UNIVERSIDAD DE SEVILLA

Trabajo de Fin de Grado

Teorema de puntos fijos con aplicaciones

Presentado por:

Manuel Bernardino del Pino

Dirigido por:

María Ángeles Japón Pineda

Departamento de Análisis Matemático

JUNIO DE 2019

"Si la gente no cree que las matemáticas son simples, es solo porque no se dan cuenta de lo complicada que es la vida".

J. L. von Neumann.

Abstract

The main aim of this work is to develop several remarkable Fixed Point Theorems and apply them to different fields in Mathematical Analysis. In this context, we first consider complete metric spaces and we study the classical Banach Contraction Principle as well as several of its extensions. We apply some of these results to differential equations, fractal theory, Banach algebras and finally to convex metric spaces.

In the next chapter, we switch to the Banach space framework and more generally to locally convex vectors spaces. We prove Schauder-Tychovoff Theorem, where the main features to obtain a fixed point are now the compactness and convexity of the domain together with the continuity of the mapping. This will lead us to the study of invariant measures for a continuous function defined on a compact Hausdorff space and to approach one of the most famous open problems inside Operator Theory: the Invariant Subspace Problem.

In the third chapter, we study Markov-Kakutani Theorem to achieve a common fixed point for a commutative family of affine continuous functions. As applications, we prove the existence of a Haar measure on a compact topological group as well as the existence of invariant means on every commutative semigroup.

Índice general

Prólogo

La teoría métrica del punto fijo trata de establecer bajo qué condiciones una aplicación T definida en un espacio métrico (M, d) tiene punto fijo, es decir, el objetivo es hallar un elemento $x \in M$ tal que $Tx = x$. En este ámbito podemos decir que el teorema de punto fijo más conocido es el de Banach (1892 - 1945), el cual asegura que toda contracción definida en un espacio métrico completo en sí mismo tiene un único punto fijo. Además la demostración de este resultado es constructiva, ya que nos permite obtener tal punto fijo mediante iteraciones con una estimación del error.

Si abandonamos la noción de distancia y nos centramos en la topología del espacio podemos encontrar una nueva teoría de punto fijo basada en la continuidad de la aplicación y compacidad del dominio. En este contexto, el Teorema de Schauder-Tychonoff nos establece que para un espacio localmente convexo X, cualquier aplicación continua definida de un convexo en sí mismo con rango compacto tiene punto fijo.

Si nos planteamos el problema más general de encontrar puntos fijos comunes a una familia conmutativa de aplicaciones continuas de un compacto y convexo en sí mismo encontramos que el Teorema de Schauder-Tychonoff no se cumple ni siquiera para dos aplicaciones con estas características en el intervalo [0,1] (ver [2]). Resulta que si añadimos la hipótesis extra de que las aplicaciones sean afines entonces podemos encontrar punto fijo común, hecho que se conoce como Teorema de Markov-Kakutani.

En el primer capítulo de esta memoria presentaremos el Teorema de punto fijo de Banach y veremos algunas variantes de éste. Además veremos su relación con la moderna teoría de fractales, ecuaciones diferenciales, raíces cuadradas en álgebras de Banach y espacios métricos convexos.

En el segundo capítulo probaremos el Teorema de Schauder-Tychonoff y veremos cómo aplicarlo en el problema del subespacio invariante y para obtener funciones que conserven la medida en ciertos espacios.

En el tercer, y último capítulo, estudiaremos el Teorema de Markov-Kakutani y lo aplicaremos para obtener una medida invariante bajo la acción de un grupo (medida de Haar) y "medias" invariantes en semigrupos.

No quiero pasar la ocasión sin agradecerle a mi tutora, la Prof. M.Ángeles Japón Pineda todo su tiempo, paciencia y dedicación.

Y por último, y no menos importante, quiero agradecer a mi familia y amigos por sus apoyos, cariño y ánimos en los momentos difíciles. Gracias por todo.

El autor

Capítulo 1

Teorema de punto fijo de Banach. Aplicaciones.

El teorema de punto fijo de Banach aparece de forma explícita en la tesis de Banach en 1922 donde se usaba para establecer la existencia de solución de una ecuación integral. Debido a su simpleza y utilidad, ha llegado a ser una herramienta muy popular para resolver problemas de existencia en muchas ramas del análisis matemático. En este capítulo probaremos el teorema de punto fijo de Banach, veremos algunas variantes de éste y daremos algunas aplicaciones tales como:

1. Resolución de ecuaciones diferenciales.

2. Conjuntos autosimilares (fractales).

3. Raíces cuadradas en álgebras de Banach.

4. Espacios métricos convexos.

En el desarrollo de este tema hemos seguido principalmente la referencia de [5].

1.1. Teorema de punto fijo de Banach. Extensiones.

Comencemos con un espacio métrico (M, d) siendo d una función distancia (métrica). A menudo escribiremos sólo M para abreviar notación.

Definición 1.1.1. Se dice que una aplicación $T : M \to M$ es *lipschitziana* si existe una constante $k > 0$ tal que:

$$d(Tx, Ty) \leq kd(x, y), \quad \forall\, x, y \in M. \tag{1.1}$$

La k más pequeña que verifica la desigualdad (1.1) se denomina constante de Lipschitz de T.

Nota 1.1.2. A menudo denotaremos por $k(T)$ a la constante de Lipschitz de una aplicación T ó $k_d(T)$ para hacer referencia a la métrica d.

Proposición 1.1.3. Sean $S, T : M \to M$ dos aplicaciones. Se tiene que:

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

(b) $k(T^n) \leq k^n(T)$, para todo $n \in \mathbb{N}$.

Definición 1.1.4. Una aplicación $T : M \to M$ lipschitziana se dice que es una *contracción* si $k(T) < 1$. Más precisamente, T es una contracción con respecto a la métrica d si $k_d(T) < 1$.

Teorema 1.1.5 (Banach)**.** Sea (M, d) un espacio métrico completo y sea $T : M \to M$ una contracción. Entonces T tiene un único punto fijo en M, y para cada $x_0 \in M$ la sucesión de iteradas $\{T^n x_0\}$ converge a dicho punto fijo.

Demostración. Denotaremos $k = k_d(T)$. Tomemos ahora $x_0 \in M$ y definamos la sucesión $\{x_n\}$ como $x_{n+1} = Tx_n$ (equivalentemente, $x_n = T^n x_0$), $n \in \mathbb{N}$. Observar que para cualesquiera índices n, $p \in \mathbb{N}$,

$$\begin{aligned} d(x_n, x_{n+p}) &= d(T^n x_0, T^{n+p} x_0) = d(T^n x_0, T^n \circ T^p x_0) \leq k(T^n) d(x_0, T^p x_0) \\ &\leq k^n [d(x_0, Tx_0) + d(Tx_0, T^2 x_0) + \cdots + d(T^{p-1} x_0, T^p x_0)] \\ &\leq k^n (1 + k + k^2 + \cdots + k^{p-1}) d(x_0, Tx_0) \\ &= k^n \left(\frac{1 - k^p}{1 - k} \right) d(x_0, Tx_0) \leq \frac{k^n}{1 - k} d(x_0, Tx_0). \end{aligned} \tag{1.2}$$

Esto muestra que $\{x_n\}$ es una sucesión de Cauchy, y como M es completo, existe un $x \in M$ tal que $\lim_{n \to \infty} x_n = x$.

Para ver que x es el único punto fijo de T, observar que

$$x = \lim_{n\to\infty} x_n = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} Tx_n \overset{(*)}{=} Tx$$

donde en $(*)$ se ha usado que T es continua (por ser una contracción).
Además, $x = Tx$ e $y = Ty$ implican que

$$d(x, y) = d(Tx, Ty) \leq kd(x, y)$$

obteniendo $d(x, y) = 0$, es decir, $x = y$, y por tanto T tiene un único punto fijo en M. Para terminar, haciendo $p \to \infty$ en (1.2) se obtiene la siguiente cota del error de estimación:

$$d(x_n, x) = d(T^n x_0, x) \leq \frac{k^n}{1-k} d(x_0, Tx_0). \tag{1.3}$$

□

Nota 1.1.6. El Teorema 1.1.5 no es cierto si el espacio métrico (M, d) no es completo. Basta tomar $M = (0, 1]$ (que no es completo pues la sucesión $\{1/n\}$ con $n \in \mathbb{N}$ es de Cauchy pero no es convergente en M) con la métrica euclídea y $T : M \to M$ definida por $Tx = x/2$. Entonces T es una contracción con constante $k(T) = 1/2$ pero no tiene puntos fijos en M.

Veamos un ejemplo de aplicación:

Ejemplo 1.1.7. Consideremos $T : \ell^2 \to \ell^2$ definida por $Tx = x/2+(1, 1/2, 1/3, \ldots)$. Entonces es fácil ver que T es una contracción con constante $k(T) = 1/2$ y que su único punto fijo es $(2/1, 2/2, 2/3, \ldots)$.

Nota 1.1.8. Un análisis de la prueba del Teorema 1.1.5 muestra que la hipótesis $k(T) < 1$ es más fuerte de lo necesario. Es suficiente con asumir que $k(T^n) < 1$ para al menos un $n \in \mathbb{N}$. Esto implica que T^n es una contracción y por el teorema de Banach tiene un único punto fijo x. Pero $Tx = T^{n+1}x = T^n \circ Tx$, así que Tx es tambíen un punto fijo de T^n. Por lo tanto $Tx = x$, y como consecuencia x también es un punto fijo de T (y el único).
No es difícil encontrar encontrar ejemplos de aplicaciones tales que $k(T^n) < 1$ mientras que $k(T) \geq 1$.

Ejemplo 1.1.9. Consideremos la aplicación $T : \mathcal{C}[0,1] \to \mathcal{C}[0,1]$ definida como $(T\varphi)(t) = 1 + \int_0^t 2s\varphi(s)\,ds$ donde $\mathcal{C}[0,1]$ es el espacio de las funciones continuas de variable real en $[0,1]$ con la norma del supremo. Veamos que $k(T) = 1$.
Sean $\varphi_1, \varphi_2 \in \mathcal{C}[0,1]$, entonces $|T\varphi_1(t) - T\varphi_2(t)| = |1 + \int_0^t 2s\varphi_1(s)\,ds - 1 - \int_0^t 2s\varphi_2(s)\,ds| \leq \int_0^t 2s|\varphi_1(s) - \varphi_2(s)|ds \leq 2\,\|\varphi_1 - \varphi_2\| \int_0^t s\,ds = t^2\,\|\varphi_1 - \varphi_2\| \leq \|\varphi_1 - \varphi_2\|$. Por tanto, $\|T\varphi_1 - T\varphi_2\| \leq \|\varphi_1 - \varphi_2\|$ y $k(T) = 1$.
Veamos que, sin embargo, $k(T^2) < 1$:
$(T^2\varphi)(t) = T(T\varphi)(t) = 1 + 2\int_0^t s(T\varphi)(s)\,ds = 1 + 2\int_0^t s(1 + 2\int_0^s \sigma\varphi(\sigma)\,d\sigma)\,ds = 1 + 2\int_0^t s\,ds + 4\int_0^t s\int_0^s \sigma\varphi(\sigma)\,d\sigma\,ds = 1 + t^2 + 4\int_0^t\int_0^s s\,\sigma\varphi(\sigma)\,d\sigma\,ds =$ (Teorema de Fubini) $= 1 + t^2 + 4\int_0^t\int_\sigma^t s\,\sigma\varphi(\sigma)\,ds\,d\sigma = 1 + t^2 + 2\int_0^t \sigma\varphi(\sigma)(t^2 - \sigma^2)\,d\sigma$.
Sean $\varphi_1, \varphi_2 \in \mathcal{C}[0,1]$, entonces $|T^2\varphi_1(t) - T^2\varphi_2(t)| = 2|\int_0^t \sigma(t^2 - \sigma^2)(\varphi_1(\sigma) - \varphi_2(\sigma))\,d\sigma| \leq 2\,\|\varphi_1 - \varphi_2\| \int_0^t \sigma(t^2 - \sigma^2)\,d\sigma = \frac{t^4}{2}\,\|\varphi_1 - \varphi_2\| \leq \frac{1}{2}\,\|\varphi_1 - \varphi_2\|$. Luego, $\left\|T^2\varphi_1 - T^2\varphi_2\right\| \leq \frac{1}{2}\,\|\varphi_1 - \varphi_2\|$ y $k(T^2) = \frac{1}{2} < 1$.

Nota 1.1.10. Sea $T : M \to M$ una aplicación lipschitziana, y fijemos $x_0 \in M$, y sea $x_n = T^n x_0$. Obervermos que $d(x_n, x_{n+p}) \leq \sum_{i=n}^{n+p-1} d(T^i x_0, T^{i+1} x_0) \leq \sum_{i=n}^{n+p-1} k(T^i) d(x_0, Tx_0) \leq \sum_{i=n}^{\infty} k(T^i) d(x_0, Tx_0)$. Así $\{x_n\}$ es una sucesión de Cauchy si

$$\sum_{i=0}^{\infty} k(T^i) < +\infty. \tag{1.4}$$

Usando que $k(T^n)$ es multiplicativa, es decir, $k(T^{n+m}) \leq k(T^n)k(T^m)$ es fácil ver que existe un número $k_\infty(T)$ que satisface:

$$k_\infty(T) = \lim_{n\to\infty} [k(T^n)]^{1/n} = \inf\{[k(T^n)]^{1/n} : n = 1, 2, \ldots\}. \tag{1.5}$$

Veamos que (1.4) se tiene si y solo si $k_\infty(T) < 1$:
La implicación hacia la izquierda ($k_\infty(T) < 1 \Rightarrow$ (1.4)) es una consecuencia inmediata del criterio de la raíz. Para ver la otra implicación, basta probar que $a_n := [k(T^n)]^{1/n} < 1$ para algún $n \in \mathbb{N}$. En efecto, si se verifica (1.4), por la condición necesaria de convergencia se tiene que $\lim_{n\to\infty} k(T^n) = 0$, lo cual implica que existe un $n_0 \in \mathbb{N}$ tal que $k(T^n) < 1$ para todo $n \geq n_0$. Es decir se tiene $\inf\{[k(T^n)]^{1/n}\} < 1$ o equivalentemente que $k_\infty(T) < 1$.

Así la hipótesis $k(T) < 1$ en el Teorema 1.1.5 puede ser reemplazada por $k_\infty(T) < 1$.

Por supuesto, la pregunta sigue siendo si la hipótesis más débil $k_\infty(T) < 1$ produce una versión más fuerte del Teorema 1.1.5.

Definición 1.1.11. Diremos que dos métricas d y r definidas en un conjunto dado M son equivalentes si existen dos constantes positivas a y b tales que para todos $x, y \in M$, $ar(x,y) \leq d(x,y) \leq br(x,y)$.

Proposición 1.1.12. Sea M un espacio métrico completo. Sea $T : M \to M$ una aplicación para la cual $k_\infty(T) < 1$. Entonces existe una métrica equivalente para la cual T es una contracción.

Demostración. No es difícil probar que que para dos métricas equivalentes d y r, cualquier sucesión de Cauchy respecto de r es también de Cauchy con respecto de d (y viceversa). Como consecuencia, (M, d) es completo si y sólo si, (M, r) lo es.

Para una aplicación d-lipschitziana $T : M \to M$, la Definición 1.1.11 implica que

$$r(Tx, Ty) \leq \frac{1}{a} d(Tx, Ty) \leq \frac{1}{a} k_d(T) d(x, y) \leq \frac{b}{a} k_d(T) r(x, y)$$

y por tanto $k_r(T) \leq \frac{b}{a} k_d(T)$. De forma similar, $k_d(T) \leq \frac{b}{a} k_r(T)$ y así para cualquier $n \in \mathbb{N}$,

$$\frac{a}{b} k_d(T^n) \leq k_r(T^n) \leq \frac{b}{a} k_d(T^n).$$

Por consiguiente, $\lim_{n\to\infty}[k_r(T^n)]^{1/n} = \lim_{n\to\infty}[k_d(T^n)]^{1/n}$, y esto muestra que $k_\infty(T)$ es constante respecto a métricas equivalentes. Además, en vista de (1.5), $k_\infty(T) \leq k_r(T)$ para todas las métricas r equivalentes a d.

Por otro lado, hemos de observar que para cualquier $\lambda \in [0, \frac{1}{k_\infty(T)})$ la serie $r_\lambda(x, y) = \sum_{n=0}^{\infty} \lambda^n d(T^n x, T^n y)$ converge y genera una métrica r_λ, equivalente a d:

$$d(x, y) \leq r_\lambda(x, y) \leq (\sum_{n=0}^{\infty} k_d(T^n)\lambda^n) d(x, y).$$

Así

$$r_\lambda(Tx, Ty) = \sum_{n=0}^{\infty} \lambda^n d(T^{n+1}x, T^{n+1}y) = \frac{1}{\lambda}\left[r_\lambda(x, y) - d(x, y)\right] \leq \frac{1}{\lambda} r_\lambda(x, y).$$

Por tanto, $k_{r_\lambda}(T) \leq 1/\lambda$. Tomando $\lambda = 1/\left[k_\infty(T) + \varepsilon\right], \varepsilon > 0$, tenemos $k_{r_\lambda}(T) \leq k_\infty(T) + \varepsilon$ y así concluimos que $k_\infty(T) = \text{ínf}\, k_r(T)$ donde el ínfimo es sobre todas las métricas r equivalentes a d. $\square$

En principio, la hipótesis $k_\infty(T) < 1$ no genera una versión más fuerte del Teorema 1.1.5. Sin embargo, como ya veremos, la elección de una métrica apropiada es a menudo útil en algunas aplicaciones ya que proporciona mejores estimaciones en la velocidad de convergencia de la sucesión de iteradas.
La pregunta más obvia que surge cuando se estudian las contracciones es la siguiente: ¿Qué ocurre si $k(T) = 1$?. El simple ejemplo $Tx = x + 1$ (una traslación) para $x \in \mathbb{R}$ muestra que la tesis del Teorema de punto fijo de Banach no se cumple. Podemos definir por tanto una clase de aplicaciones que caen entre las contracciones y las aplicaciones que verifican $k(T) = 1$.

Definición 1.1.13. Una aplicación $T : M \to M$ se dice que es *contractiva* si se verifica que para todo $x, y \in M$ con $x \neq y$ se tiene que $d(Tx, Ty) < d(x, y)$.

Teorema 1.1.14 (Edelstein)**.** Sea (M, d) un espacio métrico compacto y sea $T : M \to M$ una aplicación contractiva. Entonces T tiene un único punto fijo en M, y para cada $x_0 \in M$ la sucesión de iteradas $\{T^n x_0\}$ converge a dicho punto fijo.

Demostración. Consideremos la función $\varphi : M \to \mathbb{R}^+$ definida por $\varphi(y) = d(y, Ty)$ que es continua en M y por tanto, por compacidad tiene mínimo, digamos $x \in M$. Si $x \neq Tx$ entonces $\varphi(Tx) = d(Tx, T^2x) < d(x, Tx)$ y llegamos a una contradicción. Así que ha de ser $x = Tx$. Tomemos ahora $x_0 \in M$ y la sucesión $a_n := d(T^n x_0, x)$. Como $a_{n+1} = d(T^{n+1}x_0, x) = d(T^{n+1}x_0, Tx) < d(T^n x_0, x) = a_n$, tenemos que la sucesión $\{a_n\}$ es decreciente y está acotada inferiormente ($a_n \geq 0, \forall n \in \mathbb{N}$), por tanto tiene límite, llamémoslo a.
Por otro lado, por compacidad , $\{T^n x_0\}$ tiene una subsucesión convergente $\{T^{n_k} x_0\}$; digamos $\text{lím}_{k\to\infty} T^{n_k} x_0 = z$. Obviamente $d(z, x) = a$. Si $a > 0$, entonces obtenemos una contradicción ya que:

$$a = \lim_{k\to\infty} d(T^{n_k+1}x_0, x) = d(Tz, x) = d(Tz, Tx) < d(z, x) = a.$$

Así se deduce que $a = 0$. Además cualquier sucesión convergente de $T^n x_0$ debe converger a x, así por compacidad, $\lim_{n\to\infty} T^n x_0 = x$.
La unicidad se prueba de forma análoga como en el Teorema 1.1.5. $\square$

Nota 1.1.15. La hipótesis de compacidad del Teorema 1.1.14 no puede ser reemplazada por completitud y acotación. Veámoslo en el siguiente ejemplo:

Ejemplo 1.1.16. Recordemos que $\mathcal{C}[0,1]$ es el espacio de las funciones continuas de variable real en $[0,1]$ con la norma del supremo, es decir, para $x \in \mathcal{C}[0,1]$

$$\|x\| = \sup\{|x(t)| : t \in [0,1]\}.$$

Sea $M = \{x \in \mathcal{C}[0,1] : 0 = x(0) \leq x(t) \leq x(1) = 1\}$. M es cerrado y acotado, y como $\mathcal{C}[0,1]$ es completo con la métrica inducida por la norma anterior, M también es completo. Se puede comprobar que la aplicación $T : M \to M$ definida por $(Tx)(t) = tx(t), x \in M, t \in [0,1]$ es contractiva pero sin puntos fijos.

La simpleza y el uso del Teorema de punto fijo de Banach ha inspirado a muchos autores a analizarlo con más detalle. Estos estudios han llevado a generalizaciones y modificaciones de dicho teorema, uno de los resultados más profundos es debido a Caristi. Nosotros vamos a derivar este teorema a partir de la siguiente reformulación de un resultado de Brezis y Browder (1976) que usa la noción de orden parcial. Empecemos recordando dicho concepto:

Definición 1.1.17. Sea X un conjunto. Diremos que una relación $\leq$ en X es un *orden parcial* si se verifican las siguientes condiciones:

(a) Reflexividad: $a \leq a$ para todo $a \in X$.

(b) Transitividad: $a \leq b$ y $b \leq c$ implican $a \leq c$ para todos $a, b, c \in X$.

(c) Antisimetría: $a \leq b$ y $b \leq a$ implican $a = b$ para todos $a, b \in X$.

Diremos que $(X, \leq)$ es un *conjunto parcialmente ordenado.*

Teorema 1.1.18. Sea X un conjunto parcialmente ordenado, y para $x \in X$, definamos $S(x) := \{y \in X : y \geq x\}$. Sea $\psi : X \to \mathbb{R}$ una función que satisface:

(a) $x \leq y$ y $x \neq y$ implica $\psi(x) < \psi(y)$.

(b) Para cualquier sucesión creciente $\{x_n\}$ en X que verifique $\psi(x_n) \leq C < +\infty$ para todo n, existe algún $y \in X$ tal que $x_n \leq y$ para todo n.

(c) Para cada $x \in X$, $\psi(S(x))$ está acotado superiormente.

Entonces para cada $x \in X$ existe $x' \in S(x)$ tal que x' es maximal, es decir, $\{x'\} = S(x')$.

Demostración. Para $a \in X$, sea $\rho(a) = \sup\{\psi(b) : b \in S(a)\}$. Supongamos que la conclusión del teorema no es cierta para algún $x \in X$ y definamos por inducción una sucesión $\{x_n\}$ tal que $x_1 = x$ y $x_{n+1} \in S(x_n)$ y satisfaga $\rho(x_n) \leq \psi(x_{n+1}) + 1/n$ para todo $n \in \mathbb{N}$.
Como $\psi(x_{n+1}) \leq \rho(x) < +\infty$, se tiene por (b) que existe algún $y \in X$ tal que $x_n \leq y$ para todo n. También, por hipótesis, y no es maximal en $S(x)$, así que existe $u \in X$ tal que $y \leq u$ y $\psi(y) < \psi(u)$. Como $x_n \leq u$, entonces se cumple que $\psi(u) \leq \rho(x_n)$ para todo n. También $x_{n+1} \leq y$; así $\psi(x_{n+1}) \leq \psi(y)$. Por tanto, $\psi(u) \leq \rho(x_n) \leq \psi(x_{n+1}) + 1/n \leq \psi(y) + 1/n$ para todo n, así que $\psi(u) \leq \psi(y)$ y llegamos a una contradicción. $\square$

Antes de enunciar el Teorema de Caristi, veamos una clase de funciones que nos serán de utilidad para dicho resultado.

Definición 1.1.19. Sea M un espacio métrico y $f : M \to M$. Diremos que f es *semicontinua inferiormente* si para cualquier sucesion $\{x_n\}$ en M y $x \in M$ con $x_n \to x$ se cumple que $f(x) \leq \underline{\lim} f(x_n)$.

Teorema 1.1.20 (Caristi)**.** Sea (M, d) un espacio métrico completo y sea $\varphi : M \to \mathbb{R}$ una función semicontinua inferiormente y acotada inferiormente. Supongamos que $T : M \to M$ es una aplicación arbitraria que satisface:

$$d(u, Tu) \leq \varphi(u) - \varphi(Tu), \quad u \in M.$$

Entonces T tiene punto fijo.

Demostración. Para comenzar definamos el siguiente orden parcial:

$$\text{Dados} \quad x, y \in M,\, x \leq y \quad \text{si} \quad d(x, y) \leq \varphi(x) - \varphi(y).$$

Vamos a aplicar el Teorema 1.1.18 con $\psi = -\varphi$. Notemos que por hipótesis $u \leq T(u)$ para todo $u \in M$. Debemos verificar entonces las condiciones (a), (b) y (c) de dicho teorema. El apartado (a) es obvio, y para ver (b) observar que si $\{x_n\}$ es cualquier sucesión creciente entonces $\{\varphi(x_n)\}$ es decreciente y acotada inferiormente; por tanto $\{\varphi(x_n)\}$ converge, digamos a $r \in \mathbb{R}$. Esto a su vez implica que $\{x_n\}$ es una sucesión de Cauchy. Así $\{x_n\}$ converge a un punto $y \in M$, y como φ es semicontinua inferiormente se sigue que:

$$d(x_n, y) \leq \varphi(x_n) - r \leq \varphi(x_n) - \varphi(x).$$

Por lo tanto, $x_n \leq y$ para todo $n \in \mathbb{N}$ y así se tiene (b).
Como (c) se tiene del hecho de que φ está acotada inferiormente, concluímos que para cada $x \in M$ existe $x' \geq x$ tal que $T(x') = x'$. □

Nota 1.1.21. A partir del Teorema 1.1.20 es posible probrar el Teorema de punto fijo de Banach tomando como función $\varphi(x) = (1-k)^{-1}d(x, Tx)$. Para más detalles ver [5]

1.2. Ejemplos y aplicaciones.

En esta sección vamos a dar algunas aplicaciones del Teorema de Banach en distintas ramas del análisis y veremos una aplicación del Teorema de Caristi.

1.2.1. Ecuaciones diferenciales.

Vamos a empezar con el poblema de Cauchy clásico de existencia y unicidad de solución de una ecuación diferencial satisfaciendo una condición inicial.

Ejemplo 1.2.1. Sea $f(t, s)$ una función continua de variable real definida para $t \in [0, T], T > 0$ y $s \in \mathbb{R}$. El problema de Cauchy consiste en encontrar una función continuamente diferenciable x en $[0, T]$ satisfaciendo la ecuación diferencial:

$$\begin{cases} x'(t) = f(t, x(t)), \quad t \in [0, T] \\ x(0) = \xi \end{cases} \tag{1.6}$$

El clásico resultado establece que si f es lipschitziana con respecto a s, es decir si existe una constante $L > 0$ tal que para todos $s_1, s_2 \in \mathbb{R}$, $|f(t, s_1) - f(t, s_2)| \leq L|s_1 - s_2|, t \in [0, T]$, entonces existe una única solución de (1.6). Nosotros vamos a ver tres enfoques donde podamos aplicar el Teorema de Punto fijo de Banach.

Consideremos el espacio $\mathcal{C}[0, T]$ de las funciones continuas de variable real en $[0, T]$ con la norma del supremo. Integrando a ambos lados en (1.6) obtenemos $x(t) = \xi + \int_0^t f(s, x(s))ds$. Denotaremos por Fx a la función definida en el lado derecho de la ecuación anterior, es decir,

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

Así $F : \mathcal{C}[0, T] \to \mathcal{C}[0, T]$ y una solución de (1.6) son los puntos fijos de F.

Enfoque A (Este es el más común en los libros de ecuaciones diferenciales). Observemos que para cualquier $x, y \in \mathcal{C}[0, T]$,

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

Se sigue entonces que $\|Fx - Fy\| \leq LT\|x - y\|$ y por tanto $k(F) \leq LT$. Si $LT < 1$, entonces el resultado es inmediato por el Teorema 1.1.5. Sin embargo, si $LT \geq 1$ tenemos que dar algunos pasos más.

Tomemos $h > 0$ tal que $Lh < 1$ y consideremos el espacio $\mathcal{C}[0, h]$. Reemplazando T por h en el argumento anterior obtenemos una "solución local" de (1.6), digamos x_0 en $\mathcal{C}[0, h]$. Ahora consideremos el problema de Cauchy en $[h, 2h]$:

$$\begin{cases} x_1'(t) = f(t, x_1(t)), \\ x_1(h) = x_0(h) \end{cases} \tag{1.7}$$

Al aplicar la técnica utilizada al comienzo de este problema, obtenemos una única solución x_1 de (1.7) y, como $x_1(h) = x_0(h)$, x_1 extiende x_0 de $[0, h]$ a $[0, 2h]$. Esta extensión es diferenciable en h porque el problema de Cauchy tiene una única solución en un entorno de h. Ahora está claro que el procedimiento que se acaba de describir puede repetirse en el intervalo $[2h, 3h]$, y después de un número finito de pasos obtenemos una solución de (1.6) válida en $[0, T]$.

Enfoque B Si repetimos el cálculo inicial obtenemos:

$$\begin{aligned}
|(F^2x)(t) - (F^2y)(t)| &\leq \int_0^t |f(s, (Fx)(s)) - f(s, (Fy)(s))| \, ds \\
&\leq \int_0^t L|(Fx)(s) - (Fy)(s)| \, ds \\
&\leq \int_0^t L \int_0^s L|x(u) - y(u)| du \, ds \\
&\leq L^2 \int_0^t \int_0^s \|x - y\| \, du \, ds \\
&\leq \frac{L^2T^2}{2} \|x - y\|.
\end{aligned}$$

Repetiendo de nuevo el proceso:

$$|(F^3x)(t) - (F^3y)(t)| \leq \frac{L^3T^3}{3!} \|x - y\|,$$

y de forma general

$$|(F^nx)(t) - (F^ny)(t)| \leq \frac{L^nT^n}{n!} \|x - y\|.$$

Así $k(F^n) \leq (LT)^n/n!$ y la serie $\sum_{n=0}^{\infty} k(F^n)$ es convergente. Por lo visto en la Nota 1.1.8, esto implica que F tiene un único punto fijo x y para cualquier $x_0 \in \mathcal{C}[0, T]$,

$$\|F^nx_0 - x\| \leq \sum_{i=0}^{\infty} \|F^{n+i}x_0 - F^{n+i+1}x_0\| \leq$$

$$\sum_{i=0}^{\infty} k(F^{n+i}) \|x_0 - Fx_0\| \leq R_n \|x_0 - Fx_0\|.$$

donde $R_n = \sum_{i=0}^{\infty} (LT)^{n+i}/(n + i)!$.

La ventaja de este enfoque es que damos existencia de una solución en el intervalo completo $[0,T]$ con una buena estimación de la velocidad de convergencia de la sucesión de iteradas $\{F^n x_0\}$ hacia la solución.

Enfoque C (Remetrización) Observemos que como $k(F^n) \leq (LT)^n/n!$, entonces $k_\infty(F) = 0$. Así para cualquier $k > 0$ existe una métrica equivalente a la de $\mathcal{C}[0,T]$ para la cual F es una k-contracción. (Tales métricas fueron primero introducidas por A.Bielecki (1956) y posteriormente ampliamente utilizadas por otros autores para estudiar una variedad de ecuaciones). Más precisamente, para cualquier, $\omega \geq 0$, una nueva norma en $\mathcal{C}[0,T]$ se puede definir como sigue:

$$\|x\|_\omega = \max\left\{e^{-\omega Lt}|x(t)| : t \in [0,T]\right\}.$$

Entonces $\|\cdot\|_0$ es la norma original y como,

$$e^{-\omega LT}\|x\|_0 \leq \|x\|_\omega \leq \|x\|_0 \tag{1.8}$$

todas las $\omega-$normas son equivalentes. También:

$$\begin{aligned}
|(Fx)(t) - (Fy)(t)| &\leq \int_0^t L|x(s) - y(s)|\, ds \\
&= \int_0^t e^{\omega Ls} e^{-\omega Ls} L|x(s) - y(s)|\, ds \\
&\leq \int_0^t Le^{\omega Ls}\|x - y\|_\omega\, ds \\
&= \frac{1}{\omega}(e^{\omega Lt} - 1)\|x - y\|_\omega \\
&\leq \frac{1}{\omega}e^{\omega Lt}\|x - y\|_\omega.
\end{aligned}$$

Multiplicando a ambos lados por $e^{-\omega Lt}$ y tomando máximo en el lado izquierdo se obtiene que

$$\|Fx - Fy\|_\omega \leq \frac{1}{\omega}\|x - y\|_\omega.$$

Así para $\omega > 1$, F es una contracción con respecto a la norma $\|\cdot\|_\omega$, y se tiene que $\lim_{\omega\to\infty} k_{\|\cdot\|_\omega}(F) = 0$.

Lo que hemos visto en este enfoque C también puede ser usado para evaluar la velocidad de convergencia de las iteradas. Si x, $x_0 \in \mathcal{C}[0,T]$ y $x = Fx$, entonces (para $\omega > 1$) en vista de (1.3):

$$\|F^n x_0 - x\|_\omega \leq \frac{\omega^{-n}}{1 - \omega^{-1}}\|x_0 - Fx_0\|_\omega$$

y usando (1.8),

$$\|x - F^n x_0\| \le e^{\omega LT} \frac{\omega^{-n}}{1-\omega^{-1}} \|x_0 - Fx_0\|. \tag{1.9}$$

Tomando $\omega = n/LT$ generamos una mejor evaluación de la velocidad de convergencia en (1.9):

$$\|x - F^n x_0\| \le \left[\frac{eLT}{n}\right]^n \frac{n}{n-LT} \|x_0 - Fx_0\|.$$

1.2.2. Fractales: conjuntos autosimilares.

Sean (M, d) un espacio métrico completo, $\mathcal{M}$ la familia de todos los subconjuntos no vacíos, acotados, cerrados de M, y sea $\mathcal{N}$ la subfamilia de $\mathcal{M}$ formada por los conjuntos compactos. Para $X, Y \in \mathcal{M}$, tomemos:

$$d(X,Y) = \sup\{\mathrm{dist}(y,X) : y \in Y\},$$

$$d(Y,X) = \sup\{\mathrm{dist}(x,Y) : x \in X\},$$

y sea

$$D(X,Y) = \text{máx}\{d(X,Y), d(Y,X)\}.$$

D proporciona una métrica para $\mathcal{M}$ (y por tanto para $\mathcal{N}$) comúnmente conocida como métrica de Hausdorff. Se puede probar que la completitud de M implica la completitud de $(\mathcal{M}, D)$ y $(\mathcal{N}, D)$ (ver [14]).

Supongamos ahora que $T_1, T_2, \ldots, T_n$ es una familia finita de d-contracciones en M. Estas aplicaciones generan una aplicación $\mathcal{F} : \mathcal{N} \to \mathcal{N}$ definida por:

$$\mathcal{F}(X) = \bigcup_{i=1}^{n} T_i(X), \qquad X \in \mathcal{N}$$

y no es difícil probar que dicha aplicación $\mathcal{F}$ es una contracción en $\mathcal{N}$ relativa a D con $k(\mathcal{F}) \le \text{máx}\{k(T_i) : i = 1, \ldots, n\}$. Por tanto, vamos a tener el siguiente resultado.

Teorema 1.2.2. Sea M un espacio métrico completo y sea $T_i : M \to M$, $i = 1, \ldots, n$ una familia de contracciones. Entonces existe un único subconjunto compacto y no vacío X de M tal que:

$$X = \bigcup_{i=1}^{n} T_i(X). \tag{1.10}$$

Este teorema tiene aplicaciones en la teoría de fractales. En particular, si M es un espacio euclídeo $\mathbb{R}^p$ y las aplicaciones T_i verifican $k_i < 1$, entonces el conjunto satisfaciendo (1.10) es llamado autosimilar con respecto a $\{T_1, T_2, \ldots, T_n\}$.
Para un caso especial de lo anterior, consideremos la rectal real $\mathbb{R}$ y dos aplicaciones definidas como siguen:

$$T_1(x) = \frac{1}{3}x, \qquad T_2(x) = \frac{1}{3}x + \frac{2}{3}.$$

La aplicación $\mathcal{F}$ se define asociando a cada compacto $X \subset \mathbb{R}$ el conjunto

$$\mathcal{F}(X) = \frac{1}{3}X \cup (\frac{1}{3}X + \frac{2}{3}).$$

Como $\mathcal{F}$ es una contracción con respecto a la métrica de Hausdorff D podremos obtener su punto fijo por iteración. Tomemos $X_0 = [0,1]$, entonces $X_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$, $X_2 = [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{3}{9}] \cup [\frac{6}{9}, \frac{7}{9}] \cup [\frac{8}{9}, 1]$, etc. La sucesión $\{X_n\}$ converge en D al bien conocido conjunto de Cantor.

Ejemplo 1.2.3. Veamos a continuación un ejemplo de fractal en el plano. Hemos escogido el Triángulo de Sierpinski.
Consideremos el espacio métrico completo formado por $\mathbb{R}^2$ y la métrica euclídea, $(\mathbb{R}^2, d_e)$ y las siguientes aplicaciones definidas de $\mathbb{R}^2$ en sí mismo:

$$T_1(x,y) = \frac{1}{2}(x,y), \quad T_2(x,y) = \frac{1}{2}(x,y) + (\frac{1}{2}, 0), \quad T_3(x,y) = \frac{1}{2}(x,y) + (\frac{1}{4}, \frac{\sqrt{3}}{4})$$

Notemos que estas aplicaciones están escogidas para que el Triángulo de Sierpinski sea el de vértices $(0,0), (1,0), (\frac{1}{2}, \frac{\sqrt{3}}{2})$.
Las aplicaciones T_1, T_2, T_3 son contracciones con constante $k_{d_e} = \frac{1}{2}$. La aplicación $\mathcal{F}$ se define asociando a cada compacto $X \subset \mathbb{R}^2$ el conjunto

$$\mathcal{F}(X) = T_1(X) \cup T_2(X) \cup T_3(X).$$

Como $\mathcal{F}$ es una contracción con respecto a la métrica de Hausdorff D podremos obtener su punto fijo por iteración. Vamos a tomar como conjunto inicial X_0 el círculo de centro $(-2, 0)$ y radio 1. A continuación se adjuntan, usando el programa informático MATHEMATICA, las primeras iteraciones y se observa que la sucesión de iteradas $\{X_n\}$ convergerá en D al Triángulo de Sierpinski.

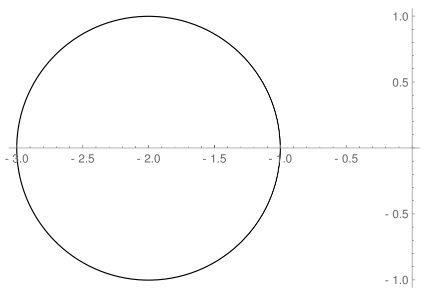

Figura 1.1: Conjunto inicial.

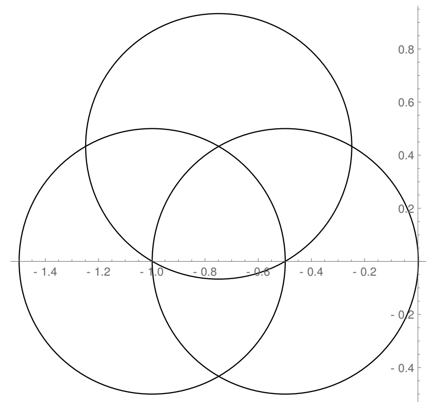

Figura 1.2: Primera iteración.

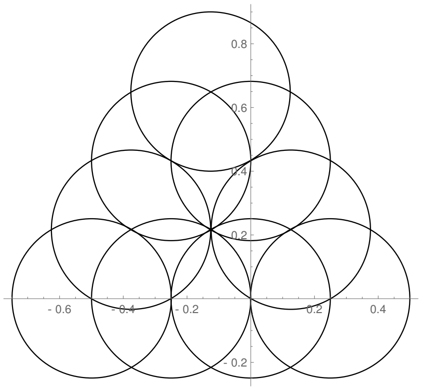

Figura 1.3: Segunda iteración.

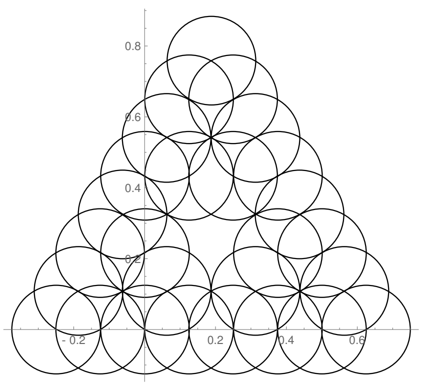

Figura 1.4: Tercera iteración.

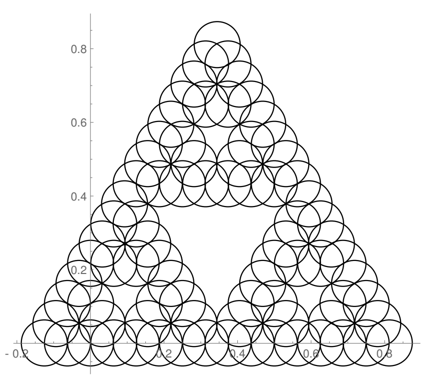

Figura 1.5: Cuarta iteración.

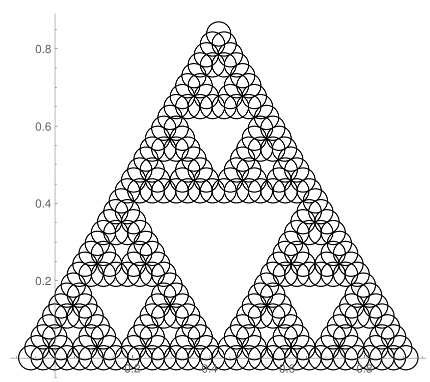

Figura 1.6: Quinta iteración.

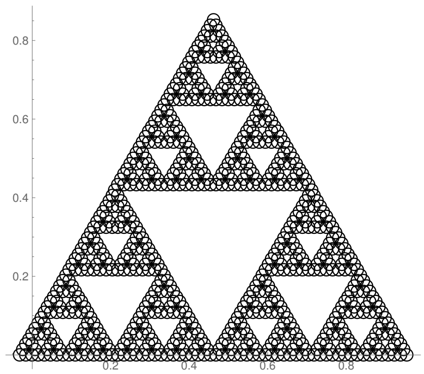

Figura 1.7: Sexta iteración.

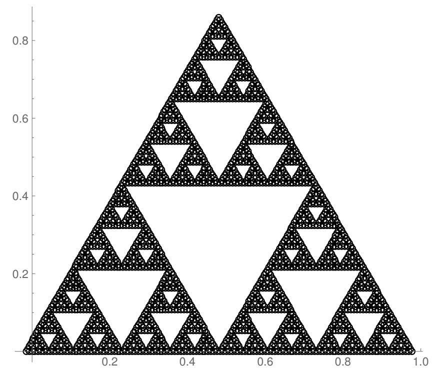

Figura 1.8: Séptima iteración.

Debemos notar de al ser una aplicación del Teorema de Banach, la convergencia al Triángulo de Sierpinski no depende del conjunto inicial X_0 escogido.

1.2.3. Raíces cuadradas en álgebras de Banach.

Definición 1.2.4. Diremos que X es un *álgebra de Banach* si $(X, \|\cdot\|)$ es un espacio de Banach en conjunción con una operación producto que satisface que para $x, y, z \in X, \alpha \in \mathbb{K}$ siendo $\mathbb{K}$ un cuerpo ($\mathbb{R}$ ó $\mathbb{C}$),

$$x(yz) = (xy)z, \quad x(y+z) = xy+xz, \quad (y+z)x = yx+zx, \quad (\alpha x)y = \alpha(xy) = x(\alpha y)$$

Además se verifica que $\|xy\| \leq \|x\| \, \|y\|$

Para cualquier $z \in X$, denotaremos por $X(z)$ a la subálgebra generada por z (la menor subálgebra cerrada de X que contiene a z). Es fácil comprobar que el álgebra $X(z)$ es siempre conmutativa: $xy = yx, \forall x, y \in X(z)$.

Proposición 1.2.5. Sea X un álgebra de Banach. Entonces, para cualquier $z \in X$ con $\|z\| < 1$ existe un único elemento $x \in X(z)$ tal que $\|x\| < 1$ y $x^2 - 2x + z = 0$

Demostración. Tomemos cualquier número d verificando $\|z\| < d < 1$ y consideremos la aplicación T definida como:

$$Tx = \frac{1}{2}(x^2 + z), \quad x \in B(0, d) \cap X(z).$$

Como $\|Tx\| \leq \frac{1}{2}(\|x\|^2 + \|z\|) \leq \frac{1}{2}(d^2 + d) < d$, entonces $T : B(0,d) \cap X(z) \to B(0,d) \cap X(z)$. Además como $X(z)$ es conmutativo, se tiene

$$\begin{aligned}
\|Tx - Ty\| &= \frac{1}{2}\left\|x^2 - y^2\right\| = \frac{1}{2}\|(x-y)(x+y)\| \\
&\leq \frac{1}{2}(\|x\| + \|y\|)\|x - y\| \\
&\leq d\|x - y\|.
\end{aligned}$$

Esto prueba que T es una contracción en $B(0,d) \cap X(z)$, y por tanto por el Teorema 1.1.5, tiene un único punto fijo $x \in B(0,d) \cap X(z)$. Como $d < 1$ puede ser elegido tan cerca de uno como se quiera, x es la única solución de $x^2 - 2x + z = 0$ en el interior de $B(0,1)$. □

Si X tiene elemento unidad e, el cual satisface que $ex = xe = x$ para todo $x \in X$, entonces $x^2 - 2x + z = 0$ puede ser escrito de la forma $(e - x)^2 = e - z$ y la proposición anterior puede reescribirse como:

Proposición 1.2.6. Sea X un álgebra de Banach con unidad. Entonces, para cualquier elemento de X de la forma $e - z$ con $\|z\| < 1$ existe un único elemento $y = e - x$ con $x \in X(z)$ y $\|x\| < 1$ tal que $y^2 = e - z$.

Nota 1.2.7. Esta proposición establece que $e - z$ tiene una "raíz cuadrada".

Nota 1.2.8. Otra aplicación de lo que hemos visto es la siguiente. Sea S un conjunto y denotemos por X un álgebra de funciones de variable real acotadas definidas en S con la norma del supremo. (La operación producto en X es la multiplicación punto a punto). Se puede probar que si X es completo y contiene la función unidad, cualquier función no negativa $f \in X$ tiene una raíz cuadrada $f^{1/2} \in X$. En efecto, podemos escribir $f = 1 - (1 - f)$. Entonces si $\|1 - f\| < 1$, aplicando la Proposición 1.2.6 existe una función $g \in X$ tal que $g^2 = f$. Si no se cumpliera la condición $\|1 - f\| < 1$, entonces basta tomar $a > 0$ para que $\|1 - af\| < 1$ y así se obtiene una raíz cuadrada de af y por tanto de f.
Consecuentemente para $f, g \in X$ las siguintes funciones pertenecen a X:

- $|f| = (f^2)^{1/2}$.

- $\text{máx}\{f,g\} = \frac{1}{2}(f+g+|f-g|)$.
- $\text{mín}\{f,g\} = \frac{1}{2}(f+g-|f-g|)$.

Observar que esto es válido sin pedirle a f propiedades especiales o incluso sin que S tenga una topología. Sin embargo, estos hechos son cruciales para probar el Teorema de Stone-Weierstrass, mostrando que el Teorema de punto fijo de Banach está indirectamente involucrado en uno de los teoremas más usados del análisis matemático.

Teorema 1.2.9 (Stone-Weierstrass)**.** Sea S un espacio topológico de Hausdorff compacto y sea $\mathscr{C}(S)$ el álgebra de todas las funciones reales continuas y acotadas en S. Supongamos que X es un subálgebra de $\mathscr{C}(S)$ que contiene la función unidad y separa puntos de S. Entonces X es denso en $\mathscr{C}(S)$.

Demostración. Seguiremos la prueba de [5].
Como X separa puntos de S, entonces para cualesquiera $x, y \in S$ con $x \neq y$ existe $g \in X$ tal que $g(x) \neq g(y)$.
Tomemos ahora $f \in \mathscr{C}(S)$, $\varepsilon > 0$, y fijemos $x \in S$. Como X separa puntos de S y contiene a las funciones constantes, para cada $y \in S$ existe una función $g_y \in X$ tal que $g_y(x) = f(x)$ y $g_y(y) = f(y)$. Para ver esto, tomemos $g \in X$ con $g(x) \neq g(y)$ y para $u \in S$, sea:

$$g_y(u) = f(x) + (f(y) - f(x))(g(y) - g(x))^{-1}(g(u) - g(x)).$$

Como g_y es continua, existe un entorno U_y de y tal que si $t \in U_y$, $g_y(t) < f(t) + \varepsilon$. Además se tiene que $\{U_y\}$ cubre S y como S es compacto, entonces existe $\{y_1, \cdots, y_n\}$ tal que $S \subset \bigcup_{i=1}^{n} U_{y_i}$. Definamos ahora:

$$h_x(t) = \text{mín}\{g_{y_1(t)}, \ldots, g_{y_n(t)}\} \in X.$$

Entonces $h_x(t) < f(t) + \varepsilon$ para todo $t \in S$, y $h \in \overline{X}$ donde $\overline{X}$ denota la menor subálgebra completa de $\mathscr{C}(S)$ que contiene a X. (Notemos que no estamos exigiendo que X sea completo).
Ahora para cada $x \in S$ elijamos un entorno V_x de x verificando que $h_x(t) > f(t) - \varepsilon$

para todo $t \in V_x$, y sea $\{V_{x_1}, \cdots, V_{x_{m(t)}}\}$ un subrecubrimiento finito de $\{V_x\}$ de S. Sea

$$h(t) = \max\{h_{x_1(t)}, \ldots, h_{x_m(t)}\} \in X$$

y observemos que $h \in \overline{X}$ con $\|f - h\| < \varepsilon$. Como $\varepsilon > 0$ es arbitrario, se concluye la prueba. □

1.2.4. Espacios métricos convexos.

El concepto de convexidad está estrechamente ligado a la existencia de segmentos y a la estructura vectorial del espacio. Sin embargo, en ciertos espacios métricos es posible definir una noción similar a la convexidad. En 1928 K. Menger introdujo el siguiente concepto de métrica convexa.

Definición 1.2.10. Diremos que un espacio métrico (M, d) es *métricamente convexo* si para cualesquiera $x, y \in M$ con $x \neq y$ existe $z \in M$, $x \neq z \neq y$, tal que $d(x, z) + d(z, y) = d(x, y)$.
Si z es como antes diremos que z se encuentra entre x e y y lo denotaremos por el símbolo (xzy).

Ejemplo 1.2.11. La esfera en el plano con la métrica inducida por la longitud.

El teorema fundamental en la métrica convexa es el llamado Teorema de Menger. Antes de enunciarlo y dar su correspondiente prueba veremos dos lemas que usaremos en dicha prueba.

Lema 1.2.12. Sea (M, d) un espacio métrico completo con $x, y \in M, x \neq y$ y supongamos que $0 < \lambda < d(x, y)$. Sea

$$B(x, y) = \{z \in M : (xzy)\},$$

$$S = S(x, y, \lambda) = \{z \in B(x, y) : d(x, z) \leq \lambda\} \cup \{x\}.$$

Entonces existe un punto $z_\lambda \in M$ tal que:

(a) $z_\lambda \in S(x, y, \lambda)$.

(b) $u \in B(x,y)$ y $(xz_\lambda u)$ implica $d(x,u) > \lambda$.

Demostración. Distinguimos dos casos:

Caso 1. Si existe $z' \in S$ con $d(x,z') < \lambda$ tal que $(xz'u)$, entonces $u \notin S$. En este caso, tomar $z_\lambda = z'$.

Caso 2. Para cada $z \in S$ con $d(x,z) < \lambda$ existe y_z tal que (xzy_z). En este caso, definimos $G : S \to S$ por $G(z) = y_z$ si $d(x,z) < \lambda$ y $G(z) = z$ en otro caso. Definamos también $\varphi : S \to \mathbb{R}^+$ por $\varphi(z) = \lambda - d(x,z)$. Es claro que φ es continua y para $z \in S$,

$$\begin{aligned} d(z,G(z)) &= d(x,G(z)) - d(x,z) = \lambda - d(x,z) - (\lambda - d(x,G(z))) = \\ &= \varphi(z) - \varphi(G(z)). \end{aligned}$$

Como S es cerrado (por tanto completo), por el Teorema de Caristi $G(z') = z'$ para algún $z' \in S$. Esto implica que $d(x,z') = \lambda$ y así $z' = z_\lambda$ satisface (a) y (b). $\square$

Nota 1.2.13. Necesitaremos ahora la siguiente noción de relación transitiva en espacios métricos que afirma que para $p,q,r,s \in M$,

$$((pqr) \quad y \quad (prs)) \quad \Leftrightarrow \quad ((pqs) \quad y \quad (qrs))$$

Lema 1.2.14. Sea M un espacio completo y métricamente convexo, $x,y \in M$ con $x \neq y$ y supongamos que $0 < \lambda < d(x,y)$. Entonces existe $z' \in M$ tal que $(xz'y)$ y $d(x,z') = \lambda$.

Demostración. Por el Lema 1.2.12, existe $z_\lambda \in M$ tal que:

(a) $z_\lambda \in S(x,y,\lambda)$.

(b) $u \in B(x,y)$ y $(xz_\lambda u)$ implica $d(x,u) > \lambda$.

Sea $\lambda' = d(x,y) - \lambda$ y de nuevo por el Lema 1.2.12 se tiene que $y_{\lambda'} \in M$ y:

(a') $y_{\lambda'} \in S(y,z_\lambda,\lambda')$.

(b') $u \in B(y,z_\lambda)$ y $(yy_{\lambda'}u)$ implica $d(y,u) > \lambda'$.

Distinguimos dos casos:

Caso 1. $z_\lambda = y_{\lambda'}$. Entonces, como $d(x,y) = d(x,z_\lambda) + d(z_\lambda, y)$ se sigue que $d(x,z_\lambda) = \lambda$.

Caso 2. $z_\lambda \neq y_{\lambda'}$. En este caso, como M es métricamente convexo, existe $w \in M$ tal que $(z_\lambda w y_{\lambda'})$. Por hipótesis se tiene que $(xz_\lambda y), (z_\lambda y_{\lambda'} y)$ y $(z_\lambda w y_{\lambda'})$. Por la propiedad de transitividad se tiene que $(xwy), (xz_\lambda w), (ywz_\lambda)$ y $(yy_{\lambda'}w)$. Ahora bien por (b), (xwy) y $(xz_\lambda w)$ implican que $d(x,w) > \lambda$ y por (b') (ywz_λ) y $(yy_{\lambda'}w)$ implican $d(y,w) > \lambda$. Por tanto, $d(x,y) = d(x,w) + d(w,y) > \lambda + \lambda' = d(x,y)$. Esto es una contradicción, y por consiguiente este caso no se puede dar. $\square$

El teorema de Menger prueba que en un espacio métrico convexo y completo M, dados $x, y \in M$, existe todo segmento que los une. El enunciado es el siguiente:

Teorema 1.2.15. (Menger) Si (M,d) es un espacio métrico completo y métricamente convexo, entonces cualesquiera $x, y \in M$ están unidos por un segmento métrico, es decir, existe una isometría $\varphi : [0, d(x,y)] \to M$ con $\varphi(0) = x$ y $\varphi(d(x,y)) = y$.

La prueba que damos es básicamente la misma que la original (Menger, 1928) excepto por la demostración del Lema 1.2.12. Debemos notar que una prueba diferente de este teorema, debida a Aronszajn, puede ser encontrada, por ejemplo, en [3].

Demostración. Sean $x_0, x_1 \in M, x_0 \neq x_1$. Por el Lema 1.2.14 existe $x_{1/2} \in M$ tal que $d(x_0, x_{1/2}) = d(x_{1/2}, x_1) = \frac{1}{2} d(x_0, x_1)$ (i.e, $x_{1/2}$ es “punto medio” del par (x_0, x_1)). Sea $\rho = d(x_0, x_1)$ y definamos la aplicación F de la siguiente forma:

$$F(0) = x_0, \quad F(\rho/2) = x_{1/2}, \quad F(\rho) = x_1.$$

De nuevo por el Lema 1.2.14 existen puntos $x_{1/4}, x_{3/4}$ que son los puntos medios de los pares $(x_0, x_{1/2})$ y $(x_{1/2}, x_1)$. Pongamos:

$$F(\rho/4) = x_{1/4}, \quad F(3\rho/4) = x_{3/4}.$$

y usando la propiedad transitiva concluímos que F es una isometría en el conjunto $\{0, \rho/4, \rho/2, 3\rho/4, \rho\}$. Por inducción es posible obtener puntos $\{x_{p/2^n}\}, 1 \leq p \leq 2^n - 1 (n = 1, 2, \ldots)$, en M tales que la aplicación $F : p\rho/2^n \mapsto x_{\rho/2^n}$ es una isometría. Como $\{p\rho/2^n\}$ es un subconjunto denso de $[0, \rho]$ y M es completo, es posible extender F de la forma habitual al intervalo $[0, \rho]$ y así obtener un segmento métrico en M que una x_0 y x_1. $\square$

Capítulo 2

Teorema de Schauder-Tychonoff. Aplicaciones.

En este capítulo probaremos el Teorema de Schauder-Tychonoff que no es más que una extensión del Teorema de Schauder en espacios localmente convexos. Como aplicación de este resultado veremos las siguientes aplicaciones:

1. La existencia de medidas invariantes para toda aplicación continua en un compacto con imagen en sí misma.
2. Una respuesta parcial al problema del subespacio invariante en espacios de Hilbert.

En el desarrollo de este capítulo hemos seguido fundamentalmente las referencias de [1], [6] y [8].

2.1. Preliminares.

Dado que en este capítulo y en el siguiente se usan nociones de teoría de la medida, y topologías débiles, vamos a empezar viendo estos conceptos de modo introductorio.

2.1.1. Topologías débiles.

La topología de la norma en un espacio normado tiene una gran cantidad de abiertos y eso es una ventaja porque supone que hay una gran cantidad de funciones continuas definidas en el espacio. Pero como contrapartida tiene, lógicamente pocos compactos. De hecho la bola unidad cerrada no es compacta salvo que el espacio sea de dimensión finita. Además el dual de un espacio normado está determinado por dicha topología, por tanto es natural considerar topologías en X que nos permitan asegurar que los funcionales del dual sigan siendo continuos para dichas topologías.

Definición 2.1.1. Sea X un espacio normado. Se llama *topología débil* de X, y se denota por $\sigma(X, X^*)$ o simplemente por ω, a la topología inicial en X para la familia X^* de los funcionales lineales y continuos (respecto de la topología de la norma) de X en $\mathbb{K}$. Es decir, la topología débil de X es la topología menos fina sobre X que hace continuos a los elementos de X^*.

A partir de este momento, τ_X denotará la topología de la norma en X. Puesto que la topología débil de X es la topología menos fina sobre X que hace continuos a los elementos de X^* se sigue, de la propia definición de X^*, que $\sigma(X, X^*) \subset \tau_X$.
Si X es un espacio normado, sobre X^* tenemos dos topologías, la topología de la norma y su topología débil $\sigma(X^*, X^{**})$. No obstante, en este caso podemos introducir una nueva topología en X^*.
Antes de continuar, recordemos que para un espacio de Banach X, existe una inclusión natural $i : X \to X^{**}$ definida por $i(x)(x^*) = x^*(x)$, que es una isometría lineal. Entonces, a menudo identificamos X con $i(X)$ y tratamos a X como un subespacio de X^{**}.

Definición 2.1.2. Sea X un espacio normado. Se llama *topología débil-** de X^*, y se denota por $\sigma(X^*, X)$ o simplemente por ω^*, a la topología inicial en X^* para la familia $i(X)\,(\subseteq X^{**})$. Es decir, la topología débil-* de X^* es la topología menos fina sobre X^* que hace continuos a los elementos de $i(X)$.

Como $i(X)$ es un subespacio de X^{**} y la topología $\sigma(X^*, X^{**})$ hace continuos a todos los elementos de X^{**}, podemos asegurar que también hace continuos a los elementos de $i(X)$. Como la topología débil-* de X^* es la topología inicial en X para este último subconjunto es claro que:

$$\sigma(X^*, X) \subset \sigma(X^*, X^{**}) \subset \tau_{X^*}.$$

Veamos ahora algunas propiedades básicas de las topologías débil y débil-*.

Proposición 2.1.3. Sea X un espacio de Banach.

(a) Sea $x_0 \in X$. Una base de entornos de x_0 para la topología débil de X está formada por los conjuntos

$$U(x_0; \varepsilon, x_1^*, \dots, x_n^*) = \{x \in X : |x_j^*(x) - x_j^*(x_0)| < \varepsilon \quad \forall\, j = 1, \dots, n\}$$

para todo $\varepsilon > 0$, $n \in \mathbb{N}$ y $x_1^*, \dots, x_n^* \in X^*$.

(b) Sea $x_0^* \in X^*$. Una base de entornos de x_0^* para la topología débil-* de X^* está formada por los conjuntos

$$U(x_0; \varepsilon, x_1, \dots, x_n) = \{x^* \in X^* : |x^*(x_j) - x_0^*(x_j)| < \varepsilon \quad \forall\, j = 1, \dots, n\}$$

para todo $\varepsilon > 0$, $n \in \mathbb{N}$ y $x_1, \dots, x_n \in X$.

(c) Una sucesión $(x_n)_{n=1}^{\infty} \subset X$ converge en la topología débil a un punto $x \in X$, denotado por $x_n \xrightarrow{w} x$, si y sólo si $\forall\, x^* \in X^*$, $x^*(x_n) \xrightarrow{n\to\infty} x^*(x)$.

(d) Una sucesión $(x_n^*)_{n=1}^{\infty} \subset X^*$ converge en la topología débil-* a un punto $x^* \in X^*$, denotado por $x_n^* \xrightarrow{w^*} x^*$, si y sólo si $\forall\, x \in X$, $x_n^*(x) \xrightarrow{n\to\infty} x^*(x)$.

(e) La topología $\sigma(X, X^*)$ es compatible con la estructura vectorial de X. Esto es, la suma y el producto por escalares son continuos cuando en X se considera la topología $\sigma(X, X^*)$, en $\mathbb{K}$ la topología usual y en $X \times X$ y $\mathbb{K} \times X$ las correspondientes topologías producto.

(f) La topología $\sigma(X^*, X)$ es compatible con la estructura vectorial de X^*.

(g) Las topologías $\sigma(X, X^*)$ y $\sigma(X^*, X)$ son siempre Hausdorff.

El concepto de dual topológico de un espacio normado puede ser generalizado a duales topológicos de espacios más generales que los espacios normados. Recordemos que un *espacio vectorial topológico* es un espacio vectorial X que posee además una estructura de espacio topológico (para nosotros Hausdorff) verificando que la suma y el producto por escalares son continuos para dicha topología. Si

además existe una base de entornos formada por conjuntos convexos, diremos que el espacio es *localmente convexo*. Si X es un espacio normado entonces X equipado con la topología de la norma es un espacio localmente convexo, pues las bolas son convexas. Los apartados (e) y (f) de la Proposición 2.1.3 nos aseguran que X equipado con su topología débil es también un espacio vectorial topológico y que lo mismo sucede con X^* cuando consideramos su topología débil-*. Además ambos apartados definen una topología localmente convexa en X y en X^*.
Por último vamos a ver uno de los resultados más importantes, el Teorema de Banach-Alaoglu. Este resultado nos va a asegurar que la topología débil-$*$ de un espacio de Banach dual hace compacta a la bola unidad cerrada de dicho espacio, propiedad que habíamos perdido para la topología de la norma en espacios infinito dimensionales.
La historia de este resultado comienza en 1932 con Banach. Dicho autor probó que la bola unidad del dual de un espacio de Banach separable (i.e, que posee un subconjunto denso numerable) es ω^*-compacta. En 1940, Alaoglu suprime la hipótesis de separabilidad estableciendo el resultado que hoy día conocemos como Teorema de Banach-Alaoglu (una prueba puede ser encontrada en [9]).

Teorema 2.1.4 (Banach-Alaoglu)**.** La bola unidad cerrada B_{X^*} del dual de un espacio normado X es ω^*-compacta.

2.1.2. Nociones de teoría de la medida.

Para comenzar vamos a recordar algunas nociones básicas de teoría de la medida que necesitaremos posteriormente.
Partimos de un conjunto $X \neq \emptyset$. Por $\mathcal{P}(X)$ denotaremos el conjunto de partes de X. Recordemos que, si $A \subset X$, el complementario de A se define como $A^c = \{x \in X : x \notin A\}$.

Definición 2.1.5. Sea X un conjunto no vacío y $\mathcal{M} \subset \mathcal{P}(X)$. Decimos que $\mathcal{M}$ es una *σ-álgebra* sobre X cuando se verifican las siguientes condiciones:

(a) $X \in \mathcal{M}$.

(b) Si $A \in \mathcal{M}$ entonces $A^c \in \mathcal{M}$.

(c) Si $\{A_n\}_{n\geq 1} \subset \mathcal{M}$ entonces $\bigcup_{n=1}^{\infty} A_n \in \mathcal{M}$.

Se llama *espacio medible* al par $(X, \mathcal{M})$, y *conjunto medible* a cada miembro de $\mathcal{M}$.

De propiedades elementales de álgebra de conjuntos (incluyendo las leyes de De Morgan) se deduce que, si $\mathcal{M}$ es una σ-álgebra sobre X, entonces están en $\mathcal{M}$ los siguientes conjuntos: $\emptyset$, uniones finitas de miembros de $\mathcal{M}$, intersecciones finitas o numerables de miembros de $\mathcal{M}$, diferencia de dos miembros de $\mathcal{M}$. Como ejemplos triviales y extremos, se tiene que $\{\emptyset, X\}$ y $\mathcal{P}(X)$ son σ-álgebras sobre X.

Definición 2.1.6. Sea $(X, \mathcal{M})$ un espacio medible. Por definición, una *medida positiva* o simplemente una medida sobre $(X, \mathcal{M})$ es una aplicación $\mu : \mathcal{M} \to [0, +\infty]$ tal que $\mu(\emptyset) = 0$ y es numerablemente aditiva, es decir, si $\{A_n\}_{n\geq 1} \subset \mathcal{M}$ y los A_n son disjuntos dos a dos, entonces $\mu\big(\bigcup_{n=1}^{\infty} A_n\big) = \sum_{n=1}^{\infty} \mu(A_n)$. La terna $(X, \mathcal{M}, \mu)$ se denomina espacio de medida.

Si $\mu(X) < +\infty$, μ se dice finita. Si $\mu(X) = 1$, μ es una medida de probabilidad.

Nota 2.1.7. Si X es un conjunto no vacío, es fácil probar que la intersección de una familia de σ-álgebras es también una σ-álgebra. Si $\mathcal{A} \subset \mathcal{P}(X)$, se llama σ-álgebra generada por $\mathcal{A}$ a la intersección de todas las σ-álgebras que contienen a $\mathcal{A}$. Es claro que es la menor σ-álgebra que contiene a $\mathcal{A}$. Si ahora X es un espacio topológico y $\mathcal{A}$ es la familia de los abiertos, la σ-álgebra $\mathcal{B}$ que genera se conoce como *σ-álgebra de Borel* de X, y sus conjuntos se denominan borelianos o conjuntos de Borel.
Una medida de Borel es simplemente una medida en los conjuntos de Borel de un espacio topológico. Se dice que dicha medida es regular si para cualquier conjunto de Borel E:

$$\mu(E) = \text{ínf}\{\mu(U) : E \subset U,\, U \text{ abierto}\} = \sup\{\mu(K) : K \subset E,\, K \text{ compacto}\}.$$

2.2. Teorema de Schauder-Tychonoff.

Uno de los teoremas de punto fijo más conocido y usado dentro de la teoría de punto fijo es debido a Schauder en espacios de Banach o su penalización en espacios localmente convexos. Se dice que es un teorema de tipo topológico porque ya no depende de una distancia, sino de la topología del espacio y las hipótesis requeridas para la existencia de punto fijo se basan en la continuidad de la aplicación y compacidad del dominio.

Empezaremos enunciando un resultado, debido a Brouwer, que es un versión finito-dimensional del Teorema de Schauder.

El Teorema de Brouwer es un resultado que ha tenido una enorme influencia en el desarrollo de importantes campos de las matemáticas, sobre todo en la topología algebraica. Nosotros no daremos su prueba pero se puede encontrar, por ejemplo, en [1].

Teorema 2.2.1 (Brouwer)**.** Sea $f : \overline{B} \to \overline{B}$ una función continua de la bola unidad cerrada en sí misma. Entonces, existe un punto $\overline{x} \in \overline{B}$ tal que $f(\overline{x}) = \overline{x}$.

Veamos la siguiente proposición que nos será útil para probar posteriormente un corolario de este teorema.

Proposición 2.2.2. Sea H un espacio de Hilbert, K un subconjunto, convexo, cerrado y no vacío de H y $x \in H$. Entonces, existe un único $x_0 \in K$ tal que:

$$\|x - x_0\| = \operatorname{dist}(x, K) \equiv \operatorname{ínf}\{\|x - k\| : k \in K\}.$$

No damos la prueba de este resultado pues es una consecuencia de la identidad del paralelogramo.

Corolario 2.2.3. Sea X un espacio de Banach de dimensión finita. Sea $K \subseteq X$ un subconjunto no vacío, convexo y compacto. Entonces para cualquier función continua $f : K \to K$ existe $\overline{x} \in K$ tal que $f(\overline{x}) = \overline{x}$.

Demostración. Como X es un espacio normado de dimensión finita, es homeomorfo a $\mathbb{R}^N$ para algún $N \in \mathbb{N}$. Por tanto, podemos asumir, sin pérdida de generalidad, que $X = (\mathbb{R}^N, \|\cdot\|_2)$, que es un espacio de Hilbert. Además, por ser K acotado podemos suponer que $K \subseteq \overline{B} = \{x \in \mathbb{R}^N : \|x\| \leq r\}$ para algún $r > 0$. Sea $\phi : \overline{B} \to K$ la función dada por $\phi(x) = y_K$, donde y_K es el único punto de K que satisface que $\|x - y_K\| = \text{dist}(x, K)$. La aplicacion ϕ está bien definida por la Proposición 2.2.2 y cumple que $\phi(x) = x$ para todo $x \in K$. Veamos que ϕ es continua en $\overline{B}$. Sean $x_j, x \in \overline{B}$ tales que $x_j \to x$. Tomamos $j \in \mathbb{N}$:

$$\begin{aligned} \|x - \phi(x)\| &\leq \|x - \phi(x_j)\| \leq \|x - x_j\| + \|x_j - \phi(x_j)\| \\ &= \|x - x_j\| + \inf_{k \in K} \|x_j - k\| \\ &\leq \|x - x_j\| + \inf_{k \in K} (\|x_j - x\| + \|x - k\|) \\ &= \|x - x_j\| + \|x_j - x\| + \inf_{k \in K} \|x - k\| . \end{aligned}$$

Tomando límite cuando $j \to \infty$ llegamos a que

$$\|x - \phi(x)\| \leq \lim_{j \to \infty} \|x - \phi(x_j)\| \leq \|x - \phi(x)\| .$$

Por tanto todas las desigualdades son igualdades, y en particular $\|x - \phi(x)\| = \lim_{j \to \infty} \|x - \phi(x_j)\|$. Es decir $x - \phi(x_j)$ es una sucesión minimizante en $x - K$. Por último basta notar que como $\phi(x_j) \subseteq K$ y a K es compacto existe una subsucesión $\phi(x_{j_s}) \subseteq K$ tal que $\phi(x_{j_s}) \to u \in K$. Ahora bien:

$$\|x - \phi(x)\| \leq \lim_{s \to \infty} \|x - \phi(x_{j_s})\| = \|x - u\| .$$

Pero por unicidad $u = \phi(x)$ y por tanto $\phi(x_j) \to \phi(x)$.
Como ϕ es continua, la composición $f \circ \phi : \overline{B} \to K \subseteq \overline{B}$ es una función continua. Entonces, por el Teorema de Brouwer, existe $x \in \overline{B}$ tal que $(f \circ \phi)(x) = x$, pero como $(f \circ \phi)(\overline{B}) \subseteq K$ tenemos que $x \in K$. Luego, $x = f(\phi(x)) = f(x)$ □

Una vez que tenemos probado la existencia de un punto fijo para toda aplicación continua $f : K \to K$ con K convexo y compacto de un espacio de Banach finito dimensional, nos dirigimos a probar una versión generalizada en espacios localmente convexos. Las particiones de la unidad serán una herramienta fundamental en el desarrollo de la prueba.

Definición 2.2.4. Sea Q un espacio topológico compacto, y sea $\mathcal{F}$ un recubrimiento por abiertos finito de Q. Una *partición de la unidad* en Q subordinada a $\mathcal{F}$ es un subconjunto $\{\phi_V : V \in \mathcal{F}\}$ de $\mathcal{C}(Q, [0,1])$ tal que $\sum_{V\in\mathcal{F}} \phi_V \equiv 1$ en Q y $\operatorname{sop} \phi_V \subset V$ para cada $V \in \mathcal{F}$.

La existencia de una partición de la unidad es una consecuencia del lema de Urysohn (ver [12]).

Teorema 2.2.5 (Shauder-Tychonoff)**.** Sea X un espacio localmente convexo, $K \subseteq X$ un conjunto convexo y no vacío, y $K_0 \subseteq K$ compacto. Si $f : K \to K_0$ es una función continua, entonces existe un punto $\overline{x} \in K_0$ tal que $f(\overline{x}) = \overline{x}$.
Nótese que es suficiente con suponer la hipótesis de compacidad para el rango de f.

Demostración. Denotemos por $\mathcal{B}$ la base local formada por conjuntos convexos y simétricos para la topología de X generada por la familia separadora de seminormas $\mathcal{P}$ en X (para más detalles consultar (Capítulo 3,[10]). Sea $U \in \mathcal{B}$ (notemos que $0 \in U$). Sin pérdida de generalidad podemos suponer que $U = -U$ (en otro caso tomamos $U\cap(-U)$). Ahora bien, por la compacidad de K_0, existen $x_1, \ldots, x_n \in K_0$ tal que

$$K_0 \subseteq \bigcup_{j=1}^{n} (x_j + U).$$

Sea $\varphi_1, \ldots, \varphi_n \in \mathcal{C}(K_0)$ una partición de la unidad para K_0 subordinada al recubrimiento por abiertos $\{x_j + U\}$, y definamos

$$f_U(x) = \sum_{j=1}^{n} \varphi_j(f(x))x_j\,, \quad \forall x \in K.$$

Entonces,

$$f_U(K) \subseteq K_U := \operatorname{co}(\{x_1, \ldots, x_n\}) \subseteq K$$

y por el Corolario 2.2.3 existe $x_U \in K_U$ tal que $f_U(x_U) = x_U$. Entonces

$$x_U - f(x_U) = f_U(x_U) - f(x_U) = \sum_{j=1}^{n} \varphi_j(f(x_U))(x_j - f(x_U)) \in U \tag{2.1}$$

para $\varphi_j(f(x_U)) = 0$ cuando $x_j - f(x_U) \notin U$. Usando de nuevo la compacidad de K_0, existe

$$\overline{x} \in \bigcap_{W \in \mathcal{B}} \overline{\{f(x_U) : U \in \mathcal{B}, U \subset W\}} \subseteq K_0. \tag{2.2}$$

Elijamos ahora $p \in \mathcal{P}$ y $\varepsilon > 0$, y sea

$$V = \{x \in X : p(x) < \varepsilon\} \in \mathcal{B}.$$

Como f es continua en K, existe $W \in \mathcal{B}$ con $W \subseteq V$ tal que $f(x) - f(\overline{x}) \in V$ cuando $x - \overline{x} \in W, x \in K$. Además, existe $W_1 \in \mathcal{B}$ tal que $W_1 + W_1 \subseteq V$. Ahora bien por (2.2), existe $U \in \mathcal{B}, U \subseteq W_1$ tal que

$$\overline{x} - f(x_U) \in W_1 \subseteq V. \tag{2.3}$$

Uniendo ahora (2.1) y (2.3), se tiene que

$$x_U - \overline{x} = x_U - f(x_U) + f(x_U) - \overline{x} \in U + W_1 \subseteq W_1 + W_1 \subset W.$$

y por tanto

$$f(x_U) - f(\overline{x}) \in V. \tag{2.4}$$

Luego por (2.3) y (2.4) se obtiene que

$$p(\overline{x} - f(\overline{x})) \leq p(\overline{x} - f(x_U)) + p(f(x_U) - f(\overline{x})) < 2\varepsilon.$$

Como p y ε son arbitrarios, concluimos que $p(\overline{x} - f(\overline{x})) = 0$ para cada $p \in \mathcal{P}$, lo cual implica que $f(\overline{x}) = \overline{x}$. □

2.3. Aplicaciones conservando la medida en espacios topológicos de Hausdorff compactos.

En esta sección Q denotará un espacio topológico de Hausdorff compacto y $P(Q)$ el conjunto de todas las medidas regulares de probabilidad de Borel en X.

Se puede probar que si el espacio es métrico entonces todas las medidas de Borel son regulares.

Comenzamos dando unos resultados sin demostración (para más detalles consultar [10]) que nos va a permitir identificar el dual del espacio de las funciones continuas definidas sobre Q.

Teorema 2.3.1 (Representación de Riesz)**.** Sea Q un espacio topológico de Hausdorff compacto y sea Λ un funcional lineal positivo en $\mathcal{C}(Q)$. Entonces, existe una $\sigma-$álgebra $\mathcal{M}$ en Q que contiene a todos los conjuntos de Borel en Q, y existe una única medida positiva de Borel μ en $\mathcal{M}$ tal que:

$$\Lambda_\mu(f) = \int_Q f d\mu.$$

Teorema 2.3.2. Sea Q un espacio topológico de Hausdorff compacto y sea Λ un funcional lineal y acotado en $\mathcal{C}(Q)$. Entonces, existe una única medida compleja regular de Borel μ tal que:

$$\Lambda_\mu(f) = \int_Q f d\mu.$$

Más generalmente, el dual del espacio $\mathcal{C}(Q)$ puede ser identificado con el espacio $M(Q)$ de las medidas (complejas) de Borel regulares en Q. Notemos que la norma $\|\mu\|$ de un elemento $\mu \in M(Q)$ viene dada por la variación total de μ. $P(Q)$, definido anteriormente, es convexo y está contenido en la bola unidad del espacio $M(Q)$. Dotemos a $M(Q)$ con la topología débil-* $\sigma(M(Q), \mathcal{C}(Q))$. No es difícil probar que $P(Q)$ es $\sigma(M(Q), \mathcal{C}(Q))$-cerrado. En efecto, sea $(\nu_\alpha) \subseteq P(Q)$ una red tal que $\nu_\alpha \xrightarrow{w^*} \nu$. Como $\nu \in M(Q)$ tenemos que probar que ν es una medida positiva y $\nu(Q) = 1$. Por la convergencia débil estrella, $\nu(f) \geq 0$ para toda $f \in \mathcal{C}(Q)$, $f \geq 0$. Además por ser ν regular se tiene que $\mathcal{C}(Q)$ es denso en $L^1(\nu)$ (es consecuencia del Lema de Urysohn, para más detalles consultar [10]) y por tanto $\nu(A) \geq 0$ para cualquier conjunto A de Borel con $\nu(A) < \infty$. Finalmente:

$$\nu(Q) = \int_Q 1\, d\nu = \lim_\alpha \int_Q 1\, d\nu_\alpha = 1.$$

Además, $P(Q)$ es ω^*-compacto. En efecto, es un subconjunto ω^*-cerrado de la bola unidad del espacio $M(Q)$, el cual es ω^*-compacto por el Teorema de Banach-Alaoglu.

Definición 2.3.3. Sea $\mu \in P(Q)$. Una aplicación μ-medible $f : Q \to Q$ se dice que conserva la medida si $\mu(B) = \mu(f^{-1}(B))$ para cada conjunto de Borel $B \subset Q$. En tal caso se dice que μ es una *medida invariante* para la función f.

Nota 2.3.4. Si f es una aplicación μ-medible, la medida $\tilde{f}\mu$ definida por

$$\tilde{f}\mu(B) = \mu(f^{-1}(B)), \quad \text{para cualquier } B \text{ Borel}$$

pertenece a $P(Q)$. En particular, si f es continua (y por tanto medible con respecto a cada $\mu \in P(Q)$), tenemos una aplicación $\tilde{f} : P(Q) \to P(Q)$ definida por $\mu \mapsto \tilde{f}\mu$. Además, se puede probar que:

$$\int_Q g\, d(\tilde{f}\mu) = \int_Q g \circ f\, d\mu, \quad \forall\, g \in \mathcal{C}(Q)$$

En efecto:

1. Es trivial comprobar que se cumple la propiedad para las funciones características $g = \chi_A$ de conjuntos medibles.

2. Por linealidad, la propiedad es válida para funciones simples medibles $g = \sum_{j=1}^{N} \alpha_j \chi_{A_j}$.

3. Dada ahora una $g \in \mathcal{C}(Q)$, en particular, es μ-medible y también lo son $g^+ := \max\{g, 0\}$ y $g^- := \max\{-g, 0\}$ (pues son continuas). Notar que $g = g^+ - g^-$. Como g^+ y g^- son μ-medibles y no negativas, entonces existen sucesiones $\{\varphi_n\}_{n=1}^{\infty}$ y $\{\psi_n\}_{n=1}^{\infty}$ de funciones simples, μ-medibles y no negativas tales que $\varphi_n(x) \xrightarrow{n\to\infty} g^+(x)$ y $\psi_n(x) \xrightarrow{n\to\infty} g^-(x)$ en casi todo $x \in Q$ y $\{\varphi_n(x)\}_{n=1}^{\infty}$, $\{\psi_n(x)\}_{n=1}^{\infty}$ crecientes en casi todo $x \in Q$. Aplicando ahora el teorema de la convergencia monótona, se tiene que:

$$\lim_{n\to\infty} \int_Q \varphi_n\, d(\tilde{f}\mu) = \int_Q g^+\, d(\tilde{f}\mu)$$

$$\lim_{n\to\infty} \int_Q \psi_n\, d(\tilde{f}\mu) = \int_Q g^-\, d(\tilde{f}\mu)$$

Por tanto, se tiene que:

$$\begin{aligned}
\int_Q g\, d(\tilde{f}\mu) &= \int_Q (g^+ - g^-)\, d(\tilde{f}\mu) = \int_Q g^+\, d(\tilde{f}\mu) - \int_Q g^-\, d(\tilde{f}\mu) \\
&= \lim_{n\to\infty} \int_Q \varphi_n\, d(\tilde{f}\mu) - \lim_{n\to\infty} \int_Q \psi_n\, d(\tilde{f}\mu) \\
&= \lim_{n\to\infty} \int_Q \varphi_n \circ f\, d\mu - \lim_{n\to\infty} \int_Q \psi_n \circ f\, d\mu \\
&= \int_Q \lim_{n\to\infty} (\varphi_n \circ f)\, d\mu - \int_Q \lim_{n\to\infty} (\psi_n \circ f)\, d\mu \\
&= \int_Q (g^+ \circ f)\, d\mu - \int_Q (g^- \circ f)\, d\mu = \int_Q (g^+ - g^-) \circ f\, d\mu \\
&= \int_Q g \circ f\, d\mu.
\end{aligned}$$

Lema 2.3.5. Sea $f : Q \to Q$ una aplicación continua. Entonces la aplicación $\tilde{f} : P(Q) \to P(Q)$ es continua para la topología débil estrella.

Demostración. Sea $\{\mu_i\}_{i\in I} \subset P(Q)$ una red convergiendo a algún $\mu \in P(Q)$. Entonces, para cada $g \in \mathcal{C}(Q)$ se tiene que $g \circ f \in \mathcal{C}(Q)$ y por tanto:

$$\lim_{i\in I} \int_Q g\, d(\tilde{f}\mu_i) = \lim_{i\in I} \int_Q g \circ f\, d\mu_i = \int_Q g \circ f d\mu = \int_Q g\, d(\tilde{f}\mu).$$

Por tanto $\tilde{f}$ es continua para dicha topología. □

Ahora estamos interesados en encontrar elementos de $P(Q)$ que sean medidas invariantes para f. Esto es equivalente a encontrar un punto fijo para la aplicación $\tilde{f}$.

Teorema 2.3.6. Sea $f : Q \to Q$ continua. Entonces existe $\mu \in P(Q)$ para la cual f es una aplicación conservando la medida μ.

Demostración. Por el Lema 2.3.5, $\tilde{f}$ es una aplicación continua de un subconjunto convexo y compacto de $M(Q)$ en sí mismo, y la existencia de un punto fijo viene dada por el Teorema 2.2.5. □

2.4. El problema del subespacio invariante.

El problema del subespacio invariante es probablemente "el problema" en la teoría de operadores. La cuestión, que atrajo la atención de una gran cantidad de matemáticos, es simple de formular:

Dado un espacio de Banach X y un operador $T \in L(X)$(es decir, lineal y continuo) se pretende encontrar un subespacio cerrado y no trivial M de X (es decir, $M \neq X$ y $M \neq \{0\}$ tal que $TM \subset M$). Diremos que M es un *subespacio invariante* para T. Es conocido que no todos los operadores lineales y continuos en espacios de Banach tienen subespacios invariantes. Por ejemplo, C. J. Read ([4]) encuentra un contraejemplo a este problema en ℓ_1. En el caso de un espacio de Hilbert H de dimensión infinita podemos encontrar algunos resultados. Uno de ellos, debido a J. von Neumann y probado por Aronszajn y Stmith ([7]) en 1954 establece que cualquier operador $T \in L(H)$ compacto posee un subespacio invariante. El problema está aún abierto para T arbitrario sobre H.

Uno de los resultados más generales al respecto es el Teorema de Lomonosov, que da la existencia de un subespacio hiperinvariante para una amplia clase de operadores. Aunque la prueba del resultado de Lomonosov no es muy extensa, el Teorema de Schauder-Tychonoff juega un papel fundamental. Antes de enunciar dicho resultado veamos la siguiente definición.

Definición 2.4.1. Sea X un espacio de Banach. Un subespacio invariante M para $T \in L(X)$ se dice *hiperinvariante* si es invariante para todos los operadores que conmutan con T (esto es, para todo $T' \in L(X)$ tal que $TT' = T'T$).

Nota 2.4.2. Si $T \in L(X)$ es no escalar, es decir, no es un múltiplo del operador identidad, y tiene un autovalor λ, entonces el espacio de autovectores M asociado a λ es hiperinvariante para T. En efecto, si $x \in M$ y T' conmuta con T, tenemos que

$$TT'x = T'Tx = T'\lambda x = \lambda T'x$$

Por tanto, $TT'x = \lambda T'x$ y $T'x \in M$.

Los operadores compactos van a jugar un papel importante en el Teorema de Lomonosov.

Definición 2.4.3. Sean X, Y espacios normados. Un operador lineal $T : X \to Y$ se dice *compacto* si $T(K)$ es relativamente compacto en norma para cada subconjunto acotado K de X.

Teorema 2.4.4 (Lomonosov)**.** Sea X un espacio de Banach. Sea $T \in L(X)$ un operador no escalar conmutando con un operador compacto no trivial $S \in L(X)$. Entonces T tiene un subespacio hiperinvariante no trivial.

Veamos una observación que será usada en la prueba.

Nota 2.4.5. Sea $S \in L(X)$ un operador compacto. Si $\lambda \neq 0$ es un autovalor de S, entonces el espacio de autovectores asociado a λ

$$F = \{x \in X : Sx = \lambda x\}$$

tiene dimensión finita. En efecto, la restricción de S a F es un múltiplo (no cero) del operador identidad en F, y la identidad es compacta si y sólo si el espacio es de dimensión finita.

Demostración. La haremos por reducción al absurdo. Sea $\mathcal{A}$ el álgebra de los operadores conmutando con T, es decir,

$$\mathcal{A} = \{A \in L(X) : AT = TA\}.$$

Notemos que $\mathcal{A} \neq \emptyset\,(S \in \mathcal{A})$ y que es efectivamente un álgebra ya que si $A_1, A_2 \in \mathcal{A}$, entonces $A_1 \circ A_2 \circ T = A_1 \circ T \circ A_2 = T \circ A_1 \circ A_2$, y por tanto $A_1 \circ A_2 \in \mathcal{A}$.
El objetivo de la prueba es encontrar un operador compacto en $\mathcal{A}$ con un autovalor (en este caso el subespacio de autovectores es de dimensión finita e hiperinvariante para T).
Veamos en primer lugar que si T no tiene subespacios hiperinvariantes, entonces $\overline{\mathcal{A}x} = X$ para cada $x \in X, x \neq 0$. Basta probar que $\mathcal{A}x$ es invariante para cualquier $A_0 \in \mathcal{A}$. Sea pues $y \in \mathcal{A}x$, lo que implica $y = Ax$ para $A \in \mathcal{A}$. Entonces $A_0 y = A_0 A x$ y por tanto, $A_0 y \in \mathcal{A}x$ y $\overline{\mathcal{A}x}$ es hiperinvariante para todo x.
Sin pérdida de generalidad, podemos suponer que $\|S\|_{L(X)} \leq 1$. Podemos elegir entonces $x_0 \in X$ tal que $\|Sx_0\| > 1$ (lo que implica que $\|x_0\| > 1$) y tomemos $B = \overline{B}_X(x_0, 1)$.
Sea $x \in \overline{SB}$. Notemos que x no puede ser el vector nulo. En efecto, sea $y \in SB$,

entonces existe $x \in B$ tal que $y = Sx$. Como $\|Sx\| = \|S(x_0) + S(x - x_0)\| \geq \|S(x_0)\| - \|S(x - x_0)\| \geq \|S(x_0)\| - 1 > 0$, entonces $x \neq 0$.

Entones existe $T' \in \mathcal{A}$ tal que $\|T'x - x_0\| < 1$. Por tanto, para cada $x \in \overline{SB}$ existe $T' \in \mathcal{A}$ y un un entorno abierto V_x tal que $T'V_x \subset B$. Por la compacidad de $\overline{SB}$, encontramos un recubrimiento finito $V_1, \ldots, V_n$ y $T'_1, \ldots, T'_n \in \mathcal{A}$ tal que

$$T'_j V_j \subset B, \quad \forall j = 1, \ldots, n$$

y

$$\overline{SB} \subseteq \bigcup_{j=1}^{n} V_j$$

Sea $\varphi_1, \ldots, \varphi_n \in C(\overline{SB})$ una partición de la unidad para $\overline{SB}$ subordinada al recubrimiento $\{V_j\}$, y definamos, para $z \in B$,

$$f(z) = \sum_{j=1}^{n} \varphi_j(Sz) T'_j Sz.$$

Entonces f es una función continua de B en B. En efecto, como $z \in B$, $Sz \in \overline{SB} \subseteq \bigcup_{j=1}^{n} V_j$. Fijado $j \in \{1, \ldots, n\}$ pueden pasar dos cosas:

1. $Sz \in V_j$ lo que implica que $T'_j(Sz) \subseteq B$
2. $Sz \notin V_j$ lo cual implica que $\varphi_j(Sz) = 0$ (pues $\operatorname{sop} \varphi_j \subseteq V_j$)

Como $T'_j S$ es una aplicación compacta para cada j, es fácil ver que $f(B)$ es relativamente compacto. Luego, por el Teorema 2.2.5, existe $\overline{x} \in B$ tal que $f(\overline{x}) = \overline{x}$. Definiendo el operador $\widetilde{T} \in \mathcal{A}$ como

$$\widetilde{T} = \sum_{j=1}^{n} \varphi_j(S\bar{x}) T'_j$$

obtenemos la relación $\widetilde{T}S\overline{x} = \overline{x}$. Ahora bien, como $\widetilde{T} \in \mathcal{A}$ entonces $\widetilde{T}S \in \mathcal{A}$ y al ser S compacto, el operador $\tilde{F}S$ también es compacto. Por tanto, el espacio de autovectores T de $\tilde{T}S$ asociado al autovalor 1 es de dimensión finita. Como $\tilde{T}S$ conmuta con T, concluímos que T es invariante para T, lo cual significa que T tiene un autovalor, y por tanto un subespacio hiperinvariante, contrariamente a nuestra suposición. $\square$

Capítulo 3

Teorema de Markov-Kakutani. Aplicaciones.

Recordemos que el Teorema de Schauder establece que toda aplicación continua definida en un compacto y convexo de un espacio de Banach en sí mismo tiene un punto fijo. Sin embargo, dicho teorema no puede extenderse cuando se estudia la existencia de un punto fijo común para una familia conmutativa de aplicaciones continuas. Por ejemplo, en [2] se puede encontrar un ejemplo de dos aplicaciones definidas de $[0,1]$ en $[0,1]$ las cuales son continuas, conmutan pero no tienen un punto fijo común. Otro ejemplo sería el siguiente:

Ejemplo 3.0.1. Sea $S = \{1,2,3,4,5\}$ y $\Phi = \{\varphi, \psi\}$ donde φ deja fijo al 3,4,5 e intercambia el 1 y el 2, mientras que φ deja fijo al 1 y al 2 y manda el 3 al 4, el 4 al 5 y el 5 al 3. En el lenguaje de las permutaciones serían los ciclos: $\varphi = (\begin{array}{cc} 1 & 2 \end{array})$ y $\psi = (\begin{array}{ccc} 3 & 4 & 5 \end{array})$ y siendo disjuntos conmutan bajo composición. Tenemos pues que S es compacto para la métrica discreta y Φ es una familia conmutativa de aplicaciones continuas en la que cada elemento tiene puntos fijos pero no hay un punto fijo común.

En el caso de que las aplicaciones sean además afines, el Teorema de Markov-Kakutani prueba la existencia de un punto fijo común para una familia conmutativa de funciones continuas definidas en un compacto convexo de un espacio vectorial topológico localmente convexo.

En este capítulo probaremos el Teorema de Markov-Kakutani y daremos algunas aplicaciones:

1. Mostraremos que cada grupo compacto abeliano tiene una "medida de Haar": una medida de probabilidad de Borel regular invariante bajo la acción de un grupo.

2. Medias invariantes en semigrupos.

En el desarrollo de este tema hemos seguido principalmente las referencias de [8] y [11].

3.1. Teorema de Markov-Kakutani.

Antes de enunciar y probar el teorema de Markov-Kakutani necesitamos algunos resultados previos.

Definición 3.1.1. Sea V un espacio vectorial complejo, C un subconjunto convexo de V y $f : C \to V$. Diremos que f es una *aplicación afín* si:

$$f(tx + (1-t)y) = tf(x) + (1-t)f(y), \quad \forall x, y \in C, \quad t \in [0,1]$$

Notemos que la restricción de una aplicación lineal sobre un conjunto convexo es trivialmente afín.

Lema 3.1.2. Sea V un espacio vectorial topológico y supongamos que U es un entorno del vector cero. Entonces $V = \bigcup_{n\in\mathbb{N}} nU$.

Demostración. Sea $x_0 \in V$ y $\phi : \mathbb{R} \times V \longrightarrow V$ tal que $\phi(0, x_0) = 0$. Notemos que $0 \in U$. Como dicha aplicación es continua tenemos que $(0, x_0) \in \phi^{-1}(U)$ es un abierto. Entonces existe $\varepsilon > 0$ y G abierto tal que $\phi((-\varepsilon, \varepsilon) \times G) \subseteq U$.
Para terminar notemos que existe un n suficientemente grande tal que $(1/n, x_0) \in (-\varepsilon, \varepsilon) \times G$, lo cual implica que $\frac{1}{n}x_0 \in U$ y por tanto $x_0 \in nU$. □

Lema 3.1.3. Si K es un subconjunto compacto de un espacio vectorial topológico localmente convexo V con $0 \in K$, entonces $\bigcap_{n\in\mathbb{N}} n^{-1}K = \{0\}$.

Demostración. Supongamos que U es un entorno convexo del vector cero en V. Por el Lema 3.1.2, los conjuntos $\{nU : n \in \mathbb{N}\}$ cubren V, por tanto cubren K. Como dicho conjunto K es compacto existe un subrecubrimiento finito. Además como los conjuntos nU crecen con n, existe $n \in \mathbb{N}$ tal que $K \subset nU$, es decir, $n^{-1}K \subset U$. Por tanto, $\bigcap_{n\in\mathbb{N}} n^{-1}K \subset U$ para cada entorno U del cero en V. El resultado se deduce del hecho de que la topología de V es Hausdorff, así que la intersección de todos los entornos del cero es $\{0\}$ □

Proposición 3.1.4. Sea K un subconjunto convexo y compacto de un espacio vectorial topológico localmente convexo V y $A : K \to K$ una aplicación continua y afín. Entonces, A tiene un punto fijo en K. Además, el conjunto de dichos puntos fijos es compacto y convexo.

Nota 3.1.5. La Proposición 3.1.4 se deduce del Teorema 2.2.5 aunque vamos a dar una prueba alternativa más simple.

Demostración. Sea $\mathbb{N}^* = \mathbb{N} \cup \{0\}$ el conjunto de todos los enteros no negativos. Para cada $n \in \mathbb{N}^*$, denotemos $A^n = A \circ \cdots \circ A$, es decir, la composición de A con sí misma n veces (donde A^0 es la identidad en K). Entonces cada aplicación A^n es afín y continua de K en K, al igual que cada media aritmética M_n definida por:

$$M_n x = \frac{1}{n+1}\sum_{j=0}^{n} A^j x \qquad (x \in K, n \in \mathbb{N}^*).$$

Sea $S = \bigcap_{n\in\mathbb{N}^*} M_n(K)$. Veamos, en primer lugar, que $S \neq \emptyset$.
Para ello, sea $\mathscr{M} = \{M_n : n \in \mathbb{N}^*\}$, así $\mathscr{M}(K) = \{M(K) : M \in \mathscr{M}\}$ es una familia de subconjuntos cerrados del compacto K, con $\bigcap \mathscr{M}(K) = S$. Si podemos probar que cada subfamilia finita de $\mathscr{M}(K)$ tiene intersección no vacía, entonces por la propiedad de intersección finita de conjuntos compactos, lo mismo será cierto de $\mathscr{M}(K)$, acabándose la prueba.
Sea pues $\mathscr{F}$ una subfamilia de $\mathscr{M}$ y sea F la composición de aplicaciones $\mathscr{F}$, cada aplicación ocurriendo exactamente una vez en la composición. Como todas las aplicaciones en $\mathscr{M}$ conmutan bajo la composición, en la definición de F pueden aparecer en cualquier orden. En consecuencia, para cada $M \in \mathscr{F}$ tenemos que

$F = M \circ H$ donde H es una aplicación de K en sí misma, y por tanto $M(K) \supset M(H(K)) = F(K)$.

Conclusión: $\bigcap_{M \in \mathscr{F}} M(K) \supset F(K) \neq \emptyset$. Como S es una intersección de conjuntos compactos y convexos, S es también compacto y convexo. Veamos ahora que S es el conjunto de puntos fijos de A, es decir, $S = \{x \in A : Ax = x\}$.

($\supset$)

Evidentemente cada punto fijo de A pertenece a S.

($\subset$)

Tomemos $y \in S$. Queremos probar que $Ay = y$. Por la definición de S, para cada $n \in \mathbb{N}^*$ hay un vector $x_n \in K$ tal que $y = M_n x_n$. Como la aplicación A es afín, respeta las sumas convexas; en particular, $AM_n = M_n A$ para cada n. Por tanto,

$$\begin{aligned} Ay - y &= AM_n x_n - M_n x_n = M_n A x_n - M_n x_n = \frac{1}{n+1} \sum_{j=0}^{n} (A^{j+1} x_n - A^j x_n) \\ &= \frac{1}{n+1}(A^{n+1} x_n - x_n) \in \frac{1}{n+1}(K - K), \end{aligned}$$

donde $K - K = \{x - y : x, y \in K\}$. Como V es un espacio vectorial topológico, la aplicación $V \times V \to V$ definida por $(v, w) \mapsto v + (-1)w$ es continua, así que la imagen bajo esta aplicación del conjunto compacto K es compacta. En los cálculos anteriores n era un entero no negativo arbitrario, lo que implica que $Ay - y \in \bigcap_{n \in \mathbb{N}^*} \frac{1}{n+1}(K - K)$, el cual por el Lema 3.1.3 (y el hecho de que $0 \in K - K$) consiste sólo del vector cero. Así $Ay = y$.

Hasta ahora lo que hemos visto es que el subconjunto compacto y convexo $S = \bigcap_{n \in \mathbb{N}^*} M_n(K)$ de K es el conjunto de puntos fijos de A. $\square$

Teorema 3.1.6 (Markov-Kakutani)**.** Sea V un espacio vectorial topológico localmente convexo y K un subconjunto no vacío, compacto y convexo. Supongamos que $\mathscr{A}$ es una familia conmutativa de aplicaciones continuas y afines tomando valores de K en sí mismo. Entonces, existe un punto $p \in K$ tal que $Ap = p$ para cada $A \in \mathscr{A}$.

Demostración. Para $A \in \mathscr{A}$ sea $S_A = \{x \in K : Ax = x\}$, el conjunto de puntos fijos de A. Por la Proposición 3.1.4, sabemos que S_A es un subconjunto convexo y compacto de K que es no vacío. Lo que deseamos probar es que $\bigcap_{A\in\mathscr{A}} S_A$, el conjunto de puntos fijos comunes de $\mathscr{A}$, es no vacío. Para ello, es suficiente (de nuevo por la propiedad de intersección finita de compactos) con probar que:

$$\bigcap_{A\in\mathscr{F}} S_A \neq \emptyset. \tag{3.1}$$

para cada subfamilia finita $\mathscr{F}$ de $\mathscr{A}$.

Procedemos por inducción en el número n de elementos de $\mathscr{F}$. El caso $n = 1$ es exactamente la Proposición 3.1.4. Supongamos que (3.1) es cierto para algún $n \geq 1$, y veamos que es cierto para un subconjunto $\{A_1, \ldots, A_{n+1}\}$ de $\mathscr{F}$ formado por $n+1$ aplicaciones. Cojamos una aplicación A de $\mathscr{F}$ y denotemos por S al conjunto de puntos fijos comunes de las n aplicaciones que quedan. Entonces S, siendo la intersección de n subconjuntos compactos y convexos de K, es de nuevo compacto y convexo en K; por nuestra hipótesis de inducción $S \neq \emptyset$. Por conmutatividad, tenemos que $A(S_T) \subset S_T$ para cada $T \in \mathscr{F} \setminus \{A\}$, por tanto A va de S, la intersección de esos conjuntos, en S. Por la Proposición 3.1.4, A tiene un punto fijo en S, el cual es además punto fijo común para $\mathscr{F}$.

Conclusión: (3.1) se cumple para cada subfamilia $\mathscr{F}$ formada por $n+1$ aplicaciones, así que por inducción, se cumple para cada subfamilia finita de $\mathscr{A}$. $\square$

3.2. Medida de Haar como punto fijo.

Nuestro objetivo ahora es aplicar el Teorema de Markov-Kakutani para producir una medida invariante bajo la acción de un grupo compacto y conmutativo G.

3.2.1. Nociones básicas.

En este capítulo consideraremos sólo medidas (positivas) de Borel regulares y de probabilidad. Esto lo denotaremos por MBRP.

Dado que nosotros para construir una “medida de Haar” trabajaremos con grupos topológicos vamos a dar la definición en este momento.

Definición 3.2.1. Sea G un conjunto y $\cdot$ una operación sobre G. Diremos que $(G, \cdot)$ es un *grupo* si se verifican las siguientes propiedades:

(a) La operación es asociativa.

(b) La operación tiene elemento neutro. Es decir, existe un elemento $e \in G$ tal que para todo $x \in G$ se tiene que $x \cdot e = e \cdot x = x$.

(c) Cada elemento de G posee un simétrico. Es decir, para cada $x \in G$ existe un elemento $x^{-1} \in G$ tal que $x \cdot x^{-1} = x^{-1} \cdot x = e$.
Denotaremos, por simplifación, la operación del grupo por simple yuxtaposición.

Un *grupo topológico* es un grupo G dotado con una topología de Hausdorff que hace que las operaciones del grupo sean continuas. Es decir la aplicación $(x, y) \longmapsto xy^{-1}$ es continua para cada $x, y \in G$.
Para cualquier $y \in G$ las aplicaciones $x \mapsto xy$, $x \mapsto yx$, $x \mapsto x^{-1}$ son homeomorfimos de G en G. Por tanto, la topología de G está únicamente determinada por cualquier base local del elemento unidad e. En efecto, si U es un entorno de $e \in G$, los conjuntos $xU = \{xy : y \in U\}$ y $Ux = \{yx : y \in U\}$ son entornos de x

Definición 3.2.2. Sea G un grupo topológico compacto. Una *medida de Haar* para G es una MBRP que es invariante bajo la acción del grupo, es decir, $\mu(gB) = \mu(B)$ para cada $g \in G$ y $B \subset G$, siendo B un conjunto de Borel.

Resulta que cada grupo compacto tiene una (única) medida de Haar. En este tema usaremos teoremas de punto fijo para probarlo.

Nota 3.2.3. Para grupos compactos no conmutativos tenemos que distinguir entre medida de Haar por la izquierda (esta es la que hemos puesto en la definición anterior) y por la derecha (en este caso tenemos una MBRP tal que $\mu(Bg) = \mu(B)$ para cada $g \in G$ y $B \subset G$, siendo B un conjunto de Borel). En el capítulo 12 de [11] se prueba que para grupos compactos ambos conceptos son el mismo y que la medida de Haar es única.

3.2.2. Planteamiento del problema.

Nuestro objetivo es hallar una medida μ tal que $\mu(gB) = \mu(B)$ para cada $g \in G$ y $B \subset G$, siendo B un conjunto de Borel. Esto es equiquivalente a:

$$\int \chi_B \, d\mu = \int \chi_{gB} \, d\mu = \int \chi_B(g^{-1}x) \, d\mu \quad \forall \, B \subseteq G \text{ Borel.} \tag{3.2}$$

Si (3.2) se cumple, la condición de la integral es cierta, por linealidad, para funciones simples y por tanto para cualquier función medible. En particular si f es continua aproximándola por funciones medibles se cumple:

$$\int f \, d\mu = \int f(g^{-1}x) \, d\mu \quad \forall g \in G$$

o de manera equivalente (porque estamos en un grupo):

$$\int f \, d\mu = \int f(gx) \, d\mu \quad \forall g \in G. \tag{3.3}$$

Definamos para cada $g \in G$ la aplicación $L_g : \mathcal{C}(G) \to \mathcal{C}(G)$ mediante:

$$(L_g f)(x) = f(gx) \quad (f \in \mathcal{C}(G)).$$

Entonces (3.3) puede ser escrito como:

$$\int f \, d\mu = \int L_g f \, d\mu \quad \forall g \in G.$$

Recordemos ahora el teorema de representación de Riesz para el dual de $\mathcal{C}(G)$.

Teorema 3.2.4 (Representación de Riesz)**.** Sea Q un espacio topológico de Hausdorff compacto y sea Λ un funcional lineal positivo en $\mathcal{C}(X)$. Entonces, existe una $\sigma-$álgebra $\mathcal{M}$ en Q que contiene a todos los conjuntos de Borel en X, y existe una única medida positiva de Borel μ en $\mathcal{M}$ tal que:

$$\Lambda_\mu(f) = \int_Q f d\mu.$$

Con la notación del Teorema 3.2.4, buscamos un funcional $\Lambda_\mu \in \mathcal{C}(G)^*$ tal que:

$$\Lambda_\mu(f) = \Lambda_\mu \circ L_g(f) \quad \forall f \in \mathcal{C}(G).$$

Es decir, hemos deducido que la condición (3.2) es equivalente a encontrar una medida μ tal que:

$$\Lambda_\mu = \Lambda_\mu \circ L_g \quad \forall g \in G.$$

Vamos a usar ahora el adjunto del operador L_g para reescribir de nuevo nuestro problema como un resultado de punto fijo:

$$(L_g)^* : \mathcal{C}(G)^* \to \mathcal{C}(G)^*$$

Entonces $(L_g)^*(\Lambda_\mu)(f) = \Lambda_\mu(L_g f) = \Lambda_\mu(f)$, para todo $g \in G$, y por tanto buscamos un punto fijo común de la familia de operadores $\{(L_g)^* : g \in G\}$.
Como dijimos, para producir una medida de Haar para un grupo compacto y conmutativo G vamos a aplicar el Teorema 3.1.6. Tomemos $V = \mathcal{C}(G)^*$ y K el conjunto de los funcionales lineales Λ_μ en $\mathcal{C}(G)^*$, donde μ es una RBPM sobre G. La familia $\mathscr{A}$ de aplicaciones afines tomando valores de K en sí mismo será la colección de adjuntos $(L_g)^* : \mathcal{C}(G)^* \to \mathcal{C}(G)^*$ para $g \in G$ (notemos que al ser lineales son afines). No es difícil probar que $\mathscr{A}$ hereda la conmutatividad de G. Tenemos además que las aplicaciones $(L_g)^*$ toman valores de K en K. Por tanto, para aplicar el Teorema de Markov-Kakutani sólo falta encontrar una topología en V que haga K compacto y cada $(L_g)^*$ continua. Entonces, si dotamos a V de la topología débil estrella, es fácil probar que efectivamente cada aplicación $(L_g)^*$ es ω^*-continua. Además en la Sección 2.3 del Capítulo 2 probamos que $(B_{\mathcal{C}(G)^*}, \omega^*)$ es ω^*-compacto, donde $B_{\mathcal{C}(G)^*}$ representa la bola unidad del espacio $\mathcal{C}(G)^*$.
Como conclusión, dicho teorema nos garantiza que $\mathscr{A}$ tiene punto fijo en K y por el Teorema de Representación de Riesz se deduce que existe una medida de Haar para G.
Para terminar, veamos una interesante aplicación del Teorema de Markov-Kakutani.

3.3. Medias invariantes en semigrupos.

Consideremos el espacio de Banach (real) de las funciones acotadas de variable real definidas sobre un semigrupo (un conjunto dotado con una operación binaria interna asociativa) S, es decir,

$$\ell^\infty(S) = \left\{ f : S \to \mathbb{R} : \|f\| := \sup_{s \in S} |f(s)| < \infty \right\}.$$

Un elemento $f \in \ell^\infty(S)$ es positivo si $f(s) \geq 0$ para todo $s \in S$. Diremos que un funcional lineal $\Lambda : \ell^\infty(S) \to \mathbb{R}$ es positivo si $\Lambda f \geq 0$ para todo elemento positivo $f \in \ell^\infty(S)$. Denotaremos una función constante en S por el valor de dicha constante.

Veamos a continuación un resultado que usaremos más adelante.

Lema 3.3.1. Sea $\Lambda \in \ell^\infty(S)^*$, con $\|\Lambda\| = \Lambda 1 = 1$. Entonces Λ es positiva.

Demostración. Supongamos que existe $f \in \ell^\infty(S), f \geq 0$, tal que $\Lambda f = \beta < 0$. Para $\varepsilon > 0$ adecuado, tenemos que:

$$\|1 - \varepsilon f\| = \sup_{s \in S} |1 - \varepsilon f(s)| \leq 1$$

Por tanto,

$$1 < 1 - \varepsilon\beta \leq |1 - \varepsilon\beta| = |\Lambda(1 - \varepsilon f)| \leq \|1 - \varepsilon f\| \leq 1$$

obteniéndose una contradicción. □

Si $t \in S$, podemos definir el operador t-traslación por la izquierda $L_t : \ell^\infty(S) \to \ell^\infty(S)$ de la siguiente forma:

$$(L_t f)(s) = f(ts), \quad \forall s \in S.$$

De manera análoga, podemos definir el operador t-traslación por la derecha.

Definición 3.3.2. Una *media invariante* (por la izquierda) en S es un funcional lineal y positivo Λ en $\ell^\infty(S)$ verificando las siguientes condiciones:

(a) $\Lambda 1 = 1$.

(b) $\Lambda(L_s f) = \Lambda f$ para cada $s \in S$ y cada $f \in \ell^\infty(S)$.

Cuando tal funcional existe, diremos que S es amenable (por la izquierda).

Claramente, podemos dar la definición anterior reemplazando izquierda por derecha. Esta distinción es fundamental si S no es abeliano.

Ejemplo 3.3.3 (Banach)**.** Sea $S = \mathbb{N}$. Entonces $\ell^\infty(\mathbb{N}) = \ell^\infty$. En este caso, una medida invariante se denomina límite de Banach generalizado. La razón es que si Λ es una medida invariante en $\mathbb{N}$ y $x = \{x_n\}_{n\in\mathbb{N}} \in \ell^\infty$ es tal que $\lim_{n\to\infty} x_n = \alpha \in \mathbb{R}$, entonces $\Lambda x = \alpha$. En efecto, para cualquier $\varepsilon > 0$, podemos elegir n_0 tal que $\alpha + \varepsilon \leq x_n \leq \alpha + \varepsilon$ para cada $n \geq n_0$. Por tano, si definimos $y = \{y_n\}_{n\in\mathbb{N}} \in \ell^\infty$ por $y_n = x_{n+n_0}$, tenemos $\Lambda x = \Lambda y$, y

$$\alpha - \varepsilon = \Lambda(\alpha - \varepsilon) \leq \Lambda y \leq \Lambda(\alpha + \varepsilon) \leq \alpha + \varepsilon$$

la cual genera la igualdad $\Lambda x = \alpha$.
Para probar la existencia de una medida invariante, vamos a considerar el subespacio $\mathcal{M}$ de ℓ^∞ dado por:

$$\mathcal{M} = \left\{ x = \{x_n\}_{n\in\mathbb{N}} \in \ell^\infty : \lim_{n\to\infty} \frac{x_0 + \cdots + x_n}{n+1} = \alpha_x \in \mathbb{R} \right\}$$

y definir el funcional lineal Λ_0 en $\mathcal{M}$ como $\Lambda_0 x = \alpha_x$. Tomando para cada $x \in \ell^\infty$

$$p(x) = \limsup_{n\to\infty} \frac{x_0 + \cdots + x_n}{n+1}$$

es posible, usando el Teorema de Hahn-Banach, extender Λ_0 a un funcional Λ definido en todo el espacio, de tal manera que $-p(-x) \leq \Lambda x \leq p(x)$ para cada $x \in \ell^\infty$. En particular, Λ es continua (ya que $|\Lambda(x)| \leq |p(x)| \leq \|x\|_\infty$). Una consecuencia notable de este hecho es que no todo funcional lineal y continuo en ℓ^∞ admite una representación de la forma:

$$\{c_n\}_{n\in\mathbb{N}} \mapsto \sum_{n=0}^{\infty} c_n x_n.$$

para alguna sucesión numérica c_n. En efecto, para cada $k \in \mathbb{N}$, tomando $e_k = \{\delta_{nk}\}_{n\in\mathbb{N}}$, tenemos $\Lambda e_k = 0$. Por tanto si Λ tiene la representación anterior, todos los c_n deben ser cero, es decir, Λ es el funcional nulo, contradiciendo el hecho de que $\Lambda 1 = 1$.

El siguiente resultado, basado en el Teorema de Markov-Kakutani, da una elegante generalización del Ejemplo 3.3.3.

Teorema 3.3.4 (Day)**.** Sea S un semigrupo abeliano. Entonces S es amenable.

Demostración. Consideremos el siguiente conjunto,

$$K = \{\Lambda \in \ell^\infty(S)^* : \|\Lambda\| = \Lambda 1 = 1\}.$$

En particular, si $\Lambda \in K$ entonces Λ es positiva. Ahora bien K es convexo, y por el Teorema de Banach-Alaoglu es compacto en la topología débil-* de $\ell^\infty(S)^*$. Definamos la familia de operadores lineales $T_s : \ell^\infty(S)^* \longrightarrow \ell^\infty(S)^*, s \in S$, como

$$(T_s\Lambda)(f) = \Lambda(L_s f), \quad \forall f \in \ell^\infty(S).$$

Primero veamos que T_s es continua en la topología débil-* para cada $s \in S$. Por supuesto, es suficiente con probar la continuidad en el cero. Sea entonces V una base local del cero para dicha topología, esto es,

$$V = \{\Lambda \in \ell^\infty(S)^* : |\Lambda f_j| < \varepsilon_j,\, j = 1, \ldots, n\}$$

para ciertos $\varepsilon_1, \ldots, \varepsilon_n > 0$ y $f_1, \ldots, f_n \in \ell^\infty(S)$. Entonces,

$$\begin{aligned} T_s^{-1}(V) &= \{\Lambda \in \ell^\infty(S)^* : |(T_s\Lambda)(f_j)| < \varepsilon_j, j = 1, \ldots, n\} \\ &= \{\Lambda \in \ell^\infty(S)^* : |\Lambda(L_s f_j)| < \varepsilon_j, j = 1, \ldots, n\} \end{aligned}$$

es un entorno abierto del cero.
El segundo paso consiste en probar que $T_s K \subset K$. En efecto,

$$(T_s\Lambda)(1) = \Lambda(L_s 1) = \Lambda 1 = 1$$

y

$$\|T_s\Lambda\| = \sup_{\|f|\leq 1} |(T_s\Lambda)(f)| = \sup_{\|f|\leq 1} |\Lambda(L_s f)| \leq \sup_{\|f\|\leq 1} |\Lambda f| = \|\Lambda\| = 1$$

como $\|L_s f\| \leq \|f\|$.
Finalmente, para cada $s, t \in S$,

$$T_s T_t \Lambda = T_s(\Lambda \circ L_t) = \Lambda \circ L_t \circ L_s = \Lambda \circ L_{st} = \Lambda \circ L_{ts} = T_t T_s \Lambda.$$

Por tanto, por el Teorema de Markov-Kakutani, existe $\Lambda \in K$ tal que $T_s\Lambda = \Lambda$ para cada $s \in S$, lo cual significa que $\Lambda(L_s f) = \Lambda f$ para cada $s \in S$ y $f \in \ell^\infty(S)$. $\square$

Bibliografía

[1] AYERBE, J.M. & DOMÍNGUEZ, T. & LÓPEZ, G., Measures of noncompactness in metric fixed point theory, Birkhäuser, 1997.

[2] BOYCE, W.M., *Commuting functions with no common fixed point*, American Mathematical Society, 137 (1967), 77-92.

[3] BLUMENTHAL, L., Theory and applications of distance geometry, Oxford University Press 1953.

[4] C.J. READ, *A solution to the invariant subspace problem on the space* ℓ_1, Bull. London Math. Soc. 17 (1985), 305-317.

[5] GOEBEL, K. & KIRK, W.A., Topics in metric fixed point theory, Cambridge University Press 1990.

[6] KRANTZ, S., A guide to functional analysis, American Mathematical Society, 2013.

[7] N. ARONSZAJN & K. T. STMITH, *Invariant subspaces of completely continuous operator*, Annals of Mathematics, 60 (1954), 345-350.

[8] PATA, V., Fixed point theorems and applications, Politecnico di Milano.

[9] RUDIN, W., *Functional analysis*, McGraw-Hill Book Company, 1973.

[10] RUDIN, W., *Real and complex analysis*, McGraw-Hill Book Company, 1987.

[11] SHAPIRO, J.H, A fixed-point farrago, Springer, 1990.

[12] WILANSKY, A., Topology for analysis, Dover Publications, 1983.

[13] WOJTASZCZYK, P., Banach spaces for analysts, Cambridge University Press, 1991.

[14] YAGMAGUTI, M. & HATA, M. & KIGAMI, J., Mathematics of fractals, American Mathematical Society, 1993.