de Baire en X. El complementario $X - K_0$ es de segunda categoría de Baire en X y es denso en X.

(c) Si $r(A) = 1$, y si A tiene un valor propio $\lambda \in \mathbb{C}$ con $|\lambda| = 1$ y con el correspondiente vector propio y, entonces las dos sucesiones $\{x_n\}$ y $\{y_n\}$, donde $x_n = Ax_{n-1}$, $x_0 = 0$ y $y_n = Ay_{n-1}$, $y_0 = y$, no pueden converger simultáneamente al mismo punto en X, es decir, (42) no tiene un método iterativo estable.

Los conjuntos K_0 de primera categoría de Baire son distintamente raros en comparación con X y $X - K_0$. Por ejemplo, en $X = \mathbb{R}$, hay un número contable de puntos o líneas de primera categoría. Así, el Corolario 1.19 (b) dice a grandes rasgos que cuando $r(A) > 1$, en general hay que contar con la divergencia del método iterativo.

Demostración del Teorema 1.16(a):

(I) Si $\|A\| < 1$, entonces

$$\|I\| + \|A\| + \|A^2\| + \cdots \leqslant 1 + \|A\| + \|A^2\| + \cdots ,$$

y la serie mayor es una serie geométrica convergente. Así, la serie definida por $F(A) = I + A + A^2 + \cdots$ es absolutamente convergente en el B-espacio $L(X, X)$. Puesto que $I = F(A)(I - A) = (I - A)F(A)$, de aquí se deduce que $F(A) = (I - A)^{-1}$.

(II) Fijando $Tx = Ax + b$, tenemos

$$\|Tx - Ty\| = \|Ax - Ay\| \leqslant \|A\| \, \|x - y\|, \quad \text{para todo } x \in X;$$

es decir, T es k-contractiva con $k = \|A\|$. El Teorema 1.4 implica entonces (46), (47) y (48).

$\square$

Demostración del Corolario 1.19(a): Podemos introducir normas equivalentes $\|\cdot\|_\varepsilon$ en X para cada $\varepsilon > 0$, es decir,

$$m(\varepsilon)\|x\| \leqslant \|x\|_\varepsilon \leqslant M(\varepsilon)\|x\|, \quad \text{para todo } x \in X, \tag{52}$$

de modo que el correspondiente operador norma $\|A\|_\varepsilon$ satisface la desigualdad
$r(A) \leqslant \|A\|_\varepsilon \leqslant r(A) + \varepsilon$. Como $r(A) < 1$, también lo es $r(A) + \varepsilon$, para ε suficientemente pequeño, y por tanto $\|A\|_\varepsilon$. Por el Teorema 1.16(a), la ecuación $x = Ax + b$ tiene exactamente una solución x para cada $b \in X$, y para $x_n = Ax_{n-1} + b$, tenemos

$$\begin{aligned}\|x - x_n\|_\varepsilon &\leqslant \|A\|_\varepsilon^n (1 - \|A\|_\varepsilon)^{-1} \|x_1 - x_0\|_\varepsilon \leqslant \\ &\leqslant (r(A) + \varepsilon)^n (1 - r(A) - \varepsilon)^{-1} \|x_1 - x_0\|_\varepsilon.\end{aligned}$$

La desigualdad (52) implica entonces (51). $\square$

Demostración del Corolario 1.19(b): Sea $r(A) > 1$. Puesto que $\|A^n\|^{1/n} \to r(A)$ cuando $n \to \infty$, tenemos

$$\|A^n\|^{1/n} \geqslant q > 1,$$

de modo que $\|A^n\| \geqslant q^n$ para n grande. En consecuencia, $\|A^n\| \to \infty$ cuando $n \to \infty$. Para $b \in K_0$ y $x_{n+1} = Ax_n + b$, $x_0 = 0$, tenemos que

$$x_{n+1} = b + Ab + A^2 b + \cdots + A^n b, \quad n \geqslant 1.$$

Por la definición de K_0, tenemos la convergencia de $\{x_n\}$, por lo que

$$x_{n+1} - x_n = A^n b \to 0 \text{ cuando } n \to \infty$$

para todo $b \in K_0$. Si K_0 no es de primera categoría, entonces debe ser de segunda categoría. Por el principio de acotación uniforme, que nos dice que:

Sean X e $Y \in \mathbb{K}$ B-espacios. Sea $A_\gamma : X \to Y$ una aplicación lineal y continua para todo γ en un conjunto de indices Γ. Supongamos que $\sup_{\gamma \in \Gamma} \|A_\gamma x\| < \infty$ para todo $x \in M$ donde M es de segunda categoría de Baire en X (por ejemplo $X = M$). Entonces $\sup_{\gamma \in \Gamma} \|A_\gamma\| < \infty$,

tenemos entonces que $sup_n \|A^n\| < \infty$, contradiciendo $\|A^n\| \to \infty$ cuando $n \to \infty$. Si K_0 es de primera categoría, entonces $X - K_0$ es de segunda categoría y densa en X. $\square$

Demostración del Corolario 1.19(c). De $y_n - x_n = A(y_{n-1} - x_{n-1})$ obtenemos que

$$y_n - x_n = A^n (y_0 - x_0) = A^n y = \lambda^n y.$$

Dado que $|\lambda| = 1$ y $y \neq 0$ no podemos tener $\{x_n\}$ y $\{y_n\}$ convergiendo al mismo punto. $\square$

Demostración del Teorema 1.16(b). Si $r(A) < 1$, entonces existe un método iterativo estable para $x = Ax + b$, por el Corolario 1.19(a). Combinando esto con $x_n = Ax_{n-1} + b$, obtenemos

$$x - x_n = A(x - x_{n-1})$$

y por la aplicación repetida

$$x - x_n = A^n(x - x_0),$$

de modo que $\|x - x_n\| \leqslant \|A^n\| \, \|x - x_0\|$. Para $r(A) > 1$, la ecuación $x = Ax + b$ no tiene un método iterativo estable, por el Corolario 1.19(b). $\square$

Veamos algunas aplicaciones de este teorema:
Tomamos el sistema lineal dado por

$$\sum_{j=1}^{N} b_{ij}\xi_j = c_i, \quad i = 1, \ldots, N, \tag{53}$$

donde $b_{ii} \neq 0$ para todo i, y reescribiendo en la forma equivalente tenemos que

$$\xi_i = \sum_{j=1}^{N} a_{ij}\xi_j + b_i, \quad i = 1, \ldots, N, \tag{54}$$

donde $b_i = c_i/b_{ii}$ y $a_{ij} = -b_{ij}/b_{ii}$, para $i \neq j$ y $a_{ii} \neq 0$. El método iterativo correspondiente se denomina método de paso total y tiene la siguiente expresión:

$$\xi_i^{(n+1)} = \sum_{j=1}^{N} a_{ij}\xi_j^{(n)} + b_i, \quad n = 0, 1, 2 \ldots, i = 1, \ldots, N. \tag{55}$$

Ejemplo 1.20 (Matriz diagonal por filas o por columnas para el método del paso total). Sea (b_{ij}) una matriz compleja $(N \times N)$ que satisface una de las dos condiciones siguientes para cada $i = 1, \ldots, N$:

$$\sum_{j=1, j\neq i}^{N} |b_{ij}| < |b_{ii}| \quad \text{(suma de filas);} \tag{56}$$

$$\sum_{k=1, k\neq i}^{N} |b_{ki}| < |b_{ii}| \quad \text{(suma de columnas).} \tag{57}$$

Este tipo de matrices, es decir, la matriz (b_{ij}) de diagonal estrictamente dominante, tienen mucha importancia porque aparecen frecuentemente cuando resolvemos por diferencias finitas ciertas ecuaciones diferenciales.

Para un $c_i \in \mathbb{C}$ dado pero fijo, $i = 1,...,N$, la ecuación (53) tiene exactamente una solución $\xi_i \in \mathbb{C}$, $i = 1,...,N$, y el método iterativo (55) converge a esta solución a partir de valores iniciales arbitrarios $\xi_i^{(0)} \in \mathbb{C}$, $i = 1,...,N$.

Demostración del Ejemplo 1.20: Sea $x = (\xi_1,...,\xi_N)$, $b = (b_1,...,b_N)$, y $X = \mathbb{C}^{\mathbb{N}}$. Entonces (54) se convierte en la ecuación $x = Ax + b$, $A \in L(X,X)$. Como normas en X elegimos $\|\cdot\|_\infty$ y $\|\cdot\|_1$, de modo que obtenemos las siguientes normas para A, respectivamente:

$$\|A\|_\infty = max_{1\leqslant i\leqslant N} \sum_{j=1}^{N} |a_{ij}| < 1 \quad \text{(para (56));} \tag{58}$$

$$\|A\|_1 = max_{1\leqslant j\leqslant N} \sum_{i=1}^{N} |a_{ij}| < 1 \quad \text{(para (57))} \tag{59}$$

El Teorema 1.16(a) establece la conclusión. Obsérvese que la convergencia con respecto a $\|\cdot\|_\infty$, respectivamente $\|\cdot\|_1$, es equivalente a $\xi_i^{(n)} \to \xi_i$ cuando $n \to \infty$ para todo i. □

La siguiente proposición es de importancia general.

Proposición 1.21 *Sea $A = (a_{ij})$ una matriz compleja $(N \times N)$. Entonces (54) tiene exactamente un método iterativo estable cuando todos los valores propios λ de A satisfacen que $|\lambda| < 1$.*

Demostración de la Proposición 1.21: El radio espectral satisface $r(A) = max_i|\lambda_i|$, con λ_i un valor propio de A. Por el Teorema 1.16(b) y el Corolario 1.19(c), la ecuación (54) tiene exactamente un método iterativo estable cuando $r(A) < 1$. □

Observación 1.22 El Teorema 1.16(b) proporciona las estimaciones del error para el Ejemplo 1.20. El radio espectral es $r(A) \leqslant min(\|A\|_1, \|A\|_\infty)$. Observando que $R_\infty = log_{10} r(A)^{-1}$ y usando (58 y 59) con $a_{ij} = -b_{ij}/b_{ii}$ para $i \neq j$, llegamos a esta conclusión:
Cuanto mayor sea la diagonal de elementos principales b_{ii} en comparación con los elementos restantes b_{ij}, $i \neq j$, mayor es la tasa de convergencia asintótica R_∞, es decir, más rápido converge el método iterativo.

Estudiamos ahora otra aplicación de este teorema a ecuaciones integrales lineales.
Comencemos con la ecuación integral lineal de Fredholm de segundo tipo,

$$x(t) = \mu \int_a^b K(t,s)x(s)ds + f(t), \quad \text{para todo } t \in [a,b], \tag{60}$$

con el método iterativo correspondiente

$$x_n(t) = \mu \int_a^b K(t,s)x_{n-1}(s)ds + f(t), \quad n = 1,2,\ldots . \tag{61}$$

Sea $X = C([a,b],\mathbb{K})$, $\mathbb{K} = \mathbb{R},\mathbb{C}$, donde X denota el B-espacio de todas las funciones continuas $x:[a,b] \to \mathbb{K}$, con norma máxima $\|x\| = \text{máx}_{a\leqslant t\leqslant b} |x(t)|$.

Ejemplo 1.23 Dados los puntos reales $a < b$, y sean las funciones continuas $f:[a,b] \to \mathbb{K}$, y $K:[a,b]\times[a,b] \to \mathbb{K}$. Establecemos

$$c = \max_{a\leqslant t,s\leqslant b} |K(t,s)|$$

y supongamos $c > 0$. Entonces la ecuación (60) tiene exactamente una solución $x \in X$ para cada $\mu \in \mathbb{K}$ con

$$(b-a)|\mu|c < 1,$$

y $\{x_n\}$ converge a x a partir de un valor inicial arbitrario $x_0 \in X$. Por definición de $\| \,.\, \|$, la convergencia $x_n \to x$ cuando $n \to \infty$, significa que converge uniformemente en $[a,b]$.

Demostración: Escribimos (60) en la forma $x = Ax + f$, $x \in X$, donde

$$(Ax)(t) = \mu \int_a^b K(t,s)x(s)ds, \quad \text{para todo } t \in [a,b].$$

Como K es continua, A es una aplicación lineal de X en si mismo y además,

$$\|A\| = sup_{\|x\|\leqslant 1} \|Ax\|$$

y

$$\|Ax\| = \max_{a\leqslant t\leqslant b} \left| \mu \int_a^b K(t,s)x(s)ds \right| \leqslant |\mu|(b-a)c\|x\|.$$

Dado que se cumple la condición $\|A\| < 1$ del Teorema 1.16(a), la conclusión se cumple. □

Ahora, definiremos lo que significa un operador compacto:

Definición 1.24 Se dice que un operador T es *compacto* si $T(B_X)$ es relativamente compacto, donde X es un espacio de Banach.

Observación 1.25 Cuando $\mathbb{K} = \mathbb{C}$, los resultados del Ejemplo 1.20 puede mejorarse. Llamaremos $\mu \in \mathbb{C}$ un número característico para la ecuación integral (60) si y solo si (60) con $f(t) \equiv 0$ tiene una solución $x(t) \not\equiv 0$. Aceptamos ahora que el operador $A: X \to X$ es compacto (esto lo veremos en el Ejemplo 2.18 del capítulo 2). Por tanto, (60) tiene a lo más una cantidad numerable de números $\mu_1, \mu_2, \dots$ con $0 < |\mu_1| \leqslant |\mu_2| \leqslant \dots$. Por el radio espectral de A tenemos entonces que $r(A) = |\mu|/|\mu_1|$. Por tanto, solo necesitamos

$$|\mu| < |\mu_1|,$$

en lugar de $(b-a)|\mu|c < 1$, para satisfacer la condición $r(A) < 1$ del Teorema 1.16(b), y obtener los resultados del Ejemplo 1.23 para $\mu \in \mathbb{C}$.

El hecho de que esto es realmente una mejora se deduce de la relación general $r(A) \leqslant \|A\|$.

En contraste con (60), la *ecuación lineal integral de Volterra*

$$x(t) = \mu \int_a^t K(t,s)x(s)ds + f(t), \quad \text{para todo } t \in [a,b] \tag{62}$$

con límite superior variable en la integral y con el correspondiente método iterativo

$$x_n(t) = \mu \int_a^t K(t,s)x_{n-1}(s)ds + f(t), \quad n = 0,1,2,..., \tag{63}$$

tiene una solución para todo $\mu \in \mathbb{K}$.

Ejemplo 1.26 Sea de nuevo $\mathbb{K} = \mathbb{R}$, (respectivamente $\mathbb{C}$), y $X = C([a,b],\mathbb{K})$. Bajo las mismas condiciones que en el Ejemplo 1.20, la ecuación (62) tiene exactamente una solución $x \in X$ para cada $\mu \in \mathbb{K}$, y $\{x_n\}$ converge a $x \in X$ para cada $x_0 \in X$.

Demostración: Sea $\mathbb{K} = \mathbb{C}$. El caso $\mathbb{K} = \mathbb{R}$ puede tratarse análogamente. Escribimos (62) como $x = Ax + f$, $x \in X$, donde

$$(Ax)(t) = \mu \int_a^t K(t,s)x(s)ds, \quad \text{para todo } t \in [a,b].$$

El operador A se aplica de X en si mismo. Tenemos $\|A^n\| \leqslant (|\mu|(b-a)c)^n/n!$, donde c es como en el Ejemplo 1.20. Esto es cierto porque

$$\begin{aligned} |(Ax)(t)| &= \left|\mu \int_a^t K(t,s)x(s)ds\right| \leqslant |\mu|c\|x\| \int_a^t ds = \\ &= |\mu|c\|x\|(t-a) \\ |(A^2x)(t)| &= \left|\mu \int_a^t K(t,s)(Ax)(s)ds\right| \leqslant |\mu|^2c^2\|x\| \int_a^t (s-a)ds = \\ &= |\mu|^2c^2\|x\|(t-a)^2/2!. \end{aligned}$$

Haciendo el proceso reiteradamente, obtenemos

$$|(A^nx)(t)| \leqslant (|\mu|c(t-a))^n\|x\|/n!,$$

y por tanto,

$$\|A^nx\| = max_{a\leqslant t\leqslant b}|(A^nx)(t)| \leqslant (|\mu|c(b-a))^n\|x\|/n!,$$

Así encontraremos que el radio espectral $r(A) = li_{n\to\infty}\|A^n\|^{1/n} = 0$. Ahora la conclusión se sigue del Teorema 1.16(b), que proporciona estimaciones del error. □

En todos estos resultados y aplicaciones que se deducen del teorema principal, el Teorema del punto fijo de Banach, se garantiza la existencia, la unicidad y un método estable de aproximación a la solución. Pero, las hipótesis a cumplir por el operador pueden ser demasiado restrictivas. En el capítulo siguiente veremos teoremas de punto fijo menos exigentes.

Capítulo 2

Teorema del punto fijo de Schauder y compacidad

La demostración del teorema del punto fijo de Banach solo utiliza las propiedades más sencillas de los espacios métricos. En cambio, la demostración del teorema del punto fijo de Schauder depende de un profundo resultado topológico, a saber, que no existe ninguna aplicación continua y sobreyectiva de una bola cerrada en un n-espacio que deje cada punto de la frontera fijo (el principio de repliegue negativo). La demostración de esta afirmación intuitivamente atractiva no es en absoluto trivial. Una vez establecido este teorema topológico, se deduce inmediatamente el teorema del punto fijo de Brouwer; un proceso de aproximación que implica el teorema del punto fijo de Schauder. Los operadores compactos desempeñan un papel fundamental en esta teoría.
Haremos una aplicación inmediata de los teoremas del punto fijo de Brouwer y Schauder a sistemas no lineales de ecuaciones, a sistemas de desigualdades, a ecuaciones integrales no lineales y a ecuaciones diferenciales. Seguiremos la refencia [1]

2.1. Teorema de extensión

Primero vamos a dar la definición de cierre convexo.

Definición 2.1 Sea X un espacio métrico, se define el *cierre convexo* como la intersección de todos los conjuntos convexos que contienen a X.

El siguiente teorema de extensión lo utilizaremos frecuentemente.

Teorema 2.2 (Tietze (1915), Dugundji (1951)). *Sea $T : M \subset X \to Y$ un operador continuo sobre un subconjunto cerrado y no vacío M del espacio métrico (X, d) al espacio normado Y. Entonces*

T tiene una extension continua $\overline{T} : X \to co(T(M))$, siendo $co(T(M))$ el cierre convexo de $T(M)$.

Ejemplo 2.3 *En el caso especial de $X = \mathbb{R}$ y $M = [a,b]$, el Teorema 2.2 dice que toda función real continua $T : [a,b] \to [c,d]$ puede extenderse a una función continua $\overline{T} : \mathbb{R} \to [c,d]$.*

Demostración:

(I) A cada $y \in X - M$ le asignamos una bola abierta, U_y, con $diam(U_y) < d(U_y, M)$. Esto nos da un recubrimiento (U_y) de $X - M$. Para este recubrimiento, hay una partición de la unidad,

$$\sum_{i \in J} f_i(x) = 1, \quad \text{para todo } x \in X - M, \tag{63}$$

donde $f_i : X - M \to [0,1]$ es continua para todo $i \in J$ y es cero fuera de $U_{y(i)}$ para un $y(i)$ adecuado, mientras que cada $x \in X - M$ tiene un entorno de $V(x)$ tal que todo menos un número finito de f_i son idénticamente cero en $V(x)$.

(II) Para cada $y \in X - M$ elegimos un m_y tal que $d(m_y, U_y) < 2d(M, U_y)$ y definimos

$$\overline{T}(x) = \begin{cases} T(x) & \text{si } x \in M, \\ \sum_{i \in J} f_i(x) T(m_{y(i)}) & \text{si } x \in X - M. \end{cases}$$

(III) De (63) se deduce que $\overline{T}(x) \subseteq co(T(M))$.

(IV) $\overline{T}$ es claramente continuo en $int(M)$ y $X - M$.

(V) Demostramos que $\overline{T}$ es continua en ∂M. Sea $x_0 \in \partial M$. Entonces, $\overline{T}(x_0) = T(x_0)$.

(V-1) Si $x \in X - M$ y $f_i(x) \neq 0$, entonces por la construcción de f_i, tenemos que $x \in U_y$, donde fijamos $y = y(i)$. Aplicando la desigualdad triangular se obtiene que

$$d(m_y, x) \leqslant d(m_y, U_y) + diam(U_y) \leqslant 3d(M, U_y) \leqslant 3d(x_0, x),$$

y por lo tanto

$$d(m_y, x_0) \leqslant d(m_y, x) + d(x, x_0) \leqslant 4d(x_0, x).$$

Dado que $0 \leqslant f_i(x) \leqslant 1$ y $f_i(x) = 0$ para $d(m_{y(i)}, x_0) > 4d(x_0, x)$, tenemos que

$$\|\overline{T}(x) - \overline{T}(x_0)\| = \left\| \sum_{i \in J} f_i(x)(T(m_{y(i)}) - T(x_0)) \right\| \leqslant$$
$$\leqslant supA(x) \quad \text{para todo } x \in X - M.$$

Aquí, definimos el conjunto

$$A(x) = \{\|T(m_{y(i)}) - T(x_0)\| : i \in J, d(m_{y(i)}, x_0) \leqslant 4d(x_0, x)\}.$$

(V-2) Si $x_n \in X - M$ para todo n y $x_n \to x_0$ cuando $n \to \infty$, entonces la continuidad de T implica que $supA(x_n) \to 0$. Por tanto $\overline{T}(x_n) \to \overline{T}(x_0)$ cuando $n \to \infty$.

(V-3) Ahora $x_n \in M$ para todo n y $x_n \to x_0$ cuando $n \to \infty$, entonces $T(x_n) = \overline{T}(x_n)$, por lo que $\overline{T}(x_n) \to \overline{T}(x_0)$. Por tanto, $\overline{T}$ es continua en $x_0 \in \partial M$.

□

2.2. Retractos

Definición 2.4 Supongamos que X es un espacio topológico (por ejemplo, el subconjunto de un B-espacio), y $r : X \to M$ es una aplicación continua con $M \subseteq X$. La aplicación r se denomina *retracción* si y solo si $r(x) = x$ para todo $x \in M$. En ese caso, el conjunto M se denomina retracto de X.

Una retracción colapsa el espacio X en M continuamente, mientras que deja cada punto de M inmóvil.

Ejemplo 2.5 Sea $X = \mathbb{R}^N$, $R > 0$ y $M = \{x \in X : ||x|| \leqslant R\}$. Entonces M es un retracto de X. Una retracción viene dado por la aplicación:

$$r(x) = \begin{cases} x & \text{si } ||x|| \leqslant R, \\ Rx/||x|| & \text{si } ||x|| > R. \end{cases}$$

Para nosotros, la importancia de la retracción está en la capacidad de reducir la cuestión de punto fijo para conjuntos complicados en una cuestión de punto fijo para subconjuntos más simples.

Enunciamos los siguientes principios de retracción:

Proposición 2.6 *(a) Todo subconjunto cerrado y convexo M de un espacio normado X es un retracto de X.*

(b) La frontera $\partial U(x_0, R)$ de una bola cerrada $\overline{U}(x_0, R)$ N-dimensional, con radio $R > 0$ y $N \geqslant 1$, no es un retracto de $\overline{U}(x_0, R)$.

2.3. El teorema del punto fijo de Brouwer

Teorema 2.7 (Teorema del punto fijo de Brouwer (1912)). *Supongamos que M es un subconjunto no vacío, convexo y compacto de $\mathbb{R}^N$, donde $N \geqslant 1$, y que $f : M \to M$ es una aplicación continua. Entonces f tiene un punto fijo.*

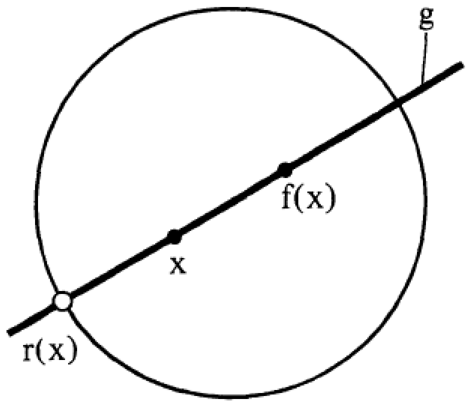


Figura 2.1: Representación geométrica Teorema 2.7(a)

Demostración:

(I) Supongamos en primer lugar que $M = B$, una bola de radio $R > 0$, y que f es una aplicación sin punto fijo, de modo que $f(x) \neq x$ para todo $x \in B$. Entonces podemos construir una retracción $r : B \to \partial B$ como sigue. Para cada x, construimos la recta que une $f(x)$ y x, y la continuamos hasta su intersección con ∂B, y sea el punto de intersección $r(x)$, como en la Figura 2.1. Esto contradice la Proposición 2.6(b).

(II) Para M general, elegimos una bola cerrada B que contenga a M. Por la Proposición 2.6(a), existe una retracción $r : B \to M$. Luego la composición de las aplicaciones

$$B \xrightarrow{r} M \xrightarrow{f} M \subseteq B$$

tiene un punto fijo $x = f(r(x))$ por (I). Los puntos en $B - M$ no son fijos y para $x \in M$, $r(x) = x$, por lo que $x = f(x)$.

□

Ejemplo 2.8 Sea $f : [0,1] \to [0,1]$ continua. Por el Teorema 2.7, la gráfica de la aplicación f intersecta la diagonal como se ve en la Figura 2.2. Todos estos puntos de intersección son exactamente puntos fijos de f. También está claro que los puntos fijos no tienen por qué ser únicos.

En el ejemplo sencillo que hemos puesto no nos hace falta acudir al Teorema del punto fijo de Brouwer general. Fijemos $g(x) = f(x) - x$. Entonces

$$g(0) \geqslant 0 \quad y \quad g(1) \leqslant 0.$$

Luego, por el Teorema de Bolzano, existe un punto $x \in [0,1]$ tal que $g(x) = 0$, es decir, $f(x) = x$.

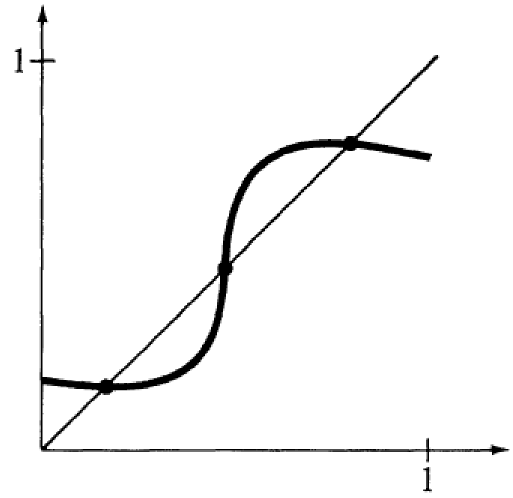


Figura 2.2: Intersección de f con la diagonal

2.3.1. Resultados equivalentes al Teorema de punto fijo de Brouwer

Veamos ahora algunos resultados que son aparentemente diferentes al Teorema de Brouwer, pero que en realidad son equivalentes. Para enunciar estos resultados hemos utilizado la referencia [4].
Empecemos por enunciar un teorema clásico.

Teorema 2.9 (Teorema de Aproximación polinomial de Weierstrass). *Sea f una función de valores reales o complejos, continua sobre el intervalo cerrado* $[0,1]$. *Entonces, existe una sucesión de polinomios p_n que converge cuando $n \to \infty$ a f uniformemente en* $[0,1]$. *(Véase [2]).*

A partir de ahora, usaremos la notación B^n y $\mathbb{S}^n$ para representar la bola cerrada centrada en el origen de radio la unidad y de dimension n y la esfera centrada en el origen de radio la unidad y dimension $n+1$ respectivamente, es decir,

$$B^n = \{x \in \mathbb{R}^n : \|x\| \leqslant 1\}$$

y

$$\mathbb{S}^n = \{x \in \mathbb{R}^{n+1} : \|x\| = 1\}.$$

Teorema 2.10 *El Teorema 2.7 es equivalente a que no existe un retracto continuo y diferenciable de B^n en $\mathbb{S}^{n-1}$.*

Demostración:
Supongamos que se verifica el Teorema 2.7 y que existe un retracto $r : B^n \to \mathbb{S}^{n-1}$ continuo y diferenciable. Tomamos $r_1(x) = -r(x)$, r_1 es una función continua de B^n en B^n que no tiene puntos fijos, lo que contradice el Teorema 2.7.
Supongamos ahora que no existe ningún retracto continuo y diferenciable de B^n en $\mathbb{S}^{n-1}$ y

veamos que se verifica el Teorema 2.7. Supongamos que no se verifica el Teorema 2.7, es decir, existe una función continua f de B^n en B^n que no tiene puntos fijos y consideramos la línea que une x y $f(x)$ intersecada con $\mathbb{S}^{n-1}$, lo que nos da un punto que llamaremos $g(x)$. La función $x \mapsto g(x)$ está bien definida y verifica que $g(x) = x$ para todo $x \in \mathbb{S}^n$. Probaremos que g es continua y diferenciable, para ello, la línea que une x y $f(x)$ tiene la forma
$\alpha(x)x + (1 - \alpha(x))f(x)$. Para probar la afirmación tenemos que ver que α es continua y diferenciable:

$$\begin{aligned}
&\langle \alpha(x)x + (1-\alpha(x))f(x), \alpha(x)x + (1-\alpha(x))f(x)\rangle = \\
&= \alpha(x)^2\|x\|^2 + 2\alpha(x)(1-\alpha(x))\langle x, f(x)\rangle + (1-\alpha(x))^2\|f(x)\|^2 = 1,
\end{aligned}$$

para todo $x \in \mathbb{S}^{n-1}$. Por tanto $\alpha(x)$ es la solución de una ecuación cuyos coeficientes son continuos y diferenciables, así la solución $\alpha(x)$ tiene la misma propiedad y es continua y diferenciable. En consecuencia, el punto fijo de g existe.
Por continuidad de f en el compacto B^n, el Teorema 2.9 nos asegura que podemos encontrar una sucesión $\{f_n\}_{n\in\mathbb{N}}$ de funciones continuas y diferenciables que convergen uniformemente a f. Tomamos ahora $\{x_n\}_{n\in\mathbb{N}}$ dada por $x_n = f_n(x_{n-1})$. El punto al que converge dicha sucesión será un punto fijo de f. Contradicción con que no se verifica el Teorema 2.7. □

Otro resultado equivalente es el Teorema de Henri Poincaré, 24 años antes que el Teorema 2.7.

Teorema 2.11 (Henri Poincaré (1886)). *Sea $f : B^n \to B^n$ una aplicación continua y supongamos que para algún $r > 0$ y todo $\lambda > 0$, $f(u) + \lambda u \neq 0$ para todo u con $\|u\| = r$. Entonces existe un punto u_0, con $\|u_0\| < r$, tal que $f(u_0) = 0$.*

A continuación, enunciaremos otro Teorema equivalente muy útil para las ecuaciones diferenciales.

Teorema 2.12 (Teorema de Poincaré-Miranda). *Si denotamos por*
$D^n = \{x = (x_1, \dots, x_n) \in \mathbb{R}^n : a_i \leqslant x_i \leqslant b_i\}$ y dada $f_i : D^n \to \mathbb{R}$ continua y verificando:

$$\begin{aligned}
f_i(x_1, \dots, x_{i-1}, a_i, x_{i+1}, \dots, x_n) \geqslant 0, \\
f_i(x_1, \dots, x_{i-1}, b_i, x_{i+1}, \dots, x_n) \leqslant 0,
\end{aligned}$$

para todo $i = 1, \dots, n$. Entonces existe un punto $x_0 \in D^n$ tal que $f_i(x_0) = 0$ para todo $i = 1, 2, \dots, n$.

Demostración:
Veamos que es equivalente al Teorema 2.7, para ello, en primer lugar, supongamos que se verifica el Teorema 2.7 y veamos que se cumple dicho Teorema. Por ser f_i una función continua definida en un compacto, alcanza el máximo y el mínimo en dicho compacto. Denotemos

por m_i y M_i el mínimo y el máximo de f_i en D^n y sean δ_i' y δ_i'' las distancias del conjunto de puntos de D^n tales que $f_i(x) < 0$ y $f_i(x) > 0$ a los hiperplanos $x_i = a_i$ y $x_i = b_i$, respectivamente. Sabemos que $m_i < 0$ y que $M_i > 0$, por tanto, podemos encontrar n números ε_i tales que

$$0 < \varepsilon_i \leqslant -\delta_i'/m_i,$$
$$0 < \varepsilon_i \leqslant \delta_i''/M_i.$$

Definimos así las funciones $g_i(x_1, ...x_n) = x_i + \varepsilon_i f_i(x_1, ..., x_n)$, las cuales cumplen, si $f(x_1, ..., x_n) = 0$, que

$$a_i \leqslant x_i \leqslant g_i(x_1, ..., x_n) \leqslant b_i.$$

Si $f_i(x_1, ..., x_n) > 0$, tenemos

$$a_i \leqslant x_i < g_i(x_1, ..., x_n) \leqslant b_i - \delta_i'' + \varepsilon_i M_i \leqslant b_i.$$

En el caso de $f(x_1, ..., x_n) < 0$, se tiene

$$b_i \geqslant x_i > g_i(x_1, ..., x_n) \geqslant a_i + \delta_i' + \varepsilon_i m_i \geqslant a_i.$$

Si tomamos $G = (g_1, ..., g_n)$, por el Teorema 2.7 sabemos que G admite un punto fijo, es decir, existe $x = (x_1, ..., x_n) \in D^n$ tal que $G(x) = x$, en particular, $x_i = x_i + \varepsilon_i f_i(x)$, es decir, $f_i(x) = 0$. Recíprocamente, supongamos ahora que se verifica el Teorema 2.12. Veamos que entonces se cumple el Teorema 2.7. Sea $f_i : D^n \to \mathbb{R}$ continua y verificando:

$$f_i(x_1, ..., x_{i-1}, a_i, x_{i+1}, ..., x_n) \geqslant 0,$$
$$f_i(x_1, ..., x_{i-1}, b_i, x_{i+1}, ..., x_n) \leqslant 0,$$

$\forall i = 1, ..., n$. Entonces existe un punto $x_0 \in D^n$ tal que $f_i(x_0) = 0$ $\forall i = 1, 2, ..., n$. Tomando $g_i^*(x) = f_i(x) - x_i$, por hipótesis sabemos que existe $x_0 \in D^n$ tal que $g_i(x_0) = 0$, por tanto, $F = (f_1, ..., f_n)$ tiene un punto fijo. □

Otra forma de enunciar el Teorema 2.7 es:

Teorema 2.13 *Si C es un subconjunto compacto y convexo de un espacio vectorial de dimensión finita y $f : C \to C$ es una aplicación continua, entonces f admite al menos un punto fijo.*

La demostración es inmediata, se sigue de que todo subconjunto compacto y convexo de un espacio vectorial de dimension finita es homeomorfo a la bola cerrada en dicha dimension y, como tener un punto fijo se hereda topológicamente, se obtiene el resultado.

Una de las aplicaciones importantes del Teorema 2.7 es el Teorema Fundamental del Álgebra, que establece que todo polinomio de grado n tiene exactamente n raíces complejas. Veamos que una de sus demostraciones es una consecuencia del Teorema 2.7:

Teorema 2.14 (Teorema Fundamental del Álgebra). *Sea $p(z) = a_0+a_1z+\cdots+a_nz^n$ un polinomio con coeficientes complejos. Entonces existe $z_0 \in \mathbb{C}$ tal que $p(z_0) = 0$. (Véase [4])*

Demostración: Supongamos, sin pérdida de generalidad, que $a_n = 1$, de no ser así, bastaría con dividir todos los coeficientes del polinomio entre a_n.
Fijemos $z = re^{i\theta}$, con $0 \leqslant \theta < 2\pi$, y sea $R = 2+|a_0|+\cdots+|a_{n-1}|$. Definimos la función $g : \mathbb{C} \to \mathbb{C}$ dada por:

$$g(z) = \begin{cases} z - p(z)/(Re^{i(n-1)\theta}), \text{ si } |z| \leqslant 1, \\ z - p(z)/(Rz^{n-1}), \text{ si } |z| > 1. \end{cases}$$

Es claro que g es una función continua (por construcción).
Consideramos el conjunto $A = \{z \in \mathbb{C} : |z| \leqslant R\}$ compacto y convexo en el plano complejo. Veamos que es invariante por g. En efecto, dado $z \in A$, distinguimos dos casos:
Si $|z| \leqslant 1$, entonces:

$$\begin{aligned} |g(z)| = |z - p(z)/Re^{i(n-1)\theta}| &\leqslant |z| + |p(z)|/R = \\ &= |z| + |a_0 + a_1z + \cdots + a_{n-1}z^{n-1} + z^n|/R \leqslant \\ &\leqslant 1 + (|a_0| + |a_1z| + \cdots + |a_{n-1}z^{n-1}| + |z^n|)/R \leqslant \\ &\leqslant 1 + (|a_0| + |a_1| + \cdots + |a_{n-1}| + 1)/R \leqslant 1 + 1 = 2 \leqslant R. \end{aligned}$$

Si $|z| \geqslant 1$, tenemos que

$$\begin{aligned} |g(z)| = |z - p(z)/Rz^{n-1}| &= |z - (a_0 + a_1z + \cdots + a_{n-1}z^{n-1} + z^n)/(Rz^{n-1})| = \\ &= |z - z/R - (a_0 + \cdots + a_{n-1}z^{n-1})/(Rz^{n-1})| \leqslant \\ &\leqslant |z - z/R| + |(a_0 + \cdots + a_{n-1}z^{n-1})/(Rz^{n-1})| \leqslant \\ &\leqslant (R-1) + (|a_0| + \cdots + |a_{n-1}z^{n-1}|)/|Rz^{n-1}| \leqslant \\ &\leqslant (R-1) + |a_0|/|Rz^{n-1}| + |a_1z|/|Rz^{n-1}| + \cdots + |a_{n-1}z^{n-1}|/|Rz^{n-1}| = \\ &= (R-1) + |a_0|/|Rz^{n-1}| + |a_1|/|Rz^{n-2}| + \cdots + |a_{n-1}|/|R| \leqslant \\ &\leqslant (R-1) + |a_0|/|R| + |a_1|/|R| + \cdots + |a_{n-1}|/|R| = \\ &= (R-1) + (|a_0| + |a_1| + \cdots + |a_{n-1}|)/R = R - 1 + (R-2)/R \leqslant R. \end{aligned}$$

Por tanto, A es invariante mediante g. Aplicando el Teorema 2.7 a la aplicación $g_{|A} : A \to A$, se deduce que existe $z_0 \in A$ tal que $g_{|A}(z_0) = z_0$, es decir, se satisface que $p(z_0) = 0$. □

2.4. Principio de existencia para sistemas de ecuaciones

Volvemos a seguir la referencia [1] para los siguientes resultados.
Como una aplicación sencilla del teorema del punto fijo de Brouwer, demostraremos un importante teorema de existencia para el sistema de ecuaciones siguiente:

$$g_i(x) = 0, \quad \text{donde } x = (\xi_1, ..., \xi_N) \in \mathbb{R}^N, \quad i = 1, ..., N. \tag{64}$$

Este teorema de existencia jugará un papel decisivo en la discusión del método de Galerkin para operadores monótonos. La clave es la condición de contorno

$$\sum_{i=1}^{N} g_i(x)\xi_i \geqslant 0, \quad \text{para todo } x \text{ con } \|x\| = R. \tag{65}$$

Proposición 2.15 (Solución de (65)). *Sea* $\overline{U}(0,R) = \{x \in \mathbb{R}^N : \|x\| \leqslant R\}$ *para* $R > 0$ *fijo y* $\|\cdot\|$ *una norma en* $\mathbb{R}^N$. *Sea* $g_i : \overline{U}(0,R) \to \mathbb{R}$ *continua para* $i = 1, ..., N$.

Si (65) se satisface, entonces (64) tiene una solución x *con* $\|x\| \leqslant R$.

Demostración: Lo hacemos por Reducción al Absurdo. Sea $g(x) = (g_1(x), ..., g_N(x))$ y supongamos que $g(x) \neq 0$ para todo $x \in \overline{U}(0,R)$. Entonces definimos

$$f(x) = -Rg(x)/\|g(x)\|.$$

Ahora, f es una aplicación continua de un conjunto compacto y convexo $\overline{U}(0,R)$ en sí mismo. Por el Teorema 2.7, existe un punto fijo $x = f(x)$, de donde $\|x\| = \|f(x)\| = R$. Por tanto, $f(x) = x$ y $\|x\| = R$. Además,

$$\begin{aligned}\sum_i g_i(x)\xi_i &= -R^{-1}\|g(x)\| \sum_i f_i(x)\xi_i = \\ &= -R^{-1}\|g(x)\| \sum_i \xi_i^2 < 0,\end{aligned}$$

lo que contradice (65). □

2.5. Operadores compactos

Nuestro objetivo es generalizar el Teorema del punto fijo de Brouwer a B-espacios de dimensión infinita mediante un proceso de aproximación, obteniendo así el Teorema del punto fijo de Schauder.

Definición 2.16 Sea X e Y B-espacios, y $T : D(T) \subseteq X \to Y$ un operador. T se denomina *compacto* si y solo si:

(i) T es continuo;

(ii) La aplicación T va de conjuntos acotados en conjunto relativamente compactos.

Los operadores compactos desempeñan un papel fundamental en el Análisis Funcional no lineal. Su importancia se debe a que muchos resultados sobre operadores continuos en $\mathbb{R}^N$ se trasladan a B-espacios cuando se sustituye “continuo” por “compacto”.

Ejemplo 2.17 Para B-espacios de dimensión finita, los operadores continuos y compactos son lo mismo siempre que el dominio $D(T)$ sea cerrado. Pues si M es acotado, entonces $\overline{M}$ es compacto, ya que $dim(X) < \infty$. Entonces $f(\overline{M})$ es compacto, y por tanto $f(M)$ es relativamente compacto.

Los ejemplos típicos de operadores relativamente compactos en B-espacios de dimensión infinita son los operadores integrales con integrando suficientemente regular. Tenemos la ecuación de tipo Fredholm

$$(Tx)(t) = \int_a^b K(t,s,x(s))ds,$$

y la ecuación de tipo Volterra

$$(Sx)(t) = \int_a^t K(t,s,x(s))ds \quad \text{para todo } t \in [a,b].$$

Ejemplo 2.18 (Operadores integrales compactos T y S). Supongamos que tenemos una función continua

$$K : [a,b] \times [a,b] \times [-R,R] \to \mathbb{K},$$

donde $-\infty < a < b < \infty$, $0 < R < \infty$, y $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Sea

$$M = \{x \in C([a,b],\mathbb{K}) : ||x|| \leqslant R\}$$

donde $||x|| = \text{máx}_{a\leqslant s\leqslant b} |x(s)|$ y $C([a,b],\mathbb{K})$ es el espacio de aplicaciones continuas $x : [a,b] \to \mathbb{K}$. Entonces los operadores integrales S y T van de M en $C([a,b],\mathbb{K})$ y son compactos.

Para la demostración usaremos el siguiente teorema:

Teorema 2.19 (Teorema de Ascoli-Arzelá.) *Sea G un conjunto abierto no vacío en $\mathbb{R}^n$, el conjunto M en $C(\overline{G})$ es relativamente compacto si y solo si:*

1. *(Uniformemente acotado)* $\sup_{f\in M}(\sup_{x\in\overline{G}} |f(x)|) < \infty$

2. *(Equicontinua) Para todo $\varepsilon > 0$ hay un $\delta(\varepsilon) > 0$ tal que*

$$\sup_{f\in M} |f(x) - f(y)| < \varepsilon \quad \text{cuando} \quad x,y \in \overline{G}, \text{ y } |x-y| < \delta(\varepsilon).$$

Aquí, $\delta(\varepsilon)$ es independiente de x, y y f.

Demostración: Consideremos S para $\mathbb{K} = \mathbb{R}$. Los demás casos se hacen de manera similar.

(I) El conjunto $A = [a,b] \times [a,b] \times [-R,R]$ es compacto, por lo que K es acotado y uniformemente continuo en A. Por tanto, existe un número α de modo que $|K(t,s,x)| \leqslant \alpha$, para todo $(t,s,x) \in A$, y para todo $\varepsilon > 0$ existe un $\delta(\varepsilon) > 0$ tal que

$$|K(t_1,s_1,x_1) - K(t_2,s_2,x_2)| < \varepsilon$$

para todo (t_i,s_i,x_i) en A, $i = 1,2$, satisfaciendo que $|t_1 - t_2| + |s_1 - s_2| + |x_1 - x_2| < \delta(\varepsilon)$.

(II) Sea $z = Sx$ y $x \in M$. Entonces

$$|z(t)| \leqslant \left| \int_a^t K(t,s,x(s))ds \right| \leqslant (b-a)\alpha, \quad \text{para todo } t \in [a,b].$$

Además, para $|t_1 - t_2| \leqslant min(\delta(\varepsilon),\varepsilon)$, tenemos la desigualdad

$$\begin{aligned} |z(t_1) - z(t_2)| &= \left| \int_a^{t_1} K(t_1,s,x(s))ds - \int_a^{t_2} K(t_2,s,x(s))ds \right| = \\ &= \left| \int_a^{t_1} K(t_1,s,x(s)) - K(t_2,s,x(s))ds - \int_{t_1}^{t_2} K(t_2,s,x(s))ds \right| \leqslant \\ &\leqslant (b-a)\varepsilon + |t_1 - t_2|\alpha \leqslant ((b-a) + \alpha)\varepsilon. \end{aligned}$$

(III) Las desigualdades de (II) se satisfacen uniformemente para todo $z = Sx$ con $x \in M$ arbitrario. Por el teorema de Ascoli-Arzela, el conjunto $S(M)$ es relativamente compacto.

(IV) El operador S es continuo en M. Para ver esto, cogemos una sucesión $\{x_n\}$ en M con $\|x_n - x\| \to 0$ cuando $n \to \infty$, es decir, las funciones $x_n(\cdot)$ converge uniformemente en $[a,b]$ a $x(\cdot)$. Fijemos $z_n = Sx_n$ y $z = Sx$. Entonces

$$\begin{aligned} \|z - z_n\| &= \max_{a \leqslant t \leqslant b} |z(t) - z_n(t)| = \\ &= \max_{a \leqslant t \leqslant b} \left| \int_a^t K(t,s,x(s)) - K(t,s,x_n(s))ds \right| \to 0 \quad n \to \infty. \end{aligned}$$

Obsérvese la continuidad uniforme de K y la convergencia uniforme de las funciones de $x_n(\cdot)$ a $x(\cdot)$

(III) y (IV) implican conjuntamente la compacidad de S. □

Teorema 2.20 (Teorema de aproximación para operadores compactos). *Sea X e Y B-espacios, con M un subconjunto acotado no vacío de X. Sea $T : M \subseteq X \to Y$ un operador dado. Entonces T es compacto si y solo si se satisface la siguiente condición:*
Para cada $n \in \mathbb{N}$ existe un operador compacto $P_n : M \to Y$ tal que

$$\sup_{x \in M} \|T(x) - P_n(x)\| \leqslant 1/n \quad y \quad dim(spanP_n(M)) < \infty. \tag{66}$$

En la siguiente demostración utilizaremos esencialmente la caracterización de conjuntos relativamente compactos en B-espacios por ε redes finitas, es decir,

$$M \text{ es relativamente compato } \Leftrightarrow M \text{ es secuencialmente relativamente compacto } \Leftrightarrow$$
$$\Leftrightarrow \text{ para todo } \varepsilon > 0 \text{ hay una } \varepsilon \text{ red finita para } M.$$

Demostración:

$(\Rightarrow)$ Sea T compacto. Entonces $T(M)$ es relativamente compacto, por lo que para cada n existen elementos $y_i \in T(M)$, $i = 1, \ldots, N$ tal que

$$\min_i \|Tx - y_i\| < 1/n, \quad \text{para todo } x \in M, \tag{67}$$

por lo dicho anteriormente. El llamado *operador de Schauder*, definido por

$$P_n(x) = \frac{\sum_{i=1}^N a_i(x) y_i}{\sum_{i=1}^N a_i(x)}, \tag{68}$$

donde $a_i(x) = max(n^{-1} - \|Tx - y_i\|, 0)$, cumple todas las propiedades requeridas: los a_i son continuos y no se anulan simultáneamente para $x \in M$ por (67), y

$$\begin{aligned}\|P_n x - Tx\| &= \left\| \frac{\sum_i a_i(x)(y_i - Tx)}{\sum_i a_i(x)} \right\| \leqslant \\ &\leqslant \frac{\sum_i a_i(x) n^{-1}}{\sum_i a_i(x)} \leqslant \\ &\leqslant n^{-1} \quad \text{para todo } x \in M.\end{aligned}$$

La acotación de $T(M)$ implica la acotación de $P_n(M)$. Dado que el conjunto $P_n(M)$ se encuentra en un espacio de dimension finita, $P_n(M)$ es relativamente compacto, es decir, el operador P_n es compacto.

$(\Leftarrow)$ Supongamos que (66) es cierta. Como límite uniforme de operadores continuos, P_n, el operador T es a su vez continuo, luego por (66) tenemos que

$$\|Tx - Ty\| \leqslant 1/n + \|P_n x - P_n y\| + 1/n < 3\varepsilon$$

para $n \in \mathbb{N}$ fijo suficientemente grande y $\|x - y\| < \delta(\varepsilon)$. Además, $T(M)$ es relativamente compacto, ya que (66) implica que para cada $n \in \mathbb{N}$, el conjunto $T(M)$ tiene una red $2/n$, es decir, un recubrimiento finito. Obsérvese que el conjunto relativamente compacto $P_n(M)$ tiene una red $1/n$. □

2.6. Teorema del punto fijo de Schauder

Teorema 2.21 (Teorema del punto fijo de Schauder(1930)). *Sea M un subconjunto no vacío, cerrado, acotado y convexo de un B-espacio X, y supongamos que $T : M \to M$ es un operador compacto. Entonces T tiene un punto fijo.*

Corolario 2.22 (Versión alternativa del Teorema del punto fijo de Schauder). *Sea M un subconjunto no vacío, compacto y convexo de un B-espacio X, y supongamos que $T : M \to M$ es un operador continuo. Entonces T tiene un punto fijo.*

El corolario es una traslación directa del Teorema del punto fijo de Brouwer a los B-espacios. La primera version (Teorema 2.21) se utiliza con más frecuencia en aplicaciones donde M suele ser una bola. Como todo B-espacio complejo puede considerarse también como un B-espacio real, podemos suponer que X es real en la siguiente demostración.

Demostración del Corolario 2.22: Sea $M_n = co(y_1, ..., y_n)$, donde y_i y P_n se eligen como en (67), (68). La convexidad de M implica que

$$M_n \subseteq co(T(M)) \subseteq M.$$

Por lo tanto,

$$P_n : M_n \to M_n$$

es continua. Además, M_n es compacto y convexo, y $M_n \subseteq \mathbb{R}^N$. Aquí, identificamos $span\{y_1, ..., y_N\}$ con un subespacio de $\mathbb{R}^N$.
Por el Teorema 2.7 (Principio del punto fijo de Brouwer), existe un punto fijo

$$x_n = P_n(x_n), \quad \text{donde } x_n \in M_n \subseteq M.$$

Dado que M es compacto, hay una subsucesión convergente, de nuevo denotada por $\{x_n\}$, tal que $x_n \to x$ cuando $n \to \infty$. Este x es el punto fijo deseado, pues

$$\|x_n - Tx\| \leqslant \|P_n x_n - Tx_n\| + \|Tx_n - Tx\|,$$

y el lado derecho desaparece cuando $n \to \infty$, por (66), luego $x = Tx$. □

Demostración del Teorema 2.21: Sea $A = \overline{co}(T(M))$. Entonces $A \subseteq M$ y el conjunto A es compacto y convexo (debido al Teorema de Mazur que nos dice que :
Si M es relativamente compacto en un B-espacio X, entonces el cierre convexo $co(M)$ es relativamente compacto y el cerrado $\overline{co}(M)$ es compacto).
Además, $T(A) \subseteq A$. Por tanto, la restricción

$$T : A \to A$$

tiene un punto fijo, por el Corolario 2.22, y este punto fijo es automáticamente un punto fijo de T en M. □

2.7. Teorema de Peano

De forma paralela a la Sección 1.4, consideramos el problema del valor inicial

$$\begin{cases} x'(t) = f(t, x(t)), \\ x(t_0) = y_0 \end{cases} \tag{69}$$

Comparando con el Teorema de Picard-Lindelön (Teorema 1.11), aquí f es meramente continua.

Teorema 2.23 (Peano (1890)). *Dados los números reales t_0 e y_0, y el rectángulo*

$$Q_b = \{(t, x) \in \mathbb{R}^2 : |t - t_0| \leqslant a, |x - y_0| \leqslant b\},$$

donde a y b son números positivos fijos. Supongamos que $f : Q_b \to \mathbb{R}$ es continua y acotada con

$$|f(t, x)| \leqslant K \quad \text{para todo } (t, x) \in Q_b,$$

y $K > 0$ fijo. Sea $c = min(a, K/b)$.
Entonces el problema del valor inicial (69) tiene una solución continuamente diferenciable en $[t_0 - c, t_0 + c]$.

Demostración: Como en la demostración del Teorema 1.11, sustituimos primero (69) por la ecuación integral

$$x(t) = y_0 + \int_{t_0}^{t} f(s, x(s)) ds,$$

y a continuación, lo escribimos como la ecuación del operador

$$x = Tx, \quad x \in M \subseteq X,$$

donde $X = C([t_0 - c, t_0 + c])$, $M = \{x \in X : \|x - y_0\| \leqslant b\}$, y $\|x\| = \text{máx}_{t_0 - c \leqslant t \leqslant t_0 + c} |x(t)|$.
El conjunto M es cerrado y acotado en X. Como en la demostración del Teorema 1.11, en (I-2), se deduce que $T(M) \subseteq M$. El operador T es compacto por el Ejemplo 2.18. Ahora, el Teorema 2.21 implica la existencia de una solución $x = Tx$, $x \in M$. $\square$

2.8. Ecuaciones integrales con parámetros pequeños

Estudiamos la ecuación integral no lineal

$$x(t) = \mu \int_a^b F(t, s, x(s)) ds + \int_a^b G(t, s, x(s)) ds + \alpha g(t). \tag{70}$$

Queremos utilizar los teoremas del punto fijo de Banach y de Schauder para establecer la existencia de soluciones cuando se cumplan las siguientes condiciones:

(i) μ y α son parámetros pequeños reales;

(ii) El término no lineal sin parámetros es de orden superior a 1 en x (véase (71)).

Proposición 2.24 (Solución de (70)). *Dado un conjunto*

$$Q = \{(t,s,x) \in \mathbb{R}^3 : a \leqslant t, s \leqslant b, |x| \leqslant r_0\},$$

donde a, b, r_0 son números positivos fijos. Supongamos que hay funciones continuas $F, G : Q \to \mathbb{R}$ y $g : [a,b] \to \mathbb{R}$, con

$$|G(t,s,x)| \leqslant K|x|^\rho \quad \text{en } Q \quad \text{para } \rho > 1 \text{ y } K > 0 \text{ fijos.} \tag{71}$$

(a) Entonces existen los números μ_0, $\alpha_0 > 0$, de modo que la ecuación (70) tiene una solución para μ fijo con $|\mu| \leqslant \mu_0$, y α fijo con $|\alpha| \leqslant \alpha_0$.

(b) Si F y G tienen las primeras derivadas parciales continuas respecto de x tales que

$$G_x(t,s,0) = 0 \quad \text{para todo } (t,s,0) \in Q, \tag{72}$$

entonces existen números μ_0, α_0, $r > 0$, de modo que para μ fijo con $|\mu| \leqslant \mu_0$, y α fijo con $|\alpha| \leqslant \alpha_0$, (70) tiene exactamente una solución $x(\cdot)$ en $[a,b]$ que es continua y tiene $\max_{t\in[a,b]} |x(t)| \leqslant r$.
Esta solución puede obtenerse mediante aproximaciones sucesivas en (70) a partir del elemento inicial $x_0(t) \equiv 0$.

Demostración:

(a) Elegimos $X = C([a,b])$ con norma máxima $\|x\| = \max_{t\in[a,b]} |x(t)|$ y

$$M = \{x \in X : \|x\| \leq r\}.$$

Escribimos (70) como una ecuación del operador

$$x = Tx, \quad x \in M,$$

y comprobamos las hipótesis del Teorema del punto fijo de Schauder. Todos los máximos son con respecto a $[a,b]$.

(I) El conjunto M es cerrado, acotado y convexo. El operador T es compacto por el Ejemplo 2.18.

(II) T es una aplicación de M en sí mismo. Para ver esto, fijamos que $\|x\| \leq r$. Para un valor fijo suficientemente pequeño $r > 0$, de (71) se deduce que

$$\max_t \left| \int_a^b G(t,s,x(s))ds \right| \leq (b-a)Kr^p \leq r/2.$$

Si elegimos μ_0 y α_0 suficientemente pequeños, entonces para μ y α fijos con $|\mu| \leq \mu_0$ y $|\alpha| \leq \alpha_0$, se tiene que

$$\max_t \left| \mu \int_a^b F(t,s,x(s))ds \right| \leq r/4 \quad \text{y} \quad \max_t |\alpha g(t)| \leq r/4.$$

De estas dos condiciones, se deduce que $\|Tx\| \leq r/2 + r/4 + r/4 = r$. Por el Teorema del punto fijo de Schauder (Teorema 2.21) podemos asegurar una solución de $x = Tx$, $x \in M$, y por tanto se tiene (70).

(b) En este caso demostraremos que se cumplen las hipótesis del Teorema del punto fijo de Banach. El Teorema del valor medio, con $A = F, G$, establece que

$$|A(t,s,x(s)) - A(t,s,y(s))| \leq |A_x(t,s,\xi(s))||x(s) - y(s)|,$$

donde $\xi(s) = x(s) + \vartheta(y(s) - x(s))$, $0 < \vartheta < 1$. Si aplicamos (72) con $x, y \in M$ y r pequeño, llegamos a que

$$\max_t \left| \int_a^b G(t,s,x(s)) - G(t,s,x(s))ds \right| \leq \max_t |x(t) - y(t)|/4 = \|x - y\|/4.$$

Razonando de manera análoga para F se demuestra que para μ_0 y α_0 pequeños positivos y μ fijo con $|\mu| \leq \mu_0$ y α con $|\alpha| \leq \alpha_0$, tenemos que

$$\|Tx - Ty\| \leq \|x - y\|/2 \quad \text{para todo } x, y \in M,$$

es decir, el operador T es k-contractivo. La conclusión se deduce entonces del Teorema del punto fijo de Banach (Teorema 1.4)

□

Ejemplo 2.25 (Oscilación no lineal). Consideremos el problema de contorno

$$\begin{cases} x''(t) + \omega^2 x(t) = \mu f(t, x(t)), \\ x(a) = x(b) = 0 \end{cases} \tag{73}$$

Aquí, ω y μ son parámetros reales fijos. La ecuación (73) describe un muelle oscilante con una fuerza lineal externa μf, y una deflexión $x(t)$ en un tiempo t.

Si la única solución de (73), $\mu = 0$, es la trivial, $x(t) \equiv 0$, entonces se puede usar una función de Green continua G con f continua para convertir (73) en la ecuación integral equivalente

$$x(t) = \mu \int_a^b G(t,s) f(s, x(s)) ds, \tag{74}$$

y aplicamos la Proposición 2.24 con μ pequeño.

2.9. Sistemas de ecuaciones integrales y ecuaciones diferenciales semilineales

Los Teoremas de punto fijo de Banach y Schauder pueden ser aplicados sin dificultad a los sistemas de ecuaciones integrales haciendo uso de los *espacios producto*. Por ejemplo, consideramos el problema de valor inicial

$$\begin{cases} \xi_i'(t) = f_i(t, x(t)), \\ \xi_i(t_0) = \eta_{i0}, \quad i = 1, ..., n, \end{cases} \tag{75}$$

donde $x(t) = (\xi_1(t), ..., \xi_n(t))$. Para f_i continua obtenemos el sistema equivalente de ecuaciones integrales

$$\xi_i = \eta_{i0} + \int_{t_0}^{t} f_i(s, x(s))ds, \quad i = 1, ..., n. \tag{76}$$

Como B-espacio elegimos el conjunto de todos los $x(\cdot)$ para los que las componentes $\xi_i(\cdot)$ son continuas en $[t_0 - c, t_0 + c]$. Para la norma, usaremos

$$||x|| = \max_{1 \le i \le n} (\max_{a \le t \le b} |\xi_i(t)|).$$

Entonces, análogamente con el Teorema 1.11 y la Proposición 2.24, se demuestra la existencia de soluciones para f_i continuamente Lipschitziana o continuas en un intervalo $[t_0 - c, t_0 + c]$, para c pequeño.

Las ecuaciones integrales no lineales son también una herramienta para investigar las ecuaciones diferenciales parciales semilineales. Por ejemplo, si L es un operador diferencial lineal de segundo orden, un problema de valor inicial o de frontera para

$$Lu(x) = f(x, u(x))$$

puede convertirse, análogamente con (74), en

$$u = Jf(\cdot, u),$$

donde J es un operador integral y este último puede resolverse con ayuda de los Teoremas de punto fijo (el de Banach para f continuamente Lipschitziana y el de Schauder para f continua).

La reducción a una ecuación integral es el enfoque clásico. Podemos verlo también como un operador $L : X \to Y$ entre dos B-espacios, y tener estimaciones sobre L de la forma

$$||Lu|| \ge \gamma ||u||, \quad \gamma > 0,$$

que garantiza entonces la existencia de un operador lineal y continuo $L^{-1} : Y \to X$. Entonces $J = L^{-1}$.

2.10. Estrategia general

Un estudio cuidadoso de la demostración de la Proposición 2.24 descubre las siguiente observaciones sobre la aplicación de los Teoremas de punto fijo de Banach y de Schauder a las ecuaciones integrales.

(a) Para obtener una aplicación de una bola en sí misma, con centro el origen, en un B-espacio de funciones continuas, podemos hacer una de las dos opciones siguientes:

- Elegir los parámetros dados suficientemente pequeños.

- Elegir el radio de la bola suficientemente pequeño si los parámetros son de orden superior a 1.

(b) Si la región de integración es compacta, y si todas las funciones dentro de la integral son continuas, entonces obtenemos operadores compactos sobre B-espacio de funciones continuas, y podemos aplicar el Teorema del punto fijo de Schauder.

(c) Si la región de integración es compacta, y si todas las funciones dentro de la integral son continuamente integrables, o más generalmente, continuamente Lipschitziana, entonces podemos aplicar el Teorema del punto fijo de Banach. Para forzar la k-contractividad, podemos elegir una de las siguientes opciones:

- Elegir los parámetros dados suficientemente pequeños.

- Elegir el radio r de la bola suficientemente pequeño, si la constante de Lipschitz para $r \to 0$ también va a 0 (parámetros libres de orden mayor que 1).

Para el caso de una ecuación integral de Volterra con un límite superior de integración variable, o más generalmente, con una región de integración variable, la situación se simplifica sustancialmente. En (a) y (c) basta elegir la región de integración suficientemente pequeña. Este método ya se utilizó en la demostración de los Teoremas de Picard-Lindelöf y Peano en las Secciones 1.4 y 2.7. Este sencillo truco puede aplicarse también a la ecuación integral (70), uno se conforma con los resultados de existencia para un intervalo suficientemente pequeño $[a,b]$. En este caso, se prescinde de las condiciones restrictivas (71) y (72). Para ecuaciones diferenciales semilineales, las consideraciones de la Sección 2.9 muestran que esta técnica aportará resultados de existencia para regiones suficientemente pequeñas.

2.11. Principio de existencia de sistemas de inecuaciones

Queremos encontrar un $u \in K$ que sea solución de la inecuación

$$\langle f - Tu, v - u \rangle_X \geq 0, \quad \text{para todo } (v, f) \in M. \tag{77}$$

Proposición 2.26 (Debrunner,Flor (1964)). *Sea K un subconjunto convexo compacto no vacío de un B-espacio X y sea M un subconjunto monótono de $K \times X^*$, lo que significa que*

$$\langle f - g, v - w \rangle_X \geq 0, \quad \text{para todo } (v, f), (w, g) \in M. \tag{78}$$

Además, sea la aplicación $T : K \to X^*$ continua. Entonces (77) tiene una solución $u \in K$.

Veamos el significado de esta proposición con un ejemplo.

Ejemplo 2.27 Sea $X = \mathbb{R}$, $K = [a, b]$, $-\infty < a$, $b < \infty$, y sea la función $\varphi : [a, b] \to \mathbb{R}$ monótona creciente. Entonces $X = X^*$. Ahora la gráfica de φ es monótona, es decir, el conjunto $M = \{(v, \varphi(v)) : v \in [a, b]\}$ es monótono, ya que

$$\langle f - g, v - w \rangle_X = (f - g)(v - w) \geq 0 \quad \text{para todo } (v, f), (w, g) \in M.$$

El problema (77) es equivalente a

$$(\varphi(v) - T(u))(v - u) \geq 0 \quad \text{para todo } v \in [a, b].$$

La Proposición 2.26 dice que este problema tiene una solución $u \in [a, b]$, para cada función continua $T : [a, b] \to \mathbb{R}$. La Figura 2.3 da una interpretación geométrica de la proposición para el ejemplo anterior.

Demostración de la Proposición 2.26: Supongamos, por el contrario, que (77) no tiene solución $u \in K$.
Para la primera parte de la demostración usaremos:

Teorema 2.28 (Teorema de partición de la unidad). *Sea $f : M \subseteq X \to \mathbb{R}$ una función de valores reales. El soporte de f, denotado por $sop(f)$, es el cierre del conjunto $\{x \in M : f(x) \neq 0\}$.*
Sea X un espacio paracompacto (por ejemplo, un espacio métrico o un espacio topológico compacto). Sea $\{U_\alpha\}$ cualquier recubrimiento abierto de X. Entonces hay una partición de la unidad $\{f_\alpha\}$ contenida en $\{U_\alpha\}$.
Más concretamente, hay un sistema de funciones continuas de valores reales $\{f_\alpha\}$ con las siguientes propiedades:

(i) $0 \leq f_\alpha \leq 1$ *para todo* $x \in X$ *y todo* α.

(ii) $\sum_\alpha f_\alpha = 1$ *para todo* $x \in X$.

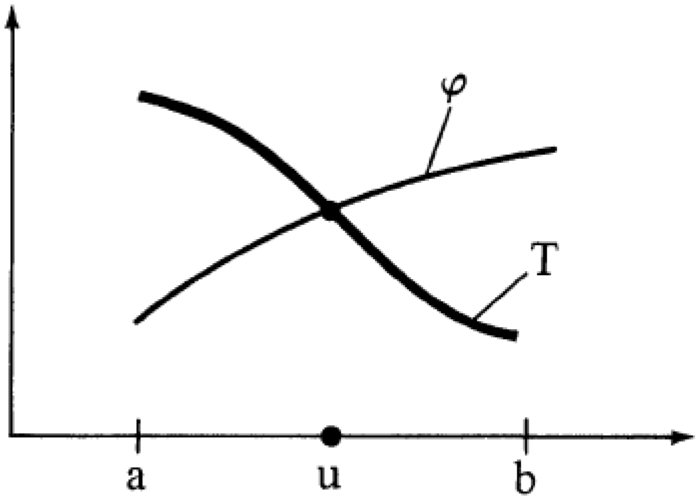


Figura 2.3: Interpretación geométrica de la Proposición 2.26 en el Ejemplo 2.27

(iii) *El recubrimiento* $\{sop(f_\alpha)\}$ *de* X *es localmente finito, es decir, para cada* $x \in X$ *un entorno* $U(x)$ *tal que un número finito de* f_α *no son idénticamente nulos en* $U(x)$.

(iv) *El recubrimiento* $\{supp(f_\alpha)\}$ *de* X *es una afinación de* $U(x)$*, es decir, para cada* α *existe un* β *con* $sop(f_\alpha) \subseteq U_\beta$*. Si sólo hay un número finito de* $U_1, \ldots, U_n$*, entonces* $f_1, \ldots, f_n$ *puede ser elegido para que* $sop(f_i) \subseteq U_i$ *para todo* i.

Y además, vamos a definir lo que es un conjunto relativamente abierto:

Definición 2.29 Sea X un espacio métrico y sea $N \subseteq M \subset X$, se dice que N es *relativamente abierto* en M si y solo si existe O abierto en X tal que $N = M \cap O$.

(I) *Partición de la unidad.* Definimos el conjunto

$$U(v,f) = \{u \in K : \langle f - Tu, v - u\rangle_x < 0\}.$$

$U(v,f)$ es relativamente abierto en K, ya que $\langle f - Tu, v - u\rangle_x$ cambia ligeramente solo si $\|f - Tu\|$ y $\|v - u\|$ cambian ligeramente. Como (77) no tiene solución por hipótesis, la colección de $U(v,f)$, para todo $(v,f) \in M$, forma un recubrimiento del conjunto K. Pero K es compacto, por lo que se puede extraer un subrecubrimiento finito de K, $U(v_i, f_i)$, $i = 1, \ldots, m$. Existe entonces, por el Teorema de partición de la unidad, una subpartición ordenada de la unidad, es decir, existen funciones continuas

$$\begin{gathered}\beta_i : K \to \mathbb{R}, \text{ con } supp\beta_i \subseteq U(v_i, f_i), \text{ tales que} \\ \sum_{i=1}^{m} \beta_i(u) = 1,\ 0 \le \beta_i(u) \le 1, \text{ para todo } u \in K, i = 1, \ldots, m.\end{gathered} \tag{79}$$

(II) *Teorema del punto fijo de Brouwer.* Sea $K_1 = co(\{v_1, ..., v_m\})$, y definimos

$$p(u) = \sum_{i=1}^{m} \beta_i(u) v_i, \quad q(u) = \sum_{i=1}^{m} \beta_i(u) f_i, \quad \text{para todo } u \in K_1.$$

Por (79), p aplica el conjunto K_1 en sí mismo continuamente. Por el Teorema 2.7 p tiene un punto fijo, $u^* = p(u^*)$, $u^* \in K_1$.

(III) *Construcción de la contradicción.* Sea

$$\Delta_{ij} = \langle f_i - Tu^*, v_j - u^* \rangle_x.$$

Entonces

$$\Delta_{ij} + \Delta_{ji} = \Delta_{ii} + \Delta_{jj} + \langle f_i - f_j, v_j - v_i \rangle_x.$$

Por (78), $\langle f_i - f_j, v_j - v_i \rangle \leq 0$, de modo que

$$\Delta_{ij} + \Delta_{ji} \leq \Delta_{ii} + \Delta_{jj}.$$

De $p(u^*) = u^*$ y de la definición de p y q llegamos a que

$$\begin{aligned} 0 &= \langle q(u^*) - Tu^*, p(u^*) - u^* \rangle = \\ &= \sum_{i,j=1}^{m} \beta_j(u^*) \beta_j(u^*) \Delta_{ij} = \\ &= \sum_{i,j=1}^{m} \beta_j(u^*) \beta_j(u^*) (\Delta_{ij} + \Delta_{ji})/2, \end{aligned}$$

por lo que

$$0 \leq \sum_{i,j=1}^{m} \beta_j(u^*) \beta_j(u^*) (\Delta_{ii} + \Delta_{jj})/2, \tag{80}$$

Si $\beta_i(u^*)\beta_j(u^*) \neq 0$, entonces $u^* \in U(v_i, f_i) \cap U(v_j, f_j)$. Por construcción de $U(v, f)$, esto significa que $\Delta_{ii} < 0$ y $\Delta_{jj} < 0$. En vista de (80), esto significa que $\beta_i(u^*) = 0$ para todo $i = 1, ..., m$. Puesto que $u^* \in K$, esto contradice (79). $\square$

Bibliografía

[1] Zeidler E., *Nonlinear Functional Analysis and its Applications, Fixed-Point Teorems*, Springer-Verlag, New York (1986).

[2] Yosida K., *Functional Analysis*, Springer-Verlag, Berlin (1971).

[3] Smart D.R., *Fixed Point Theorems*, 1º edition, Cambridge University Press, Cambridge, UK (1974).

[4] Intrătescu V.I., *Fixed Point Theory, An Introduction*, D. Reidel Pubishing Company, Dordrecht, Holland (1981).