Universidad Nacional del Centro de la Provincia de Buenos Aires

Facultad de Ciencias Exactas

Departamento de Matemática

SOBRE TEORIA ERGODICA

Trabajo Final para optar por el título de Licenciado en Ciencias Matemáticas

Ulises Chialva Director: Dr. Carlos Peña

Marzo de 2015

Índice general

Capítulo 1

Introducción

Así como el estudio de la topología se supone que comenzó con el problema de cómo atravesar los puentes de Koenisberg, de manera análoga, la teoría ergódica surge de problemas sugeridos por la mecánica estadística de los gases. Y de manera similar, mientras que el problema de los puentes dió lugar al estudio de los grafos y sus posibles recorridos, el problema de la mecánica de gases derivó en el estudio de ciertos objetos llamados "transformaciones preservadoras de medidas", que resultaron ser el ejemplo principal de lo que luego serían los sistemas dinámicos discretos y sobre la cual von Neumann demostró la primera versión del teorema ergódico medio.

Si tratamos de precisar el concepto de *sistemas dinámicos*, podríamos decir que se trata del estudio de sistemas deterministas, es decir, consideramos situaciones que dependan de algún parámetro dado, que frecuentemente suponemos es el tiempo, y que varían de acuerdo a leyes establecidas. De manera que el conocimiento de la situación en un momento dado nos permite reconstruir el pasado y predecir el futuro. En general, dado un espacio de fases X, que representa el conjunto de todos los posibles estados en los que se encuentre un sistema u objeto, consideremos una aplicación $T_t : X \to X$ continua. El significado físico que le daremos a la aplicación es sencillo: dado $x_0 \in X$, el elemento $x_t = T_t(x_0)$ describe el estado en que, luego de un tiempo t, se encontrará el sistema que originalmente estaba en el estado x_0. Así a cada instante de tiempo t, que en general se toma positivo, le asignamos un automorfismo T_t definido en el espacio de fases X. Además, por obvias consideraciones físicas, debe ocurrir que $T_s(T_t(x_0)) = T_{s+t}(x_0)$. El conjunto $\{T_t\}$ forma así un *grupo uniparamétrico de automorfismos*. Formalmente, esto se sistematiza del siguiente modo:

Definición 1.0.1. Dado un espacio topológico X y un semigrupo S, un sistema dinámico es una aplicación

$$T : S \times X \to X \text{ , } T(s, x) = T_s(x)$$

continua en la variable x para cada $s \in S$ fijo, de modo que $T_0(x) = x$ para cada $x \in X$

y además $T_t \circ T_s = T_{st}$ si $s, t \in S$ (propiedad de semigrupo).[1]

Ejemplo 1.0.1. Supongamos que deseamos solicitar un préstamo al banco. La tasa de interés que cobra el banco es de 2 por ciento mensual, y nuestra capacidad de pago real es de dos mil pesos mensuales máximo. ¿Cuánto dinero queremos que nos preste el banco? Ingenuamente podríamos responder "todo lo que se pueda", pero vamos a analizar este sencillo ejemplo un poco más. Denotemos por x_0 la cantidad de dinero que queremos pedir prestado al banco, y por x_n nuestra deuda después de n meses. Entonces tenemos que

$$\begin{aligned} x_1 &= (1,02)x_0 - 2000 \\ x_2 &= (1,02)x_1 - 2000 = (1,02)^2 x_0 - 2000(1+1,02) \\ x_n &= (1,02)^n x_0 - 2000 \sum_{k=0}^{n-1} (1,02)^k \end{aligned}$$

Haciendo $T(x) = (1,02)x - 2000$ tendremos que $x_n = T^n(x_0)$. En este caso el semigrupo $\mathcal{S}$ son los naturales, y el espacio de fases $\mathbb{R}_{\geq 0}$.
Si el préstamo inicial es de 100000 pesos, resultará que $x_n = 100000$ para cada $n \in \mathbb{N}$. Es decir, $x_0 = 100000$ es un punto fijo del sistema dinámico.
Podemos entonces concluir:

a) Si el banco nos presta menos de un 100000, algún día terminaremos de pagarle.

b) Si el banco nos presta más de 100000, no importa cuanto tiempo paguemos, jamás alcanzará el tiempo para cancelar la deuda.

c) Si el banco nos presta exactamente 100000, simplemente le pagaremos dos mil pesos mensuales de interés por el resto de nuestra vida, y 'en el infinito' cancelaremos la deuda.

Así- que conocer las leyes que rigen el sistema, nos permite predecir el futuro de la deuda.
Expresiones del tipo presentado, donde tenemos una sucesión de valores $\{x_n\}$, n un natural, tales que el valor de x_n está determinado por los valores anteriores x_{n-1}, x_{n-2}, son ejemplo de sistemas dinámicos discretos, donde la palabra discreto significa que el parámetro "tiempo" toma valores discretos (cada mes, cada año, cada hora, etc). Los sistemas dinámicos discretos, más sencillos surgen mediante la iteración de funciones. En este ejemplo particular, $T(x) = (1,02)x - 2000$.

[1] Los sistemas dinámicos en que $S = \mathbb{R}_{\geq 0}$ se suelen denominar *flujos* (flows).

Lo que deseamos rescatar del ejemplo anterior, es que cada valor x_n está completamente determinado por los anteriores, y gracias a ello podemos llegar a calcularlo. Sin embargo, realmente no importa saber qué valor exacto tendrá x_n, lo que importa saber es si $\lim_{n\to\infty} x_n$ tiende a 0 o no (es decir, si nuestra deuda termina un día o no).

En general, por cómo como se definen los sistemas dinámicos, dado un *estado* x_0 en un *momento* $t_0 = 0$ los instantes siguientes están completamente determinados. Sin embargo el estudio de la dinámica *no consiste en la determinación exacta de la posición* x_t. Lo que se procura describir es el comportamiento *asintótico* o *recurrente* de la transformación T_t. Básicamente la pregunta *¿qué posición exacta ocupará el elemento* x_t*?* resulta sustituída por *¿cuál es la probabilidad de que* x_t *pertenezca a cierto subconjunto de* X*?*. Dar respuestas a la última pregunta es lo que se entiende por estudio del *comportamiento asintótico o recurrente del sistema*. Quien introdujo el estudio del comportamiento recurrente de transformaciones fué el matemático francés Henri Poincaré (1854-1912), y lo hizo en el contexto de sus investigaciones sobre el comportamiento de soluciones de sistemas hamiltonianos. Inspirado en resultados de la mecánica como el teorema de Louville, Poincaré introdujo el concepto de *medidas invariantes*, que sería el primer objeto sobre el cual pudo estudiarse el comportamiento recurrente. Sin embargo sus resultados aplican sobre el caso en que el tiempo es discreto en vez de continuo, y esto constituye la primera simplificación del problema: si $T_0 = T$ y el tiempo es indexado naturalmente, la propiedad de semigrupo implica que $T_n = T^n$. De esta manera el estudio de tal sistema dinámico queda reducido al de las sucesivas potencias de la transformación T.

En particular, el interés por el teorema ergódico se debió al estudio de ciertos sistemas dinámicos de la física matemática en los cuales resultaban fundamentales las medidas invariantes introducidas por Poincaré. Asimismo, era relevante justificar matemáticamente la llamada *hipótesis ergódica* debida al físico Ludwig Boltzmann (1844-1906): para ciertas medidas invariantes μ, ciertos flujos T_t y ciertas funciones medibles $f : X \to \mathbb{C}$, existen infinitos estados $\overline{x} \in X$ tales que existe el *promedio temporal continuo*

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f(T_t\overline{x}) \tag{1.1}$$

y coincide con el *promedio espacial* $\frac{1}{\mu(X)} \int_X f d\mu$.[2]

Al mismo tiempo, la utilización del enfoque introducido por Poincaré de simplificar el semigrupo de transformaciones $\{T_t\}$ del flujo sustituyéndolo por las potencias sucesivas de una única transformación $T = T_0$, llevó a estudiar, en relación a la fórmula 1.1, condiciones bajo las cuales tuviera sentido y quedara garantizada la convergencia de la

[2]Señalemos que esta no es la formulación original dada por Boltzmann.

expresión denominada *promedio temporal discreto*

$$\frac{1}{N}\sum_{n=0}^{N-1} f(T^n x). \tag{1.2}$$

Fue VonNeumann en 1931 quien consiguió dar condiciones de convergencia de la expresión 1.2 para operadores de composición en espacios $L^2(\mu)$. Su afamado resultado, conocido como *teorema ergódico medio,* daría origen a una fructífera búsqueda de la comunidad matemática contemporánea por generalizarlo en nuevos contextos y expandir su rango de aplicación.

Así, a través de los estudios posteriores de Birkhoff, Yosida, Kakutani, Eberlein, Lorch, etc. se demostraron e introdujeron los denominados *teoremas ergódicos*, que es la nomenclatura asignada a aquellos resultados que brindan condiciones de convergencia para la expresión 1.2 en distintos espacios (espacios de Hilbert, de Banach, espacios vectoriales localmente convexos, etc.).

El objetivo del presente trabajo es dar cuenta de algunos hitos de tal búsqueda, mediante la exposición de aquellos aportes individuales que sistematizaron los principales desarrollos de la comunidad matemática al respecto; desarrollos realizados en su mayoría, durante el período 1930-1970. En el Capítulo 2 presentamos el teorema de recurrencia de Poincaré.

Capítulo 2

El teorema de H. Poincaré

Como ya señalamos, el introductor del estudio del comportamiento recurrente fue Henri Poincaré. Sin embargo sus resultados aplican sobre el caso en que el tiempo es discreto en vez de continuo, y esto constituye la primera simplificación del problema: si $T_0 = T$ y el tiempo es indexado naturalmente, la propiedad de semigrupo de la definición 1.0.1 implica que $T_n = T^n$ por lo que el estudio del sistema discreto queda reducido al de las sucesivas potencias de la transformación $T = T_0$. De esta manera, los sistemas dinámicos discretos quedan abarcados en la siguiente definición:

Definición 2.0.2. Un sistema dinámico topológico es un par $T = (X, T)$, donde X es un espacio topológico y T un homeomorfismo sobre X.

2.1. El teorema de recurrencia

Definición 2.1.1. Dados un espacio de medida $(X, \mathfrak{B}, \mu)$ y $T : X \to X$ una aplicación medible, diremos que μ es T-invariante si para todo $A \in \mathfrak{B}$ es $\mu(T^{-1}A) = \mu(A)$. Al conjunto de medidas borelianas T-invariantes sobre X lo indicaremos $M_T(X)$.

Teorema 2.1.2. *(Poincaré) Sean $(X, \mathfrak{B}, \mu)$ espacio de probabilidad, $T : X \to X$ función medible y asumamos que μ es T-invariante. Si $A \in \mathfrak{B}$ y $\mu(A) > 0$ para $x \in A$ μ-c.t.p. existe $n \in \mathbb{N}$ tal que $T^n(x) \in A$. En particular, existe $N \in \mathbb{N}$ tal que $\mu(A \cap T^{-N}A) > 0$.*

Demostración. Para todo $k \in \mathbb{N}$ sea $B_k =: \{x \in A : T^j(x) \notin A \, \forall j \geq k\}$. Llamemos $E = A - \bigcup_{k=1}^{\infty} B_k$. Si resulta $\mu(E) = \mu(A)$ tendremos probada la primer parte del teorema. Veamos entonces que $\mu(B_k) = 0$ si $k \in \mathbb{N}$. Tenemos

$$B_k = A - \bigcup_{j=k}^{\infty} T^j A \subset \bigcup_{j=0}^{\infty} T^{-j} A - \bigcup_{j=k}^{\infty} T^{-j} A.$$

En particular
$$\mu(B_k) \leq \mu(\bigcup_{j=0}^{\infty} T^{-j}A) - \mu(\bigcup_{j=k}^{\infty} T^{-j}A).$$
Pero siendo μ una medida T-invariante resulta
$$\mu(\bigcup_{j=k}^{\infty} T^{-j}A) = \mu(T^{-k}(\bigcup_{j=0}^{\infty} T^{-j}A)) = \mu(\bigcup_{j=0}^{\infty} T^{-j}A)$$
con lo cual, $\mu(B_k) \leq 0$. Sea ahora $F =: \{x \in A : \exists n \in \mathbb{N}/T^n(x) \in A\}$. Por lo anterior $F \subset A$ y $\mu(F) = \mu(A)$. Por otro lado $F = A \cap \bigcup_{j=1}^{\infty} T^{-j}A$ y $\mu(F) = \mu(\bigcup_{j=1}^{\infty} A \cap T^{-j}A)$. Si ocurriese que $\forall j\ \mu(A \cap T^{-j}A) = 0$ entonces
$$\mu(A) = \mu(E) \leq \sum_j \mu(A \cap T^{-j}A) = 0,$$
lo cual es absurdo. Luego debe existir N tal que $\mu(A \cap T^{-N}A) > 0$. ■

Observación 2.1.3. Notemos que el resultado 2.1.2 afirma que para μ-c.t.p. $x \in X$, dado cualquier medible A tal que $x \in A$ y $\mu(A) > 0$ tendremos que existe una sucesión de índices n_k tal que $T^{n_k}(x) \in A$. A los puntos que satisfacen tal condición se los denomina T-recurrentes. De esta forma, una formulación usual y alternativa del teorema 2.1.2 es la siguiente:

Proposición 2.1.4. *Sea $(X, \mathfrak{B}, \mu)$ un espacio de medida finito T-invariante. Entonces casi todo punto es T-recurrente.*

Capítulo 3

Teorema ergódico de J. Von Neumann

En todo sistema dinámico topológico (X, T) hay inducido un *operador de composición* $C_T : C_b(X) \to C_b(X)$ vía $C_T(f) = f \circ T$, donde $C_b(X)$ es el espacio de Banach de las funciones complejas continuas acotadas sobre X consideradas con la norma uniforme $\|f\|_\infty = \sup_{x \in X} |f(x)|$. Es inmediato que este operador es lineal, continuo y contractivo. Si X es compacto, en virtud del *teorema de representación de Riesz* tendremos que $C(X)^* = M(X)$, donde $M(X)$ denota el conjunto de todas las medidas complejas regulares de Borel sobre X. Entonces, el operador adjunto tiene la forma $C_T^* : M(X) \to M(X)$ vía $C_T^*(\mu)(A) = \mu(T^{-1}A)$ para todo boreliano A. En efecto:

$$\begin{aligned}
\langle C_T(f); \mu \rangle &= \int_X f(T(x)) d\mu(x) \\
&= \int_X f\,(y)\, d\mu(T^{-1}(y)) \text{ (T es homeomorfismo)} \\
&= \langle f; C_T^*(\mu) \rangle.
\end{aligned}$$

Si μ fuere T-invariante por lo anterior, resultará

$$\int_X f d\mu = \int_X (f \circ T)\, d\mu \tag{3.1}$$

De esta manera, hemos obtenido la siguiente

Proposición 3.0.5. *Si μ es una medida de probabilidad, X es compacto y $\mu \in M(X)$, entonces μ es T-invariante sii $C_T^*(\mu) = \mu$.*

Además, asumiendo siempre que X es compacto, si $\mu \in M(X)$ entonces $C(X)$ se sumerje continuamente en $L^p(\mu)$ cuando $p \geq 1$ y $\|f\|_{L^p(X,\mu)} \leq \|\mu\|_{M(X)}^{1/p} \|f\|_\infty$. Por otra parte, si $f \neq 0$ entonces existe U abierto tal que $f\,|_U$ es no nula en todo punto. Si además μ fuese tal que $\mu(U) > 0$ para cualquier abierto no vacío, por la regularidad de μ obtenemos

$$\int_X |f|^p\, d\mu \geq \int_U |f|^p\, d\mu > 0$$

i.e. hay una inmersión $C(X) \hookrightarrow L^p(\mu)$. Así, por la densidad de $C(X)$ en $L^p(\mu)$ y la completitud de $L^p(\mu)$, sigue que C_T es extendible a un operador $\widetilde{C}_T \in \mathcal{B}(L^p(\mu))$ cuya norma no excede $\|\mu\|^{1/p}$. Si $\mu \in M(X)$ es T-invariante, $\widetilde{C}_T$ deviene isométrico como consecuencia de (3.1). Recordemos que notamos con M_T al conjunto de las medidas de probabilidad T-invariantes.

Proposición 3.0.6. *(cf. [11]) Sean X compacto, $T : X \to X$ homeomorfismo y $\mu \in M_T$. Entonces $\widetilde{C}_T$ es un operador unitario sobre $L^2(\mu)$.*

Demostración. Si $f, g \in C(X)$ tenemos

$$\begin{aligned}\left\langle \widetilde{C}_T(f), g \right\rangle &= \langle C_T(f), g\rangle \\ &= \int_X f(T(x))\,\overline{g(x)}\,d\mu(x) \\ &= \int_X f(y)\,\overline{g(T^{-1}(y))}\,d\mu(y) \quad \text{(por (3.1))} \\ &= \left\langle f, \widetilde{C}_{T^{-1}}(g) \right\rangle.\end{aligned}$$

Podemos inferir que $\left(\widetilde{C}_T\right)^* = \widetilde{C}_{T^{-1}} = \left(\widetilde{C}_T\right)^{-1}$ y sigue la afirmación. ■

Definición 3.0.7. Se dirá que una medida μ es T-ergódica si los únicos puntos fijos de $\widetilde{C}_T$ son las funciones constantes.

Proposición 3.0.8. *Con la notación anterior, son equivalentes:*

(a) *μ es ergódica.*

(b) *Para todo boreliano A, si $T^{-1}(A) = A$ entonces $\mu(A) = 1$ ó 0.*

Demostración.

$(a \Rightarrow b)$ Sea A boreliano tal que $T^{-1}(A) = A$ y hagamos $f = \chi_A$. Sea $\{f_n\} \subseteq C(X)$ tal que $\|f - f_n\|_2 \to 0$. Luego

$$\widetilde{C}_T f = \lim_{n\to\infty} \widetilde{C}_T f_n = \lim_{n\to\infty} f_n \circ T.$$

Como T es biyectiva y $T(A) = A$ deducimos que $f \circ T = f$ y, como μ es T-invariante usando (3.1) tenemos

$$\begin{aligned}\left\|\widetilde{C}_T f_n - f \circ T\right\|_2 &= \|(f_n - f) \circ T\|_2 \\ &= \left(\int_X |(f_n - f) \circ T|^2 \, d\mu\right)^{1/2} \\ &= \left(\int_X |f_n - f|^2 \, d\mu\right)^{1/2} \\ &= \|f - f_n\|_2 \to 0 \text{ si } n \to \infty,\end{aligned}$$

i.e. $\widetilde{C}_T f = f \circ T = f$. Como μ se asume ergódica la condición es necesaria.

$(b \Rightarrow a)$ Sea $f \in L^2(\mu)$ tal que $\widetilde{C}_T f = f$. Podemos suponer que f es finita en todo punto y considerar $\{f_n\} \subseteq C(X)$ tal que $\|f - f_n\|_2 \to 0$. Más aún, podemos suponer que $f_n \to f$ en todo punto, sino consideramos alguna subsucesión en el que la convergencia es puntual en casi todo punto. Tampoco hay pérdida de generalidad si suponemos que f es real. Dado $a \in \mathbb{R}$ escribamos $E_a = \{f \leq a\}$. Así

$$x \in T^{-1}(E_a) - E_a \text{ si y solo si } f(T(x)) \leq a < f(x).$$

Pero

$$f(T(x)) = \lim_{n\to\infty} f_n(T(x)) = \widetilde{C}_T f(x) = f(x) \text{ a.e. } \mu,$$

i.e. $\mu(T^{-1}(E_a) - E_a) = 0$. Análogamente, $\mu(E_a - T^{-1}(E_a)) = 0$ e inferimos que $E_a = T^{-1}(E_a)$ a.e. μ. Por hipótesis, $\mu(E_a) \in \{0, 1\}$. Hagamos

$$A = \{a \in \mathbb{R} : \mu(E_a) = 1\}.$$

Si $A = \varnothing$ entonces $f > a$ a.e. μ cualquiera sea $a \in \mathbb{R}$, de donde sigue fácilmente que $f = +\infty$ a.e. μ, lo cual es imposible. Asimismo sería $f = -\infty$ a.e. μ en caso que A no sea acotado inferiormente, lo que tampoco es posible. Indiquemos $\alpha = \inf A$. Como $\{f < \alpha\} = \cup_{n=1}^{\infty} E_{\alpha - 1/n}$ vemos que $\mu(\{f < \alpha\}) = 0$. Además $\{f > \alpha\} = \cup_{n=1}^{\infty} \{f > \alpha + 1/n\}$ y para $n \in \mathbb{N}$ es

$$\mu(\{f > \alpha + 1/n\}) = 1 - \mu\left(E_{\alpha+1/n}\right) = 0,$$

o sea $\mu(\{f > \alpha\}) = 0$. En definitiva, $f = \alpha$ a.e. μ. El caso general se deduce considerando las partes real e imaginaria de f.

■

El siguiente resultado es conocido como ***Teorema Ergódico Medio***[1]:

Teorema 3.0.9. *(von Neumann, cf. [18]) Sean $\mu \in M(X)$ una medida T-invariante, X espacio compacto, $U = \widetilde{C}_T$ y si $N \in \mathbb{N}$ hagamos $U_N = \frac{1}{N}\sum_{n=0}^{N-1} U^n$.*

(i) *La sucesión de promedios $\{U_N\}$ convergen fuertemente en $L^2(\mu)$ a la proyección P sobre el subespacio $\ker(U - I_{L^2(\mu)})$.*

(ii) *Si T es μ-ergódica resulta $\ker(U - I) = \mathbb{C} \cdot I_{L^2(\mu)}$ y $P(f) = \left(\int_X f d\mu\right) \cdot 1_X$.*

Demostración.

(i) Primero observemos que dado $f \in \ker(U - I)$ entonces vale $U_N f = f$ cualquiera sea N. Por otra parte, si $f \in \overline{\mathrm{Ran}(U - I)}$ veamos que $U_N f \to 0$. Si $f \in \overline{\mathrm{Ran}(U - I)}$ dado $\epsilon > 0$ existen $g, h \in \mathcal{H}$ tal que $\|h\| \leq \epsilon$ y $f = Ug - g + h$. Entonces, para N suficientemente grande obtenemos

$$\|U_N f\| = \|U_N(Ug - g + h)\| \leq \|U_N(Ug - g)\| + \|U_N h\| \leq \frac{2\|g\|}{N} + \epsilon \leq 2\epsilon.$$

De lo anterior se sigue $U_N f \longrightarrow 0$. Bastará ver entonces que

$$\overline{\mathrm{Ran}(U - I)} \oplus \ker(U - I) = \mathcal{H}.$$

Para ello, como U es unitario basta señalar las identidades

$$\overline{\mathrm{Ran}(U - I)} = \ker(U^* - I)^{\perp} \quad \text{y} \quad \ker(U^* - I) = \ker(U - I).$$

(ii) Es inmediato.

■

[1]La demostración original que Von Neumann dio del teorema 3.0.9 involucraba teoría espectral y recurría a resultados sobre la convergencia de la integral de Stieltjes. Sin embargo, Riesz dio una demostración elemental y generalizada que sólo involucra un espacio de Hilbert $\mathcal{H}$ arbitrario y un operador unitario U. Para una discusión acerca de las implicaciones físicas del teorema ergódico remitimos al lector al artículo de Von Neumann [19].

Capítulo 4

Teorema ergódico de G. Birkhoff

Observación 4.0.10. Sabemos que $L^2(\mu) \subset L^1(\mu)$ si μ es finita. Además, tomando una función $f \in L^2(\mu)$ tendremos por la desigualdad de Cauchy-Schwartz

$$\|f\|_1 = \int_X |f| d\mu = \langle |f|; 1 \rangle \leq \|f\|_2$$

por lo que la inclusión $L^2(\mu) \hookrightarrow L^1(\mu)$ es continua, y por la densidad de funciones simples, se obtiene la densidad del rango en el codominio. En este contexto, resulta natural intentar extender el Teorema Ergódico Medio 3.0.9 a $L^1(\mu)$, de manera que los promedios $\frac{1}{N}\sum_{n=0}^{N-1} f \circ T^n$ converjan para toda $f \in L^1(\mu)$.

Teorema 4.0.11. *(cf.[2], [3]) Sea $(X, \mathfrak{B}, \mu)$ espacio de probabilidad, $T : X \to X$ una transformación que preserva medidas. Dada $f \in L^1(\mu)$ existe en μ-c.t.p.*

$$\widehat{f}(x) =: \lim_{N\to\infty} (1/N) \sum_{n=0}^{N-1} f(T^n x)$$

y además $\widehat{f} \in L^1(X, \mathfrak{B}, \mu)$ verificándose $\int_X \widehat{f} d\mu = \int_X f d\mu$.

Antes de demostrar el teorema anterior, deberemos introducir ciertas funciones especiales y dar algunos resultados técnicos. Para cada $f \in L^1(X, \mu)$ real, definiremos

$$f^+(x) = \limsup_N (1/N) \sum_{n=0}^{N-1} f(T^n x) \quad \text{y} \quad f^-(x) = \liminf_N (1/N) \sum_{n=0}^{N-1} f(T^n x).$$

Proposición 4.0.12. *Sea $f \in L^1(\mu)$ real, entonces $f^+ \circ T = f^+$ y $f^- \circ T = f^-$*

Demostración. Si $N \in \mathbb{N}$ sea $f_N(x) =: (1/N) \sum_{n=0}^{N-1} f(T^n x)$. Por construcción tendremos que

$$f_{N+1}(x) = (N/N+1) f_N(Tx) + (1/N) f(x).$$

Así por la subaditividad del límite superior valdrá que $f^+(x) \leq f^+(Tx)$. Por otro lado, teniendo que

$$f_N(Tx) = (N + 1/N)f_{N+1}(x) - f(x)/N,$$

y utilizando nuevamente la subaditividad, tendremos la otra desigualdad. Resultará pues $f^+(x) = f^+(Tx)$. Utilizando la superaditividad del límite inferior y razonando de manera análoga se demuestra la proposición para f^-. ∎

El siguiente resultado, conocido como *Teorema Ergódico Maximal* es de fundamental importancia:

Teorema 4.0.13. *Dada $f \in L^1(\mu)$ sea $E(f) = \{x : \sup_N \sum_{n=0}^{N} f(T^n x) > 0\}$. Entonces*

$$\int_{E(f)} f d\mu \geq 0.$$

Demostración. Escribamos

$$f_n(x) = \max\{0, f(x), f(x) + f(Tx), ..., \sum_{i=0}^{n-1} f(T^i x)\}, n \in \mathbb{N}.$$

Notemos que para toda $x \in X$ tendremos que valiendo $f_n(x) \leq f_{n+1}(x)$ resulta que $\{f_n > 0\} \subset \{f_{n+1} > 0\}$. Así rápidamente podemos ver que $E(f) = \bigcup_n \{f_n > 0\}$. Sea $A_n = \{f_n > 0\}$. Entonces $|\chi_{A_n} f| \leq |f|$, donde f integrable. Además $X_{A_n} f \to \chi_{E(f)} f$ puntualmente (pues A_n es una sucesión creciente y su unión da $E(f)$). Por el Teorema de Convergencia Dominada tendremos que

$$\int_{\{f_n>0\}} f d\mu = \int_{A_n} f d\mu \longrightarrow \int_{E(f)} f d\mu.$$

En consecuencia, para demostrar el teorema bastará probar que $\int_{\{f_n>0\}} f d\mu \geq 0$. En efecto, si $0 \leq n < m$ tendremos que

$$f_n(Tx) + f(x) \geq \sum_{i=0}^{m} f(T^i Tx) + f(x) = \sum_{i=0}^{m+1} f(T^i x).$$

Así $f_n(Tx) + f(x) \geq f_n(x)$ si $f_n(x) > 0$. Entonces

$$\begin{aligned}
\int_{\{f_n>0\}} f d\mu &\geq \int_{\{f_n>0\}} f_n - \int_{\{f_n>0\}} f_n \circ T d\mu \\
&= \int_X f_n - \int_{\{f_n>0\}} f_n \circ T \\
&\geq \int_X f_n - \int_X f_n \circ T = 0,
\end{aligned}$$

donde la igualdad del medio sale porque $f_n = 0$ sobre $X - \{f_n > 0\}$ y la igualdad del último término por la T-invariancia de μ. ■

Corolario 4.0.14. *Si $A \subset E(f)$, A es medible y $T^{-1}A = A$, entonces* $\int_A f d\mu \geq 0$.

Demostración. Como $T^{-1}A = A$ tendremos que $E(\chi_A f) = A$. En efecto: si $x \in E(\chi_A f)$ entonces $\sup_n\{\sum_{i=0}^n f(T^i x)\} > 0$, de lo que sigue $\chi_A(x) > 0$ y resulta $x \in A$.
Recíprocamente: si $\chi_A(x) = 1$, como $A \subset E(f)$ obtenemos

$$\sup\{\sum_{i=0}^{n-1} f(T^i x)\} = \sup\{\sum_{i=0}^{n-1} \chi_A(x) f(T^i x)\} > 0,$$

de donde $x \in E(\chi_A f)$. Luego, por el Teorema 4.0.13 resulta

$$\int_A f d\mu = \int_{E(\chi_A f)} \chi_A f d\mu \geq 0.$$

■

Ahora, con los resultados previos, estamos en condiciones de dar la demostración de 4.0.11.
Demostración. Pretendemos probar en primer lugar que $f^+(x) = f^-(x)$ c.t.p. o, equivalentemente, que $\mu(\{f^+ > f^-\}) = 0$. Intentaremos expresar $\{f^+ > f^-\}$ como unión numerable de conjuntos de medida nula. Notemos que $f^+(x) > f^-(x)$ si y sólo si existe $\beta \in \mathbb{R}$ tal que $f^+ > \beta > f^-(x)$. Entonces tenemos la igualdad

$$\{f^+ > f^-\} = \bigcup_{\beta \in \mathbb{R}} \{f^+ > \beta > f^-\}.$$

Sin embargo, tal unión no es numerable. Para corregirlo, tomamos $\{a_N\}_{N \in \mathbb{N}}$ denso en $\mathbb{R}$. Entonces podemos escribir

$$\{f^+ > f^-\} = \bigcup_{a_N > a_M} \{f^+ > a_N > a_M > f^-\}$$

y esta unión si resulta numerable. Para obtener lo que deseamos, sólo quedará ver que

$$\mu(\{f^+ > a_N\} \cap \{a_M > f^-\}) = 0.$$

Así resulta que debemos restringir nuestra atención a los conjuntos $E_\alpha^+ =: \{f^+ > \alpha\}$ y $E_\alpha^- =: \{f^- < \alpha\}$. En virtud de 4.0.12 vemos que los conjuntos $E_\alpha^+(f)$ y $E_\alpha^-(f)$ son T-invariantes. De manera sencilla puede corroborarse que $E_0^+(f - \alpha) = E_\alpha^+(f)$ y $E_\alpha^-(f) = E_{-\alpha}^+(-f)$. Nos interesa ver qué ocurre con $\mu(E_\alpha^+(f) \cap E_\beta^-(f))$ cuando $\alpha > \beta$.

Sea $A = E_\alpha^+(f) \cap E_\beta^-(f)$. Como $\int_A f d\mu = \int_A (f-\alpha)d\mu + \alpha\mu(A)$, si logramos probar que

$$\int_A (f-\alpha)d\mu \geq 0 \text{ donde } A \subset E_0^+(f-\alpha) \tag{4.1}$$

tendremos que $\int_A f d\mu \geq \alpha\mu(A)$. Análogamente, si conseguimos probar que

$$\int_A (\beta - f)d\mu \geq 0 \text{ donde } A \subset E_0^+(\beta - f) \tag{4.2}$$

tendremos que $\int_A f d\mu \leq \beta\mu(A)$. Finalmente, de las expresiones anteriores resultará lo deseado. Como $A = E_\alpha^+(f) \cap E_\alpha^-(f)$, $T^{-1}A = A$ y $A \subset E_\alpha^+(f) \subset E(f-\alpha)$. Al mismo tiempo $A \subset E_\beta^-(f) \subset E(\beta - f)$, obtenemos en virtud del corolario 4.0.14 y por 4.1 y 4.2 tendremos

$$\alpha\mu(A) \leq \int_A f d\mu \leq \beta\mu(A).$$

Como $\alpha > \beta$ resultará $\mu(E_\alpha^+(f) \cap E_\beta^-(f)) = 0$. Luego $\mu(\{f^+ > f^-\}) = 0$. De esta manera queda bien definida $\widehat{f} = \lim_N \sum_{n=0}^{N-1} f(T^i x)$ y por ser el límite puntual de una sucesión de funciones medibles, $\widehat{f}$ resulta medible. Considerando que $\int_X |f_N| d\mu = \int_X |f| d\mu$ tendremos, en virtud del Lema de Fatou,

$$\int_X |\widehat{f}| d\mu = \int_X \liminf |f_N| d\mu \leq \liminf \int_X |f_N| d\mu = \int_X |f| d\mu.$$

Así resulta que $\widehat{f} \in L^1(\mu)$ y que $\int_X |\widehat{f}| d\mu \leq \int_X |f| d\mu$. Ahora veremos la otra desigualdad: si $f \in L^\infty(\mu) \Rightarrow \|f_N\|_\infty \leq \|f\|_\infty$. Así $\|\widehat{f}\|_\infty \leq \|f\|_\infty$ y $\|f_N - \widehat{f}\|_\infty \leq 2\|f\|_\infty$. Por Convergencia Dominada tendremos:

$$\lim_{N\to\infty} \int_X |\widehat{f} - f_N| d\mu = \int_X \lim_{N\to\infty} |\widehat{f} - f_N| d\mu = 0.$$

Si $f \in L^1(\mu)$, dado $\epsilon > 0$ tomamos $g \in L^\infty(\mu)$ tal que $\int_X |f-g| d\mu \leq \epsilon$. Sea $M \in \mathbb{N}$ tal que $\int_X |g - g_N| d\mu \leq \epsilon$ si $N \geq M$. Entonces, si $N > M$ obtenemos

$$\int_X |\widehat{f} - f_N| d\mu \leq \int_X |\widehat{f} - \widehat{g}| d\mu + \int_X |\widehat{g} - g_N| d\mu + \int_X |(f-g)_N| d\mu. \tag{4.3}$$

Pero en el primer término tendremos

$$\int_X |\widehat{f} - \widehat{g}| d\mu = \int_X |\widehat{f-g}| d\mu \leq \int_X |f-g| d\mu \leq \epsilon.$$

Como $N > M$ tendremos que el segundo término también es menor que ϵ. En cuanto al tercer término

$$\int_X |(f-g)_N| d\mu \leq \int_X |f-g| d\mu \leq \epsilon.$$

Así, por 4.3 si $N > M$ es $\int_X |f_N - \widehat{f}| d\mu \leq 3\epsilon$, i.e. $f_N \to \widehat{f}$ en $L^1(\mu)$. Entonces

$$\int_X |\widehat{f}| d\mu = \lim_{N\to\infty} \int_X |f_N| d\mu = \int_X |f| d\mu.$$

Resta ver que $\int_X f d\mu = \int_X \widehat{f} d\mu$. Pero en c.t.p. es

$$\lim_N f_N^+(x) = \widehat{f}^+(x) \quad \text{y} \quad \lim_N f_N^-(x) = \widehat{f}^-(x).$$

Como $|f^+| = f^+$ y $|f^-| = f^-$ para toda f, tendremos que

$$\int_X f d\mu = \lim_N \int_X f_N d\mu = \int_X (f_N^+ - f_N^-) d\mu = \int_X \widehat{f}^+ d\mu - \int_X \widehat{f}^- d\mu = \int_X \widehat{f} d\mu.$$

$\blacksquare$

Observación 4.0.15. Notemos que por 4.0.11 $f_N \to \widehat{f}$ en $L^1(X, \mathfrak{B}, \mu)$ y puesto que

$$\widehat{f} = \lim_N f_N = \limsup_N f_N = f^+$$

resultará $\widehat{f} = \widehat{f} \circ T$. De esta manera, el teorema 4.0.11, al igual que el 3.0.9, permite una aplicación $f \to \widehat{f}$ donde $\widehat{f} \in \{f : f \circ T = f\}$. Así podemos reformular el teorema 4.0.11 de la siguiente manera

Teorema 4.0.16. *Sea $(X, \mathfrak{B}, \mu)$ espacio de probabilidad, $T : X \to X$ una función que preserva medidas. Para toda $f \in L^1(\mu)$ los promedios*

$$\frac{1}{N} \sum_{n=0}^{N-1} f \circ T^n$$

convergen puntualmente a Pf, donde P es la proyección sobre el subespacio de puntos fijos $\{f \in L^1 : \widetilde{C}_T f = f\}$.

Observación 4.0.17. Notemos que el teorema de Birkhoff sólo requiere que T preserve medidas, en cambio, el teorema de Von Neumann 3.0.9 fue demostrado sólo para operadores de composición inducidos por una aplicación inversible. También en este aspecto el teorema 4.0.11 generaliza los resultados del capítulo anterior.

Es inmediato que si $\widetilde{C}_T$ posee al conjunto de las funciones constantes como únicos puntos fijos entonces, restringiendo el operador a $L^2(\mu)$, la medida de probabilidad μ resulta ergódica tal como lo definimos en 3.0.7. Recíprocamente, tomando $f \in L^1(\mu)$ en la demostración de $(b) \Rightarrow (a)$ de 3.0.8, podemos ver que dada μ ergódica, las constantes siguen siendo los únicos puntos fijos de $\widetilde{C}_T$ cuando el operador es extendido a $L^1(\mu)$. Así hemos probado que

Proposición 4.0.18. *Una medida $\mu \in M_T$ es ergódica sii dada $f \in L^1(\mu)$ verificando que $f \circ T = f$, resulta $f = cte.$ μ-c.t.p.*

La interacción entre el teorema de Birkhoff y las medidas ergódicas queda establecida en la siguiente proposición:

Proposición 4.0.19. *Sea μ ergódica, $f \in L^1(\mu)$ entonces, para μ-c.t.p.*

$$\frac{1}{N}\sum_{n=0}^{N-1} f \circ T^n(x) \to \int_X f d\mu.$$

Demostración. Por el teorema 4.0.11 sabemos que existe $\widehat{f} \in L^1(\mu)$ tal que $f_N \to \widehat{f}$ puntualmente y $\int_X f = \int_X \widehat{f}$. Y además $\widehat{f} \circ T = \widehat{f}$, entonces $\widehat{f}$ resulta constante μ-c.t.p. Así existe $c \in \mathbb{C}$ tal que $f = c$ en μ-c.t.p. Tendremos pues $\widehat{f} = c\chi_{\{\widehat{f}=c\}}$. En consecuencia, para μ-c.t.p.

$$\frac{1}{N}\sum_{n=0}^{N-1} f \circ T^n(x) = \widehat{f}(x) = \int_X \widehat{f} d\mu = \int_X f d\mu.$$

∎

Definición 4.0.20. Dado A medible y $x \in A$ llamamos *promedio de permanencia de x en A* al límite

$$\tau_A(x) = \lim_N \frac{1}{N}\{i \in \mathbb{N} : i = 0, 1...N-1 \text{ } y \text{ } T^i(x) \in A\},$$

siempre que tal límite exista.

Observación 4.0.21. Notemos que $\tau_A(x) = \lim_N \frac{1}{N}\sum_{n=0}^{N-1} \chi_A(T^n x)$. Por otro lado, si $\mu \in M_T(X)$ y es ergódica, en virtud del teorema 4.0.11 tendremos que $\tau_A(x) = \mu(A)$ μ-c.t.p. en A, por lo que obtenemos el siguiente corolario:

Corolario 4.0.22. *Si $\mu \in M_T(X)$ es ergódica, entonces para todo medible A, el promedio de permanencia de μ-c.t.p. en A coincide con $\mu(A)$. Notemos que el valor $\tau_A(x)$ tal como lo definimos* ***sólo depende de*** *T. Por lo que, en virtud de lo anterior, concluímos que las únicas medidas de $M_T(X)$ ergódicas son aquellas que miden los promedios de permanencia de T.*

Por otro lado, observando que el conjunto M_T es cerrado y convexo, podemos enunciar el siguiente teorema:

Teorema 4.0.23. *μ es ergódica sii μ es un punto extremal de $M_T(X)$.*

Demostración.

($\Rightarrow$) Supongamos que μ puede descomponerse como $\mu = t\mu_1 + (1-t)\mu_2$ con $0 < t < 1$. Como μ es real por ser T-ergódica, cada μ_i es real. Por otro lado, es claro que para todo boreliano A, $\mu(A) = 0$ implica que $\mu_i(A) = 0$. Así $\mu_i \ll \mu$. Por el teorema de Radon-Nikodym tendremos que existe $f_i \in L^1(\mu)$ tal que para cada boreliano $\mu_i(A) = \int_A f_i d\mu$. Como μ y μ_i son T-invariantes, tendremos que $f_i \circ T = f_i$, y por la ergodicidad de μ, resulta que f_i es constante y $\mu = \mu_1 = \mu_2$.

($\Leftarrow$) Sea μ un punto extremal, y supongamos que existe A boreliano tal que $T^{-1}A = A$ y $\mu(A) \neq 0$ o 1. Para B medible hacemos

$$\mu_A(B) = \mu(A \cap B)/\mu(A) \quad y \quad \mu_{A^c}(B) = \mu(B/A)/(1 - \mu(A)).$$

Es inmediato que ambas son medidas T-invariantes, es decir μ_A $\mu_{A^c} \in M_T$. Y finalmente tendremos que $\mu = \mu(A)\mu_A + (1 - \mu(A))\mu_{A^c}$, lo cual es una contradicción.

■

Capítulo 5

Teoremas ergódicos individuales

Hemos visto que a través de la demostración de Riesz dada en 3.0.9, el teorema ergódico medio de von Neumann puede reformularse de la siguiente manera:
"Sea T un operador unitario en un espacio de Hilbert $\mathcal{H}$, entonces para cualquier $x \in \mathcal{H}$ la sucesión $T_N x$ converge fuertemente (i.e. en norma) a un punto $\overline{x} \in \mathcal{H}$".

De esta manera, el estudio de la convergencia de los promedios del operador $\widetilde{C}_T$ sobre espacios $L^p(\mu)$ dio pie a la búsqueda de teoremas ergódicos para operadores lineales definidos en espacios más generales. Así en 1938 C. Visser generaliza el teorema de von Neumann demostrando el siguiente resultado:
"Sea T un operador lineal sobre $\mathcal{H}$ tal que $\|T^n\| \leq C$ para todo $n \in \mathbb{N}$, entonces para todo $x \in \mathcal{H}$ la sucesión $T_N x$ converge débilmente a un punto $\overline{x} \in \mathcal{H}$"[17].

Al año siguiente, E. Lorch demuestra la validez del Teorema Ergódico en espacios de Banach reflexivos para transformaciones con potencias acotadas.

Simultáneamente S. Kakutani ampliaba el resultado de convergencia débil de Visser para aplicaciones lineales ahora definidas sobre un espacio de Banach general y K. Yosida conseguía recuperar la convergencia en norma de la sucesión $T_N x$ pero imponiendo ciertas condiciones sobre la aplicación T, lo cual condujo a la introducción de una nueva clase de operadores lineales.

5.1. Teorema ergódico de E. Lorch

Definición 5.1.1. Sea $\mathcal{V}$ un espacio normado. Decimos que $T : \mathcal{V} \to \mathcal{V}$ lineal es de potencias acotadas si existe $K > 0$ tal que $\|T^n\| \leq K$ con $n \in \mathbb{N}$.

Lema 5.1.2. *Sea $\mathcal{B}$ un espacio de Banach reflexivo, T un endomorfismo lineal acotado. Entonces*

a) $\ker(T^*) = \overline{\operatorname{ran}}(T)^\circ$

b) $\overline{\mathrm{ran}}(T^*) = \ker(T)^\circ$.

Demostración.

a) Supongamos que $f \in \ker(T^*)$. Dado $x \in \overline{\mathrm{ran}}(T)$ existe una sucesión $y_n \in \mathcal{B}$ tal que $Ty_n \to x$. Así $f(Ty_n) \to f(x)$. Pero $f(Ty_n) = T^*f(y_n) = 0$. Así resulta que $f(x) = 0$. Tenemos entonces que $\ker(T^*) \subset \overline{\mathrm{ran}}(T)^\circ$.
Veamos la otra inclusión: sea $f \in \overline{\mathrm{ran}}(T)^\circ$. Entonces para todo elemento $x \in \mathcal{B}$ valdrá $T^*f(x) = f(Tx) = 0$. Así obtenemos $f \in \ker(T^*)$.

b) Tomemos $f \in \mathcal{B}^*$. Entonces si $x \in \ker(T)$ tendremos $T^*f(x) = f(Tx) = 0$. Así $T^*f \in \ker(T)^\circ$. Tenemos así que $\mathrm{ran}(T^*) \subset \ker(T)^\circ$. Como $\ker(T)^\circ$ es cerrado se sigue $\overline{\mathrm{ran}}(T^*) \subset \ker(T)^\circ$.
Veamos ahora la otra inclusión: supongamos que existe $f \in \ker(T)^\circ$ y además $f \notin \overline{\mathrm{ran}}(T^*)$. Por el teorema de Hahn-Banach existe $\mathcal{F} \in \mathcal{B}^{**}$ tal que $\mathcal{F}(f) = 1$ y $\mathcal{F}$ se anula en $\overline{\mathrm{ran}}(T^*)$. Como $\mathcal{B}$ es reflexivo, existe $x \in \mathcal{B}$ tal que $\mathcal{F}$ es la evaluación en x. Entonces, dado $g \in \mathcal{B}^*$ tendremos $g(Tx) = T^*g(x) = \mathcal{F}(T^*g) = 0$. Así $Tx = 0$, entonces $x \in \ker(T)$, y eso implica $\mathcal{F}(f) = f(x) = 0$, lo cual es absurdo.

■

Teorema 5.1.3. *(cf. [12])Sea $\mathcal{B}$ un espacio de Banach reflexivo, T un endomorfismo lineal de potencias acotadas por K, y P la proyección sobre los puntos fijos de T. Entonces para toda $x \in \mathcal{B}$ el promedio $T_N x = \frac{1}{N}\sum_{n=0}^{N-1} T^n x$ converge en norma a Px.*

Demostración. Demostraremos el teorema a través de los siguientes pasos:

Paso 1) Es inmediato que si $x \in \ker(T - I)$ entonces $T_N x = x$.

Paso 2) Dado $x \in \overline{\mathrm{Ran}(T - I)}$ veamos que $T_N x \to 0$. Si $x \in \overline{\mathrm{Ran}(T - I)}$ entonces dado $\epsilon > 0$ existen $y, z \in \mathcal{B}$ tal que $\|z\| \leq \epsilon$ y $x = Ty - y + z$. Entonces, para N suficientemente grande

$$\|T_N x\| = \|T_N(Ty - y + z)\| \leq \|T_N(Ty - y)\| + \|T_N z\| \leq \frac{2\|y\|}{N} + \epsilon \leq 2\epsilon.$$

De lo anterior se sigue $T_N x \longrightarrow 0$.

Demostraremos ahora que $\mathcal{B} = \overline{\mathrm{ran}}(T - I) \oplus \ker(T - I)$ y el teorema quedará probado.

Paso 3) Veamos primero que $\overline{\mathrm{ran}}(T-I)\cap\ker(T-I)=0$. En efecto, si $x\in\ker(T-I)$ y $y\in\overline{\mathrm{ran}}(T-I)$, entonces $T_N(x+y)\to f$. Por otro lado, debido a que T es de potencias acotadas por K, tendremos $\|T_N(x+y)\|\leq K\|x+y\|$. Así resulta que $\|x\|\leq K\|x+y\|$. Si $x\in\overline{\mathrm{ran}}(T-I)\cap\ker(T-I)$ por lo anterior tendremos que $\|x\|\leq K\|x+(-x)\|=0$. De lo que sigue $x=0$.

Finalmente notemos que como T es de potencias acotadas, también lo es T^*, por lo que el razonamiento anterior también aplica sobre T^*. Así obtendremos que además se cumple $\overline{\mathrm{ran}}(T^*-I)\cap\ker(T^*-I)=0$.

Paso 4) Veamos que $\mathcal{B}=\overline{\mathrm{ran}}(T-I)\oplus\ker(T-I)$. Si existe un elemento $x\in\mathcal{B}$ tal que $x\notin\overline{\mathrm{ran}}(T-I)\oplus\ker(T-I)=\mathcal{S}$ por el teorema de Hahn-Banach, existirá $f\in\mathcal{B}^*$ tal que $f(x)=1$ y es nula sobre $\mathcal{S}$. Entonces, por el lema 5.1.2 tendremos $f\in\ker(T-I)^\circ=\overline{\mathrm{ran}}(T^*-I)$ y $f\in\mathrm{ran}(T-I)^\circ=\ker(T^*-I)$. Entonces $f\in\overline{\mathrm{ran}}(T^*-I)\cap\ker(T^*-I)=0$. Así $f=0$, lo cual es absurdo.

■

5.2. Aportes de K.Yosida y S.Kakutani

En 1938, el matemático japonés Kosaku Yosida, generaliza el teorema ergódico demostrando el siguiente resultado:

Teorema 5.2.1. *(Yosida, cf. [20] y [21]) Sea T un operador lineal sobre un espacio de Banach $\mathcal{B}$ satisfaciendo las siguientes condiciones*

(i) *T es de potencias acotadas por K;*

(ii) *para toda $x\in\mathcal{B}$ el conjunto $\{T_Nx\}$ es relativamente w-compacto.*

Entonces, para todo $x\in\mathcal{B}$, la sucesión T_Nx converge en norma a un punto fijo $\overline{x}$.

Demostración. Sea $x\in\mathcal{B}$ y por comodidad llamemos $x_N=T_Nx$ por (i) tendremos que $\|x_N\|\leq K\|x\|$ para $n=1,2,\ldots$. Entonces, por (ii) la sucesión $\{x_N\}$ es relativamente w-compacta. Así existe una subsucesión $x_{N'}$ y un elemento $\overline{x}$ tal que

$$\lim_{N'\to\infty} f(T_{N'}x)=f(\overline{x}) \tag{5.1}$$

para cualquier funcional lineal f definido en $\mathcal{B}$. Entonces tendremos

$$T\overline{x}=\overline{x} \tag{5.2}$$

pues

$$\|Tx_{N'} - x_{N'}\| = \|\frac{T^{N'}x - x}{N'}\| \leq \frac{2K\|x\|}{N'}$$

Pondremos $x = \overline{x}+(x-\overline{x})$. Por 5.2 tendremos $T^n\overline{x} = \overline{x}$ y entonces resultará $x_N = \overline{x}+z_N$ donde $z_N = T_N(x - \overline{x})$. De esta manera, vemos que resulta suficiente probar que z_N converge *fuertemente* a 0. Consideremos el subespacio lineal *cerrado* $\mathcal{B}_0$ de $\mathcal{B}$ generado por los elementos de la forma $(y - Ty)$, $y \in \mathfrak{B}$. Afirmamos que para cualquier $w \in \mathcal{B}_0$, $T_N w$ converge fuertemente a cero. En efecto: si w es de la forma $(y - Ty)$ para algún $y \in \mathcal{B}$, tendremos

$$\|T_N w\| = \|\frac{y - T^N y}{N}\| \leq \frac{2K}{N}y$$

y la expresión converge fuertemente a cero. Supongamos ahora que w no es de la forma $(y - Ty)$, entonces para cualquier $\epsilon > 0$ existe $y \in \mathcal{B}$ tal que $\|w - (y - Ty)\| \leq \epsilon$. Así

$$\|T_N w - T_N(y - Ty)\| \leq K\epsilon.$$

De esta manera, vemos que para cualquier $w \in \mathcal{B}_0$, $T_N w$ converge fuertemente a cero. Asumamos que $(x - \overline{x})$ no pertenece a $\mathcal{B}_0$. Entonces por el teorema de Hahn-Banach existe un funcional lineal f_0 tal que $f_0(x - \overline{x}) = 1$, $f_0(z) = 0$ para cualquier $z \in \mathcal{B}_0$. Como $(T^m x - T^{m+1}x) \in \mathcal{B}_0$ tendremos que $f_0(T^m x) = f_0(T^{m+1}x)$. Entonces

$$f_0(\frac{I + T + T^2 + ... + T^{N-1}}{N}x) = f_0(x)$$

para $N = 1, 2, 3....$ Así, por 5.1, $f_0(\overline{x}) = f_0(x)$ contradiciendo que $f_0(x - \overline{x}) = 1$. De esta manera, resulta que $(x - \overline{x}) \in \mathfrak{B}_0$ y entonces x_N converge fuertemente a $\overline{x}$. ∎

Notemos el teorema 5.1.3 de Lorch utiliza de manera fundamental una condición impuesta al espacio $\mathcal{B}$ (la reflexividad), en cambio el teorema 5.2.1 de Yosida sólo requiere hipótesis sobre el operador T. Por lo que resulta natural, a partir de este último resultado, enfocarse en condiciones sobre el operador más que imponerlas al espacio. Y fue precisamente el estudio de la condición *ii)* de 5.2.1, lo que llevó a K. Yosida y a S. Kakutani a introducir la siguiente

Definición 5.2.2. Sea un operador $T : \mathcal{B} \to \mathcal{B}$ con $\mathcal{B}$ un espacio de Banach y S la bola unitaria, se dice w-compacto si $T(S)$ es relativamente w-compacto.

Observación 5.2.3. Notemos que si T es w-compacto automáticamente el conjunto $\{T_N x\}$ resulta relativamente w-compacto, por lo que la condición *ii)* es automáticamente satisfecha. En consecuencia, si T también es de potencias acotadas, se verifica el resultado de 5.2.1.

Al mismo tiempo, la relación entre el teorema 5.1.3 de Lorch y el 5.2.1 de Yosida resulta ser un simple corolario del siguiente resultado de Kakutani (que posee interés propio dentro del análisis funcional):

Teorema 5.2.4. *(Kakutani cf. [5] Th. 4.2) Un espacio de Banach $\mathcal{B}$ es reflexivo si y solo si la bola unitaria S es w-compacta.*

Es inmediato el siguiente

Corolario 5.2.5. *Si $\mathcal{B}$ es un espacio de Banach reflexivo, entonces todo operador acotado T es w-compacto.*

Así obtenemos que las hipótesis del teorema 5.1.3 de Lorch implican las del teorema 5.2.1 de Yosida. En consecuencia, el primero no es más que un caso particular del segundo.

Ahora retomamos el estudio de la condición *ii)* del 5.2.1 y su relación con los operadores w-compactos.

Teorema 5.2.6. *(Yosida y Kakutani, cf.[21] y [8] Th.6.1) Sea T un operador cumpliendo*

i) *T es de potencias acotadas por K;*

ii) *existe $k \in \mathbb{N}$ y un operador w-compacto V tal que $\|T^k - V\| < 1$.*

Entonces para toda $x \in \mathcal{B}$ la sucesión $\{T_N x\}$ es w-compacta.

Demostración. La herramienta básica para probar el teorema será la identidad

$$T_N = (I - D)^{-1}(T_N - T^k T_N) + (I - D)^{-1} V T_N \tag{5.3}$$

donde $D = (T^k - V)$. Como $\|D\| < 1$ por la hipótesis *ii)*, tendremos que existe

$$(I - D)^{-1} = \sum_{n=0} D^n.$$

Así la expresión 5.3 se verifica. Por otro lado, resulta inmediata la comprobación de que los operadores w-compactos forman un ideal bilátero de $L(\mathcal{B}, \mathcal{B})$.

Ahora, dado $x \in \mathcal{B}$ buscaremos exhibir un punto de w-acumulación de $\{T_N x\}$: observemos en primer lugar que por *i)* tendremos

$$\lim_N |T_N - T^k T_N| = \lim_N |\frac{1}{N}(\sum_{n=0}^{k-1} T^n) + \frac{1}{N}\sum_{n=N}^{N+k-1} T^n| \leq \lim_N \frac{2K(k+1)}{N} = 0.$$

Debido a ello, el primer término de la expresión 5.3 tiende a 0 cualquiera sea $x \in \mathcal{B}$. Por eso, nos bastará encontrar un punto de w-acumulación $\overline{x}$ para $\{(I - D)^{-1}VT_N x\}$. Pero tal punto $\overline{x}$ existe debido a la acotación de $T_N x$ y a que el operador $(I - D)^{-1}V$ es w-compacto. ■

Observación 5.2.7. Sea T un operador verficando las condiciones de 5.2.6, y P la proyección definida por $Px = \overline{x}$. Entonces tendremos

$$PT = \lim_N T_N T = \lim_N (T_{N+1} - \frac{1}{N} I) = P.$$

Y razonando de manera análoga finalmente tendremos $PT = TP = P$.

Así, multiplicando a derecha por P en 5.3, tendremos $P = (I - D)^{-1}VP$, entonces P resulta w-compacto.

5.3. Teorema ergódico medio individual

Hasta aquí, los teoremas exhibidos estudian condiciones de convergencia de los promedios T_N asociados a un único operador T. Razón por la que tales resultados suelen ser denominados “teoremas ergódicos individuales”, para distinguirlos de aquellos que involucran la convergencia de ciertos promedios generados por familias de operadores (como estudiaremos en las próximas secciones).

Para finalizar este capítulo, exhibimos una formulación usual del teorema ergódico medio individual, que además nos facilitará compararlo con próximos resultados:

Teorema 5.3.1. *Sea $\mathcal{B}$ un espacio de Banach y $T \in L(\mathcal{B}, \mathcal{B})$ de potencias acotadas, entonces:*

1) $\ker(T - I) \oplus \overline{\operatorname{ran}}(T - I) =: \mathcal{M}$*;*

2) $x \in \mathcal{M}$ *sii existe* $\lim_N T_N x$*;*

3) Sea $T'_N = T_N|_M$. Entonces T'_N converge puntualmente a P la proyección de $\mathcal{M}$ sobre $\ker(T - I)$.

Demostración. Son los pasos 1, 2 y 3 de la demostración de 5.1.3. ■

Así podemos ver que las hipótesis adicionales de los teoremas 5.1.3 de Lorch y 5.2.1 de Yosida, **sólo son utilizadas para garantizar que $\mathcal{M}$ coincida con todo el espacio**.

Teorema 5.3.2. *(cf.5.1.3) Si $\mathcal{B}$ es reflexivo, entonces $\mathcal{M} = \mathcal{B}$.*

Teorema 5.3.3. *(cf.5.2.1 y 5.2.6) Si $\{T_N x\}$ es relativamente w-compacto, o en particular, si existe V operador w-compacto y $n \in \mathbb{N}$ tal que $\|T^n - V\| < 1$, entonces $\mathcal{M} = \mathcal{B}$.*

Capítulo 6

Teorema ergódico de W. Eberlein

Los resultados mencionados hasta ahora consistieron en teoremas que buscaban establecer convergencia de los promedios $T_N x$ a un punto fijo. O lo que es equivalente, estudiar las órbitas descriptas por la acción del semigrupo de iteraciones $\{T^n\}$ actuando sobre un espacio de Banach $\mathcal{B}$. Fueron Alaoglu y Birkhoff [1] quienes reemplazaron el semigrupo de iteraciones por un semigrupo $\mathcal{S}$ de transformaciones lineales y mostraron que la convergencia de ciertos promedios de los transformados de un punto $x \in \mathcal{B}$ resulta equivalente a la existencia y unicidad de un punto fijo x_0 en la cápsula convexa de la órbita de x vía la acción de $\mathcal{S}$.

Definición 6.0.4. Sea $\mathcal{B}$ un espacio de Banach, y $\mathcal{S}$ un semigrupo de endomorfismos lineales. Notaremos con $\widehat{\mathcal{S}} = \{\sum a_j T_j | a_j \geq 0 / \sum a_j = 1,\ T_j \in \mathcal{S}\}$ a la envoltura convexa de $\mathcal{S}$, y además escribiremos $O(x) = \{Tx | T \in \widehat{\mathcal{S}}\}$. Por conveniencia *siempre* asumiremos que la identidad I está en el semigrupo $\mathcal{S}$.

Los semigrupos $\mathcal{S}$ de interés en esta sección serán aquellos que satisfagan la propiedad que denominaremos 'ergódica'.

Definición 6.0.5. Diremos que una familia dirigida de funciones $\{T_j\} \subset L(\mathcal{B}, \mathcal{B})$ es un **sistema de promedios** para $\mathcal{S}$ si:

I) para todo $x \in \mathcal{B}$ $T_j x \in \overline{O(x)}$;

II) para todo índice $\|T_j\| \leq M$;

III) para $x \in \mathcal{B}$, $S \in \mathcal{S}$

   a) $\lim_j T_j(S - I)x = 0$ (en norma);
   b) $\lim_j (S - I)T_j x = 0$ (en norma).

Definición 6.0.6. Un semigrupo $\mathcal{S} \subset L(\mathcal{B}, \mathcal{B})$ se dirá **ergódico** si admite un sistema de promedios.

Análogamente a lo realizado en el capítulo anterior, definimos

$$M_0(\mathcal{S}) = \{x \in \mathcal{B} : Sx = x \ \forall S \in \mathcal{S}\} = \cap_{S \in \mathcal{S}} \ker(S - I)$$

y

$$M_1(\mathcal{S}) = \overline{\langle x - Sx : x \in \mathcal{B}, S \in \mathcal{S} \rangle} = \overline{\langle \cup_{S \in \mathcal{S}} \operatorname{ran}(S - I) \rangle}$$

y tendremos $M(\mathcal{S}) = M_0(\mathcal{S}) + M_1(\mathcal{S})$.

Observación 6.0.7. Si $\{T_j\}$ es un sistema de promedios, entonces, para toda $S \in \widehat{\mathcal{S}}$ tendremos que $\lim_j T_j(S - I) = \lim_j (S - I)T_j = 0$.

También observemos que si $x \in M_0(\mathcal{S})$, por definición tendremos que $O(x) = \{x\}$ y necesariamente $T_j x = x$.

Ahora enunciamos y demostraremos el resultado más importante de este capítulo:

Teorema 6.0.8. *(Eberlein 1949, cf.[8]) Si $\mathcal{S}$ es ergódico y T_j un sistema de promedios, entonces:*

Parte 1 Son equivalentes:

- *i) Existe* $\lim_j T_j x =: px$ *(en norma);*
- *ii) Existe* $w - \lim_j T_j x =: px$*;*
- *iii) Hay un punto de w-acumulación de $\{T_j x\}$;*
- *iv) Existe un punto fijo en $\overline{O(x)}$;*
- *v)* $x \in M(\mathcal{S})$.

Parte 2 Tendremos que:

- *i)* $px = x$ *sii* $x \in M_0(\mathcal{S})$*;*
- *ii)* $px = 0$ *sii* $x \in M_1(\mathcal{S})$*;*
- *iii)* $M(\mathcal{S}) = M_0(\mathcal{S}) \oplus M_1(\mathcal{S})$*;*
- *iv) p es la proyección de $M(\mathcal{S})$ sobre $M_0(\mathcal{S})$.*

Parte 3

- *i) Para todo $S \in \widehat{\mathcal{S}}$ tendremos $Sp = pS = p$;*
- *ii) Para todo $x \in M(\mathcal{S})$ resulta $\{px\} = \overline{O(x)} \cap M_0(\mathcal{S})$.*

Demostración.
Parte 1: $i) \Rightarrow ii)$ y $ii) \Rightarrow iii)$ son triviales.

$iii) \Rightarrow iv)$ Sea $x_0 = w - \text{lím}_{j'} T_{j'}x$. Como $T_{j'}x \in \overline{O(x)}$, y $\overline{O(x)}$ es w-cerrado [1], así obtenemos $x_0 \in \overline{O(x)}$.
Probemos que $x_0 \in M_0(\mathcal{S})$ tomando $S \in \mathcal{S}$ y viendo que $Sx_0 - x_0 = 0$. Dado $f \in \mathcal{B}^*$ probaremos que $f(Sx_0 - x_0) = 0$. Como la convergencia en la topología fuerte implica la convergencia en la débil, tendremos que por la condición IIIb) de 6.0.5, existe j_0 tal que si $j' \geq j_0$ vale

$$|f(x_0 - Sx_0)| \leq |f(x_0 - T_{j'}x_0)| + |f((I - S)T_{j'}x_0)| + |f(ST_{j'}x_0 - Sx_0)| \leq \epsilon.$$

$iv) \Rightarrow v)$ Sea x_0 el punto fijo. Podemos escribir $x = x_0 + (x - x_0)$. Veamos ahora que $x - x_0 \in M_1(\mathcal{S})$: como $x_0 \in \overline{O(x)}$ dado $\epsilon > 0$ existen $\{\alpha_1 ... \alpha_k\}$ reales positivos tales que $\sum \alpha_i = 1$ y $\{S_1 ... S_k\} \subset \mathcal{S}$ verificando $\|x_0 - \sum \alpha_i S_i x\| < \epsilon$. Luego:

$$\|(x - x_0) - \underbrace{(x - \sum \alpha_i S_i x)}_{\in M_1(\mathcal{S})}\| < \epsilon.$$

$v) \Rightarrow i)$ Sea $x \in M(\mathcal{S})$ y escribamos $x = x_0 + x_1$ con $x_i \in M_i(\mathcal{S})$.
Por definición de $M_1(\mathcal{S})$ dado $\epsilon > 0$ tendremos que existe una combinación lineal $\widetilde{S} = \sum \alpha_i S_i$ tal que $\|x_1 - (x - \widetilde{S}x)\| < \epsilon$. Así resultará que

$$\|x_0 - \widetilde{S}x\| = \| - x_1 + (x_1 + x_0) - \widetilde{S}x\| < \epsilon.$$

También por IIIa) existirá j_0 tal que si $j \geq j_0$ tendremos que $\|T_j(\widetilde{S} - I)x\| \leq \epsilon$. Así podremos escribir:

$$\|x_0 - T_j x\| = \|T_j(x_0 + \widetilde{S}x) + T_j(\widetilde{S} - I)x\| \leq M\epsilon + \|T_j(\widetilde{S} - I)x\| \leq (M + 1)\epsilon$$

Resulta pues $x_0 = \text{lím}_j T_j x$.

Parte 2:

$i)$ $\Leftarrow)$ es inmediato. Veamos $\Rightarrow)$: si $px = x$ tendremos que $x \in M(\mathcal{S})$ por la Parte 1 demostrada anteriormente. Así podemos volver a escribir $x = x_0 + x_1$. Por la demostración $v) \Rightarrow i)$ del paso anterior, tendremos que $x_0 = \text{lím}_j T_j x = px = x$. Así $x \in M_0(\mathcal{S})$.

[1] Recordemos el ***Teorema de Mazur:*** *en un espacio vectorial localmente convexo si un conjunto es cerrado y convexo, entonces es w-cerrado* (cf.[7])

$ii)$ $\Rightarrow$) Supongamos $x \neq 0$. Si $x = x_0 + x_1$ entonces, por el paso anterior, $x_0 = px = 0$ y $x = x_1 \in M_1(\mathcal{S})$.
$\Leftarrow$) Si $x \in M_1\mathcal{S}$, dado $\epsilon > 0$ existe $\widetilde{S} = \sum_i^k \alpha_i S_i$ tal que $\|x - (S - I)x\| \leq \epsilon$.
Así

$$\lim_j T_j x = \lim_j T_j x - \lim_j T_j(S - I)x = \lim_j T_j(x - (S - I)x).$$

Entonces $\lim_j \|T_j x\| \leq M\epsilon$. Resulta pues $\lim_j T_j x = 0$.

$iii)$ Es consecuencia inmediata de $i)$ y $ii)$.

$iv)$ Como $px \in M_0(\mathcal{S})$, tendremos $p(px) = px$. Así $p^2 = p$.

Parte 3

$i)$ Por IIIa) tendremos

$$pSx = \lim_j T_j Sx = \lim_j T_j(Sx - x + x) = \lim_j T_j(S - I)x + \lim_j T_j x = px.$$

Por IIIb) tendremos

$$Spx = \lim_j ST_j x = \lim_j (S - I)T_j x + \lim_j T_j x = px.$$

$ii)$ Puesto que $px \in M_0(\mathcal{S})$ por $2i)$ y $px = lim_j T_j x$, teniendo además que $T_j x \in \overline{O(x)}$, resultará $px \in \overline{O(x)}$.
Supongamos que $y_0 \in \overline{O(x)} \cap M_0(\mathcal{S})$, así dado $\epsilon > 0$ existirá un operador $S \in \widehat{\mathcal{S}}$ tal que $\|y_0 - Sx\| \leq 0$. Luego, por el inciso anterior

$$y_0 - px = \lim_j T_j y_0 - T_j x = \lim_j T_j y_0 - T_j Sx = \lim_j T_j(y_0 - Sx).$$

Entonces $\|y_0 - px\| \leq M\epsilon$. Así resulta $y_0 = px$.

■

Observación 6.0.9. La condición IIIa) se utilizó al demostrar $v) \Rightarrow i)$ de la Parte 1, y que $pS = p$ para todo $S \in \mathcal{S}$ en la Parte 3)i). Por otro lado, la condición IIIb) se utilizó al probar $iii \Rightarrow iv)$ de la Parte 1 (pero sólo se requirió la convergencia débil) y al ver que $Sp = p$ en 3)i).

6.1. El subespacio ergódico

Definición 6.1.1. Si $\mathcal{S}$ es ergódico, diremos que un punto x es ergódico si $x \in M(\mathcal{S})$.

Observación 6.1.2. Notemos que la definición de $M(\mathcal{S})$ (y por ende de la proyección p) es independiente de cualquier sistema de promedios que logremos definir sobre $\mathcal{S}$. Así la propiedad de ser ergódico es instrínseca del semigrupo $\mathcal{S}$, y no depende de la familia $\{T_j\}$ que usemos para probarlo.

Observación 6.1.3. Uno de los desafíos que el Teorema 6.0.8 inmediatamente plantea es el de garantizar condiciones suficientes para asegurar que un punto $x \in \mathcal{B}$ sea ergódico. Para esto, el criterio $1iii)$ es el más utilizado, y la w-compaccidad la condición más útil. En efecto, desde que $\{T_\alpha x\} \subset \overline{O(x)}$ si $O(x)$ es relativamente w-compacto, inmediatamente existe el punto de acumulación requerido.

Otra clase de puntos de interés inmediato son:

Definición 6.1.4. Un elemento $x \in \mathcal{B}$ se dice 'casi periódico' (resp. casi w-periódico) por $\mathcal{S}$ si la órbita $\{Sx : S \in \mathcal{S}\}$ es relativamente compacta (resp. w-compacta). Notaremos $AP(\mathcal{S})$ al conjunto de los puntos casi periódicos bajo la acción de $\mathcal{S}$, y $WAP(\mathcal{S})$ al conjunto de los puntos casi w-periódicos.

Proposición 6.1.5. *Sea $\mathcal{S}$ un grupo ergódico, entonces $AP(\mathcal{S}) \subset WAP(\mathcal{S}) \subset M(\mathcal{S})$.*

La demostración nos demandará recurrir a un profunda caracterización de la débil compaccidad de conjuntos en espacios vectoriales.

Teorema 6.1.6. *(Eberlein-Smulian) Dado $\mathcal{B}$ espacio de Banach y $A \subset \mathcal{B}$ no vacío, son equivalentes:*

i) *A es relativamente w-compacto.*

ii) *Toda sucesión de puntos de A tiene un punto de w-acumulación (i.e. A es secuencialmente w-compacto).*

Ahora demostramos 6.1.5:
Demostración. Tomemos $x \in AP(\mathcal{S})$, y sea $\{S_n x\}$ una sucesión de $\{Sx | S \in \mathcal{S}\}$. Por la relativa compaccidad de tal conjunto, existe un punto de acumulación x_0 en la topología normada, que a su vez, será punto débil de acumulación. Entonces en virtud del teorema de Eberlein-Smulian, $\{Sx : S \in \mathcal{S}\}$ es relativamente w-compacto. Así $x \in WAP(\mathcal{S})$. Veamos ahora que $WAP(\mathcal{S}) \subset M(\mathcal{S})$. Dado $x \in WAP(\mathcal{S})$ probemos que existe x_0

satisfaciendo alguna de las condiciones de la Parte 1 de 6.0.8. Sea x_0 un punto de w-acumulación de $\{Rx : R \in \mathcal{S}\}$, entonces existe una red $R_\beta x$ tal que $w - \text{lím}_\beta R_\beta x = x_0$. Sea T_j un sistema de promedios sobre $\mathcal{S}$, probaremos que $w - \text{lím}_j T_j x = x_0$

$$\begin{aligned}
\text{lím}_j T_j x - x_0 &= \text{lím}_j \text{lím}_\beta (T_j x - R_\beta T_j x + R_\beta T_j x - R_\beta x) \\
&= \text{lím}_\beta \text{lím}_j ((T_j x - R_\beta T_j x) + (R_\beta T_j x - R_\beta x)) \\
&= \text{lím}_\beta (\text{lím}_j T_j x - R_\beta T_j x) + \text{lím}_j (R_\beta T_j x - R_\beta x) \\
&= \text{lím}_\beta 0 = 0.
\end{aligned}$$

■

6.2. Ejemplos y aplicaciones

En esta sección justificaremos por qué el teorema 6.0.8 recibe el nombre de 'ergódico' obteniendo los teoremas ergódicos individuales ya mencionados en 5. El procedimiento es sencillo y consiste (1) en verificar la ergodicidad del semigrupo $\mathcal{S}$ que estemos considerando en particular y (2) verificar la ergodicidad de un elemento x mostrando la existencia del elemento x_0 requerido (en general, lo haremos recurriendo a un argumento que involucre la w-compacidad de algún conjunto).

6.2.1. El caso usual

Ejemplo 6.2.1. Sea $\mathcal{S}$ consistiendo en las iteraciones $\{T^n\}$ de una transformación lineal acotada T, y como sistema de promedios tomemos

$$T_N = \frac{1}{N} \sum_{n=0}^{N-1} T^n.$$

Es inmediato que $T_N \in L(\mathcal{B}, \mathcal{B})$ y la condición I de la definición 6.0.5 es satisfecha inmediatamente. Para intentar verificar III, asumamos que $\|T^n\| \leq M$ asi obtenemos que si $N > m \geq 1$

$$\begin{aligned}
\|T^m T_N - T_N\| &= \|\frac{1}{N}(\sum_{n=m}^{N+m-1} T^n - \sum_{0}^{N-1} T^n)\| \\
&= \|\frac{1}{N}(\sum_{n=m}^{N+m-1} T^n - \sum_{n=0}^{m-1} - \sum_{n=m}^{N-1} T^n)\| \\
&= \|\frac{1}{N}(\sum_{n=N}^{N+m-1} T^n - \sum_{n=0}^{m-1} T^n)\| \leq \frac{2m}{N}.
\end{aligned}$$

Y sigue la convergencia uniforme de T_N a 0. Vimos así que la hipótesis usual de solicitar $\|T^n\| \leq M$ para todo n asegura la condición II y de III de manera uniforme. Sin embargo, esto es muy restrictivo: por ejemplo, en el caso que $\mathcal{B} = \mathcal{H}$ un espacio de Hilbert, pedir $\|T^n\|$ no es más general que pedir que T sea unitario (cf. [16]). Por ello, solicitaremos condiciones más generales, y el teorema 6.0.8 se especializa de la siguiente manera:

Teorema 6.2.1. *Sea $T \in L(\mathcal{B}, \mathcal{B})$ un operador tal que $\forall n$ $\|T_N\| \leq M$ con $M > 0$ y $\lim_n T^n x/n = 0$ para todo $x \in \mathcal{B}$. Si el conjunto $\{T_N x\}$ es relativamente w-compacto, entonces existe $y = \lim_N T_N x$ y $Ty = y$.*

Demostración. Por hipótesis la familia $\{T_N\}$ es acotada y se cumple II. La condición I es inmediata. Para ver que se satisface la condición III retomamos la expresión ya antes desarrollada y tendremos

$$\lim_N T^m T_N x - T_N x = \lim_N \frac{1}{N}(\sum_{n=N}^{N+m-1} T^n x - \sum_{n=0}^{m-1} T^n x) = 0.$$

Así obtenemos un sistema de promedios. A su vez, la relativa w-compacidad de $\{T_N x\}$ garantiza la existencia de un punto x_0 de w-acumulación. Entonces por la Parte $1iii)$ del teorema 6.0.8 tendremos la convergencia al punto fijo x_0. ■

6.2.2. Semigrupos abelianos acotados

Ejemplo 6.2.2. Consideremos un semigrupo arbitrario $\mathcal{S}$ abeliano y acotado de transformaciones sobre un espacio $\mathcal{B}$, esto es, $\|T\| \leq M$ para todo $T \in \mathcal{S}$. Ordenamos el semigrupo $\widehat{\mathcal{S}}$ de la siguiente manera: definimos $V \ll U$ si existe $W \in \widehat{\mathcal{S}}$ tal que $U = WV$. Por la conmutatividad del semigrupo se sigue que $UV = VU$ es un sucesor común de U y V, así $(\widehat{\mathcal{S}}, \ll)$ resulta dirigido considerando a cada elemento como su propio índice. Veremos que $\widehat{\mathcal{S}}$ así ordenado resulta un sistema de promedios para $\mathcal{S}$. Las propiedades I, II de 6.0.5 son satisfechas de manera inmediata.

Para demostrar III observemos primero que $T_N T - T_N = TT_N - T_N = \frac{1}{N}(T^N - I)$, así, dado $x \in \mathcal{B}$ y $\epsilon > 0$ existe $N_0 \in \mathbb{N}$ tal que, para $N \geq N_0$ vale

$$\|T_N T - T_N\| = \frac{1}{N}\|T^N x - x\| \leq \frac{M+1}{N}\|x\| < \frac{\epsilon}{M}.$$

Si $T^* = VT_{N_0}$ es cualquier sucesor de T_{N_0} tendremos que

$$\|TT^*x - T^*x\| = \|T^*Tx - T^*x\| = \|V(TT_N - T_N)x\| < \epsilon.$$

De esta manera hemos probado

Proposición 6.2.2. *Sea $\mathcal{S}$ un semigrupo abeliano y acotado de transformaciones, entonces es ergódico.*

6.2.3. Sumabilidad de series de Fourier

Continuando con el ejemplo anterior, ahora exhibiremos el caso de un semigrupo abeliano, no acotado pero que aún así resulta ergódico.

Ejemplo 6.2.3. Sea $C = \{x : \mathbb{R} \to \mathbb{C} : x$ es continua, 2π-periódica y acotada$\}$. Podemos considerar

$$C \hookrightarrow L^2(\frac{\mathbb{R}}{[0, 2\pi)}) \equiv L^2(S^1).$$

Así, dada $f \in C$ podemos escribir

$$f(x) = (\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)dt),1 + \sum_{n=1}^{\infty}[\langle f, e^{int}\rangle e^{inx} + \langle f, e^{-int}\rangle e^{-int}].$$

Sea

$$S_N f = \sum_{n=-(N-1)}^{N-1} \langle f; e^{int}\rangle e^{int}.$$

Utilizando el desarrollo de Dirichlet (cf. sección 2.5 de [22]), tendremos:

$$S_N f(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t) \underbrace{\frac{\sin[(x-t)(N+\frac{1}{2})]}{\sin\frac{(x-t)}{2}}}_{D_N(x-t)} dt = f * D_N(x).$$

Recordemos también la identidad (véase A)

$$f * D_m * D_{m+p} = f * D_{m+p} = f * D_{m+p} * D_m. \tag{6.1}$$

Definamos ahora $U_n = I - S_n$ para todo $n \in \mathbb{N}$. Y sea $\mathcal{S}$ el semigrupo consistente en I y todas las transformaciones U_n. Afirmamos que $\mathcal{S}$ es abeliano, es decir, que se cumple $U_{m+p}U_m = U_mU_{m+p} = U_{m+p}$. En efecto, utilizando 6.1 tenemos

$$\begin{aligned} U_{m+p}U_mf &= U_mf - U_mf * D_{m+p} \\ &= f - f * D_m - (f - f * D_m) * D_{m+p} \\ &= f - f * D_{m+p} = U_{m+p}f. \end{aligned}$$

Y análogamente se ve que $U_mU_{m+p} = U_{m+p}$.

Veamos ahora que $\mathcal{S} = \{U_n\}$ es un semigrupo no acotado:

$$\begin{aligned} S_Nf(x) &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)D_N(x-t)dt \\ &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)D_N(t-x)dt \text{ (pues } D_N \text{ es par)} \\ &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x+t)D_N(t)dt \text{ (por la invariancia de la integral por traslaciones).} \end{aligned}$$

Luego

$$\begin{aligned} \|S_Nf\|_1 &= \frac{1}{2\pi}\int_{-\pi}^{\pi} |\frac{1}{\pi}\int_{-\pi}^{\pi} f(x+t)D_N(t)dt|dx \\ &\leq \frac{1}{\pi}\int_{-\pi}^{\pi} \frac{1}{2\pi}(\int_{-\pi}^{\pi} |f(x+t)|dx)|D_N(t)|dt \\ &= \|f\|_1\frac{1}{\pi}\int_{-\pi}^{\pi} |D_N(t)|dt. \end{aligned}$$

Así resulta $\|S_N\| \leq \frac{1}{\pi}\int_{-\pi}^{\pi} |D_N(t)|dt = 2\|D_N\|_1$.

Hay además una aproximación acotada de la unidad en $L^1(T)$, ie., hay una red $\{f_\nu\}_{\nu\in\Lambda}$ con $\|f_\nu\| = 1$ para todo $\nu \in \Lambda$ tal que dada $g \in L^1(T)$ vale

$$f_\nu * g \to g$$

en consecuencia $S_N(f_\nu) \to D_N$ y

$$\|S_N\| = 2\|D_N\|_1 = L_N = \frac{4}{\pi^2}\log(N) + O(1)$$

donde L_N son las constantes de Lebesgue, cuyos valores divergen cuando N crece (cf. sección 2.12 de [22]). Luego, de la desigualdad triangular inversa se deduce la no acotación de la familia $\{U_N\}$:

$$|L_N - 1| = |\|S_N\| - \|I\|| \leq \|U_N\|.$$

Pese a no ser acotado, $\mathcal{S}$ es un semigrupo ergódico. En efecto, los promedios

$$T_N = \frac{1}{N}\sum_{0}^{N-1} U_n = I - \frac{1}{N}\sum_{n=0}^{N-1} S_n$$

forman un sistema de promedios: la linealidad de U_N y la propiedad I de la definición 6.0.5 son inmediatas. Y la validez de la condición III surge de la siguiente identidad para $N > M$

$$U_M T_N - T_N = \frac{1}{N}\sum_{0}^{M-1}(U_M - U_n).$$

Para establecer la condición II hacemos

$$T_N f = f(t) - \frac{1}{N}\sum_{n=0}^{N-1} S_n = f(t) - \frac{1}{\pi}\int_{-\pi}^{\pi} f(t+u)K_N(u)dt,$$

donde

$$K_N(u) = \frac{1}{2(N+1)}[\frac{\sin(N+1)u/2}{\sin(u/2)}]^2 \geq 0.$$

Resulta $\frac{1}{\pi}\int_{-\pi}^{\pi} K_N(u)du = 1$. Entonces se sigue de la positividad de K_N (denominado *núcleo de Fejér*) y de la desigualdad triangular que $\|T_N\| \leq 2$ para todo N. Así $\mathcal{S} = \{U_N\}$ es ergódico pues admite un Sistema de Promedios.
El Teorema de Féjer afirma (cf. cap.4,Th. 3.4 de [22]):

Teorema 6.2.3. *Para toda $f \in C$, los promedios de las sumas parciales S_N convergen uniformemente a f. En consecuencia*

$$f - T_N f = \frac{1}{N}\sum_{n=0}^{N-1} S_n f \to f.$$

Así, el teorema anterior toma la forma:

Teorema 6.2.4. *Todo elemento $f \in C$ es ergódico bajo la acción de G, y $pf = 0$.*

6.2.4. Funciones casi periodicas

Ejemplo 6.2.4. Consideremos el espacio de Banach $\mathcal{B}$ de las funciones complejas acotadas y uniformemente continuas sobre $\mathbb{R}$ dotado de la norma suprema. Sea G el grupo de traslaciones $U_a x(t) = x(t+a)$, cuyos puntos fijos son únicamente las funciones constantes. Tomaremos como el sistema $\{T_\alpha\}$ a los promedios

$$T_\alpha x(t) = \frac{1}{\alpha}\int_0^\alpha x(t+s)ds \text{ con } \alpha > 0.$$

Es inmediato que T_α es lineal y que $\|T_\alpha\| \leq 1$. Así obtenemos la condicion II de la definición 6.0.5. Que se cumple III es consecuencia del hecho de que U_a y T_α conmutan y de la siguiente desigualdad que vale para $\alpha > a$:

$$\begin{aligned}
\|U_a T_\alpha x - T_\alpha x\| =& \frac{1}{\alpha}\|\int_a^{a+\alpha} x(t+s)ds - \int_0^\alpha x(t+s)ds\| \\
=& \frac{1}{\alpha}\|\int_a^{\alpha} x(t+s)ds + \int_\alpha^{\alpha+a} x(t+s)ds - \int_0^\alpha x(t+s)ds\| \\
=& \frac{1}{\alpha}\|\int_a^{\alpha} x(t+s)ds + \int_\alpha^{\alpha+a} x(t+s)ds - \int_0^a x(t+s)ds + \int_a^{\alpha+a} x(t+s)ds\| \\
=& \frac{1}{\alpha}\|\int_\alpha^{\alpha+a} x(t+s)ds - \int_0^a x(t+s)ds\| \\
\leq & 2|a|.\|x\|/\alpha.
\end{aligned}$$

En consecuencia, para todo $a \in \mathbb{R}_{\geq 0}$ $\lim_{\alpha\to\infty} U_a T_\alpha x - T_\alpha x = 0$.

Para ver I utilizamos la hipótesis de la continuidad uniforme para aproximar $T_\alpha x(t)$ por sumas de Riemann. Si $x : (-\infty;\infty) \to \mathbb{C}$ sabemos que existe N tal que

$$|\int_0^\alpha x(s)ds - \sum_0^{N-1} \frac{\alpha}{N} x(\alpha n/N)| \leq \epsilon.$$

De esta manera, existe N cumpliendo que, dado $t \in \mathbb{R}$, vale

$$\begin{aligned}
|\int_0^\alpha x(t+s)ds - \sum_0^{N-1}\frac{\alpha}{N}x(t+\frac{n\alpha}{N})| =& |\int_0^\alpha x(t+s)ds - \sum_0^{N-1}\frac{\alpha}{N}U_{\frac{n\alpha}{N}}x(t)| \\
=& \alpha|T_\alpha x(t) - \sum_0^{N-1}\frac{1}{N}U_{\frac{n\alpha}{N}}x(t)| \leq \epsilon.
\end{aligned}$$

Entonces

$$\|T_\alpha x(t) - \sum_0^{N-1} U_{\frac{n\alpha}{N}} x(t)\| \leq \epsilon.$$

Y claramente $\sum_0^{N-1} \frac{\alpha}{N} U_{\alpha n/N} x(t) \in O(x)$. Así $T_\alpha x(t) \in \overline{O(x)}$.

Recordando que los únicos puntos fijos de la aplicación son las constantes y que la topología del espacio C es la de convergencia uniforme, el teorema 6.0.8 se especializa:

Teorema 6.2.5. *Si $x(t)$ $(-\infty < t < \infty)$ es casi periódica (en el sentido de 6.1.4), entonces $\lim_{\alpha\to\infty} \frac{1}{\alpha}\int_0^\alpha x(t+s)ds$ existe uniformemente y es constante.*

6.3. El caso continuo

Ejemplo 6.3.1. Sea X un espacio de Haussdorf localmente compacto y T_t un flujo continuo sobre X. Para todo $t \in \mathbb{R}$ definimos $C_t : C_0(X) \to C_0(X)$ vía $C_t f = f \circ T_t$. Y de esta manera obtenemos un nuevo flujo sobre $C_0(X)$.

Observación 6.3.1. Notemos que $\{C_t\}$ es un semigrupo conmutativo y para todo t vale $\|C_t\| \leq 1$. Tendremos así que $\{C_t\}$ es un semigrupo ergódico por 6.4.4, exhibamos explícitamente otro sistema de promedios.

Dado $T \in \mathbb{R}$ y $f \in C_0(X)$ el 'promedio temporal' del elemento $x \in X$ se define de la siguiente manera:

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T f(T_t x)dt = \lim_{T\to\infty} \frac{1}{T}\int_0^T C_t f(x)dt. \tag{6.2}$$

Lo que deseamos saber es cuándo y a dónde converge tal promedio. A su vez, las expresiones 6.2 nos permiten definir

$$\Gamma_T f(x) =: \frac{1}{T}\int_0^T f(T_t x)dt.$$

Es claro que para todo T, la función $\Gamma_T f \in C_0(X)$.

Proposición 6.3.2. *$\{\Gamma_T\}$ es un sistema de promedios.*

Demostración. Es inmediata la linealidad de cada Γ_T

Para ver la condición I de 6.0.5 debemos probar $\Gamma_T f \in \overline{O(f)}$. Para ello, notemos que si fijamos el punto $x \in X$, obtenemos una aplicación $\Gamma_T f(x) : [0,T] \to \mathbb{R}$ vía $\Gamma_T f(x)(s) = \frac{1}{T}\int_0^s f(T_t x)dx$. Cómo la función $\Gamma_T f(x)$ es continua y está definida sobre un compacto, resulta uniformemente continua. Así, por construcción de la integral de Riemman, dado $\epsilon > 0$ existe $N \in \mathbb{N}$ tal que para cada $x \in X$ vale:

$$|\int_0^T f(T_s x)ds - \sum_{n=0}^{N-1} \frac{T}{N} f(T_{\frac{n}{N}T} x)|.$$

Cabe recalcar que $N = N(f, \epsilon)$. Esto es gracias a la ***equicontinuidad*** *de la familia* $\{\Gamma_T f(x)/x \in X\}$. En efecto:

dados $T \in \mathbb{R}$, $f \in C_0(X)$ tendremos lo siguiente:

$$\begin{aligned}|\Gamma_T f(x)(r) - \Gamma_T f(x)(r)| &= |\frac{1}{T}(\int_0^s f(T_t x)dt - \int_0^r f(T_t x)dt)| \\ &= \frac{1}{T}|\int_s^r f(T_t x)dt| \leq \frac{|r-s|\|f\|}{T}.\end{aligned}$$

Donde $s, r \in [0, T]$, y la desigualdad es independiente del punto $x \in X$. Luego, dado ϵ si $|s - r| < \frac{T\epsilon}{\|f\|} = \delta$ tendremos que $|\Gamma_T f(x)(s) - \Gamma_T f(x)(r)| \leq \epsilon$.

II) Γ_T pertenece al conjunto de automorfismos lineales acotados de $C_0(X)$, conjunto que a su vez es un espacio de Banach. Por ello basta ver la equicontinuidad de la familia de operadores $\{\Gamma_T : T \in \mathbb{R}\}$. Así:

$$\|\Gamma_T(f)\| = \sup_{x\in X} |\frac{1}{T}\int_0^T C_s f(x)ds \leq \sup_{x\in X} \frac{1}{T}\int_0^T |C_s f(x)|ds \leq \sup_{x\in X} \|f\| = \|f\|$$

Así resulta que $\|\Gamma_T\| \leq 1$.

Veamos III.b)

$$\begin{aligned}\lim_{T\to\infty} C_t \circ \Gamma_T f(x) - \Gamma_T f(x) &= \lim_{T\to\infty} \frac{1}{T}\int_t^{T+t} f(T_s x)ds - \frac{1}{T}\int_0^T f(T_s x)ds \\ &= \lim_{T\to\infty} \frac{1}{T}(\int_T^{T+t} f(T_s x)ds - \int_0^t f(T_s x)ds) \\ &= \lim_{T\to\infty} \frac{2t\|f\|}{T} \\ &= 0.\end{aligned}$$

La condición IV.a) se demuestra de manera análoga. ■

Ahora veamos condiciones para que una $f \in C(X)$ sea casi-periódica. Supongamos además que X es compacto. Será fundamental el

Teorema 6.3.3. *(Arzelá-Ascoli) Sea X compacto y $\mathcal{F} \subset C(X)$. Entonces $\mathcal{F}$ es relativamente compacto sii*

i. *$\mathcal{F}$ es equicontinuo,*

ii. *para cada $x \in X$ $\mathcal{F}_x = \{f(x) : f \in \mathcal{F}\}$ es relativamente compacto.*

Si f es tal que la familia $C_t f$ resulta equicontinua, en virtud del teorema de Arzelá-Ascoli, tendremos que el conjunto $\{C_t f\}$ es relativamente compacto en $(C(X), \|.\|)$. Y por ello podemos hallar un punto de acumulación que cumplirá la condición $1iii)$ de 6.0.8 De esta manera, podemos enunciar una versión especializada de 6.0.8:

Teorema 6.3.4. *Sea $f \in C(X)$ tal que la familia $\{C_t f\}$ es equicontinua, entonces f es casi-periódica (y así resulta ergódica).*

6.4. Semigrupos acotados

En esta sección generalizamos el ejemplo dado en 6.2.2 al caso no abeliano y la definición de punto ergódico. Luego damos un resultado que será fundamental en el teorema 7.3.2 del siguiente capítulo:

Ejemplo 6.4.1. Dado un semigrupo $\mathcal{S} \subset L(\mathcal{B}, \mathcal{B})$ acotado, ordenamos $\widehat{\mathcal{S}}$ vía $U \leq V$ sii existe $W \in \widehat{\mathcal{S}}$ tal que $V = WU$. Al mismo tiempo, dado un elemento $x \in \mathcal{B}$ diremos que $\lim_{S \in \widehat{\mathcal{S}}} Sx = y$ si dados $\epsilon > 0$ y $U \in \widehat{\mathcal{S}}$ existe V_0 tal que $U \leq V_0$ y $\|Wx - y\| < \epsilon$ siempre que $V_0 \leq W$.
Esta forma de ordenar a $\widehat{\mathcal{S}}$ generaliza la introducida en 6.2.2 para el caso abeliano. Por otro lado, introducir la nueva forma de límite mencionada, tiene sentido a la luz de la siguiente proposición:

Proposición 6.4.1. *Sea $\mathcal{S}$ acotado y ergódico. Entonces $x \in \mathcal{B}$ es ergódico sii $\lim_{S \in \widehat{\mathcal{S}}} Sx$ converge a un punto fijo x_0.*

Demostración. $\Rightarrow$) Dado $\epsilon > 0$ y $U \in \widehat{\mathcal{S}}$, por la invariancia de $M(\mathcal{S})$ por $\widehat{\mathcal{S}}$ tenemos $x_0 = px = pUx = \lim_j T_j Ux$, donde $\{T_j\}$ es un sistema de promedios.
Así $x_0 \in \overline{O(x)}$. Luego existe $V_0 \in \widehat{\mathcal{S}}$ tal que $\|x_0 - V_0 Ux\| < \frac{\epsilon}{M}$. Sea $V_1 = V_0 U$, entonces $U \leq V_1$ y si $W \in \widehat{\mathcal{S}}$ tendremos que

$$\|x_0 - WV_1 x\| = \|Wx_0 - WV_1 x\| \leq M\|x_0 - V_1 x\| \leq \epsilon.$$

$\Leftarrow$) Si Sx converge a un punto fijo x_0, entonces $x_0 \in \overline{O(x)}$ y resulta satisfecha la condición $1 - iv)$ del teorema 6.0.8. ∎

Por otro lado, **independientemente** de que el semigrupo $\mathcal{S}$ sea o no ergódico (cf. definición 6.0.6) tiene sentido enunciar la siguiente:

Definición 6.4.2. Un elemento $x \in \mathcal{B}$ se dirá ergódico bajo la acción de $\mathcal{S}$ si existe $x_0 = \lim_{S \in \widehat{\mathcal{S}}} Sx$.

Observación 6.4.3. Sea $x \in \mathcal{B}$ ergódico, entonces $x_0 \in \overline{O(x)}$. Veamos que x_0 es punto fijo: dado $\epsilon > 0$ existe $S_0 \in \widehat{\mathcal{S}}$ tal que $\|x_0 - SS_0 x\| < \epsilon/2M$ para todo $S \in \mathcal{S}$. Así tendremos que:

$$\|Sx_0 - x_0\| = \|x_0 - SS_0 x + SS_0 x - x_0\| \leq \frac{\epsilon}{2M} + \frac{\epsilon}{2} < \epsilon.$$

Al mismo tiempo, x_0 es único: si y_0 fuese otro punto fijo en $\overline{O(x)}$ tendremos que dado ϵ existe U_0 tal que $\|x_0 - U_0 x\| < \epsilon/2$ y $\|y_0 - U_0 x\| < \epsilon/2$. En consecuencia

$$\|x_0 - y_0\| = \|x_0 - U_0 x + U_0 x - y_0\| < \epsilon.$$

Así resulta que un punto x es ergódico si y solo si hay un único punto fijo en $\overline{O(x)}$

Ahora enunciamos el siguiente:

Teorema 6.4.4. *Sean $\mathcal{S}_1$ y $\mathcal{S}_2$ semigrupos acotados tales que sus elementos conmutan de a pares (es decir, si $u \in \mathcal{S}_1$ y $v \in \mathcal{S}_2$, entonces $uv = vu$) y $M_0(\mathcal{S}_1) = M_0(\mathcal{S}_2)$ (es decir, comparten los puntos fijos). Entonces*

(i) *Los puntos fijos de $\overline{O_1(x)}$ y $\overline{O_2(x)}$ se reducen a un único y común punto fijo.*

(ii) *Sea $\mathcal{W} \subset \mathcal{B}$ subespacio invariante por $\mathcal{S}_1$ y $\mathcal{S}_2$ y supongamos que para cada $x \in \mathcal{W}$ existe un punto fijo $x_0 \in \overline{O_1(x)} \cap \overline{O_2(x)}$, entonces cada $x \in \mathcal{W}$ es ergódico bajo $\mathcal{S}_1$, $\mathcal{S}_2$ y $\mathcal{S} = \mathcal{S}_1\mathcal{S}_2 = \mathcal{S}_2\mathcal{S}_1$.*

Demostración.

(i) Sean $y_i \in \overline{O_i(x)}$ puntos fijos. Veremos que $y_1 = y_2$. Sea M una cota común a ambos semigrupos. Dado $\epsilon > 0$, tomemos $V_i \in \widehat{\mathcal{S}_i}$ tales que $\|y_i - V_i x\| \leq \epsilon/M$. Como $V_1V_2 = V_2V_1$ e y_1, y_2 son puntos fijos bajo ambas acciones, tendremos $y_1 = y_2$ pues:

$$\begin{aligned}\|y_1 - y_2\| &= \|V_2 y_1 - V_2V_1 x + V_1V_2 x - V_1 y_2\| \\ &= \|V_2(y_1 - V_1 x) + V_1(V_2 x - y_2)\| \leq \epsilon.\end{aligned}$$

(ii) Si $x \in \mathcal{W}$, entonces por hipótesis y la parte anterior del teorema, $\overline{O_i(x)}$ contiene un único punto fijo. Así queda garantizada la ergodicidad del elemento x bajo las acciones de $\mathcal{S}_i$.

Queda ver la ergodicidad de $x \in \mathcal{W}$ bajo la acción de $\mathcal{S} = \mathcal{S}_1\mathcal{S}_2$. Dado $\epsilon > 0$ y $W = \sum_j a_j U_j V_j \in \widehat{\mathcal{S}}$ (con $\sum a_j = 1$, $a_j \geq 0$, $U_j \in \mathcal{S}_1$ y $V_j \in \mathcal{S}_2$) exhibiremos un W_0 tal que $W \leq W_0$ y $\|x_0 - W_0 x\| < \epsilon$. Como $x \in M(\mathcal{S}_1)$ podemos elegir $U_1'...U_n' \in \widehat{\mathcal{S}}_1$ tal que

$$\|U_j'...U_1' U_j V_j x - x_0\| < \frac{\epsilon}{M^2}.$$

Entonces resulta

$$\|U_j'...U_1' U_j V_j x - x_0\| = \|V_j U_j'...U_1' U_j x - x_0\| < \frac{\epsilon}{M}$$

Tomando $W_0 = U_n'...U_1'$ resulta

$$\|W_0 x - x_0\| = \| \sum a_j U_n'...U_{j+1}'(U_j'...U_1' U_j V_j x - x_0\| < \sum a_j M \frac{\epsilon}{M} = \epsilon.$$

■

Capítulo 7

Aportes de M. Day

En el capítulo anterior hemos trabajado con semigrupos que podían representarse como operadores acotados sobre un espacio de Banach. En este capítulo nos restringiremos a aquellas representaciones que son acotadas, y ampliaremos el concepto de 'ergodicidad' para un semigrupo dado.

Sea $\mathcal{B}$ un espacio de Banach, y $\mathcal{S}$ un semigrupo acotado contenido en el algebra de Banach $L(\mathcal{B}, \mathcal{B})$. Entonces $\sup\{\|S\| : S \in \mathcal{S}\} \leq M < \infty$. Para cada $x \in \mathcal{B}$ notaremos con $O(x) = \mathrm{co}\{Sx : S \in \mathcal{S}\}$. Definiendo $K = \{T \in L(\mathcal{B}, \mathcal{B}) : Tx \in \overline{O(x)} \; \forall x \in \mathcal{B}\}$, tendremos que

$$K = L(\mathcal{B}, \mathcal{B}) \cap \prod_{x \in \mathcal{B}} \overline{O(x)}$$

Con esta notación y en este contexto de semigrupos acotados la definició 6.0.6 se especializa del siguiente modo:

Definición 7.0.5. $\mathcal{S}$ será *ergódico* si existe una familia dirigida $\mathcal{A}$ de operadores cumpliendo:

a) $\mathcal{A} \subset K$;

b) para todo $S \in \mathcal{S}$ vale $\lim_j T_j(S - I) = 0$;

c) para todo $S \in \mathcal{S}$ vale $\lim_j (S - I)T_j = 0$.

Como en el capítulo anterior $\mathcal{A}$ será denominado Sistema de Promedios.

En el capítulo anterior, al definir ergodicidad, se recurrió a la convergencia fuerte en $b)$ y $c)$. Ahora ampliaremos el concepto de ergodicidad dado en 6.0.5.

Definición 7.0.6. El semigrupo $\mathcal{S}$ será uniformemente ergódico, fuertemente ergódico (escribiremos s^*-ergódico) o débilmente ergódico (escribiremos w^*-ergódico) si existe un conjunto de promedios $\mathcal{A}$ tal que la convergencia en $b)$ y $c)$ es uniforme, fuerte o débil respectivamente.

En el capítulo anterior nos centramos en los semigrupos s^*-ergódicos, sin embargo, como veremos, para muchos resultados es suficiente usar el w^*-límite en $c)$.

Proposición 7.0.7. *Si $\mathcal{S}$ es w^*-ergódico, las equivalencias ii) a v) de la Parte 1 del teorema 6.0.8 siguen valiendo, así como las Partes 2 y 3.*

Demostración. Para cerrar el circuito de equivalencias (omitiendo $1-i)$) sólo queda ver la implicación $v) \Rightarrow ii)$: dado $x \in M(\mathcal{S})$ escribamos $x = x_0 + x_1$. Así, tomando $f \in \mathcal{B}^*$ tendremos $f(T_j x) = f(x_0) + f(T_j x_1)$. Veamos que $\lim_j f(T_j x_1) = 0$ y obtendremos lo deseado.
Dado $\epsilon > 0$ tomemos $S \in \widehat{\mathcal{S}}$ tal que $\|Sx_1\| = \|x_1 - (x - Sx)\| < \epsilon$. Así resultará:

$$\limsup_j |f(t_j x_1)| \leq \lim_j |f(T_j(I-S)x_1| + \limsup_j |f(T_j x_1)| \leq \epsilon.$$

Por otro lado, es inmediato que en las demostraciones de $2-ii) \Leftarrow)$, $3-i)$ y $3-ii)$ puede fácilmente reemplazarse el límite en norma por el w-límite. ■

De esta manera, vemos que el teorema ergódico medio para semigrupos sigue valiendo en su totalidad al considerar solamente la ergodicidad débil.

7.1. Ergodicidad débil

En esta sección daremos una útil caracterización de la w^*-ergodicidad. Asumamos que $\mathcal{B}$ está embebido en $\mathcal{B}^{**}$ vía la inmersión natural $Qx(f) = f(x)$ para todo $x \in \mathcal{B}$ y $f \in \mathcal{B}^*$. Así podemos escribir $L(\mathcal{B},\mathcal{B}) \subset L(\mathcal{B},\mathcal{B}^{**})$. Consideremos a $\mathcal{B}^{**}$ con la w^*-topología que posee por ser un espacio dual, entonces

$$L(\mathcal{B},\mathcal{B}^{**}) \subset \prod_{x\in\mathcal{B}} \mathcal{B}^{**}.$$

Definiremos la w_*-topología de $L(\mathcal{B},\mathcal{B}^{**})$ como aquella que hereda al ser subespacio del espacio producto mencionado. Así $w_* - \lim_j T_j = T$ sii para todo $x \in \mathcal{B}$ $f \in \mathcal{B}^*$ $lim_j T_j x(f) = Tx(f)$. En esta w_*-topología cada esfera en $L(\mathcal{B},\mathcal{B}^{**})$ es compacta.
Llamemos $K_* = \overline{K}^{w_*}$.

Teorema 7.1.1. *$\mathcal{S}$ es w^*-ergódico sii existe $T_* \in L(\mathcal{B},\mathcal{B}^{**})$ cumpliendo:*

a*) *$T_* \in K_*$;*

b*) *para todo* $S \in \mathcal{S}$ $T_*(S - I) = 0$*;*

c*) *para todo* $S \in \mathcal{S}$ $(S - I)^{**}T_* = 0$.

Demostración. $\Rightarrow$) K es acotado, así tendremos que K_* es w_*-compacto. Si $\mathcal{S}$ es w^*-ergódico, sea $\mathcal{A}$ un sistema de promedios para el semigrupo. Puesto que cada $T_j \in K$, tendremos que existe un elemento $T_* \in K_*$ que es punto de w_*-acumulación de $\{T_j\}$. Sea $T_* = w_* - \lim_{j'} T_{j'}$, y veamos que se cumplen $b*)$ y $c*)$:
si $x \in \mathcal{B}$ y $f \in \mathcal{B}^*$ tendremos por la w-convergencia de $b)$ de 7.0.5

$$\begin{aligned} T_*(S - I)x(f) &= \lim_{j'} T_{j'}(S - I)x(f) \\ &= \lim_{j'} f(T_{j'}(S - I)x) = 0. \end{aligned}$$

Análogamente, por la w-convergencia de $c)$ tendremos

$$\begin{aligned} (S - I)^{**}T_*x(f) &= \lim_{j'}(S - I)^{**}T_{j'}x(f) \\ &= \lim_{j'} T_{j'}x((S - I)^*f) \\ &= \lim_{j'}(S - I)^{**}f(T_{j'}x) = \lim_{j'} f((S - I)T_{j'}x) = 0. \end{aligned}$$

$\Leftarrow$) Dado T_* cumpliendo $a*), b*), c*)$ podemos obtener un sistema de promedios de la siguiente manera: ordenamos los w_*-entornos de T_* vía $V \gg U$ sii $V \subset U$. Para cada U tomamos $T_U \in U \cap K$. Claramente $\mathcal{A} = \{T_U\} \subset K$. Veamos que se satisfacen los w-límites de $b)$ y $c)$:

$$\begin{aligned} \lim_{U} f(T_U(S - I)x) &= \lim_{U} T_U(S - I)x(f) \\ &= T_*(S - I)x(f) = 0. \end{aligned}$$

Así $w^* - \lim_U T_U(S - I) = 0$. Análogamente se prueba $c)$. ■

Observación 7.1.2. Notemos que si resultara $K_* = K$ y además tuviésemos un elemento $T_* \in K$ cumpliendo $b*)$ y $c*)$ el conjunto puntual $\{T_*\}$ será un sistema de promedios y $\mathcal{S}$ será uniformemente ergódico bajo tal sistema.

7.2. Representaciones de semigrupos y medias

7.2.1. Representaciones

A partir de esta sección $\mathcal{S}$ será un semigrupo abstracto con elementos s, u, v.

Definición 7.2.1. Una *representación derecha (izquierda) de* $\mathcal{S}$ sobre $\mathcal{B}$ será una función $F : \mathcal{S} \to L(\mathcal{B}, \mathcal{B})$ tal que $F_{ss'} = F_s F_{s'}$ ($F_{ss'} = F_{s'} F_s$).

Como queda claro en la definición anterior, ahora hay que distinguir el semigrupo $\mathcal{S}$ de una representación particular del mismo. Por ello enunciamos la siguiente:

Definición 7.2.2. Diremos que una representación $F : \mathcal{S} \to L(\mathcal{B}, \mathcal{B})$ es ergódica (uniforme, fuerte o débilmente) si el semigrupo $F(\mathcal{S}) \subset L(\mathcal{B}, \mathcal{B})$ es ergódico.
Consecuentemente, dada una representación escribiremos $M_0(F) = M_0(F(\mathcal{S}))$ y también $M_1(F) = M_1(F(\mathcal{S}))$.

Observación 7.2.3. Toda representación derecha de $\mathcal{S}$ sobre un espacio $\mathcal{B}$ genera una ***acción izquierda*** de $\mathcal{S}$ sobre $\mathcal{B}$ vía $sx =: F_s x$. Análogamente toda representación izquierda (también denominadas *antirepresentaciones*) genera una ***acción derecha*** de $\mathcal{S}$ sobre $\mathcal{B}$ vía $xs =: F_s x$.

Nos interesarán especialmente dos representaciones canónicas:

Definición 7.2.4. Definimos la *representación regular derecha* de $\mathcal{S}$ como

$$\begin{cases} r : \mathcal{S} \to L(l^\infty(\mathcal{S}), l^\infty(\mathcal{S})) \\ \\ r_s(\phi)(u) =: \phi(us) = s\phi(u) \end{cases}$$

y llamaremos $\mathcal{R} = r(\mathcal{S})$. Análogamente se define la *representación regular izquierda* vía $l_s\phi(u) =: \phi(su) = \phi s(u)$ y escribiremos $\mathcal{L} = l(\mathcal{S})$. Ambas representaciones dan origen a la *acción regular izquierda y derecha* respectivamente.

Es inmediato que los elementos de $\mathcal{R}$ y $\mathcal{L}$ conmutan, y que sus elementos tienen normas acotadas por uno. Así $\mathcal{R}$ y $\mathcal{L}$ son representaciones acotadas de $\mathcal{S}$.

7.2.2. Medias

Daremos dos definiciones y mencionaremos resultados adaptadas al caso de los semigrupos discretos. Su formulación general y las demostraciones se encuentran en el apéndice B.

Definición 7.2.5. Sea $1_\mathcal{S}$ la función unitaria (ie. $1_\mathcal{S}(s) = 1$ para todo $s \in \mathcal{S}$). Un elemento $\mu \in l^\infty(\mathcal{S})^*$ será una *media* sobre $\mathcal{S}$ si $\|\mu\| = \mu(1_\mathcal{S}) = 1$. Notaremos con $\mathfrak{m}(\mathcal{S})$ al conjunto de las medias sobre $\mathcal{S}$.

Definición 7.2.6. Un elemento $\gamma \in l^1(\mathcal{S})$ es una media finita si $\gamma \geq 0$, existen $s_1 ... s_k$ tal que $\sum_i^k \gamma(s_i) = 1$ y $\gamma(s) = 0$ si $s \neq s_i$. Notaremos con Γ al conjunto de medias finitas. Observemos que al considerar la inmersión de $l^1(\mathcal{S})$ en su doble dual, cada elemento de Γ^{**} resulta ser una combinación convexa de evaluaciones $\widehat{s_i}$.

Sea $\mathcal{P}(\mathcal{S}) =: \{f \in l^1(\mathcal{S}) : f \geq 0 \text{y} \|f\|_1 = 1\}$.

Proposición 7.2.7. *Sea $\mathcal{S}$ semigrupo discreto, entonces:*

i) $\mathfrak{m}(\mathcal{S}) = \overline{\mathcal{P}(\mathcal{S})^\wedge}^{w^*}$*;*

ii) $\mu \in \mathfrak{m}(\mathcal{S})$ *sii* $\mu(\psi) \geq 0$ *si* $\psi \geq 0$*;*

iii) Γ *es denso en* $\mathcal{P}(\mathcal{S})$*;*

iv) *si* $\mu \in \mathfrak{m}(\mathcal{S})$ *entonces existe* $\{\gamma_j\} \subset \Gamma$ *tal que* $\mu = w^* - lim_j \widehat{\gamma}_j$.

Demostración. $i)$ y $ii)$ se encuentran en el apéndice B. Para ver $iii)$ tomemos $f \in \mathcal{P}(\mathcal{S})$ que no sea media finita. Así, dado $\epsilon > 0$ tendremos $\{s_i\}$ con $i = 1...n$ cumpliendo que $1 - \epsilon < \sum_i^n |f(s_i)| < 1 = \|f\|$. Necesariamente, si $s \neq s_i$ deberá ser $|f(s)| < \epsilon$. Sea γ tal que $\gamma(s_i) = f(s_i)$ y $\gamma(s) = 0$ si $s \neq s_i$. Afirmamos que $\|f - \gamma\|_1 < \epsilon$. Si ocurriese lo contrario, existirá una colección finita $s_j \neq s_i$ tal que $\sum_{j=1}^m |f(s_j)| \geq \epsilon$. Así resultaría que

$$1 = 1 - \epsilon + \epsilon < \sum_i |f(s_i)| + \sum_j |f(s_j)| = \|f\|_1$$

lo cual es contradictorio. Luego debe ser $\|f - \gamma\| < \epsilon$. Sin embargo, γ no es una media finita. Para compensar esto, tomamos $s_0 \neq s_i$ y definimos $\gamma_0(s_i) = f(s_i)$ si $i = 1...n$, $\gamma_0(s_0) = 1 - \sum_i^n |f(s_i)|$ y $\gamma_0(s) = 0$ si $s \neq s_i, s_0$. Y tendremos que γ_0 si es una media finita. Entonces resulta:

$$\|f - \gamma_0\|_1 = \| - \gamma - |f(s_{n+1})| e_{s_0}\| \leq \|f - \gamma\| + |f(s_{n+1})| < 2\epsilon.$$

El último inciso es consecuencia de lo anterior. ■

7.3. El teorema ergódico de M. Day

Definición 7.3.1. Dado un semigrupo discreto $\mathcal{S}$, diremos que es *amenable* si existe una media μ invariante bajo la acción de $\mathcal{R}^* = \{r^* : r \in \mathcal{R}\}$ y $\mathcal{L}^* = \{l^* : l \in \mathcal{L}\}$.

Teorema 7.3.2. *(cf. [6]) Dado un semigrupo $\mathcal{S}$ son equivalentes:*

a) Toda representación acotada derecha o izquierda, es w^*-ergódica;

b) Las representaciones regulares $\mathcal{R}$ y $\mathcal{L}$ son w^*-ergódicas;

c) $1_\mathcal{S}$ está a distancia 1 de $\mathcal{M}_1(\mathcal{R}) + \mathcal{M}_1(\mathcal{L})$;

d) $\mathcal{S}$ es amenable;

e) Para cada representación acotada F de $\mathcal{S}$ existe T_* cumpliendo $a*), b*), c*)$. De hecho $T_* \in \overline{\mathrm{co}\{F(\mathcal{S})\}}^{w_*}$.

Antes de demostrar el teorema, enunciamos el siguiente resultado auxiliar:

Proposición 7.3.3. *Si las acciones regulares $\mathcal{R}$ y $\mathcal{L}$ son w^*-ergódicas, entonces*

$$M_0(\mathcal{R}) = M_0(\mathcal{L}) = \{c1_{\mathcal{S}} : c \in \mathbb{C}\}.$$

Demostración. Claramente $\{c1_{\mathcal{S}} : c \in \mathbb{C}\} \subset M_0(\mathcal{L})$. Si $x \in M_0(\mathcal{L})$ entonces para todo $s \in \mathcal{S}$ $l_s x = x$. Así, para cualquier $s, u \in \mathcal{S}$ tendremos $x(su) = l_s x(u) = x(u)$. Así $x(su) = r_u x(s) = x(u)1_{\mathcal{S}}$ cualquiera sean s, u. Así, por ergodicidad de $\mathcal{R}$

$$x = \underbrace{(x - r_u)}_{\in M_1(\mathcal{R}} + \underbrace{r_u x}_{\in M_0(\mathcal{R})} .$$

Como $r_u x$ es un punto fijo bajo la acción regular $\mathcal{R}$, si p_R es su correspondiente proyección, tendremos $p_R x = r_u x$. Por la parte 3-i) del teorema 7.3.2, tendremos que, dado $v \in \mathcal{S}$ vale $r_v p_R x = p_R r_v x$. Así

$$x(u)1_{\mathcal{S}} = r_v(x(u)1_{\mathcal{S}}) = r_v p_R x = p_R r_v x = p_R x(v)1_{\mathcal{S}} = x(v)1_{\mathcal{S}}.$$

Así $c = x(u) = x(v)$ y $x = c1_{\mathcal{S}}$. Análogamente se demuestra el caso $M_0(\mathcal{R})$. ■

Demostramos ahora 7.3.2:

Demostración. Son inmediatas las implicaciones $e) \Rightarrow a) \Rightarrow b)$.

$b) \Rightarrow c)$ Sabemos que los elementos de $\mathcal{R}$ y $\mathcal{L}$ conmutan entre sí, y por la proposición anterior tenemos que comparten los puntos fijos. Entonces se satisfacen las hipótesis del teorema 6.4.4. Sea $\mathcal{P} = \{RL : R \in \mathcal{R}, L \in \mathcal{L}\}$, así tenemos definida la proyección $p : M(\mathcal{P}) \to M_0(\mathcal{P})$, naturalmente $M(\mathcal{R}), M(\mathcal{L}) \subset M(\mathcal{P})$ y p extiende las proyecciones regulares p_R y p_L.
Si $y \in M_1(\mathcal{R}) + M_1(\mathcal{L})$, entonces tendremos $y = y_R + y_L$. Así resultará que $p(y) = p(y_R) + p(y_L) = p_R(y_R) + p_L(y_L) = 0$. Al mismo tiempo $px = \lim_{S \in \widehat{\mathcal{P}}} Sx$, y dado de que $\|S\| \leq 1$ para toda $S \in \mathcal{P}$, tendremos $\|px\| \leq \|x\|$.
Finalmente

$$1 = \|1_{\mathcal{S}}\| = \|\lim_{S \in \widehat{\mathcal{P}}} S(y + 1_{\mathcal{S}})\| \leq \|y + 1_{\mathcal{S}}\|$$

y así obtenemos c).

$c) \Rightarrow d)$ Si $1_{\mathcal{S}}$ está a distancia 1 de $M_1(\mathcal{R}) + M_1(\mathcal{L}) =: M$, por el teorema de Hahn-Banach queda garantizada la existencia de $\mu \in \mathfrak{m}(\mathcal{S})$ que se anula sobre M. Así, para todo $s \in \mathcal{S}$, $x \in l^\infty(\mathcal{S})$ tendremos

$$\mu(l_s x - x) = 0 \Rightarrow \mu(l_s x) = \mu(x) \Rightarrow l_s^* \mu = \mu$$

y análogamente obtenemos la invariancia de μ por la acción de $\mathcal{R}^*$.

$d) \Rightarrow e)$ Sea μ la media invariante por las acciones regulares. Supongamos que F es una representación izquierda de $\mathcal{S}$ sobre $\mathcal{B}$, así, como mencionamos en 7.2.3, queda definida una acción derecha sobre $\mathcal{B}$ vía $xs = F_s x$.
Para todo $x \in \mathcal{B}$ y $f \in \mathcal{B}^*$ definimos $f_x : \mathcal{S} \to \mathbb{C}$ vía $f_x(s) =: f(xs)$. Claramente $f_x \in l^\infty(\mathcal{S})$. Luego contruímos T_* de la siguente manera

$$\begin{cases} T_* : \mathcal{B} \to \mathcal{B}^{**} \\ T_* x(f) = \mu(f_x). \end{cases}$$

Para ver que T_* satisface $a*)$, tomamos una red de medias finitas γ_j cumpliendo que $\mu = w^* - \lim_j \widehat{\gamma}_j$. Así, para cada $x \in \mathcal{B}$ y $f \in \mathcal{B}^*$ tendremos

$$\begin{aligned} T_* x(f) = \mu(f_x) = \lim_j \widehat{\gamma}_j(f_x) = \lim_j \sum_{i=1}^{k_j} \gamma_j(s_i)\widehat{s}_i(f_x) &= \lim_j \sum_{i=1}^{k_j} \gamma_j(s_i) f(xs) \\ &= \lim_j \sum_{i=1}^{k_j} \gamma_j(s_i) f(F_s x). \end{aligned}$$

Finalmente obtenemos

$$T_* = w_* - \lim_j \sum_{i=1}^{k_j} \gamma_j(s_i) F_{s_i}.$$

Observemos que dada una representación izquierda F, podemos hacer el desarrollo $f_{xs}(u) = f(xsu) = f_x(su) = (sf)_x(u)$. Así resulta

$$f_{xs} = (sf)_x.$$

De esta manera, para ver $b*)$ hacemos

$$T_*(F_s - I)x(f) = T_* xs(f) - T_* x(f) = \mu(f_{xs}) - \mu(f_x) = s\mu(f_x) - \mu(f_x) = 0.$$

Análogamente se demuestra $c*)$.
Si F fuese una representación derecha, definimos $f_x(u) = f(sx)$ a través de la acción izquierda en $\mathcal{B}$ inducida por la representación. Entonces valdrá $f_{sx} = (fs)_x$, y procedemos de forma análoga al desarrollo anterior.

■

Retomando la demostracón de $c \Rightarrow d)$, obtenemos lo siguiente:

Lema 7.3.4. *Si $\mathcal{S}$ es amenable, entonces existe una familia de medias finitas $\{\gamma_j\}$ tal que, dada una representación acotada F de $\mathcal{S}$, la familia*

$$\phi_j = \sum_{i=1}^{k_j} \gamma(s_i) F_{s_i}.$$

resultará un sistema de promedios para $F(\mathcal{S})$.

El resultado anterior dista de ser trivial, significa que si el semigrupo es amenable, siempre podremos acercarnos al 'promedio final' T_* por elementos de $\mathrm{co}(F(\mathcal{S}))$.
Por otro lado, es importante rescatar que la demostración NO es constructiva. Sólo tenemos garantizada la existencia de tal familia, pero no tenemos aún medios para dar una expresión explícita de cálculo para los promedios.

7.3.1. Discusión de la amenabilidad

¿Qué significa que un semigrupo discreto sea amenable? Para verlo, definamos primero

$$\begin{cases} T_s : \mathcal{S} \to \mathcal{S} \\ \\ T_s u = us. \end{cases}$$

Análogamente definimos ${}_sTu =: su$. Entonces tendremos que para toda $\phi \in l^\infty(\mathcal{S})$ y $s \in \mathcal{S}$ valdrá que $s\phi = \phi \circ T_s$ y $\phi s = \phi \circ {}_sT$.

Si $\mu \in \mathfrak{m}(\mathcal{S})$ tenemos definida una medida finitamente aditiva sobre $\mathcal{S}$ considerada con el álgebra de partes, vía $\mu(A) = \mu(\chi_A)$. Al mismo tiempo, al ser μ invariante por las acciones de $\mathcal{R}^*$ y $\mathcal{L}^*$ tendremos que

$$\mu(A) = r_s^* \mu(A) = \mu(r_s \chi_A) = \mu(\chi_A \circ T_s) = \mu(\chi_{T_s^{-1}A}) = \mu(T_s^{-1}A)$$

y análogamente $\mu(A) = \mu({}_sT^{-1}A)$.
Para ver con más claridad la consecuencia de esto, supongamos además que el semigrupo es cancelativo, entonces para toda $s \in \mathcal{S}$ tendremos que T_s y ${}_sT$ son inyectivas. Entonces $A = T_s^{-1}(T_s A) = {}_sT^{-1}({}_sTA)$ y en consecuencia

$$\mu(A) = \mu(As) = \mu(sA).$$

En otras palabras, que un semigrupo sea amenable significa que hay una medida finita invariante por translaciones. Y esto, por 7.3.2 es equivalente a la w^*-ergodicidad de cualquier representación acotada del semigrupo.

7.4. Grupos amenables discretos

Restrinjámonos ahora al caso en que $\mathcal{G}$ es un grupo discreto. Entonces podemos definir

$$\begin{cases} T : l^{\infty}(\mathcal{G}) \to l^{\infty}(\mathcal{G}) \\ \\ T\phi(u) = \phi(u^{-1}). \end{cases}$$

Son inmediatas las siguientes identidades: $r_v T = T l_v^{-1}$, $l_v T = T r_v^{-1}$, $T^2 = I$, $T 1_{\mathcal{S}} = 1_{\mathcal{S}}$. Y además T es un operador isométrico.

Teorema 7.4.1. *Si existe una media $\lambda \in \mathfrak{m}(\mathcal{G})$ invariante por $\mathcal{R}^*$, entonces $\mathcal{G}$ es amenable.*

Demostración. A partir de λ podemos construir una media invariante por $\mathcal{L}^*$ vía $\rho = T^*\lambda$. En efecto:

$$\rho(l_u x) = T^*\lambda(l_u x) = \lambda(T l_u x) = \lambda(r_{u^{-1}} T x) = \lambda(Tx) = \rho(x).$$

Ahora, dada medias λ y ρ invariantes por $\mathcal{R}^*$ y $\mathcal{L}^*$ respectivamente, construiremos una media μ invariante por ambas acciones:
Dado $x \in l^{\infty}(\mathcal{G})$ definimos $\overline{x}(g) = \lambda(l_g x)$. Y finalmente definimos $\mu(x) = \rho(\overline{x})$. Observemos que

$$\overline{r_v x}(g) = \lambda(l_g r_v x) = \lambda(r_v l_g x) = \lambda(l_g x) = \overline{x}(g).$$

Así para cualquier $\overline{r_v x} = \overline{x}$. Luego resulta:

$$r_v^* \mu(x) = \mu(r_v x) = \rho(\overline{r_v x}) = \rho(\overline{x}) = \mu(x)$$

y μ es invariante por $\mathcal{R}^*$.
Análogamente obtenemos $\overline{l_v x} = l_v \overline{x}$ y resulta:

$$l_v^* \mu(x) = \rho(\overline{l_v x}) = \rho(l_v x) = \rho(\overline{x}) = \mu(x)$$

y obtenemos la invariancia de μ por $\mathcal{L}^*$. ∎

De manera similar puede procederse a partir de una acción ρ invariante por $\mathcal{L}^*$.

Capítulo 8

Grupos amenables y ergódicos

En la sección anterior caracterizamos la ergodicidad de todas las representaciones acotadas de un semigrupo discreto $\mathcal{S}$. Y en tal contexto se introdujo la noción de amenabilidad del semigrupo $\mathcal{S}$.
Sin embargo, tal condición es meramente algebraica: como la topología con la que se considera al semigrupo es la discreta, se pierde cualquier información adicional sobre estructuras topológicas compatibles con las operaciones del semigrupo.
En esta sección expandiremos la noción de 'amenabilidad' a grupos que poseen una cierta topología, y el objetivo será enunciar el teorema 8.3.2, que establece la relación de este nuevo concepto con la ergodicidad de ciertas representaciones particulares.

8.1. Grupos amenables

De aquí en adelante sea $\mathcal{G}$ un grupo topológico localmente compacto, y λ su medida de Haar invariante a izquierda (ie. $\lambda(A) = \lambda(gA)$ cuando $A \subset \mathcal{G}$ y $g \in \mathcal{G}$).
En este contexto podemos expandir las definiciones dadas en el capítulo anterior:

$$\mathcal{P}(\mathcal{G}) = \{\nu \in L^1(\mathcal{G}) : \nu \geq 0 \text{ y } \|\nu\| = 1\}$$

$$\mathfrak{m}(\mathcal{G}) = \{\mu \in L^\infty(\mathcal{G})^* : \|\mu\| = \mu(e) = 1\}.$$

También tenemos que (cf. B):

$$\mathfrak{m}(\mathcal{G}) = \overline{\mathcal{P}(\mathcal{G})}^{w^*}.$$

Observación 8.1.1. Recordemos que dado $\mathcal{G}$, automáticamente queda definida la acción regular derecha sobre $L^\infty(\mathcal{G})$ vía $\phi s(u) = \phi(su) = \phi \circ {}_sT$ con $\phi \in L^\infty(\mathcal{G})$. Esto genera de manera natural la acción regular izquierda sobre el dual $L^\infty(\mathcal{G})^*$ vía $s\mu(\phi) = \mu(\phi s)$ si $\mu \in L^\infty(\mathcal{G})^*$.
Análogamente podemos definir acciones regulares a izquierda y derecha sobre $L^\infty(\mathcal{G})$ y $L^\infty(\mathcal{G})^*$ respectivamente.

Definición 8.1.2. $\mathcal{G}$ se dice amenable a izquierda si existe $\mu \in \mathfrak{m}(\mathcal{G}, \lambda)$ invariante por la acción izquierda de $\mathcal{G}$ sobre $L^\infty(\mathcal{G})^*$. Notaremos con $\mathcal{L}(\mathcal{G})$ al conjunto de tales medias.
Análogamente se define la amenabilidad a derecha de un grupo $\mathcal{G}$, y el conjunto de tales medias lo notaremos con $\mathcal{R}(\mathcal{G})$.
Finalmente, $\mathcal{G}$ será amenable si lo es a izquierda y derecha y $\mathcal{I}(\mathcal{G})$ será el conjunto correspondiente de medias.

En 7.4 vimos que al tratar con grupos discretos, la amenabilidad a izquierda o derecha resultaban equivalentes. Sin embargo, en este nuevo contexto de grupos con cierta topología específica, tal equivalencia se pierde en general.
Por ello, frente a esta situación desdoblamos también la definición de ergodicidad:

Definición 8.1.3. Un semigrupo $\mathcal{S} \subset L(\mathcal{B}, \mathcal{B})$ se dice ***ergódico a izquierda*** (resp. a derecha) si hay un sistema (denominados promedios a izquierda) de $\mathcal{S}$ cumpliendo I, II y IIIa) (resp. IIIb).
Consecuentemente, una representación $F : \mathcal{S} \to L(\mathcal{B}, \mathcal{B})$ se dirá ergódica a izquierda si su imagen $F(\mathcal{S})$ lo es.

También podemos definir naturalmente una acción izquierda del grupo $\mathcal{G}$ sobre $L^1(\mathcal{G})$ vía $s * f(u) =: f(s^{-1}u)$.
Notemos que dada $\phi \in L^\infty(\mathcal{G})$ tendremos

$$\langle s * f; \phi\rangle = \int_G f(s^{-1}u)\phi(u)d\lambda(u) = \int_G f(u)\phi(su)d\lambda(u) = \langle f, \phi s\rangle.$$

Es sencillo ver que, dados $v \in \mathcal{G}$,$\nu \in L^1(\mathcal{G})$ y $\mu \in L^\infty(\mathcal{G})^*$ valen

$$s(\nu\mu) = (s * \nu)\mu \quad \text{y} \quad \widehat{s * \nu} = s\widehat{\nu}.$$

Análogamente se define la acción derecha de $\mathcal{G}$ sobre $L^1(\mathcal{G})$ vía $f * s(u) = f(us^{-1})$ y se obtienen identidades similares a las arriba expuestas.

8.2. Caracterización de la amenabilidad

Ahora exponemos caracterizaciones de la amenabilidad de un grupo localmente compacto $\mathcal{G}$.

Teorema 8.2.1. *(Condición de Reiter) Un grupo $\mathcal{G}$ es amenable a izquierda sii existe una red $\{\nu_j\} \subset \mathcal{P}(\mathcal{G})$ tal que para toda $s \in \mathcal{G}$ vale*

$$\lim_j \|\nu_j - s * \nu_j\|_1 = 0.$$

Compárese tal resultado con 7.3.4.

Demostración.

$\Leftarrow$) Sea μ un punto de w^*-acumulación de $\{\widehat{\nu}_j\}$. Es inmediato que $\mu \in \mathcal{L}(\mathcal{G})$.

$\Rightarrow$) Sea $\mu \in \mathcal{I}(\mathcal{G})$. Entonces existe $\{\xi_j\} \subset \mathcal{P}(\mathcal{G})$ tal que $\mu = w^* - \lim_j \widehat{\xi}_j$. Así, dados $v_1 ... v_n \in \mathcal{G}$ tendremos que

$$w - \lim_j (v_1 * \xi_j - \xi_j ... v_n * \xi_j - \xi_j) = 0$$

en $L^1(\mathcal{G})^n$. Por el teorema de Mazur (cf. 3.14 de [7]) existirá $\{\eta_j\} \subset \mathrm{co}\{\xi_j\}$ tal que $\|v_i * \eta_j - \eta_j\|_1 \to 0$ para $1 \leq i \leq n$. Se sigue la existencia de la red $\{\nu_j\}$.

■

Observación 8.2.2. Recordemos que si $\mathcal{G}$ es localmente compacto, entonces $L^1(\mathcal{G})$ deviene en álgebra de Banach vía la convolución

$$\gamma * \nu(s) = \int_{\mathcal{G}} \gamma(u)\nu(u^{-1}s)d\lambda(s).$$

Repitiendo los mismos argumentos de la demostración anterior, pero con la acción izquierda que $L^1(\mathcal{G})$ genera sobre sí mismo vía la convolución, obtenemos el siguiente

Teorema 8.2.3. *$\mathcal{G}$ es amenable a izquierda sii existe una red $\{\nu_j\} \subset \mathcal{P}(\mathcal{G})$ tal que para toda $\gamma \in \mathcal{P}(\mathcal{G})$ vale*

$$\lim_j \|\nu_j - \gamma * \nu_j\|_1 = 0.$$

Finalmente, tomando en la última demostración una media $\mu \in \mathcal{I}(\mathcal{G})$ obtenemos rápidamente el siguiente resultado:

Proposición 8.2.4. *$\mathcal{G}$ es amenable (a ambos lados) sii existe una red $\{\nu_j\} \subset \mathcal{P}(\mathcal{G})$ tal que para toda $\gamma \in \mathcal{P}(\mathcal{G})$ vale*

$$\lim_j \|\nu_j - \gamma * \nu_j\|_1 = \lim_j \|\nu_j - \nu_j * \gamma\|_1 = 0.$$

También mencionamos la siguiente caracterización y un corolario cuyas demostraciones escapan a los alcances del presente trabajo:

Teorema 8.2.5. *(Condición de Folner, cf. Th 4.16 de [14]) $\mathcal{G}$ es amenable a izquierda sii hay alguna red de subconjuntos compactos $\{K_j\}$ de medida positiva de manera que dado cualquier $s \in \mathcal{G}$ vale*

$$\frac{\lambda(sK_j \triangle K_j)}{\lambda(K_j)} \to 0.$$

Una red con las características enunciadas en el teorema precedente, se la denomina 'red de Folner'.
Dada una red de Folner $\{K_j\}$, definamos $\mu_j = \chi_{K_j}/\lambda(K_j)$, así $\mu_j \in \mathcal{P}(\mathcal{G})$.

Proposición 8.2.6. *(cf. Th 4.3 y 4.15 de [14]) Para $s \in \mathcal{G}$ y $\gamma \in \mathcal{P}(\mathcal{G})$ tendremos:*

$$\lim_j \|s * \mu_j - \mu_j\|_1 = \lim_j \|\gamma * \mu_j - \mu_j\|_1 = 0.$$

8.3. Antirepresentaciones ergódicas

Definición 8.3.1. Una antirepresentación de $L^1(\mathcal{G})$ será una aplicación lineal y continua $S : L^1(\mathcal{G}) \to L(\mathcal{B}, \mathcal{B})$ que además satisface $S_{\gamma * \nu} = S_\nu S_\gamma$.

En general, dada una antirepresentación nos interesará la acción inducida de $\mathcal{P}(\mathcal{G})$ sobre $\mathcal{B}$. Por ello damos la siguiente:

Definición 8.3.2. Diremos que una antirepresentación S es ergódica si lo es $S|_{\mathcal{P}(\mathcal{G})}$
Así, dada una antirepresentación S para cada elemento $x \in \mathcal{B}$ definimos el conjunto $C_x = \overline{\text{co}\{S_\nu x : \nu \in \mathcal{P}(\mathcal{G})\}}$. Naturalmente $M_0(S)$ y $M_1(S)$ serán los correspondientes subespacios generados por la acción de $S|_{\mathcal{P}(\mathcal{G})}$.

Ahora probamos, en este nuevo contexto, un resultado análogo al teorema 7.3.2, a la par que recuperamos la s^*-ergodicidad estudiada en el capítulo 6.

Teorema 8.3.3. *Sea $\mathcal{G}$ localmente compacto y λ su medida de Haar invariante a izquierda. Entonces son equivalentes:*

1) *$\mathcal{G}$ es amenable a izquierda;*

2) *Toda antirepresentación de $L^1(\mathcal{G})$ es s^*-ergódica a izquierda.*

Demostración.

1) $\Rightarrow$ 2) Si $\mu \in \mathcal{L}(\mathcal{G})$, y A es una antirepresentación sobre $\mathcal{B}$, sea $\{\nu_j\} \subset \mathcal{P}(\mathcal{G})$ tal que para cada $\gamma \in \mathcal{P}(\mathcal{G})$ $\|\gamma * \nu_j - \nu_j\| \to 0$. Entonces la familia $T_j = A_{\nu_j}$ es un sistema de promedios: I y II de 6.0.5 se cumplen inmediatamente, para obtener IIIa) observamos que para cada $x \in \mathcal{B}$ vale

$$\|T_j(A_\gamma - I)x\| = \|A_{\gamma * \nu_j - \nu_j} x\| \le M\|\gamma * \nu_j - \nu_j\|\|x\| \to 0.$$

$1) \Leftarrow 2)$ Consideramos la antirepresentación regular $\mathcal{L}$ de $L^1(\mathcal{G})$ sobre el espacio $L^\infty(\mathcal{G})$ vía $\mathcal{L}_\nu(\phi)(s) =: \phi\nu(s) = \phi(\nu * s)$. Es inmediato que $1_\mathcal{G} \in M_0(\mathcal{L})$. Puesto que $\mathcal{L}$ es ergódica a izquierda, tendremos $M_0(\mathcal{L}) \oplus M_1(\mathcal{L})$ (cf. teorema 6.0.8 y obs.6.0.9) y en consecuencia, por el teorema de Hahn-Banach, existirá $\mu \in \mathfrak{m}(\mathcal{G})$ tal que $\mu(1_\mathcal{G}) = 1$ y se anula sobre $M_1(\mathcal{L})$. Luego, dados $\phi \in L^\infty$ y $\nu \in L^1$ resultará que $\mu(\phi\nu - \phi) = 0$. Así $\nu\mu(\phi) = \mu$. Y dado $v \in \mathcal{G}$ resulta $v\mu = v(\nu\mu) = (v * \nu)\mu = \mu$.

■

Observación 8.3.4. Finalmente observemos que si $\mathcal{G}$ es amenable, entonces por 8.2.4 en la primera parte de la demostración anterior, podemos seleccionar una red ν_j que además cumpla $\|\nu_j - \nu_j * \gamma\|_1 \to 0$. Así tendremos

$$\|(A_\gamma - I)T_j x\| = \|A_{\nu_j * \gamma - \nu_j} x\| \leq M\|\nu_j * \gamma - \nu_j\|\|x\| \to 0$$

y resulta satisfecha también la condición IIIb). De esta manera representación es completamente ergódica.

Capítulo 9

Teorema ergódico de F. Greenleaf

En el capítulo 7, al caracterizar la ergodicidad de un semigrupo discreto $\mathcal{S}$ en el lema 7.3.4 quedó establecida la existencia de una familia γ_j de medias finitas que permite obtener sistemas de promedios para cualquier representación acotada de $\mathcal{S}$. Empero, la demostración de tal resultado no es constructiva. En consecuencia, se requiere información adicional para generar expresiones de tales promedios.
Por ello en esta sección, apoyándonos en la topología que un grupo amenable posea, daremos expresiones explícitas que permiten el cálculo de sistemas de promedios para una representación dada.

9.1. Productos de funcionales

Primero introduciremos ciertos productos entre funcionales lineales que son naturalmente inducidos por acciones de un semigrupo discreto $\mathcal{S}$; y el objetivo será enunciar los resultados 9.1.4 y 9.1.6 que utilizaremos en la siguiente sección al estudiar la relación entre antirepresentaciones de $\mathcal{G}$ y $L^1(\mathcal{G})$ de un grupo amenable.

Consideremos un semigrupo $\mathcal{S}$ actuando a derecha sobre un espacio de Banach $\mathcal{B}$. Entonces para toda $f \in \mathcal{B}^*$ y $x \in \mathcal{B}$ definimos

$$\begin{cases} f_x : \mathcal{S} \to \mathbb{C} \\ f_x(s) = f(xs). \end{cases}$$

Es claro que $f_x \in l^\infty(\mathcal{S})$ (compárese esto con lo realizado al ver la implicación $d \Rightarrow e$ en la demostración de 7.3.2).
Sea $\mathcal{B}(\mathcal{S}) = \langle f_x : f \in \mathcal{B}^*, x \in \mathcal{B} \rangle \subset l^\infty(\mathcal{S})$. Entonces, dados $p \in \mathcal{B}(\mathcal{S})^*$ y $f \in \mathcal{B}^*$

tendremos definido el producto

$$\begin{cases} pf : \mathcal{B} \to \mathbb{C} \\ pf(x) = p(f_x). \end{cases}$$

y claramente $pf \in \mathcal{B}^*$.

Proposición 9.1.1. *Son inmediatas las siguientes propiedades*

i) La correspondencia $(f, x) \to f_x$ *es una aplicación bilineal* $\mathcal{B}^* \times \mathcal{B} \to \mathcal{B}(\mathcal{S})$*;*

ii) La correspondencia $(p, f) \to pf$ *es una aplicación bilineal* $\mathcal{B}(\mathcal{S})^* \times \mathcal{B}^* \to \mathcal{B}^*$*;*

iii) Para todo $s \in \mathcal{S}$ *valen* $s(f_x) = (sf)_x$ *y* $(f_x)s = f_{xs}$*. Así* $\mathcal{B}(\mathcal{S})$ *resulta ser un subespacio invariante por las acciones regulares de* $\mathcal{S}$ *sobre* $l^\infty(\mathcal{S})$*;*

iv) Para todo $s \in \mathcal{S}$ $p \in \mathcal{B}(\mathcal{S})^*$ *y* $f \in \mathcal{B}^*$ *vale* $s(pf) = (sp)f$*.*

Definición 9.1.2. Dado un subespacio $Y \subset l^\infty(\mathcal{S})$ invariante a derecha (es decir, vale $\phi s(u) = \phi(su) \in Y$ si $\phi \in Y$), diremos que es *introvertido a izquierda* si $Y(\mathcal{S}) \subset Y$.

Observación 9.1.3. Si Y es un subespacio introvertido a izquierda, naturalmente obtenemos $Y^* \hookrightarrow Y(\mathcal{S})^*$, y dados $p, q \in Y^*$ tiene sentido el producto $pq(\phi) = p(q_\phi)$ donde $q_\phi(s) = q(\phi s)$.

La siguiente proposición, muestra que los espacios introvertidos a izquierda surgen naturalmente al tratar con representaciones.

Proposición 9.1.4. *Sea* $\mathcal{S}$ *un semigrupo que actúa a derecha sobre un espacio de Banach* $\mathcal{B}$*. Entonces* $\mathcal{B}(\mathcal{S}) \subset l^\infty(\mathcal{S})$ *es introvertido a izquierda y*

$$(pq)f = p(qf)$$

para $p, q \in \mathcal{B}(\mathcal{S})^*$ *y* $f \in \mathcal{B}^*$*.*

Demostración. Puesto que $f_x s = f_{xs}$ tenemos que $\mathcal{B}(\mathcal{S})$ es invariante a derecha. Por otro lado, tenemos que si $\phi = f_x$ vale

$$q_\phi(s) = q(\phi s) = q(f_x s) = q(f_{xs}) = pf(xs) = (pf)_x(s).$$

Sigue de la expresión anterior que $\mathcal{B}(\mathcal{S})$ es introvertido a izquierda. Además, tendremos:

$$((pq)f)(x) = pq(f_x) = pq(\phi) = p(q_\phi) = p((qf)_x) = p(qf)(x)$$

y resulta lo deseado. ∎

Observación 9.1.5. Sea $\mathcal{S}$ semigrupo discreto actuando a derecha sobre un espacio de Banach $\mathcal{B}$ y sea $\mathcal{E} = \mathcal{B}^*$. Naturalmente $\mathcal{S}$ actúa a izquierda sobre $\mathcal{E}$, y podemos definir $F_y(s) =: F(sy) \in l^\infty(\mathcal{S})$. Así, escribimos $\mathcal{E}(\mathcal{S}) = \langle F_y : F \in \mathcal{E}^*, y \in \mathcal{E}\rangle$. Luego, hacemos que $L^\infty(\mathcal{S})^*$ actúe a derecha sobre $\mathcal{E}^* = \mathcal{B}^{**}$ vía

$$F\mu(y) = \mu(F_y).$$

En particular, siendo $\widehat{s}$ la evaluación en s, ocurrirá

$$F\widehat{s}(y) = \widehat{s}(F_y) = F_y(s) = F(sy) = Fs(y).$$

Luego, si γ es una media finita, para todo $F \in \mathcal{E}^*$ tendremos

$$F\widehat{\gamma} = \sum_i^k \gamma(s_i)Fs_i. \tag{9.1}$$

También resultará $\widehat{x}\mu(f) = \mu(\widehat{x}_f) = \mu(f_x)$.

Lema 9.1.6. *Sea $x \in \mathcal{B}$ tal que $D_x = \overline{O(x)}$ es w-compacto. Entonces para todo $\mu \in \mathfrak{m}(\mathcal{S})$ tendremos*

$$\widehat{x}\mu \in \widehat{D_x}.$$

Demostración. Sea $\{\gamma_j\}$ tal que $\mu = w^* - \text{lím}_j\, \widehat{\gamma}_j$. Por la expresión 9.1 tendremos que

$$\widehat{x}\widehat{\gamma}_j = \sum_{i=1}^{k_j} \gamma(s_i)\widehat{x}s_i = \sum \gamma(s_i)\widehat{xs_i} \in \widehat{D}_x.$$

Por otro lado, resultará que para todo $y \in \mathcal{E} = \mathcal{B}^*$ vale

$$\widehat{x}\widehat{\gamma}_j = \widehat{\gamma}_j(\widehat{x}_y) \to \mu(\widehat{x}_y) = \widehat{x}\mu(y).$$

Así resulta $w^*\, \text{lím}_j\, \widehat{x}\widehat{\gamma}_j = \widehat{x}\mu$. Puesto que $\widehat{D}_x$ resulta w^*-compacto por hipótesis, tendremos que $\widehat{x}\mu \in \widehat{D}_x$. ■

9.2. Representaciones medibles

Ahora exploraremos la relación entre antirepresentaciones acotadas de $\mathcal{G}$ y $L^1(\mathcal{G})$. Sea $\mathcal{G}$ un grupo localmente compacto y $g \to S_g$ una antirepresentación acotada sobre un espacio de Banach $\mathcal{B}$.

Definición 9.2.1. Una antirepresentación S es **medible** si para toda $f \in \mathcal{B}^*$, $x \in \mathcal{B}$ resulta $f_x \in L^\infty(\mathcal{G})$. En otras palabras, S es medible si $\mathcal{B}(\mathcal{G}) \subset L^\infty(\mathcal{G})$.

Notemos que el hecho de tener una representación medible, depende de la naturaleza de la representación S y a su vez de la topología de $\mathcal{G}$, puesto que esta determina la medida de Haar que a su vez genera el espacio $L^\infty(\mathcal{G})$. Por lo tanto, las mesurabilidad de una antirepresentación depende tanto de la estructura algebraica como topológica de $\mathcal{G}$.

El siguiente teorema muestra que toda antirepresentación medible de $\mathcal{G}$, puede integrarse en cierto sentido dando origen a una antirepresentación de $L^1(\mathcal{G})$.
Recordemos que $D_x = \overline{O(x)}$ en la topología normada.

Teorema 9.2.2. *Sea S una antirepresentación medible y acotada por M. Supongamos que para todo $x \in \mathcal{B}$ D_x es w-compacto. Entonces existe una antirepresentación $\nu \to A_\nu$ de $L^1(\mathcal{G})$ tal que:*

i) $\|A_\nu\| \leq M\|\nu\|_1$ *para todo $\nu \in L^1(\mathcal{G})$ (es decir, A es continua);*

ii) *Si $f \in \mathcal{B}^*$, $\nu \in L^1(\mathcal{G})$ y $x \in \mathcal{B}$ entonces*

$$f(A_\nu x) = \int_{\mathcal{G}} f(xs)\nu(s)d\lambda(s) = \widehat{\nu}(f_x);$$

iii) *Para todo $g \in \mathcal{G}$ y $\nu \in L^1(\mathcal{G})$ vale $A_\nu S_g = A_{g*\nu}$ y $S_g A_\nu = A_{\nu * g}$;*

iv) *Si $\mu \in \mathcal{P}(\mathcal{G})$ entonces $A_\mu x \in D_x$.*

Demostración. Por 9.1.5, tendremos que S induce naturalmente una antirepresentación $B : L^1(\mathcal{G}) \to L(\mathcal{B}^{**}, \mathcal{B}^{**})$ vía $B_\nu F(f) = \widehat{\nu}(F_f)$, y en particular vale $B_\nu \widehat{x}(f) = \widehat{\nu}(f_x)$ para $x \in \mathcal{B}$ (cf. nuevamente 9.1.5).
Puesto que por hipótesis D_x es w-compacto, tendremos por el lema 9.1.6 que para cada $\nu \in \mathcal{P}(\mathcal{G})$ vale $B_v \widehat{x} = \widehat{x}\widehat{\nu} \in \widehat{D}_x$.
Así resulta que podemos definir $A : L^1(\mathcal{G}) \to L(\mathcal{B}, \mathcal{B})$ de manera tal que $\widehat{A_\nu x} = B_\nu \widehat{x}$. Claramente A es una antirepresentación de $L^1(\mathcal{G})$ satisfaciendo $i)$ y $iv)$. Por otro lado, dados $x \in \mathcal{B}$ y $f \in \mathcal{B}^*$ tendremos

$$f(A_\nu x) = \widehat{A_\nu x}(f) = B_\nu \widehat{x}(f) = \widehat{\nu}(f_x).$$

Así resulta $ii)$.
Para ver $iii)$ vemos que si $g \in \mathcal{G}$, $\nu \in L^1(\mathcal{G})$,$x \in \mathcal{B}$ y $f \in \mathcal{B}^*$, recordando que $\widehat{g * \nu} = g\widehat{\nu}$ tendremos

$$\begin{aligned} f(S_g A_\nu x) = f((A_\nu x)g) &= gf(A_\nu x) \\ &= B_v \widehat{x}(gf) \\ &= gB_v(\widehat{x}_f) \\ &= g\widehat{\nu}(f_x) = \widehat{g * \nu}(f_x) = f(A_{g*\nu}). \end{aligned}$$

y análogamente se demuestra la otra identidad. ■

En estas circunstancias escribimos:

$$A_\nu = \int_{\mathcal{G}} S_g \nu(g) d\lambda(g).$$

Si $E \subset \mathcal{G}$ es medible y $\nu = \chi_E$ entonces escribimos

$$\int_E S_g d\lambda(g) = \int_{\mathcal{G}} S_g \chi_E(g) d\lambda(g).$$

Sea $g \to T_g$ una *representación medible* de $\mathcal{G}$ sobre $\mathcal{B}$. Entonces definiendo $S_g = T_{g^{-1}}$ obtenemos una antirepresentación medible. Así, si D_x es w-compacto para cada $x \in \mathcal{B}$, obtenemos la antirepresentación de $L^1(\mathcal{G})$ correspondiente.

9.3. Teorema de F. Greenleaf

Observación 9.3.1. Sea $\{K_j\}$ una red de Folner y $\mu_j = \chi_{K_j}/\lambda(K_j)$, y sean A y S antirepresentaciones como en el teorema 9.2.2. Entonces, por 8.2.6 automáticamente tenemos que A_{μ_j} es un sistema de promedios a izquierda para A y S.

Finalmente enunciamos el siguiente resultado, que es el Teorema Ergódico Medio para grupos localmente compactos amenables.

Teorema 9.3.2. *(Greenleaf, cf. [9]) Sea $\mathcal{G}$ un grupo localmente compacto y amenable (a ambos lados) y $\{K_j\}$ una red de Folner. Sea $g \to T_g$ una representación acotada y medible de $\mathcal{G}$ sobre un espacio de Banach $\mathcal{B}$ tal que para cada $x \in \mathcal{B}$ tenemos que D_x es w-compacto. Entonces*

$$A_j = \lambda(K_j)^{-1} \int_{K_j} T_{g^{-1}} d\lambda(g)$$

es un sistema de promedios a izquierda para la representación T, que además resulta ser s^*-ergódica y $M(T) = \mathcal{B}$.

Demostración.
Aplicando el teorema 9.2.2 a $S_g = T_{g^{-1}}$ obtenemos la antirepresentación $\nu \to A_\nu$ de $L^1(\mathcal{G})$ sobre $\mathcal{B}$. En vistas de la observación anterior $A_j =: A_{\mu_j}$ resultan un sistemas de promedios a izquierda tanto para A como para T. Además por $iv)$ de 9.2.2 tendremos que $C_x \subset D_x$, así C_x resulta w-compacto para cada $x \in \mathcal{B}$.
Puesto que $\mathcal{G}$ es amenable, por 8.3.4 resultará que A es una antirepresentación s^*-ergódica. Así, por 6.1.3 tendremos que $M_0(A) \oplus M_1(A) = \mathcal{B}$. Así, por la parte 2 de 6.0.8, aunque

$\{A_j\}$ sólo sean promedios a izquierda, tendremos que $A_j x \to Px$ en norma, donde P es la proyeccion de $M(A) = \mathcal{B}$ sobre $M_0(A)$.
Veamos que $M_0(A) = M_0(T)$ y $M_1(A) = M_1(T)$ y el teorema quedará probado: demostremos que $M_0(A) \subset M_0(T)$: si $x \in M_0(A)$, $g \in \mathcal{G}$ y $\nu \in \mathcal{P}(\mathcal{G})$, entonces

$$T_g x = T_g A_\nu x = A_{\nu * g^{-1}} x$$

por $iii)$ de 9.2.2.
Recíprocamente: si $x \in M_0(T)$ por $ii)$ de 9.2.2 tendremos que dados $\nu \in \mathcal{P}(\mathcal{G})$ y $f \in \mathcal{B}^*$ resulta

$$f(A_\nu x) = \widehat{\nu}(f_x) = f(x).$$

Así $x \in M_0(A)$.
Demostremos ahora $M_1(T) \subset M_1(A)$: dado $x \in \mathcal{B}$ y $g \in \mathcal{G}$ notemos que partiendo del hecho que $A_j(I - T_g)x \to 0$ en norma, resultará que $P(I - T_g)x = 0$. De esto se sigue, por la parte 2 de 6.0.8, que $x \in M_1(A)$.
Finalmente, sea $\mu \in \mathcal{P}(\mathcal{G})$, $x \in \mathcal{B}$. Tomemos una red $\{\sigma_\delta\}$ de la forma $\sigma_\delta = \sum_{i=1}^{k_\delta} a_i \widehat{g}_i$ (con $\sum_i a_i = 1$) de manera tal que $w^* - \lim_\delta \sigma_\delta = \widehat{\mu}$ (cf.B). Así

$$\lim_\delta [f(x) - \sigma_\delta(f_x)] = [f(x) - \widehat{\mu}(f_x)] = f((I - A_\mu)x)$$

entonces resultará

$$[f(x) - \sigma_\delta(f_x)] = f(x) - \sum_{i=1}^{k_\delta} a_i \widehat{g}_i(f_x) = f(\underbrace{[I - \sum_{i=1}^{k_\delta} a_i T_{g_i}]}_{E_\delta} x).$$

Y E_δ no depende de x ni de f. Y así tendremos que $w - \lim_\delta E_\delta x = (I - A_\mu)x$. Luego

$$(I - A_\mu)x \in \overline{M_1(T)}^w = M_1(T).$$

■

Notemos que apoyándonos en la amenabilidad de $\mathcal{G}$ y la mesurabilidad de la representación, propiedades que dependen tanto de la estructura algebraica como topológica de $\mathcal{G}$, se consiguieron establecer expresiones explícitas para una familia de promedios de una representación dada.

Apéndice A

Conmutatividad de la convolución contra D_N

Pretendemos mostrar que

$$f * D_m * D_{m+p} = f * D_{m+p}.$$

Ahora

$$\begin{aligned} f * D_m * D_{m+p} &= \frac{1}{\pi}\int_{-\pi}^{\pi}(f * D_m)(t)D_{m+p}(x-t)dt \\ &= \frac{1}{\pi}\int_{-\pi}^{\pi}\frac{1}{\pi}\int_{-\pi}^{\pi} f(s)D_m(t-s)dsD_{m+p}(x-t)dt \\ &= \frac{1}{\pi}\int_{-pi}^{\pi} f(s)\underbrace{\frac{1}{\pi}\int_{-\pi}^{\pi} D_m(t-s)D_{m+p}(x-t)dt}_{E_{m,m+p}(x,s)}\, ds. \end{aligned} \tag{A.1}$$

Entonces

$$\begin{aligned} E_{m,m+p}(x,s) =& \frac{1}{\pi}\int_{-\pi}^{\pi}(\frac{1}{2}+\sum_{k=1}^{m}\cos[k(t-s)])(\frac{1}{2}+\frac{1}{\pi}\sum_{h=1}^{m+p}\cos[h(x-t)])dt \\ =& \frac{1}{\pi}\int_{-\pi}^{\pi}\{\frac{1}{4}+\frac{1}{2}\sum_{h=1}^{m+p}\cos[h(x-t)]+\frac{1}{2}\sum_{k=1}^{m}\cos[k(t-s)]\}dt+\frac{1}{\pi}\sum_{\substack{1\leq k\leq m\\ 1\leq h\leq m+p}}\alpha_{k,h}(s,x) \\ =& \frac{1}{2}+\frac{1}{2\pi}(\sum_{h=1}^{m+p}\frac{\sin[h(x-t)]}{-h}|_{-\pi}^{\pi}+\sum_{k=1}^{m}\frac{\sin[k(t-s)]}{k}|_{-\pi}^{\pi})+\frac{1}{\pi}\sum_{\substack{1\leq k\leq m\\ 1\leq h\leq m+p}}\alpha_{k,h}(s,x). \end{aligned} \tag{A.2}$$

Donde

$$\alpha_{k,h}(x,s) = \int_{-\pi}^{\pi}\cos[k(t-s)]\cos[h(x-t)]dt.$$

Utilizando que en general vale $\sin(p) - \sin(q) = 2\cos(\frac{p+q}{2})\sin(\frac{p-q}{2})$ tendremos:

$$E_{m,m+p}(x,s) = \frac{1}{2} + \frac{1}{\pi} \sum_{\substack{1 \le h \le m+p \\ 1 \le k \le m}} \alpha_{k,h}(s,x).$$

Trabajemos sobre la expresión

$$\begin{aligned}
\alpha_{k,h}(s,x) &= \int_{-\pi}^{\pi} \cos[k(t-s)]\cos[h(x-t)]dt \\
&= \int_{-\pi}^{\pi} [\cos(kt)\cos(ks) + \sin(kt)\sin(ks)][\cos(hk)\cos(ht) + \sin(hx)\sin(ht)]dt \\
&= \cos(ks)\cos(hk) \underbrace{\int_{-\pi}^{\pi} \cos(kt)\cos(ht)dt}_{I} + \sin(ks)\sin(hk) \underbrace{\int_{-\pi}^{\pi} \sin(kt)\sin(ht)dt}_{J}.
\end{aligned}$$

Por otra parte

$$\begin{aligned}
\frac{1}{2}I &= \frac{\sin(kt)}{k}\cos(ht)|_0^\pi + \int_0^\pi \frac{\sin(kt)}{k} h\sin(ht)dt \\
&= \frac{h}{k}\int_0^\pi \sin(kt)\sin(ht)dt \text{ (ie } I = \frac{h}{k}J) \\
&= \frac{h}{k}[-\frac{\cos(kt)}{k}\sin(ht)|_0^\pi + \int_0^\pi \frac{\cos(kt)}{k}\cos(ht)dt] \\
&= (\frac{h}{k})^2 I \text{ (o sea } (1-\frac{h^2}{k^2})I = 0).
\end{aligned}$$

Luego $I = 0$ si $h \neq k$ (y $J = 0$ si $h \neq k$).
Cuando $h = k$ tenemos

$$\frac{1}{2}I = \int_0^\pi \cos^2(kt)dt = \frac{1}{k}\int_0^{k\pi} \cos^2(u)du = \frac{1}{2k}[1+\cos(2u)]du = \frac{\pi}{2}.$$

Asimismo, $\frac{1}{2}J = \int_0^\pi \sin^2(kt) = \pi - \frac{I}{2} = \frac{\pi}{2}$.

Entonces

$$\alpha_{k,h}(s,x) = \pi[\cos(ks)\cos(hx) + \sin(ks)\sin(hx)]\delta_{k,h} = \pi\delta_{k,h}\cos(ks-hx).$$

Luego en A.2 resulta

$$E_{m,m+p}(x,s) = \frac{1}{2} + \frac{1}{\pi}\sum_{k=1}^{m} \pi\cos[k(s-x)] = D_m(s-x).$$

y por A.1 sigue 6.1. Análogamente se puede demostrar que

$$f * D_{m+p} * D_m = f * D_{m+p}.$$

Apéndice B

Medias y promedios

Sea X un espacio topológico localmente compacto, y μ una medida sobre X. Sea

$$\mathcal{P}(X) = \{f \in L^1(\mu) : f \geq 1y\|1\|_1 = 1\}.$$

Y notemos con

$$\mathfrak{m}(X) = \{\lambda \in L^\infty(X)^* : \|\lambda\| = \lambda(e) = 1\}.$$

En estas condiciones, podemos enunciar el siguiente

Teorema B.0.3. *Si μ es una medida regular, entonces*

$$\mathfrak{m}(X) = \overline{\mathcal{P}(X)^\wedge}^{w^*}.$$

Demostración. Es inmediata la inclusión $\supset$. Veamos la otra. Supongamos que existe una media $m \in \mathfrak{m}(X)$ tal que $m \notin \overline{\mathcal{P}(X)}^{w^*}$. Por el teorema de Hahn-Banach, tendremos que existe $\Lambda : L^\infty(X)^* \to \mathbb{C}$ lineal, w^*-continua y $\alpha, \epsilon > 0$ tales que

$$\mathrm{Re}\langle \widehat{f}, \Lambda\rangle \leq \alpha < \alpha + \epsilon < \mathrm{Re}\langle m, \Lambda\rangle.$$

Por otra parte existe $\phi_0 \in L^\infty(X)$ tal que, para cada $\eta \in L^\infty(X)^*$

$$\langle \eta, \Lambda\rangle = \langle \phi_0, \eta\rangle.$$

Así, si $f \in \mathcal{P}(X)$

$$\mathrm{Re}\langle f, \phi_0\rangle \leq \alpha < \alpha + \epsilon < \mathrm{Re}\langle \phi_0, m\rangle.$$

Además $\|\phi_0\|_\infty = \text{ínf} \sup_{N \in \mathcal{I}}\{|\phi_0(x)| : x \in X - N\}$ donde $\mathcal{I}$ es la clase de partes localmente nulas de X, es decir $N \in \mathcal{I}$ sii N es λ-medible y $\lambda(N \cap C) = 0$ para cada C compacto.
Como $\|\phi_0\|_\infty = \sup\{t > 0 : \lambda(\{|\phi_0| \geq t\}) > 0\}$, existe $t > \|\phi_0\| - \epsilon/2$ y algún compacto $C \subset \{|\phi_0| \geq t\}$ de medida positiva (por regularidad de λ).

Sin perdida de generalidad, podemos suponer que ϕ_0 es real valuada (cf.B.0.4) y que $C \subset \{\phi_0 \geq t\}$. Entonces, si $f = \lambda(C)^{-1}\chi_C$ tendremos

$$\langle f, \phi_0 \rangle = \frac{1}{\lambda(C)} \int_C \phi_0 d\lambda \geq t > \|\phi_0\| - \frac{\epsilon}{2}.$$

y también $\langle f, \phi_0 \rangle < \langle \phi_0, m \rangle - \epsilon/2$.
Así resulta

$$\|\phi_0\| < \langle \phi_0, m \rangle \leq \|\phi_0\|_\infty.$$

lo cual es absurdo. ■

Observación B.0.4. Sea $\phi \in L^\infty(X)$, $m \in \mathfrak{m}(X)$ tal que $\langle \phi, m \rangle = a + ib$ con a, b reales. Dado $t \in \mathbb{R}$ tendremos

$$a^2 + (b+t)^2 = |\langle \phi, m \rangle + it|^2 = |\langle \phi + it1_X, m \rangle|^2 \leq \|\phi + it1_X\|_\infty^2.$$

Si ϕ es real valuada en λ-c.t.p., será

$$a^2 + b^2 + 2bt + t^2 \leq \|\phi\|_\infty^2 + t^2$$

y debe ser $b = 0$. Así $\langle \phi, m \rangle \in \mathbb{R}$ si $\phi(X) \subset \mathbb{R}$.

Bibliografía

[1] Alaoglu, L. & Birkhoff, G.: *General ergodic theorems*. Ann. of Math. (2) vol 41, 293-309, (1940).

[2] Birkhoff, G.:*Proof of the ergodic theorem.* Proc. Natl. Acad. Sci. USA 17 (12), 656-660, (1931).

[3] Birkhoff, G.: *What is the ergodic theorem?-* The American Math. Monthly, Vol. 49, no. 4, 222-226, (1942).

[4] Brush, S. G.: *Proof of the impossibility of ergodic systems: the 1913 papers of Rosenthal and Plancherel.* En The kinetic theory of gases: An antology of classical papers with historical commentary. Imperial College Press, p. 505, (2003).

[5] Conway, J.: *A course in functional analysis.* Springer-Verlag, Inc. New York (1997).

[6] Day, Mahlon M.: *Means for the bounded functions and ergodicity of the bounded representations of semi-groups.* Trans. Amer. Math. Soc. 69, 276-291 (1950).

[7] Dunford, N. & Schwartz, J.: *Linear Operator Theory, Vol I.* Interscience Publishers, Inc., New York (1958).

[8] Eberlein, W.: *Abstract ergodic theorems and weak almost periodic functions.* Trans. Amer.Math. Soc., USA, Vol. 67, 217-224 (1948).

[9] Greenleaf,F.: *Ergodics theorems and the construction of summing sequences for amenable locally compacts groups.* Comm. Pure Applied Math 26, 26-46 (1973).

[10] Krabs, W. & Pickl, S.: *Dynamical Systems, stability, controllability and chaotic behaviour.* ©Springer-Verlag Berlin Heidelberg (2010).

[11] Koopman, B.: *Hamiltonian systems and transformations in Hilbert spaces.* Proc. Nat. Acad. Sci., USA, Vol. 5, 315-318 (1931).

[12] Lorch, E.: *Means of iterated transformations in reflexive vector spaces.* Bull. Amer. Math. Soc. Volume 45, Number 12, Part 1, 945-947 (1939).

[13] Reed, M. & Simon, B.: *Methods on modern mathematics physics, Vol 1.* Academic Press, INC., London (1970).

[14] Paterson,A.: *Amenability.* American Mathematical Society, U.S.A. (1988).

[15] Sarig, Omri: *Lectures on Ergodic Theory.* Pane Sate, Fall (2008).

[16] Sz. Nagy, B.: *On uniformly bounded linear transformations in Hilbert Spaces.* Acta Univ. Szeged., vol 11, 152-157, (1947).

[17] Visser, C.: *On the iteration of linear operators in a Hilbert Space.* Proc. Roy. Acad. Amsterdam, Vol 41, 487-495, (1938)

[18] von Neumann, J.: *Proof of the quasy-ergodic hypothesis.*Proc. Nat. Acad. Sci., USA, Vol. 18, 70-82, (1932).

[19] von Neumann, J.: *Physical applications of the ergodic hypothesis.* Proc. Nat. Acad. Sci., USA, Vol. 18, 263-266 (1932).

[20] Yosida, K.: *Mean ergodic theorem in Banach Spaces.* Proc. Imp. Acad. Tokyo, Vol. 14, 292-294, (1938).

[21] Yosida, K. & Kakutani, S.: *Applications of mean ergodic theorems to the problems of Markoff process.* Proc. Imp. Acad., Tokyo, Vol. 14, 333-339 (1938).

[22] Zygmund, A.:*Trigonometrical Series, Vol. 1.* Cambridge University Press (2002).