UNIVERSIDAD CENTRAL DE VENEZUELA
FACULTAD DE CIENCIAS
ESCUELA DE MATEMÁTICA

DILATACIONES UNITARIAS Y CADENAS DE MARKOV

Trabajo Especial de Grado presentado ante la ilustre Universidad Central de Venezuela para optar al título de Licenciada en Matemática.

Autora: Bra. Mary Ana Allen Meleán
Tutora: Dra. Marisela Domínguez

Caracas, Venezuela
Marzo de 2.005

Nosotros, los abajo firmantes, designados por la Universidad Central de Venezuela como integrantes del Jurado Examinador del Trabajo Especial de Grado titulado “**Dilataciones unitarias y cadenas de Markov**”, presentado por la **Bra. Mary Ana Allen Meleán**, titular de la Cédula de Identidad **13.992.561**, certificamos que este trabajo cumple con los requisitos exigidos por nuestra Magna Casa de Estudios para optar al título de **Licenciada en Matemática**.

Dra. Marisela Domínguez
Tutora

Dr. Ramón Bruzual
Jurado

Dr. Daniel Barráez
Jurado

"Mathematics is the art of giving the same name to different things"[As opposed to the quotation: Poetry is the art of giving different names to the same thing].

Poincaré, Jules Henri.

♦

"A mathematician who is not also something of a poet will never be a complete mathematician". **Weierstrass, Karl.**

AGRADECIMIENTO

A mis padres Milagros e Ivo, por permitirme Ser, a mis hermanos Ivo Alejandro y Aimara por ser además de hermanos, amigos; a mis abuelas Yolanda y Carmen (Mamama), y tías Ida e Ylia, por todo su apoyo y por hacer posible el venir a estudiar a esta gran ciudad.

A mi tutora, la profesora Marisela Domínguez por presentarme el artículo preciso que captara mi interés, sabiendo mi inclinación hacia el análisis y las probabilidades (y estadística), permitiéndome trabajar un poco en ambas áreas. Por su orden y planificación contagiantes, y por supuesto, por su guía y orientación en la realización de este trabajo.

Al profesor Daniel Barráez por su recomendación de referencias bibliográficas y al profesor Ramón Bruzual por su sugerencia en la redacción de una parte específica del trabajo.

A todos los profesores y preparadores que ayudaron en mi formación matemática durante la carrera, a quienes nombro en orden de aparición en los semestres que cursé: Margarita Olivares, Alberto Voza (Axel), Ana Karina Fermín, Pedro Alson, José Gregorio Mijares (Goyo), Yamilet Quintana (Vicky), María Morán (Loló), Miguel Méndez, Mercedes Arriojas, Marisela Domínguez, Wilfredo Urbina, Manuel Maia, Gisela Méndez, Rafael Díaz, Ventura Echandía, Lisandro Fermín y Daniel Barráez. Todos merecen mi admiración y aprecio.

A mis compañeros y amigos, conseguidos del convivir en la universidad, por todos los momentos compartidos, por todos los recuerdos de risas y llanto (porque es con los amigos con quienes se rie y se llora). En especial a mi amiga Carla, por su gran amistad, por su

apoyo y comprensión de mis “rollos existenciales” durante estos 5 años y porque si no le escribo nada aquí no me habla más...(mentira, una broma de amigas).

A un amigo, de esos pocos a quienes se conoce de manera “mágica”, con quienes sólo hace falta un instante para saber lo especiales que son, a Jorge Albarrán, quien desde muy lejos (México), de manera incondicional (sólo “para hacerme sonreir”), me ayudó con el formato (latex) de este trabajo.

Finalmente, a la Vida, que hoy en una de sus buenas caras me permite culminar esta carrera presentando este trabajo.

♦

ÍNDICE GENERAL

INTRODUCCIÓN

El objetivo de este trabajo es desarrollar, de una manera razonablemente comprensible, algunos de los principales teoremas contenidos en el artículo "Unitary dilations of Markov transition operators, and the corresponding integral representations for transition-probability matrices"de **David Kendall** publicado en la revista *Probability and Statistic, H. Cramer, Vol. 139-161*, en 1959.

El resultado principal de esa publicación de Kendall es que se puede encontrar una representación de Fourier-Stieltjes para la sucesión $p_{jk}, p^2_{jk}, p^3_{jk}, ...$, donde p^n_{jk} representa la probabilidad de que una cadena de Markov, homogénea en tiempo, inicialmente en el estado j, pase al estado k en la n-ésima transición, con j y k estados idénticos o distintos.

Aunque Kendall, en su publicación, presenta diversas variaciones de este resultado, sólo nos dedicaremos a estudiar una de ellas para el caso en que las cadenas de Markov son irreducibles. Una de las herramientas utilizadas para llegar a ese resultado son las medidas subinvariantes positivas. A ese respecto, Kendall establece un teorema donde se prueba que toda cadena de Markov irreducible admite una medida subinvariante positiva.

Para la comprensión de los resultados de Kendall es necesario tener un conocimiento, no sólo en probabilidades, sobre cadenas de Markov, si no también en importantes áreas del análisis matemático como teoría espectral y dilataciones unitarias. Es por ello que estos tópicos se desarrollan, como marco teórico, en los tres primeros capítulos del presente trabajo.

Así, el primer capítulo trata sobre cadenas de Makov, en el cual se presenta un breve estudio sobre estas cadenas, incluyendo la definición y algunos resultados básicos.

En el segundo capítulo se desarrolla la teoría espectral. Se comienza con el estudio de álgebras de Banach, luego con los operadores en espacios de Hilbert y el estudio de álgebras-C^* para entonces establecer el muy importante teorema espectral. Seguidamente en dicho capítulo se estudia otra versión de este teorema donde se establece la conexión con integrales empleando la medida espectral. Este teorema es de vital importancia en la demostración de uno de los resultados de Kendall.

En el tercer capítulo, se realiza un estudio sobre dilataciones unitarias de una contracción con el objetivo de establecer el teorema de Sz.-Nagy, utilizado por Kendall en su publicación.

Finalmente, en el cuarto capítulo se desarrolla la primera parte del artículo de Kendall que contiene los resultados que nos dedicamos a estudiar y presentar en este trabajo.

CAPÍTULO 1

Cadenas de Markov

1.1. Noción sobre Procesos Estocásticos

Definición 1.1 Un *proceso estocástico* es una familia de variables aleatorias $\{X_t : t \in \mathcal{T}\}$, donde t representa el tiempo, definidas en un espacio de probabilidad $(\Omega, \mathcal{F}, P)$. El conjunto donde el proceso toma sus valores se denomina *espacio de estados*, y lo denotaremos con la letra $\mathcal{E}$.

Informalmente, se puede decir que un proceso estocástico describe la historia de una variable aleatoria que evoluciona en el tiempo.

Si $\mathcal{T} = \mathbb{N}$ ó $\mathcal{T} = \mathbb{Z}$ entonces se dice que el proceso estocástico está a *tiempo discreto*. Cuando $\mathcal{T}$ es discreto, la *ley* de un proceso estocástico está caracterizado por las distribuciones en dimensión finita:

$$P\{X_t = x_t : t \in T\}$$

con $T \subseteq \mathcal{T}$ y $x_t \in \mathcal{E}$, $t \in T$.

1.2. Cadenas de Markov

Sea Ω un espacio muestral y P una medida de probabilidad en dicho espacio. Sea $X = \{X_n : n \in \mathbb{N}\}$ un proceso estocástico a tiempo discreto con un espacio de estados numerable $\mathcal{E}$; esto es, para cada $n \in \mathbb{N}$ y $\omega \in \Omega$, $X_n(\omega)$ es un elemento del conjunto numerable $\mathcal{E}$. Es usual decir que “el proceso se encuentra en el estado j en el instante n ” cuando $X_n = j$. Así, X_n se refiere al estado del proceso X en el tiempo n.

Definición 1.2 Un proceso estocástico $X = \{X_n : n \in \mathbb{N}\}$ es llamado *cadena de Markov* cuando:

$$P\{X_{n+1} = j|X_0, X_1, ..., X_n\} = P\{X_{n+1} = j|X_n\} \tag{1.1}$$

para todo $j \in \mathcal{E}$ y $n \in \mathbb{N}$.

Entonces, una cadena de Markov es una sucesión de variables aleatorias tal que: X_{n+1} es independiente condicionalmente de $X_0, X_1, ..., X_{n-1}$ dado X_n, para cualquier n. Esto es, el “siguiente” estado X_{n+1} del proceso es independiente de los estados “anteriores” $X_0, X_1, ..., X_{n-1}$, siempre que sea conocido el estado “presente” X_n.

Sin pérdida de generalidad, nos restringiremos a las cadenas de Markov para las cuales la probabilidad condicional

$$P\{X_{n+1} = j|X_n = i\} = p_{ij}, \qquad i, j \in \mathcal{E}, \tag{1.2}$$

es independiente de n. A las cadenas que cumplen esta condición se les dice que son *homogéneas en tiempo.*

Las probabilidades p_{ij} son las llamadas *probabilidades de transición* de la cadena de Markov X. Es común disponer las p_{ij} en una matriz cuadrada P llamada *matriz de transición* de la cadena de Markov X. Por ejemplo, si $\mathcal{E}= \{0, 1, ...\}$, la matriz de transición es

$$P = \begin{bmatrix} p_{00} & p_{01} & p_{02} & \cdots \\ p_{10} & p_{11} & p_{12} & \cdots \\ p_{20} & p_{21} & p_{22} & \cdots \\ . & . & . & \cdots \\ . & . & . & \cdots \\ . & . & . & \cdots \end{bmatrix}$$

Claramente, toda entrada de una matriz de transición P es no negativa, y los términos de cualquier fila suman uno. Esta última condición expresa el hecho de que ocurre alguna transición en cada paso (Por conveniencia se dice que una transición ha ocurrido aún cuando el estado continúe siendo el mismo).

Definición 1.3 Sea P una matriz cuadrada de entradas p_{ij} definidas para todo $i, j \in \mathcal{E}$. Entonces P es denominada *matriz de Markov sobre $\mathcal{E}$ o matriz estocástica* si:

(a) para todo $i, j \in \mathcal{E}$, $p_{ij} \geq 0$, y

(b) para cada $i \in \mathcal{E}$, $\sum_{j \in \mathcal{E}} p_{ij} = 1$.

Así, la matriz de transición de una cadena de Markov es una matriz de Markov. Recíprocamente, dada una matriz de Markov sobre un conjunto numerable $\mathcal{E}$, es posible construir un espacio muestral Ω, una probabilidad P sobre todos los subconjuntos de Ω, y variables aleatorias $X_0, X_1, ...$ en Ω, tomando valores en $\mathcal{E}$ tales que $X = \{X_n\}$ es una cadena de Markov cuya matriz de transición es la matriz dada. En dicha construcción es usual considerar Ω como el conjunto de todas las sucesiones $\omega = (\omega_0, \omega_1, \omega_2 ...)$ con $\omega_i \in \mathcal{E}$ para todo i y definir $X_n(\omega) = \omega_n$.

Sea X una cadena de Markov con matriz de transición $P = \{p_{ij}\}$ y espacio de estados $\mathcal{E}$.

Teorema 1.4 *Para todo $n, m \in \mathbb{N}$ con $m \geq 1$; $i_0, ..., i_m \in \mathcal{E}$,*

$$P\{X_{n+1} = i_1, ..., X_{n+m} = i_m | X_n = i_0\} = p_{i_0 i_1} p_{i_1 i_2} ... p_{i_{m-1} i_m}$$

Demostración:

Sea $n \in \mathbb{N}$ fijo. Demostremos por inducción sobre m.

Para $m = 1$ se cumple trivialmente.

Supongamos que se cumple para m. Veamos que se satisface para $m + 1$.

Sean $i_0, i_1, ..., i_m \in \mathcal{E}$ fijos. De la definición de probabilidad condicional y la propiedad de las cadenas de Markov (1.1) tenemos

$$\begin{aligned}
&P\{X_{n+1} = i_1, ..., X_{n+m} = i_m, X_{n+m+1} = i_{m+1}|X_n = i_0\} = \\
&= \frac{P\{X_n = i_0, ..., X_{n+m+1} = i_{m+1}\}}{P\{X_0 = i_0\}} \\
&= \frac{P\{X_n = i_0, ..., X_{n+m+1} = i_{m+1}\}}{P\{X_n = i_0, ..., X_{n+m} = i_m\}} \quad \frac{P\{X_n = i_0, ..., X_{n+m} = i_m\}}{P\{X_0 = i_0\}} \\
&= P\{X_{n+m+1} = i_{n+m+1}|X_n = i_0, ..., X_{n+m} = i_m\}P\{X_{n+1} = i_1, ..., X_{n+m} = i_m|X_n = i_0\} \\
&= P\{X_{n+m+1} = i_{n+m+1}|X_{n+m} = i_m\}P\{X_{n+1} = i_1, ..., X_{n+m} = i_m|X_n = i_0\}.
\end{aligned}$$

Finalmente por hipótesis inductiva nos queda

$$P\{X_{n+1} = i_1, ..., X_{n+m} = i_m, X_{n+m+1} = i_{m+1}|X_n = i_0\} = p_{i_0 i_1}...p_{i_{m-1} i_m} p_{i_m i_{m+1}}.$$

Hemos probado el resultado para todo $m \geq 1$. $\square$

Corolario 1.5 *Sea π una distribución de probabilidad en $\mathcal{E}$, y supongamos*

$$P\{X_0 = i\} = \pi(i)$$

para todo $i \in \mathcal{E}$. Entonces para todo $m \in \mathbb{N}$ y $i_0, ..., i_m \in \mathcal{E}$,

$$P\{X_0 = i_0, X_1 = i_1, ..., X_m = i_m\} = \pi(i) p_{i_0 i_1}...p_{i_{m-1} i_m}$$

Demostración:

Del teorema, haciendo $n = 0$, tenemos

$$P\{X_1 = i_1, ..., X_m = i_m|X_0 = i_0\} = p_{i_0 i_1}...p_{i_{m-1} i_m}$$

Luego, usando la definición de probabilidad condicional se obtiene el resultado

$$\begin{aligned}
P\{X_0 = i_0, X_1 = i_1, ..., X_m = i_m\} &= P\{X_1 = i_1, ..., X_m = i_m|X_0 = i_0\}P\{X_0 = i_0\} \\
&= \pi(i) p_{i_0 i_1}...p_{i_{m-1} i_m}
\end{aligned}$$

$\square$

Este corolario muestra que la distribución conjunta de $X_0, X_1, ..., X_m$ está completamente determinada, para todo m, por la distribución inicial π y la matriz de transición P.

La siguiente proposición muestra que la probabilidad de que la cadena se mueva del estado i al estado j en m transiciones está representada en la entrada (ij) de la m-ésima potencia de la matriz P, P^m, denotada por p_{ij}^m.

Proposición 1.6 *Para cualquier $m \in \mathbb{N}$,*

$$P\{X_{n+m} = j | X_n = i\} = p_{ij}^m$$

para todo $i, j \in \mathcal{E}$ y $n \in \mathbb{N}$ (en particular, para $m = 0, P^0 = I$).

Demostración:

Sean $i, j \in \mathcal{E}$ fijos. Para $m = 1$ se cumple por la condición de homogeneidad temporal (1.2). Supongamos que se cumple para m. Veamos para m+1.

$$\begin{aligned} P\{X_{n+m+1} = j | X_n = i\} &= \sum_{k \in \mathcal{E}} P\{X_{n+m} = k | X_n = i\} P\{X_{n+m+1} = j | X_{n+m} = k\} \\ &= \sum_{k \in \mathcal{E}} p_{ik}^m \cdot p_{kj} = p_{ij}^{m+1} \end{aligned}$$

Donde la última igualdad proviene del producto de matrices.

Por lo tanto hemos probado, por inducción, que se cumple para todo $m \geq 1$. $\square$

Obviamente, para $m, n \in \mathbb{N}$ tenemos la siguiente relación de matrices

$$P^{m+n} = P^m P^n,$$

para cuyas entradas tenemos la siguiente ecuación

$$p_{ij}^{m+n} = \sum_{k \in \mathcal{E}} p_{ik}^m \cdot p_{kj}^n, \qquad i, j \in \mathcal{E},$$

que se denomina *ecuación Chapman-Kolmogorov*, la cual establece la conexión entre las cadenas de Markov y la teoría de semigrupos de operadores.

1.3. Clasificación de Estados de una Cadena de Markov

Los estados de una cadena de Markov pueden ser clasificados de muchas maneras. Esta clasificación es de gran importancia en el estudio del comportamiento asintótico de las n-ésimas probabilidades de transición p_{ij}^n. A continuación se presentan diferentes clases de estados mediante definiciones y relaciones fundamentales.

Definición 1.7 Se dice que el estado j es *accesible* desde el estado i si existe algún entero $n \geq 0$, tal que $p_{ij}^n > 0$; i.e., el estado j es accesible desde el estado i si hay una probabilidad positiva de que en un número finito de transiciones, el estado j pueda ser alcanzado comenzando desde el estado i.

Definición 1.8 Se dice que dos estados i y j están *intercomunicados* cuando cada uno es accesible desde el otro. Y esto se denota por $i \leftrightarrow j$.

Nótese que si dos estados i y j no están intercomunicados, entonces $p_{ij}^n = 0$ para todo $n \geq 0$, ó $p_{ji}^n = 0$ para todo $n \geq 0$, o ambas.

Definición 1.9 Tenemos las siguientes definiciones:

(a) Se dice que un conjunto de estados es *cerrado* si ningún estado de afuera es accesible desde un estado del conjunto.

(b) Un *estado absorbente* es un estado que forma un conjunto cerrado por sí solo.

(c) Un conjunto cerrado es *irreducible* si no contiene ningún subconjunto propio cerrado.

(d) Se dice que una cadena de Markov es *irreducible* si su único conjunto cerrado es el conjunto de todos los estados. De lo contrario se dice *reducible*.

La matriz de Markov asociada a una cadena reducible puede ser escrita en forma de una matriz particionada; por ejemplo,

$$P = \begin{bmatrix} P_1 & 0 \\ 0 & P_2 \end{bmatrix}$$

donde P_1 y P_2 representan matrices de Markov que describen las transiciones dentro de dos conjuntos cerrados de estados.

La siguiente proposición es una forma simple de caracterizar la irreducibilidad de una cadena de Markov; la prueba es inmediata de la definición

Proposición 1.10 (Caracterización de las cadenas irreducibles) *Una cadena de Markov es irreducible si y sólo si todos los estados son accesibles desde cualquier otro.*

1.3.1. Periodicidad

Definición 1.11 Definimos el período del estado i, $d(i)$, como el máximo común divisor (M.C.D) de todos los enteros $n \geq 1$ para los cuales $p_{ii}^n > 0$. Si $p_{ii}^n = 0$ para todo $n \geq 1$, definimos $d(i) = 0$.

A continuación estableceremos tres propiedades básicas del período de un estado. Para una prueba remitimos al lector a [13] (pg. 82).

Teorema 1.12 *Si $i \leftrightarrow j$ entonces $d(i) = d(j)$.*

Teorema 1.13 *Si el estado i tiene período $d(i)$ entonces existe un entero N dependiente de i, tal que para todo entero $n \geq N$, $p_{ii}^{nd(i)} > 0$.*

Este teorema muestra que puede ocurrir un retorno al estado i para todos los múltiplos suficientemente grandes del período $d(i)$.

Corolario 1.14 *Si $p_{ji}^m > 0$, entonces $p_{ji}^{m+nd(i)} > 0$, para todo n(entero positivo) suficientemente grande.*

Definición 1.15 Una cadena de Markov se dice *aperiódica* cuando todos los estados tienen período uno. De lo contrario se denomina *periódica.*

1.3.2. Recurrencia

Sea $i \in \mathcal{E}$ un estado arbitrario pero fijo. Definimos para todo entero $n \geq 1$,

$$f_{ii}^n = P\{X_n = i, X_m \neq i, m = 1, 2, ..., n-1 | X_0 = i\}.$$

En otras palabras, f_{ii}^n es la probabilidad de que, comenzando desde el estado i, el primer retorno al estado i ocurra en la n-ésima transición. Claramente $f_{ii}^1 = p_{ii}$ y f_{ii}^n puede ser calculado recursivamente mediante

$$p_{ii}^n = \sum_{k=0}^{n} f_{ii}^k \, p_{ii}^{n-k}, \qquad n \geq 1, \tag{1.3}$$

donde definimos $f_{ii}^0 = 0$ para todo i.

La ecuación (1.3) proviene de la descomposición del evento en el cual p_{ii}^n es visto de acuerdo al tiempo del primer retorno al estado i. Para ello, consideremos todas las posibles realizaciones del proceso en el cual $X_0 = i, X_n = i$ y el primer retorno al estado i ocurre en la k-ésima transición. Llamemos a este evento E_k. Claramente los eventos E_k con $k = 1, 2, ...n$ son mutuamente excluyentes. La probabilidad de que el primer retorno sea en la k-ésima transición es por definición f_{ii}^k. En las transiciones restantes estamos tratando sólo con aquellas realizaciones para las cuales $X_n = i$. Usando la proposición 1.6, tenemos

$$P\{E_k\} = P\{\text{primer retorno en la } k\text{-ésima transición}|X_0 = i\} \, P\{X_n = i|X_k = i\} = f_{ii}^k \, p_{ii}^{n-k},$$

para $1 \leq k \leq n$ (recordemos que $p_{ii}^0 = 1$). Por lo tanto,

$$P\{X_n = i|X_0 = i\} = \sum_{k=1}^{n} P\{E_k\} = \sum_{k=1}^{n} f_{ii}^k \, p_{ii}^{n-k} = \sum_{k=0}^{n} f_{ii}^k \, p_{ii}^{n-k},$$

ya que por definición $f_{ii}^0 = 0$.

Definición 1.16 Un estado i se dice *recurrente* si y sólo si $\sum_{n=1}^{\infty} f_{ii}^n = 1$.

Esta definición dice que un estado i es recurrente si y sólo si, empezando desde el estado i, la probabilidad de retornar al estado i después de algún tiempo finito es uno.

Definición 1.17 Un estado j se dice *transitorio* cuando no es recurrente.

CAPÍTULO 2

Teoría Espectral

2.1. Algebras de Banach

Definición 2.1 Un *álgebra* $\mathcal{A}$ es un espacio vectorial, sobre un campo $\mathbb{K}$, tal que existe una aplicación $P : \mathcal{A} \times \mathcal{A} \longrightarrow \mathcal{A}$ llamada producto, que denotaremos mediante $P(a, b) = a \cdot b$ para todo $a, b \in \mathcal{A}$, y que verifica las siguientes propiedades para todo $a, b, c \in \mathcal{A}$, y para todo $\lambda \in \mathbb{K}$:

(i) $\lambda(a \cdot c) = (\lambda a) \cdot c$

(ii) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$

(iii) $a \cdot (b + c) = a \cdot b + a \cdot c$

(iv) $(a + b) \cdot c = a \cdot c + b \cdot c$

Definición 2.2 Un *álgebra de Banach* $\mathcal{B}$ es un álgebra sobre $\mathbb{C}$ con unidad 1, que posee una norma que lo hace un espacio de Banach y que satisface:

(a) $\|1\| = 1$,

(b) $\|fg\| \leq \|f\| \, \|g\|$ para $f, g \in \mathcal{B}$.

Notación 2.3 Es importante aclarar y que el lector tenga presente a lo largo del capítulo, que denotaremos por λ tanto al número complejo $\lambda \in \mathbb{C}$, como al elemento $\lambda = \lambda \cdot 1 \in \mathcal{B}$, donde $\mathcal{B}$ es un álgebra de Banach. Es decir, podemos tener $\lambda \in \mathbb{C}$ o $\lambda \in \mathcal{B}$ dependiendo del contexto donde se encuentre.

Veamos algunos ejemplos importantes de álgebras de Banach que usaremos a lo largo del capítulo.

Ejemplo 2.4 Sea $\mathcal{H}$ un espacio de Hilbert. El espacio de todas las transformaciones lineales acotadas de $\mathcal{H}$ en $\mathcal{H}$, denotado por $L(\mathcal{H})$, es un álgebra de Banach con la norma

$$\|T\| = \sup_{\|h\|_{\mathcal{H}}=1} \|T(h)\|_{\mathcal{H}}$$

para $T \in L(\mathcal{H})$.

Ejemplo 2.5 Sea $\mathcal{X}$ un espacio de Hausdorff compacto. El conjunto $C(\mathcal{X})$ de las funciones complejas continuas definidas en $\mathcal{X}$ es un álgebra de Banach con la norma

$$\|\varphi\|_\infty = \sup_{f\in\mathcal{X}} |\varphi(f)|$$

para $\varphi \in C(\mathcal{X})$.

2.1.1. Elementos Invertibles

Definición 2.6 Sea $\mathcal{B}$ un álgebra de Banach. Definimos $\mathcal{G}$ como la colección de los elementos invertibles de $\mathcal{B}$.

Observación 2.7 Es claro que $\mathcal{G}$ es un grupo multiplicativo y que contiene a la unidad 1.

Lo siguiente es una proposición fundamental que usaremos para probar que $\mathcal{G}$, la colección de elementos invertibles de $\mathcal{B}$, es un conjunto abierto y que la inversión es continua en la topología de la norma.

Proposición 2.8 *Sea $\mathcal{B}$ un álgebra de Banach. Si $f \in \mathcal{B}$ y $\|1 - f\| < 1$ entonces f es invertible y*

$$\|f^{-1}\| \leq \frac{1}{1 - \|1 - f\|}$$

Demostración:

Sea $\eta = \|1-f\| < 1$. Para $N \geq M$ tenemos

$$\begin{aligned}\left\|\sum_{n=0}^{N}(1-f)^n - \sum_{n=0}^{M}(1-f)^n\right\| &= \left\|\sum_{n=M+1}^{N}(1-f)^n\right\| \leq \sum_{n=M+1}^{N}\|1-f\|^n \\ &= \sum_{n=M+1}^{N}\eta^n = \frac{\eta^{M+1}-\eta^{N+1}}{1-\eta} \leq \frac{\eta^{M+1}}{1-\eta} \overset{M\to\infty}{\longrightarrow} 0.\end{aligned}$$

Luego la sucesión de sumas parciales $\{\sum_{n=0}^{N}(1-f)^n\}_{N=0}^{\infty}$ es de Cauchy, y consecuentemente, la serie converge (porque $\mathcal{B}$ es completo).

Sea $g = \sum_{n=0}^{\infty}(1-f)^n$, entonces

$$\begin{aligned}fg = [1-(1-f)]\sum_{n=0}^{\infty}(1-f)^n &= \lim_{N\to\infty}\left([1-(1-f)]\sum_{n=0}^{N}(1-f)^n\right) \\ = \lim_{N\to\infty}\left([1-(1-f)]\frac{1-(1-f)^{N+1}}{1-(1-f)}\right) &= \lim_{N\to\infty}(1-(1-f)^{N+1}) = 1\end{aligned}$$

ya que $\lim_{N\to\infty}\|(1-f)^{N+1}\| = 0$.

Análogamente $gf = 1$ y por lo tanto f es invertible, con $f^{-1} = g$.

Además,

$$\begin{aligned}\|g\| = \left\|\lim_{N\to\infty}\sum_{n=0}^{N}(1-f)^n\right\| &= \lim_{N\to\infty}\left\|\sum_{n=0}^{N}(1-f)^n\right\| \\ \leq \lim_{N\to\infty}\sum_{n=0}^{N}\|1-f\|^n &= \frac{1}{1-\|1-f\|}.\end{aligned}$$

De donde, $\|f^{-1}\| \leq \frac{1}{1-\|1-f\|}$. □

Proposición 2.9 *Si $\mathcal{B}$ es un álgebra de Banach entonces $\mathcal{G}$ es un conjunto abierto en $\mathcal{B}$.*

Demostración:

Sea $f \in \mathcal{G}$. Consideremos $B_f = \left\{g \in \mathcal{B} : \|f-g\| < \frac{1}{\|f^{-1}\|}\right\}$. Veamos que $B_f \subseteq \mathcal{G}$.

Sea $g \in B_f$. Tenemos que $\|f - g\| < \frac{1}{\|f^{-1}\|}$ y entonces

$$1 > \|f^{-1}\| \, \|f - g\| \geq \|f^{-1}(f - g)\| = \|1 - f^{-1}g\|$$

De donde, $\|1 - f^{-1}g\| < 1$. Luego, por la proposición anterior tenemos que $f^{-1}g \in \mathcal{G}$ y como $\mathcal{G}$ es un grupo $g = f(f^{-1}g) \in \mathcal{G}$.

Por lo tanto, $B_f \subseteq \mathcal{G}$. Es decir, para cada $f \in \mathcal{G}$ existe una bola abierta de radio $1/\|f^{-1}\|$ tal que está contenida en $\mathcal{G}$. Luego $\mathcal{G}$ es abierto. □

Corolario 2.10 *Si $\mathcal{B}$ es un álgebra de Banach, entonces la aplicación $f \to f^{-1}$ definida en $\mathcal{G}$ es continua. Así, $\mathcal{G}$ es un grupo topológico.*

Demostración:

Sea $f \in \mathcal{G}$. La desigualdad $\|f - g\| < 1/\,2\,\|f^{-1}\|$ implica que

$$\|1 - f^{-1}g\| < 1/2. \tag{2.1}$$

Y por la proposición 2.8 tenemos que $f^{-1}g \in \mathcal{G}$ y por lo tanto $g = f(f^{-1}g) \in \mathcal{G}$. Luego

$$\|g^{-1}\| \leq \|g^{-1}f\| \, \|f^{-1}\| = \|(f^{-1}g)^{-1}\| \, \|f^{-1}\| \leq 2\|f^{-1}\| \tag{2.2}$$

donde la última desigualdad viene dada ya que, por (2.1) y por proposición 2.8, tenemos

$$\|(f^{-1}g)^{-1}\| \leq \frac{1}{1 - \|1 - f^{-1}g\|} < 2.$$

Ahora, sea $0 < \varepsilon < \|f^{-1}\|$. Si

$$\|f - g\| < \frac{\varepsilon}{2\,\|f^{-1}\|^2} < \frac{1}{2\,\|f^{-1}\|}$$

entonces, por (2.2), tenemos

$$\begin{aligned}\|f^{-1} - g^{-1}\| &= \|f^{-1}(f - g)g^{-1}\| \leq \|f^{-1}\| \, \|f - g\| \, \|g^{-1}\| \\ &\leq \|f^{-1}\| \, \|f - g\| \, 2\,\|f^{-1}\| < 2\,\|f^{-1}\|^2 \frac{\varepsilon}{2\,\|f^{-1}\|^2} = \varepsilon\end{aligned}$$

Luego $f \to f^{-1}$ es continua. □

2.1.2. Espectro y Resolvente

Definición 2.11 Para $\mathcal{B}$ un álgebra de Banach y $f \in \mathcal{B}$ definimos el *espectro de* f como el conjunto

$$\sigma_{\mathcal{B}}(f) = \{\lambda \in \mathbb{C} : f - \lambda \text{ \underline{no} es invertible en } \mathcal{B}\}$$

y la *resolvente de* f como

$$\rho_{\mathcal{B}}(f) = \mathbb{C} - \sigma_{\mathcal{B}}(f)$$

Además, el *radio espectral de* f se define como

$$r_{\mathcal{B}}(f) = \sup\{|\lambda| : \lambda \in \sigma_{\mathcal{B}}(f)\}$$

Notación 2.12 Siempre que no se preste a confusión, omitiremos la $\mathcal{B}$ y sólo escribiremos $\sigma(f), \rho(f)$ y $r(f)$.

La siguiente proposición establece que el espectro de un elemento f es cerrado y se encuentra acotado por la circunferencia de radio igual a la norma de f en el plano complejo (Ver figura 2.1).

Proposición 2.13 *Si $\mathcal{B}$ es un álgebra de Banach y $f \in \mathcal{B}$, entonces $\sigma(f)$ es compacto y $r(f) \leq \|f\|$.*

Demostración:

Sea $J : \mathbb{C} \longrightarrow \mathcal{B}$ dada por $J(\lambda) = f - \lambda$. Así definida, J es continua. Por otro lado,

$$\lambda \in \rho(f) \Longleftrightarrow f - \lambda \text{ es invertible} \Longleftrightarrow f - \lambda \in \mathcal{G} \Longleftrightarrow J(\lambda) \in \mathcal{G} \Longleftrightarrow \lambda \in J^{-1}(\mathcal{G}),$$

de donde $\rho(f) = J^{-1}(\mathcal{G})$. Además $\rho(f)$ es abierto porque $\mathcal{G}$ es abierto y J es continua. Así, $\sigma(f)$ es cerrado.

Ahora veamos que $|\lambda| \leq \|f\|$, para todo $\lambda \in \sigma(f)$. Supongamos lo contrario, es decir, que existe $\lambda \in \sigma(f)$ tal que $|\lambda| > \|f\|$.

Si $|\lambda| > \|f\|$ entonces

$$1 > \frac{\|f\|}{|\lambda|} = \left\|\frac{f}{\lambda}\right\| = \left\|1 - \left(1 - \frac{f}{\lambda}\right)\right\|.$$

Luego, por proposición 2.8 tenemos que $1 - f/\lambda$ es invertible y por consiguiente $f - \lambda$ es invertible. Luego, $\lambda \in \rho(f)$. Contradicción.

Finalmente, $\sigma(f)$ está acotado y por lo tanto es compacto y

$$r(f) = \sup\{|\lambda| : \lambda \in \sigma(f)\} \leq \|f\|$$

$\square$

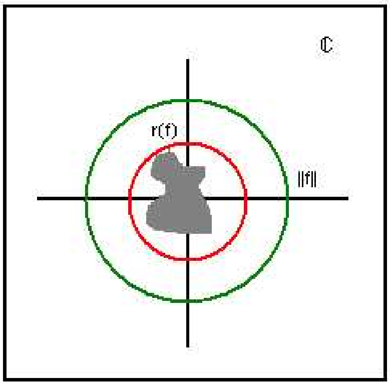

Figura 2.1: Espectro

Observación 2.14 Nótese que $0 \in \sigma(f)$ si y sólo si $f \notin \mathcal{G}$.

Teorema 2.15 *Si $\mathcal{B}$ es un álgebra de Banach y $f \in \mathcal{B}$, entonces $\sigma(f)$ es no vacío.*

Demostración:

Consideremos la función $G : \rho(f) \longrightarrow \mathcal{B}$ definida por $G(\lambda) = (f - \lambda)^{-1}$. Demostraremos que G es analítica y que es acotada. Luego usaremos el teorema de Liouville para obtener una contradicción.

Primero, como la inversión es continua (corolario 2.10), para $\lambda_0 \in \rho(f)$ tenemos que

$$\begin{aligned}
\lim_{\lambda \to \lambda_0} \frac{G(\lambda) - G(\lambda_0)}{\lambda - \lambda_0} &= \lim_{\lambda \to \lambda_0} \frac{(f-\lambda)^{-1} - (f-\lambda_0)^{-1}}{\lambda - \lambda_0} \\
&= \lim_{\lambda \to \lambda_0} \frac{[1 - (f-\lambda_0)^{-1}(f-\lambda)](f-\lambda)^{-1}}{\lambda - \lambda_0} \\
&= \lim_{\lambda \to \lambda_0} \frac{(f-\lambda_0)^{-1}[(f-\lambda_0) - (f-\lambda)](f-\lambda)^{-1}}{\lambda - \lambda_0} \\
&= \lim_{\lambda \to \lambda_0} \frac{(f-\lambda_0)^{-1}[\lambda - \lambda_0](f-\lambda)^{-1}}{\lambda - \lambda_0} \\
&= \lim_{\lambda \to \lambda_0} (f-\lambda_0)^{-1}(f-\lambda)^{-1} = (f-\lambda_0)^{-2}.
\end{aligned}$$

De donde G es analítica en $\rho(f)$. Y si $\varphi \in \mathcal{B}^*$ (espacio dual de $\mathcal{B}$), la función $\varphi(G)$ es una función analítica compleja definida en $\rho(f)$.

Además, usando proposición 2.8, tenemos que $1 - f/\lambda$ es invertible para $|\lambda| > \|f\|$ y

$$\left\| \left(1 - \frac{f}{\lambda}\right)^{-1} \right\| \leq \frac{1}{1 - \|f/\lambda\|}.$$

Luego

$$\begin{aligned}
\lim_{\lambda \to \infty} \|G(\lambda)\| = \lim_{\lambda \to \infty} \|(f-\lambda)^{-1}\| &= \lim_{\lambda \to \infty} \left\| \frac{1}{\lambda} \left(\frac{f}{\lambda} - 1 \right)^{-1} \right\| \\
&\leq \lim_{\lambda \to \infty} \frac{1}{|\lambda|} \frac{1}{1 - \|f/\lambda\|} \\
&\leq \limsup_{|\lambda| \to \infty} \frac{1}{|\lambda|} \frac{1}{1 - \|f/\lambda\|} = 0.
\end{aligned}$$

Por lo tanto, para $\varphi \in \mathcal{B}^*$ tenemos que $\lim_{\lambda \to \infty} \varphi(G(\lambda)) = 0$.

Si suponemos que $\sigma(f)$ es vacío, entonces $\rho(f) = \mathbb{C}$. Luego, para $\varphi \in \mathcal{B}^*$ tenemos que $\varphi(G)$ es una función entera que tiende a cero en el infinito. Por el teorema de Liouville (Ver [15], pg.173) tenemos que $\varphi(G) \equiv 0$.

En particular, como para $\lambda \in \mathbb{C}$ fijo tenemos que $\varphi(G(\lambda)) = 0$ para todo $\varphi \in \mathcal{B}^*$, entonces $G(\lambda) = 0$ por corolario de Hahn-Banach (Ver [9], pg. 13). Lo que es una contradicción ya que $G(\lambda)$ es, por definición, un elemento invertible de $\mathcal{B}$.

Por lo tanto $\sigma(f)$ es no vacío. $\square$

El siguiente teorema es un importante corolario del teorema anterior, el cual nos permite identificar al conjunto de múltiplos escalares de la unidad de un álgebra de Banach con el álgebra de Banach $\mathbb{C}$ de los números complejos. Este hecho es crucial en el establecimiento de las propiedades de la transformada de Gelfand, concepto que estudiaremos más adelante. Antes de enunciar el teorema recordemos una definición.

Definición 2.16 Sea $\mathcal{A}$ un álgebra. Se dice que $\mathcal{A}$ es un *álgebra con división* si para todo $a \in \mathcal{A}$ con $a \neq 0$, se tiene que a es invertible.

Teorema 2.17 (Gelfand-Mazur) *Si $\mathcal{B}$ es un álgebra de Banach con división, entonces existe un isomorfismo isométrico de $\mathcal{B}$ sobre $\mathbb{C}$.*

Demostración:

Sea $f \in \mathcal{B}$. Por el teorema anterior, $\sigma(f)$ es no vacío. Si $\lambda_0 \in \sigma(f)$, entonces $f - \lambda_0$ no es invertible por definición. Luego como $\mathcal{B}$ es un álgebra con división, $f - \lambda_0 = 0$. Además, para $\lambda \neq \lambda_0$ tenemos que $f - \lambda = \lambda_0 - \lambda$ es invertible. Así, $\sigma(f)$ consiste en exactamente un número complejo para cada $f \in \mathcal{B}$. Llamémoslo λ_f.

Consideremos la función $\psi : \mathcal{B} \to \mathbb{C}$ definida por $\psi(f) = \lambda_f$. Veamos que ψ es lineal. Sean $f, g \in \mathcal{B}$ y $\alpha \in \mathbb{C}$.

$$\psi(f+g) = \lambda_{f+g} \qquad \text{donde} \qquad f+g = \lambda_{f+g}.$$

Además,

$$\psi(f) = \lambda_f \qquad \text{donde} \qquad f = \lambda_f,$$

$$\psi(g) = \lambda_g \qquad \text{donde} \qquad g = \lambda_g.$$

Luego

$$\psi(f+g) = \lambda_{f+g} = \lambda_f + \lambda_g = \psi(f) + \psi(g.)$$

Y

$$\psi(\alpha f) = \lambda_{\alpha f} \qquad \text{donde} \qquad \alpha f = \lambda_{\alpha f}.$$

Además,

$$\psi(f) = \lambda_f \qquad \text{donde} \qquad f = \lambda_f.$$

Luego

$$\psi(\alpha\, f) = \lambda_{\alpha f} = \alpha \lambda_f = \alpha\, \psi(f).$$

Por lo tanto ψ es lineal.

Veamos que ψ es isométrica.

$$|\psi(f)| = |\lambda_f| = \|\lambda_f\| = \|f\|,$$

ya que $f = \lambda_f$. Luego ψ es isométrica.

De la linealidad y la isometría tenemos la inyectividad de ψ. Falta probar la sobreyectividad.

Sea $\lambda \in \mathbb{C}$ y sea $f = \lambda \cdot 1$, es decir, $f = \lambda \in \mathcal{B}$ (Recordar notación 2.3). Tenemos

$$f - \lambda = \lambda - \lambda = 0$$

Luego $\lambda \in \sigma(f)$. Pero $\sigma(f)$ está formado por exactamente un número complejo, entonces $\sigma(f) = \{\lambda\}$ y por lo tanto $\psi(f) = \lambda$. De donde ψ es sobreyectiva.

Hemos probado que ψ es un isomorfismo isométrico de $\mathcal{B}$ sobre $\mathbb{C}$. $\square$

2.1.3. Funcionales Lineales Multiplicativos

A continuación presentamos un concepto importante para el estudio de las álgebras de Banach.

Definición 2.18 Sea $\mathcal{B}$ un álgebra de Banach. Un funcional lineal complejo φ definido en $\mathcal{B}$ se dice *multiplicativo* si:

1. $\varphi(fg) = \varphi(f)\,\varphi(g)$, para $f, g \in \mathcal{B}$
2. $\varphi(1) = 1$

Notación 2.19 Al conjunto de todos los funcionales lineales multiplicativos definidos en $\mathcal{B}$ lo denotaremos por $\mathbf{M} = \mathbf{M}_{\mathcal{B}}$. A este conjunto también se le denomina *espacio de Gelfand de $\mathcal{B}$*.

Demostraremos que los elementos de $\mathbf{M}$ están acotados y que $\mathbf{M}$ es un subconjunto w^*-compacto de $(\mathcal{B}^*)_1$, la bola unitaria del espacio dual de $\mathcal{B}$.

Proposición 2.20 *Si $\mathcal{B}$ es un álgebra de Banach y $\varphi \in \mathbf{M}$ entonces $\|\varphi\| = 1$.*

Demostración:

Sea $K = \ker \varphi = \{f \in \mathcal{B} : \varphi(f) = 0\}$.

Como $\varphi(g - \varphi(g) \cdot 1) = 0$ para todo $g \in \mathcal{B}$, entonces $g - \varphi(g) \cdot 1 \in K$. Sea $f = g - \varphi(g) \cdot 1$, luego $g = f + \lambda$, con $\lambda = \varphi(g) \cdot 1$. De donde, todo elemento de $\mathcal{B}$ se puede escribir de la forma $\lambda + f$ para algún $\lambda \in \mathbb{C}$ y $f \in K$.

Así,

$$\begin{aligned}\|\varphi\| &= \sup_{g\neq 0}\frac{|\varphi(g)|}{\|g\|} = \sup_{f\in K,\,\lambda\neq 0}\frac{|\varphi(f+\lambda)|}{\|f+\lambda\|} = \sup_{f\in K,\,\lambda\neq 0}\frac{|\varphi(f)+\lambda\,\varphi(1)|}{\|f+\lambda\|}\\ &= \sup_{f\in K,\,\lambda\neq 0}\frac{|\lambda|}{|\lambda|\|\frac{f}{\lambda}+1\|} = \sup_{h\in K}\frac{1}{\|h+1\|} = 1.\end{aligned}$$

Demostremos la última igualdad.

Si $h = 0$ entonces $\frac{1}{\|h+1\|} = 1$ y por lo tanto $\|\varphi\| = \sup_{h\in K}\frac{1}{\|h+1\|} \geq 1$. Queremos ver que $\|\varphi\| = 1$.

Supongamos que $\|\varphi\| > 1$. Entonces $\sup_{h\in K}\frac{1}{\|h+1\|} > 1$, es decir, existe $h \in K$ tal que $\frac{1}{\|h+1\|} > 1$. Luego, $\|h+1\| < 1$, y usando proposición 2.8 tenemos que h es invertible, y por lo tanto, $h \notin K$. Contradicción. □

Proposición 2.21 *Si $\mathcal{B}$ es un álgebra de Banach entonces $\mathbf{M}$ es un subconjunto w^*-compacto de $(\mathcal{B}^*)_1$.*

Demostración:

Por proposición anterior sabemos que $\mathbf{M} \subseteq (\mathcal{B}^*)_1$. Además, $(\mathcal{B}^*)_1$ es w^*-compacto por teorema de Alaoglu (Ver [18], pg.68). Entonces basta probar que $\mathbf{M}$ es cerrado.

Sea $\{\varphi_n\}_{n=0}^{\infty} \subseteq \mathbf{M}$ una sucesión que converge a $\varphi \in (\mathcal{B}^*)_1$ en la topología-w^* de $(\mathcal{B}^*)_1$. Queremos ver que φ es multiplicativo, i.e, $\varphi \in \mathbf{M}$.

Tenemos

$$\varphi(1) = \lim_{n\to\infty}\varphi_n(1) = \lim_{n\to\infty} 1 = 1$$

Además, para $f, g \in \mathcal{B}$, tenemos

$$\begin{aligned}\varphi(fg) &= \lim_{n\to\infty}\varphi_n(fg) = \lim_{n\to\infty}\varphi_n(f)\,\varphi_n(g)\\ &= \lim_{n\to\infty}\varphi_n(f)\lim_{n\to\infty}\varphi_n(g) = \varphi(f)\,\varphi(g)\end{aligned}$$

Luego $\varphi \in \mathbf{M}$ y por lo tanto $\mathbf{M}$ es cerrado. □

Ahora recordemos algunas nociones del álgebra que nos permitirán establecer una proposición mediante la cual podremos tratar a $\mathbf{M}$ como un objeto diferente: el espacio de ideales maximales de $\mathcal{B}$.

Definición 2.22 Un subconjunto no vacío $\mathcal{I}$ de un álgebra $\mathcal{A}$ se dice que es un *ideal* (bilateral) de $\mathcal{A}$ si:

1. $\mathcal{I}$ es un subgrupo de $\mathcal{A}$ bajo la adición.

2. $ar, ra \in \mathcal{I}$, para todo $r \in \mathcal{I}$ y $a \in \mathcal{A}$.

Además, suponiendo que $\mathcal{I} \neq \mathcal{A}$, se dice que $\mathcal{I}$ es un *ideal maximal* de $\mathcal{A}$ si siempre que $\mathcal{I}'$ sea un ideal de $\mathcal{A}$ tal que $\mathcal{I} \subseteq \mathcal{I}' \subseteq \mathcal{A}$ se tiene que $\mathcal{I}' = \mathcal{A}$ ó $\mathcal{I}' = \mathcal{I}$.

Observación 2.23 Si $\mathcal{I}$ es un ideal maximal de $\mathcal{A}$, $\mathcal{A}$ es un álgebra con unidad 1 y $r \in \mathcal{I}$, entonces r no es invertible, ya que si lo fuera tendríamos que $1 = rr^{-1} \in \mathcal{I}$ y entonces $\mathcal{I} = \mathcal{A}$, lo cual es una contradicción debido a que $\mathcal{I} \neq \mathcal{A}$ por definición.

Definición 2.24 Sea $\mathcal{X}$ un espacio normado y sea $D \subset \mathcal{X}$ un subespacio cerrado. Definimos el *espacio cociente* $\mathcal{X}/D$ como

$$\mathcal{X}/D = \{x + D : x \in \mathcal{X}\}.$$

Es usual la notación $[x] = x + D$. Y para $[x] \in \mathcal{X}/D$ definimos

$$\|[x]\|_{\mathcal{X}/D} = \mathrm{dist}(x, D) = \inf_{d \in D} \|x + d\|.$$

Observación 2.25 $\mathcal{X}/D$ es un espacio vectorial y $\|\cdot\|_{\mathcal{X}/D}$ define una norma en él. Además, si $\mathcal{X}$ es un espacio de Banach entonces $\mathcal{X}/D$ es un espacio de Banach con dicha norma.

Habiendo recordado las definiciones anteriores veamos algunos preliminares para la proposición siguiente.

Sea $\mathcal{B}$ un álgebra de Banach y supongamos que $\mathcal{I}$ es un ideal (bilateral) propio y cerrado de $\mathcal{B}$. Como $\mathcal{I}$ es un subespacio cerrado de $\mathcal{B}$ podemos definir la norma usual en el álgebra cociente $\mathcal{B}/\mathcal{I}$ que lo hace un espacio de Banach. Además, como $\mathcal{I}$ es un ideal en $\mathcal{B}$ se puede ver que $\mathcal{B}/\mathcal{I}$ es un álgebra con el producto $[f] \cdot [g] = [f \cdot g]$.

Sólo faltaría probar dos hechos para asegurar que $\mathcal{B}/\mathcal{I}$ es un álgebra de Banach. Primero, debemos demostrar que $\|[1]\| = 1$. Veamos.

Como $0 \in \mathcal{I}$ tenemos que $\inf_{g\in\mathcal{I}} \|1-g\| \leq 1$. Supongamos que $\inf_{g\in\mathcal{I}} \|1-g\| < 1$. Entonces por proposición 2.8 tenemos que $g \in \mathcal{G}$, de donde $g \cdot g^{-1} = 1 \in \mathcal{I}$. Contradicción, ya que si $1 \in \mathcal{I}$ entonces $\mathcal{I} = \mathcal{B}$.

Luego

$$\|[1]\| = \inf_{g\in\mathcal{I}} \|1 - g\| = 1.$$

Segundo, para $f, g \in \mathcal{B}$ tenemos

$$\begin{aligned} \|[f][g]\| = \|[fg]\| = \inf_{h\in\mathcal{I}} \|fg - h\| &\leq \inf_{h_1,h_2\in\mathcal{I}} \|(f - h_1)(g - h_2)\| \\ &\leq \inf_{h_1\in\mathcal{I}} \|f - h_1\| \inf_{h_2\in\mathcal{I}} \|g - h_2\| = \|[f]\| \, \|[g]\| \end{aligned}$$

de donde $\|[f][g]\| \leq \|[f]\| \, \|[g]\|$.

Por lo tanto $\mathcal{B}/\mathcal{I}$ es un álgebra de Banach.

Proposición 2.26 *Si $\mathcal{B}$ es un álgebra de Banach conmutativa, entonces el conjunto* $\mathbf{M}$ *de los funcionales lineales multiplicativos está en correspondencia 1-1 con el conjunto de ideales (bilaterales) maximales de $\mathcal{B}$.*

Demostración:

Sea φ un funcional lineal multiplicativo definido en $\mathcal{B}$. Sea

$$K = \ker \varphi = \{f \in \mathcal{B} : \varphi(f) = 0\}.$$

El núcleo K de un homomorfismo es un ideal propio y si $f \notin K$, entonces

$$1 = \left(1 - \frac{f}{\varphi(f)}\right) + \frac{f}{\varphi(f)}.$$

Y como $1 - \frac{f}{\varphi(f)} \in K$, entonces el espacio generado por K y f contiene a la unidad. Así, un ideal que contiene a K y a f tendría que ser todo $\mathcal{B}$ y por lo tanto K es un ideal maximal.

Supongamos que $\mathcal{I}$ es un ideal maximal en $\mathcal{B}$. Como cada elemento $f \in \mathcal{I}$ es no invertible (ver Obs.2.23) entonces, por proposición 2.8, tenemos que $\|1 - f\| \geq 1$. Así, 1 no es punto límite de $\mathcal{I}$, y por lo tanto, no está en la clausura de $\mathcal{I}$, es decir, $1 \notin \overline{\mathcal{I}}$. Además, como $\overline{\mathcal{I}}$ es un ideal y $\mathcal{I} \subseteq \overline{\mathcal{I}} \subsetneq \mathcal{B}$ (porque $1 \notin \overline{\mathcal{I}}$), entonces $\mathcal{I} = \overline{\mathcal{I}}$ (porque $\mathcal{I}$ es maximal). Por lo tanto $\mathcal{I}$ es cerrado.

El álgebra cociente $\mathcal{B}/\mathcal{I}$ es un álgebra de Banach que, como $\mathcal{I}$ es maximal y $\mathcal{B}$ es conmutativo, es un álgebra con división. Luego por el teorema de Gelfand-Mazur (2.17), existe un isomorfismo isométrico $\psi : \mathcal{B}/\mathcal{I} \longrightarrow \mathbb{C}$. Si π denota al isomorfismo natural de $\mathcal{B}$ sobre $\mathcal{B}/\mathcal{I}$, entonces la composición $\varphi = \psi \circ \pi$ es un funcional lineal multiplicativo distinto de cero definido en $\mathcal{B}$. Veámoslo.

Sean $f, g \in \mathcal{B}$ y $\alpha \in \mathbb{C}$. Tenemos

- $$\begin{aligned}\varphi(\alpha f + g) &= \psi \circ \pi(\alpha f + g) = \psi([\alpha f + g]) = \psi(\alpha[f] + [g]) \\ &= \alpha\, \psi([f]) + \psi([g]) = \alpha\, \psi \circ \pi(f) + \psi \circ \pi(g) \\ &= \alpha\, \varphi(f) + \varphi(g)\end{aligned}$$

- $$\varphi(fg) = \psi \circ \pi(fg) = \psi([fg]) = \psi([f][g]) = \lambda_{[f][g]}$$

  donde $[f][g] - \lambda_{[f][g]} = 0$.

  Además,

  $$\begin{aligned}\psi([f]) &= \lambda_{[f]} \quad \text{donde} \quad [f] = \lambda_{[f]}, \\ \psi([g]) &= \lambda_{[g]} \quad \text{donde} \quad [g] = \lambda_{[g]}.\end{aligned}$$

  Luego

  $$\begin{aligned}\varphi(fg) &= \psi([f][g]) = \lambda_{[f][g]} = \lambda_{[f]}\lambda_{[g]} = \psi([f])\, \psi([g]) \\ &= \psi \circ \pi(f)\, \psi \circ \pi(g) = \varphi(f)\, \varphi(g).\end{aligned}$$

Hemos probado que $\varphi \in \mathbf{M}$. Entonces, $K = \ker \varphi$ es un ideal de $\mathcal{B}$.

Si $r \in \mathcal{I}$ tenemos

$$\varphi(r) = \psi \circ \pi(r) = \psi([r]) = \psi([0]) = 0,$$

ya que $0 \in \sigma([0])$. De donde $h \in K = \ker \varphi$. Luego, $\mathcal{I} \subseteq K$, y como $\mathcal{I}$ es maximal tenemos que

$$\mathcal{I} = K = \ker \varphi.$$

Finalmente queremos probar que la correspondencia $\varphi \to \ker \varphi$ es 1-1.

Sean $\varphi_1, \varphi_2 \in \mathbf{M}$ tales que $\ker \varphi_1 = \ker \varphi_2 = \mathcal{I}$. Tenemos que

$$f - \varphi_1(f) \in \ker \varphi_1 = \mathcal{I} \quad \text{y} \quad f - \varphi_2(f) \in \ker \varphi_2 = \mathcal{I},$$

entonces $(f - \varphi_2(f)) - (f - \varphi_1(f)) \in \mathcal{I}$, ya que $\mathcal{I}$ es un ideal. Por otro lado,

$$\varphi_1(f) - \varphi_2(f) \in \{\lambda,1 : \lambda \in \mathbb{C}\}.$$

Pero

$$\varphi_1(f) - \varphi_2(f) = (f - \varphi_2(f)) - (f - \varphi_1(f)).$$

Así, $\varphi_1(f) - \varphi_2(f)$ está tanto en $\mathcal{I}$ como en $\{\lambda,1 : \lambda \in \mathbb{C}\}$ para cada $f \in \mathcal{B}$. Como cada elemento en $\mathcal{I}$ es no invertible y el único elemento no invertible de $\{\lambda,1 : \lambda \in \mathbb{C}\}$ es el cero, tenemos que

$$\varphi_1(f) - \varphi_2(f) = 0.$$

Luego $\varphi_1 = \varphi_2$. De donde $\varphi \to \ker \varphi$ es 1-1. □

Observación 2.27 Dada la proposición anterior, de aquí en adelante nos referiremos, de igual forma, a $\mathbf{M}_{\mathcal{B}}$ como el espacio de ideales maximales de $\mathcal{B}$ o como el conjunto de funcionales lineales multiplicativos.

2.1.4. Transformada de Gelfand

Primero veamos una definición que nos permitirá establecer una herramienta muy útil en el estudio de álgebras de Banach: la transformada de Gelfand.

Definición 2.28 Sea $\mathcal{X}$ un espacio de Banach y $\mathcal{X}^*$ el espacio dual. Para cada $f \in \mathcal{X}$ definimos la función $\hat{f} : \mathcal{X}^* \longrightarrow \mathbb{C}$ por $\hat{f}(\varphi) = \varphi(f)$, para $\varphi \in \mathcal{X}^*$.

Observación 2.29 La *topología débil** (o topología-w^*) en $\mathcal{X}^*$ es la topología débil en $\mathcal{X}^*$ inducida por la familia de funciones $\{\hat{f} : f \in \mathcal{X}\}$.

Definición 2.30 Sea $\mathcal{B}$ un álgebra de Banach, $\mathbf{M}$ el conjunto de funcionales lineales multiplicativos en $\mathcal{B}$ y $C(\mathbf{M})$ el espacio de las funciones complejas continuas definidas en $\mathbf{M}$. La *tranformada de Gelfand* es la función $\Gamma : \mathcal{B} \longrightarrow C(\mathbf{M})$ dada por

$$\Gamma(f) = \hat{f}|_{\mathbf{M}}$$

que es, $\Gamma(f)(\varphi) = \varphi(f)$, para $\varphi \in \mathbf{M}$.

Propiedades elementales de la transformada de Gelfand

La siguiente proposición establece que Γ es un operador lineal y continuo, y que $\Gamma(fg) = \Gamma(f)\,\Gamma(g)$, para $f, g \in \mathcal{B}$.

Proposición 2.31 *Si $\mathcal{B}$ es un álgebra de Banach y Γ es la transformada de Gelfand en $\mathcal{B}$, entonces:*

1. *Γ es un homomorfismo de álgebra.*

2. *$\|\Gamma f\|_\infty \leq \|f\|$.*

Demostración:

1. Sean $f, g \in \mathcal{B}$, $\quad \varphi \in \mathbf{M} \subset (\mathcal{B}^*)_1$, $\quad \lambda \in \mathbb{C}$. Tenemos

   - $\Gamma(f+g)(\varphi) = \varphi(f+g) = \varphi(f) + \varphi(g) = \Gamma(f)(\varphi) + \Gamma(g)(\varphi)$,
   - $\Gamma(\lambda f)(\varphi) = \varphi(\lambda f) = \lambda\varphi(f) = \lambda\Gamma(f)(\varphi)$, y
   - $\Gamma(fg)(\varphi) = \varphi(fg) = \varphi(f)\varphi(g) = \Gamma(f)(\varphi)\Gamma(g)(\varphi)$,

Por lo tanto, Γ es un homomorfismo de álgebra.

2. Sea $f \in \mathcal{B}$.

$$\begin{aligned}\|\Gamma f\|_\infty = \|\hat{f}|_{\mathbf{M}}\|_\infty &= \sup_{\varphi\in\mathbf{M}} |\hat{f}(\varphi)| \\ &= \sup_{\varphi\in\mathbf{M}} |\varphi(f)| \leq \sup_{\varphi\in\mathbf{M}} \|\varphi\|\,\|f\| = \|f\|\end{aligned}$$

De donde $\|\Gamma f\|_\infty \leq \|f\|$.

$\square$

Observación 2.32 Nótese que Γ manda a todos los elementos de la forma $fg - gf$ a 0. Así, si $\mathcal{B}$ no es conmutativo, entonces la subálgebra de $C(\mathbf{M})$ que es el rango de Γ puede no reflejar las propiedades de $\mathcal{B}$ (en particular, existen ejemplos de álgebras de Banach para las cuales $\mathbf{M}$ es vacío). Sin embargo, en el caso conmutativo, $\mathbf{M}$ no es sólo no vacío, si no que es suficientemente grande como para que la invertibilidad de un elemento $f \in \mathcal{B}$ sea determinada por la invertibilidad de $\Gamma f \in C(\mathbf{M})$. Este hecho hace de la transformada de Gelfand una herramienta poderosa para el estudio de las álgebras de Banach conmutativas.

Proposición 2.33 *Si $\mathcal{B}$ es un álgebra de Banach conmutativa y $f \in \mathcal{B}$, entonces f es invertible en $\mathcal{B}$ si y sólo si $\Gamma(f)$ es invertible en $\mathcal{C}(\mathbf{M})$.*

Demostración:

($\Longrightarrow$) Si f es invertible en $\mathcal{B}$ entonces $\Gamma(f^{-1})$ es la inversa de $\Gamma(f)$, ya que

$$\Gamma(f)(\varphi)\Gamma(f^{-1})(\varphi) = \varphi(f)\varphi(f^{-1}) = \varphi(1) = 1$$

con $\varphi \in \mathbf{M}$.

($\Longleftarrow$) Si f no es invertible en $\mathcal{B}$, entonces $\mathcal{I}_0 = \{gf : g \in \mathcal{B}\}$ es un ideal propio en $\mathcal{B}$ ya que $1 \notin \mathcal{I}_0$. Como $\mathcal{B}$ es conmutativo, $\mathcal{I}_0$ está contenido en algún ideal maximal $\mathcal{I}$. Por la proposición 2.26 existe $\varphi \in \mathbf{M}$ tal que $\ker\varphi = \mathcal{I}$. Así, $\Gamma(f)(\varphi) = \varphi(f) = 0$ (porque $f \in \mathcal{I}_0 \subset \mathcal{I} \Rightarrow f \in \mathcal{I} = \ker\varphi$). Por lo tanto $\Gamma(f)$ no es invertible en $C(\mathbf{M})$. $\square$

El siguiente teorema resume los resultados anteriores para el caso conmutativo.

Teorema 2.34 (Gelfand) *Si $\mathcal{B}$ es un álgebra de Banach conmutativa,* $\mathbf{M}$ *es su espacio de ideales maximales, y* $\Gamma : \mathcal{B} \longrightarrow \mathcal{C}(\mathbf{M})$ *es la transformada de Gelfand, entonces:*

1. $\mathbf{M}$ *es no vacío;*
2. Γ *es un homomorfismo de álgebra;*
3. $\|\Gamma f\|_\infty \leq \|f\|$, *para* $f \in \mathcal{B}$; *y*
4. f *es invertible en* $\mathcal{B} \Longleftrightarrow \Gamma(f)$ *es invertible en* $\mathcal{C}(\mathbf{B})$.

Corolario 2.35 *Si $\mathcal{B}$ es un álgebra de Banach conmutativa y $f \in \mathcal{B}$, entonces*

$$\sigma(f) = \operatorname{rang} \Gamma f \quad \text{ y } \quad r(f) = \|\Gamma f\|_\infty.$$

Demostración:

($\supseteq$) Si $\lambda \notin \sigma(f)$ entonces $f - \lambda$ es invertible en $\mathcal{B}$ por definición. Esto implica que $\Gamma(f) - \lambda$ es invertible en $C(\mathbf{M})$ (por teorema de Gelfand (2.34)). Lo que implica que $(\Gamma f - \lambda)(\varphi) \neq 0$ para todo $\varphi \in \mathbf{M}$. Así, $(\Gamma f)(\varphi) \neq \lambda$ para todo $\varphi \in \mathbf{M}$. Por lo tanto $\lambda \notin \operatorname{rang} \Gamma f$.

De donde, $\operatorname{rang} \Gamma f \subseteq \sigma(f)$.

($\subseteq$) Si $\lambda \notin \operatorname{rang} \Gamma f$ entonces $(\Gamma f - \lambda)(\varphi) \neq 0$ para todo $\varphi \in \mathbf{M}$. Luego $\Gamma f - \lambda$ es invertible en $C(\mathbf{M})$ y usando el teorema de Gelfand (2.34) tenemos que $f - \lambda$ es invertible en $\mathcal{B}$. Por lo tanto, $\lambda \notin \sigma(f)$.

De donde, $\sigma(f) \subseteq \operatorname{rang} \Gamma f$.

Y ahora, habiendo probado que $\sigma(f) = \operatorname{rang} \Gamma f$, obtenemos

$$\begin{aligned} r(f) &= \sup\{|\lambda| : \lambda \in \sigma(f)\} = \sup\{|\lambda| : \lambda \in \operatorname{rang} \Gamma f\} \\ &= \sup\{|(\Gamma f)(\varphi)| : \varphi \in \mathbf{M}\} = \|\Gamma f\|_\infty. \end{aligned}$$

$\square$

Si $h(z) = \sum_{n=0}^{\infty} a_n z^n$ es una función entera con coeficientes complejos y f es un elemento del álgebra de Banach $\mathcal{B}$, entonces $h(f)$ denota el elemento $\sum_{n=0}^{\infty} a_n f^n$ de $\mathcal{B}$.

Corolario 2.36 (Teorema de la aplicación espectral) *Si $\mathcal{B}$ es un álgebra de Banach, $f \in \mathcal{B}$ y h es una función entera definida en $\mathbb{C}$, entonces*

$$\sigma(h(f)) = h(\sigma(f)) = \{h(\lambda) : \lambda \in \sigma(f)\}.$$

Demostración:

Si $h(z) = \sum_{n=0}^{\infty} a_n z^n$ es el desarrollo en serie de Taylor de h, entonces se puede ver que $h(f) = \sum_{n=0}^{\infty} a_n f^n$ converge a un elemento de $\mathcal{B}$.

Ahora, si $\mathcal{B}_0$ es la subálgebra generada por 1, f y los elementos de la forma $(f-\lambda)^{-1}$ y $(h(f)-\mu)^{-1}$ para $\lambda \in \rho(f)$ y $\mu \in \rho(h(f))$ respectivamente, entonces $\mathcal{B}_0$ es conmutativo y,

$$\sigma_{\mathcal{B}}(f) = \sigma_{\mathcal{B}_0}(f) \qquad \text{y} \qquad \sigma_{\mathcal{B}}(h(f)) = \sigma_{\mathcal{B}_0}(h(f)).$$

Estas dos últimas igualdades vienen dadas ya que si $\lambda \in \rho_{\mathcal{B}}(f)$ entonces $f-\lambda$ es invertible en $\mathcal{B}$, pero $(f-\lambda)^{-1} \in \mathcal{B}_0$, de donde $f-\lambda$ es invertible en $\mathcal{B}_0$ y por lo tanto $\lambda \in \rho_{\mathcal{B}_0}(f)$; lo cual demuestra que $\sigma_{\mathcal{B}_0}(f) \subseteq \sigma_{\mathcal{B}}(f)$, la otra inclusión es obvia ya que $\mathcal{B}_0 \subseteq \mathcal{B}$.

Así podemos asumir que $\mathcal{B}$ es conmutativo y usar la transformada de Gelfand Γ.

Como Γ es un homomorfismo de álgebra y un operador continuo (proposición 2.31), tenemos

$$\begin{aligned}\Gamma(h(f)) = \Gamma\left(\sum_{n=0}^{\infty} a_n f^n\right) &= \sum_{n=0}^{\infty} a_n \Gamma(f^n) \\ &= \sum_{n=0}^{\infty} a_n (\Gamma f)^n = h(\Gamma f),\end{aligned}$$

de donde, $\Gamma(h(f)) = h(\Gamma f)$. Luego, usando el corolario anterior

$$\sigma(h(f)) = \text{rang}\Gamma(h(f)) = \text{rang}h(\Gamma(f)) = h(\text{rang}\Gamma f) = h(\sigma(f)).$$

De donde, $\sigma(h(f)) = h(\sigma(f))$. $\square$

Ahora podemos probar un resultado básico, en el estudio de las álgebras de Banach, que se debe a Beurling y Gelfand relacionado con el radio espectral y la norma.

Teorema 2.37 (Fórmula de Gelfand) *Si $\mathcal{B}$ es un álgebra de Banach y $f \in \mathcal{B}$ entonces*

$$r(f) = \lim_{n\to\infty} \|f^n\|^{1/n}.$$

Demostración:

Sea $\mathcal{B}_0$ la subálgebra cerrada de $\mathcal{B}$, generada por $1, f$ y $\{(f^n - \lambda)^{-1} : \lambda \in \rho_{\mathcal{B}}(f^n), n \in \mathbb{Z}^+\}$. Así, $\mathcal{B}_0$ es conmutativa y $\sigma_{\mathcal{B}_0}(f^n) = \sigma_{\mathcal{B}}(f^n)$ para todo $n \in \mathbb{Z}^+$, como vimos en la prueba del corolario anterior. Además, por el mismo corolario tenemos que $\sigma_{\mathcal{B}_0}(f^n) = (\sigma_{\mathcal{B}_0}(f))^n$.

Luego

$$(r_{\mathcal{B}}(f))^n = (r_{\mathcal{B}_0}(f))^n = r_{\mathcal{B}_0}(f^n) = r_{\mathcal{B}}(f^n) \leq \|f^n\|,$$

de donde, $r_{\mathcal{B}}(f) \leq \|f^n\|^{1/n}$ y por lo tanto

$$r_{\mathcal{B}}(f) \leq \liminf_{n\to\infty} \|f^n\|^{1/n}.$$

Ahora, consideremos la función definida en $\rho(f)$ dada por

$$G(\lambda) = (f - \lambda)^{-1}.$$

En la demostración del teorema 2.15 vimos que G es analítica en $\rho(f)$.

Sea $\varphi \in \mathcal{B}^*$, entonces $\varphi(G(\lambda))$ también es analítica en $\rho(f)$, y en particular, es analítica en $\{\lambda \in \mathbb{C} : |\lambda| > r(f)\} \subseteq \rho(f)$.

Por otro lado, para $|\lambda| > \|f\|$

$$\begin{aligned} G(\lambda) = (f-\lambda)^{-1} &= \lambda^{-1}\left(\frac{f}{\lambda} - 1\right)^{-1} \\ &= -\lambda^{-1}\left(1 - \frac{f}{\lambda}\right)^{-1} = -\lambda^{-1}\sum_{n=0}^{\infty}\frac{f^n}{\lambda^n} \\ &= -\lambda^{-1}\left[1 + \sum_{n=1}^{\infty}\frac{f^n}{\lambda^n}\right] \end{aligned}$$

entonces

$$\varphi(G(\lambda)) = -\lambda^{-1}\left[\varphi(1) + \sum_{n=1}^{\infty}\frac{\varphi(f^n)}{\lambda^n}\right]$$

que es el desarrollo en serie de Laurent de $\varphi(G(\lambda))$ en el anillo $\{\lambda \in \mathbb{C} : |\lambda| > \|f\|\}$. Pero del análisis complejo sabemos que este desarrollo también es válido en $\{\lambda \in \mathbb{C} : |\lambda| > r(f)\}$.

Como la serie converge, los términos están acotados, es decir, existe $c_\varphi > 0$ tal que

$$\left|\varphi\left(\frac{f^n}{\lambda^n}\right)\right| \leq c_\varphi$$

para cada $\varphi \in \mathcal{B}^*$.

Entonces, por un corolario del Principio de Acotación Uniforme (ver [20], Teorema B, pg.240) tenemos que existe $k > 0$ tal que

$$\|f^n/\lambda^n\| \leq k, \quad \text{para todo } n \in \mathbb{N}.$$

Luego, $\|f^n\|^{1/n} \leq k^{1/n}|\lambda|$ y entonces

$$\limsup_{n\to\infty} \|f^n\|^{1/n} \leq |\lambda|, \quad \text{para } |\lambda| > r(f).$$

De donde

$$\limsup_{n\to\infty} \|f^n\|^{1/n} \leq r(f).$$

Finalmente,

$$r(f) \geq \limsup_{n\to\infty} \|f^n\|^{1/n} \geq \liminf_{n\to\infty} \|f^n\|^{1/n} \geq r(f).$$

Y por lo tanto, $r(f) = \lim_{n\to\infty} \|f^n\|^{1/n}$. $\square$

Corolario 2.38 *Si $\mathcal{B}$ es un álgebra de Banach conmutativa, entonces la transformada de Gelfand es una isometría si y sólo si $\|f^2\| = \|f\|^2$ para todo $f \in \mathcal{B}$.*

Demostración:

Usaremos los dos corolarios del teorema de Gelfand. Por corolario 2.35 tenemos que $r(f) = \|\Gamma f\|_\infty$ para todo $f \in \mathcal{B}$. Entonces Γ es una isometría si y sólo si $r(f) = \|f\|$ para todo $f \in \mathcal{B}$. Probemos que $r(f) = \|f\|$ para todo $f \in \mathcal{B}$ si y sólo si $\|f^2\| = \|f\|^2$ para todo $f \in \mathcal{B}$.

($\Longrightarrow$) Si $r(f) = \|f\|$ entonces $\|f^2\| = r(f^2) = (r(f))^2 = \|f\|^2$, por corolario 2.36. Luego $\|f^2\| = \|f\|^2$ para todo $f \in \mathcal{B}$.

($\Longleftarrow$) Si $\|f^2\| = \|f\|^2$ para todo $f \in \mathcal{B}$ tenemos

$$\|f^4\| = \|(f^2)^2\| = \|f^2\|^2 = \|f\|^4$$

y en general, tenemos $\|f^{2^k}\| = \|f\|^{2^k}$ para todo $k \in \mathbb{N}$. Ahora por la fórmula de Gelfand tenemos

$$r(f) = \lim_{n\to\infty} \|f^n\|^{1/n} = \lim_{k\to\infty} \|f^{2^k}\|^{1/2^k} = \lim_{n\to\infty} \|f\| = \|f\|,$$

de donde $r(f) = \|f\|$ para todo $f \in \mathcal{B}$. □

2.2. Operadores en Espacios de Hilbert y Algebras-C*

Adoptaremos el término *operador* para designar una transformación lineal acotada. Así, el álgebra de Banach $L(\mathcal{H})$ (ver ejemplo 2.4) es el conjunto de todos los operadores de $\mathcal{H}$ en $\mathcal{H}$, cuyos elementos son el objeto de estudio de la primera parte de esta sección. Luego, en la segunda parte, estudiaremos las álgebras-C^* para finalmente establecer el teorema espectral para operadores normales, que es uno de los objetivos principales de este capítulo.

2.2.1. Operadores

Comenzaremos con una proposición que permite la definición del adjunto de un operador, concepto muy importante que estableceremos luego de la prueba de dicha proposición.

Proposición 2.39 *Si T es un operador en el espacio de Hilbert $\mathcal{H}$ entonces existe un único operador S en $\mathcal{H}$ tal que*

$$\langle Tf, g\rangle = \langle f, Sg\rangle$$

para $f, g \in \mathcal{H}$.

Demostración:

Para $g \in \mathcal{H}$ fijo consideremos el funcional φ definido por $\varphi(f) = \langle Tf, g\rangle$ para $f \in \mathcal{H}$.

Veamos que φ es lineal. Sean $f, h \in \mathcal{H}$, $\lambda \in \mathbb{C}$.

$$\begin{aligned}\varphi(\lambda f + h) &= \langle T(\lambda f + h), g\rangle = \langle \lambda T(f) + T(h), g\rangle \\ &= \lambda\langle Tf, g\rangle + \langle Th, g\rangle = \lambda\varphi(f) + \varphi(h).\end{aligned}$$

Por lo tanto, φ es lineal.

Veamos que φ es acotado. Por la desigualdad de Cauchy-Schwartz y por ser T acotado tenemos

$$|\varphi(f)| = |\langle Tf, g\rangle| \leq \|Tf\|\,\|g\| \leq \|T\|\,\|f\|\,\|g\|$$

Por lo tanto $|\varphi(f)| \leq \|T\|\,\|g\|\,\|f\|$. De donde φ es acotado.

Luego por el teorema de representación de Riesz (Ver [2],pg.209) existe un único $h \in \mathcal{H}$ tal que $\varphi(f) = \langle f, h\rangle$ para $f \in \mathcal{H}$. Definamos $Sg = h$.

Veamos que S es lineal. Sean $g_1, g_2 \in \mathcal{H}$. Sean $h_1, h_2, h \in \mathcal{H}$ los elementos dados por el teorema de representación de Riesz tales que

$$Sg_1 = h_1, \qquad Sg_2 = h_2, \qquad S(g_1 + g_2) = h.$$

Queremos ver que $h = h_1 + h_2$. Tenemos

$$\varphi(f) = \langle f, h\rangle = \langle Tf, g_1 + g_2\rangle = \langle Tf, g_1\rangle + \langle Tf, g_2\rangle = \langle f, h_1\rangle + \langle f, h_2\rangle = \langle f, h_1 + h_2\rangle$$

Como h es único entonces $h = h_1 + h_2$. De donde

$$S(g_1 + g_2) = h = h_1 + h_2 = Sg_1 + Sg_2.$$

Sea $g \in \mathcal{H}$, $\lambda \in \mathbb{C}$. Sean $h, h_\lambda \in \mathcal{H}$ los elementos dados por el teorema de representación de Riesz tales que $Sg = h$, $S(\lambda g) = h_\lambda$. Tenemos

$$\varphi(f) = \langle f, h_\lambda\rangle = \langle Tf, \lambda g\rangle = \overline{\lambda}\langle Tf, g\rangle = \overline{\lambda}\langle f, h\rangle = \langle f, \lambda h\rangle$$

Como h_λ es único entonces $h_\lambda = \lambda h$. De donde $S(\lambda g) = h_\lambda = \lambda h = \lambda Sg$.

Entonces S es lineal y $\langle Tf, g\rangle = \langle f, Sg\rangle, \quad f, g \in \mathcal{H}$.

Haciendo $f = Sg$ obtenemos la desigualdad

$$\|Sg\|^2 = \langle Sg, Sg\rangle = \langle TSg, g\rangle \leq \|TSg\| \, \|g\| \leq \|T\| \, \|Sg\| \, \|g\|$$

para $g \in \mathcal{H}$. De donde $\|Sg\| \leq \|T\| \, \|g\|$ y por lo tanto tenemos que $\|S\| \leq \|T\|$.

Luego S es acotado.

Hemos probado que S es un operador en $\mathcal{H}$.

Para demostrar que es único, supongamos que S_0 es otro operador en $\mathcal{H}$ tal que $\langle f, S_0 g\rangle = \langle Tf, g\rangle$ para $f, g \in \mathcal{H}$. Entonces $\langle f, S_0 g - Sg\rangle = 0$ para todo $f \in \mathcal{H}$, lo que implica que $S_0 g - Sg = 0$. De donde $S = S_0$. $\square$

Definición 2.40 Sea T un operador en un espacio de Hilbert $\mathcal{H}$. Definimos el *adjunto de* T, denotado por T^*, como el único operador en $\mathcal{H}$ que satisface $\langle Tf, g\rangle = \langle f, T^* g\rangle$ para $f, g \in H$.

Con la definición anterior podemos establecer una aplicación $T \to T^*$ que cumple ciertas propiedades que la hacen una involución. Recordemos este concepto.

Definición 2.41 Sea $\mathcal{B}$ un álgebra de Banach. Una *involución* en $\mathcal{B}$ es una aplicación $f \longrightarrow f^*$ que satisface:

(i) $f^{**} = f$, para $f \in \mathcal{B}$.

(ii) $(\alpha f + \beta g)^* = \overline{\alpha} f^* + \overline{\beta} g^*$, para $f, g \in \mathcal{B}$ y $\alpha, \beta \in \mathbb{C}$.

(iii) $(fg)^* = g^* f^*$ para $f, g \in \mathcal{B}$.

Observación 2.42 En muchas situaciones la involución juega un papel análogo a la conjugación de números complejos.

Definición 2.43 Sea $\mathcal{B}$ un álgebra de Banach con involución. Sea $\mathcal{U}$ una subálgebra de $\mathcal{B}$. Se dice que $\mathcal{U}$ es *autoadjunta* si

$$f \in \mathcal{U} \text{ implica que } f^* \in \mathcal{U}$$

para todo $f \in \mathcal{U}$.

La siguiente proposición establece las propiedades del adjunto de un operador y con estas se prueba que, en efecto, la aplicación $T \to T^*$ cumple con la definición de involución, convirtiendo a $L(\mathcal{H})$ en un álgebra de Banach con involución.

Proposición 2.44 *Si $\mathcal{H}$ es un espacio de Hilbert, entonces:*

1. $T^{**} = (T^*)^* = T$, *para* $T \in L(\mathcal{H})$;
2. $\|T\| = \|T^*\|$, *para* $T \in L(\mathcal{H})$;
3. $(\alpha S + \beta T)^* = \overline{\alpha} S^* + \overline{\beta} T^*$ *y* $(ST)^* = T^* S^*$, *para* $\alpha, \beta \in \mathbb{C}$ *y* $S, T \in L(\mathcal{H})$;
4. $(T^*)^{-1} = (T^{-1})^*$, *para* $T \in L(\mathcal{H})$ *invertible, y*
5. $\|T\|^2 = \|T^*T\|$, *para* $T \in L(\mathcal{H})$.

Demostración:

1. Sean $f, g \in \mathcal{H}$, $T \in L(\mathcal{H})$. Tenemos

$$\langle f, T^{**}g \rangle = \langle T^* f, g \rangle = \overline{\langle g, T^* f \rangle} = \overline{\langle Tg, f \rangle} = \langle f, Tg \rangle$$

   de donde, $T^{**} = T$.

2. En la prueba de la proposición anterior demostramos que $\|T^*\| \leq \|T\|$, de donde, usando (1):

$$\|T\| = \|T^{**}\| \leq \|T^*\| \leq \|T\|$$

   luego, $\|T\| = \|T^*\|$.

3. Sean $f, g \in \mathcal{H}$; $\alpha, \beta \in \mathbb{C}$; $S, T \in L(\mathcal{H})$. Tenemos

$$\begin{aligned}\langle f, (\alpha S + \beta T)^* g\rangle &= \langle (\alpha S + \beta T) f, g\rangle = \langle \alpha S(f) + \beta T(f), g\rangle \\ &= \alpha\langle S(f), g\rangle + \beta\langle T(f), g\rangle = \alpha\langle f, S^* g\rangle + \beta\langle f, T^* g\rangle \\ &= \langle f, \overline{\alpha} S^* g\rangle + \langle f, \overline{\beta} T^* g\rangle = \langle f, (\overline{\alpha} S^* + \overline{\beta} T^*) g\rangle\end{aligned}$$

   por lo tanto, $(\alpha S + \beta T)^* = \overline{\alpha} S^* + \overline{\beta} T^*$.

   Sean $f, g \in \mathcal{H}$; $S, T \in L(\mathcal{H})$.

$$\langle f, (ST)^* g\rangle = \langle ST(f), g\rangle = \langle T(f), S^* g\rangle = \langle f, T^* S^* g\rangle$$

   luego, $(ST)^* = T^* S^*$.

4. Usando (3) tenemos $T^*(T^{-1})^* = (T^{-1}T)^* = I = (TT^{-1})^* = (T^{-1})^* T^*$, entonces T^* es invertible y $(T^*)^{-1} = (T^{-1})^*$.

5. Como $L(\mathcal{H})$ es un álgebra de Banach y usando (2) obtenemos

$$\|T^*T\| \leq \|T^*\|\|T\| = \|T\|^2.$$

   Sólo falta probar que $\|T\|^2 \leq \|T^*T\|$. Veamos.

$$\|Tf\|^2 = \langle Tf, Tf\rangle = \langle f, T^*Tf\rangle \leq \|f\|\|T^*Tf\| \leq \|T^*T\|\|f\|^2.$$

   Luego

$$\|T\|^2 = \sup_{\|f\|=1} \|Tf\|^2 \leq \|T^*T\|.$$

   Finalmente, $\|T\|^2 = \|T^*T\|$.

$\square$

Seguidamente, haciendo uso del adjunto de un operador (involución), definiremos otros tipos de operadores.

Definición 2.45 Si T es un operador en el espacio de Hilbert $\mathcal{H}$, entonces:

(1) T es *normal* si $TT^* = T^*T$.

(2) T es *autoadjunto o hermitiano* si $T = T^*$.

(3) T es *positivo* si $\langle Tf, f\rangle \geq 0$, para $f \in \mathcal{H}$.

(4) T es *unitario* si $TT^* = T^*T = I$.

Observación 2.46 De la definición se sigue que todo operador autoadjunto es normal, y que todo operador unitario también es normal.

Existe una analogía interesante entre $L(\mathcal{H})$ y $\mathbb{C}$ el conjunto de números complejos. Dicha analogía proviene del hecho de que cada conjunto ($L(\mathcal{H})$ y $\mathbb{C}$) es un álgebra compleja dotada con una aplicación del álgebra en sí misma que tienen propiedades similares, a saber, la involución ($T \to T^*$ y $z \to \overline{z}$). Veremos que esta analogía resulta una guía muy útil en el estudio de los operadores en $\mathcal{H}$.

El subsistema más importante en el plano complejo es la recta real, que es caracterizada por la relación $z = \overline{z}$. Por analogía, corresponde a los operadores autoadjuntos.

Otro subsistema importante en el plano complejo, después de la recta real, es la circunferencia unitaria $\mathbb{T}$, que es caracterizada por una de las siguientes identidades equivalentes $|z| = 1$ o $z\overline{z} = \overline{z}z = 1$. Consecuentemente, este subsistema corresponde a los operadores unitarios.

La relevancia de los operadores normales se verá más claramente al final del capítulo cuando establezcamos el teorema espectral.

Veamos una caracterización de los operadores normales, de los operadores unitarios y de los operadores autoadjuntos.

Lema 2.47 *Si T es un operador en un espacio de Hilbert $\mathcal{H}$ para el cual $\langle Tf, f\rangle = 0$ para todo $f \in \mathcal{H}$, entonces $T = 0$.*

Una prueba de este lema se puede ver en [20], Teorema C, pg.267.

Proposición 2.48 (Caracterización de operadores normales) *Un operador T en un espacio de Hilbert $\mathcal{H}$ es normal si y sólo si $\|T^*f\| = \|Tf\|$ para todo $f \in \mathcal{H}$.*

Demostración:

En vista del lema 2.47, la prueba se desprende del hecho de que

$$\|T^*f\| = \|Tf\| \quad \Leftrightarrow \quad \|T^*f\|^2 = \|Tf\|^2 \quad \Leftrightarrow \quad \langle T^*f, T^*f\rangle = \langle Tf, Tf\rangle \quad \Leftrightarrow$$

$$\Leftrightarrow \quad \langle TT^*f, f\rangle = \langle T^*Tf, f\rangle \quad \Leftrightarrow \quad \langle (TT^* - T^*T)f, f\rangle = 0$$

□

Proposición 2.49 (Caracterización de operadores unitarios) *Un operador T en un espacio de Hilbert $\mathcal{H}$ es unitario si y sólo si T es una isometría sobreyectiva.*

Demostración:

($\Longrightarrow$) Si T es unitario sabemos que es sobreyectivo por definición. Veamos que es una isometría.

Como $T^*T = I$ tenemos

$$\|Tf\|^2 = \langle Tf, Tf\rangle = \langle f, T^*Tf\rangle = \langle f, f\rangle = \|f\|^2$$

para todo $f \in \mathcal{H}$. De donde $\|Tf\| = \|f\|$.

($\Longleftarrow$) Si T es una isometría sobreyectiva, de la linealidad y la isometría se sigue que T es biyectivo y por lo tanto existe T^{-1}. Por otro lado, de la isometría tenemos

$$\langle f, f\rangle = \langle Tf, Tf\rangle = \langle f, T^*Tf\rangle$$

de donde

$$\langle f, (T^*T - I)f\rangle = 0$$

para todo $f \in \mathcal{H}$. Por lo tanto, por lemma 2.47, $T^*T = I$. Entonces tenemos

$$T^* = T^*I = (T^*T)T^{-1} = IT^{-1} = T^{-1}.$$

Luego, $T^*T = TT^* = I$. De donde T es unitario. □

Proposición 2.50 (Caracterización de operadores autoadjuntos) *Un operador T en un espacio de Hilbert $\mathcal{H}$ es autoadjunto si y sólo si $\langle Tf, f\rangle$ es real para todo $f \in \mathcal{H}$.*

Demostración:

($\Longrightarrow$) Si T es autoadjunto y $f \in \mathcal{H}$ entonces

$$\langle Tf, f\rangle = \langle f, T^*f\rangle = \langle f, Tf\rangle = \overline{\langle Tf, f\rangle}$$

y por lo tanto $\langle Tf, f\rangle$ es real.

($\Longleftarrow$) Si $\langle Tf, f\rangle$ es real para $f \in \mathcal{H}$ entonces

$$\langle Tf, f\rangle = \overline{\langle Tf, f\rangle} = \overline{\langle f, T^*f\rangle} = \langle T^*f, f\rangle,$$

de donde $\langle (T - T^*)f, f\rangle = 0$ para todo $f \in \mathcal{H}$. Por lema 2.47, tenemos que $T = T^*$. Luego, T es autoadjunto. □

El siguiente corolario es inmediato de la proposición anterior.

Corolario 2.51 *Si T es un operador positivo en un espacio de Hilbert $\mathcal{H}$, entonces T es autoadjunto.*

Proposición 2.52 *Si T es un operador en un espacio de Hilbert $\mathcal{H}$, entonces T^*T es un operador positivo, y por lo tanto autoadjunto.*

Demostración:

Para $f \in \mathcal{H}$ tenemos

$$\langle T^*Tf, f\rangle = \overline{\langle f, T^*Tf\rangle} = \overline{\langle Tf, Tf\rangle} = \langle Tf, Tf\rangle = \|Tf\|^2 \geq 0.$$

Luego T^*T es un operador positivo y por el corolario anterior es autoadjunto. □

A continuación presentamos unas ideas breves sobre operadores acotados inferiormente. Este concepto lo usaremos para demostrar que el espectro de un operador autoadjunto es real.

Definición 2.53 Un operador T en un espacio de Hilbert $\mathcal{H}$ es *acotado inferiormente* si existe $\varepsilon > 0$ tal que $\|Tf\| \geq \varepsilon\|f\|$ para todo $f \in \mathcal{H}$.

Para una prueba de la siguiente proposición y del corolario subsiguiente remitimos al lector a [9], pg. 83-84.

Proposición 2.54 *Si T es un operador en el espacio de Hilbert $\mathcal{H}$, entonces T es invertible si y sólo si T es acotado inferiormente y tiene rango denso.*

Corolario 2.55 *Si T es un operador en el espacio de Hilbert $\mathcal{H}$ tal que T y T^* son acotados inferiormente, entonces T es invertible.*

Considerando al operador T, definido en un espacio de Hilbert $\mathcal{H}$, como un elemento del álgebra de Banach $L(\mathcal{H})$, tiene sentido hablar de su espectro $\sigma(T)$. Se puede ver que en un espacio finito-dimensional, λ está en el espectro de T si y sólo si λ es un autovalor de T. Esto no se cumple para operadores en espacios infinito-dimensionales.

En álgebra lineal se demuestra que los autovalores de una matriz hermitiana son reales. La generalización para operadores hermitianos (autoadjuntos) toma la siguiente forma.

Proposición 2.56 *Si T es un operador autoadjunto en el espacio de Hilbert $\mathcal{H}$, entonces el espectro de T es real. Además, si T es un operador positivo entonces el espectro de T es no negativo.*

Demostración:

Para ver que $\sigma(T)$ es real, tomemos $\lambda = \alpha + i\beta$, con $\alpha, \beta \in \mathbb{R}$ y $\beta \neq 0$, y veamos que $\lambda \notin \sigma(T)$, es decir, que $T - \lambda$ es invertible.

Sea $K = (T - \alpha)/\beta$. Como T es autoadjunto, aplicando las propiedades de la involución en $L(\mathcal{H})$, tenemos que K es autoadjunto.

Además,
$$K - i = \frac{T - \alpha}{\beta} - i = \frac{T - \alpha - i\beta}{\beta} = \frac{T - \lambda}{\beta}.$$
Entonces $T - \lambda$ es invertible si y sólo si $K - i$ es invertible.

Para ver que $K - i$ es invertible probemos que

$$K - i \qquad \text{y} \qquad (K-i)^* = K^* + \overline{(-i)} = K + i$$

están acotados inferiormente. Usando que K es autoadjunto (en la última igualdad)tenemos que para $f \in \mathcal{H}$

$$\begin{aligned}\|(K \pm i)(f)\|^2 &= \langle (K \pm i)f, (K \pm i)f \rangle = \|Kf\|^2 \mp i\langle Kf, f\rangle \pm i\langle f, Kf\rangle + \|f\|^2 \\ &= \|Kf\|^2 + \|f\|^2 \geq \|f\|^2\end{aligned}$$

de donde, $K \pm i$ están acotados inferiormente.

Luego, por corolario 2.55 tenemos que $K - i$ es invertible y por lo tanto el espectro de un operador autoadjunto es real.

Si además asumimos que T es positivo y $\lambda < 0$ entonces

$$\|(T-\lambda)f\|^2 = \langle (T-\lambda)f, (T-\lambda)f\rangle = \|Tf\|^2 - 2\lambda\langle Tf, f\rangle + \lambda^2 \|f\|^2 \geq \lambda^2 \|f\|^2$$

de donde $\|(T-\lambda)f\| \geq \lambda \|f\|$ y por lo tanto, $T - \lambda$ está acotado inferiormente.

Y como $(T-\lambda)^* = (T-\lambda)$, entonces, usando nuevamente el corolario 2.55 tenemos que $T - \lambda$ es invertible. Luego, el espectro de un operador positivo es no negativo.

$\square$

Ahora consideraremos una clase especial de operadores positivos que tienen un papel muy importante en la construcción de operadores normales dada por el teorema espectral que veremos al final del capítulo.

Definición 2.57 Un operador P en el espacio de Hilbert $\mathcal{H}$ es una *proyección* si P es idempotente ($P^2 = P$) y autoadjunto.

Observación 2.58 Los operadores I y 0 son proyecciones, y son distintos si y sólo si $\mathcal{H} \neq \{0\}$.

La siguiente construcción genera una proyección, y, de hecho, todas las proyecciones se originan de esta manera.

Definición 2.59 Sea $\mathcal{M}$ un subespacio cerrado del espacio de Hilbert $\mathcal{H}$. Definimos $P_{\mathcal{M}}$ como $P_{\mathcal{M}}f = g$, donde $f = g + h$ con $g \in \mathcal{M}$ y $h \in \mathcal{M}^{\perp}$.

Teorema 2.60 *Si $\mathcal{M}$ es un subespacio cerrado de un espacio de Hilbert $\mathcal{H}$, entonces $P_{\mathcal{M}}$ es una proyección de rango $\mathcal{M}$. Además, si P es una proyección en $\mathcal{H}$, entonces existe un subespacio cerrado $\mathcal{M}(= \mathrm{rang}P)$ tal que $P = P_{\mathcal{M}}$.*

Demostración:

Primero veamos que $P_{\mathcal{M}}$ es un operador en $\mathcal{H}$.

Sean $f_1, f_2 \in \mathcal{H}$; $\lambda_1, \lambda_2 \in \mathbb{C}$. Entonces,

$$f_1 = g_1 + h_1 \qquad \text{y} \qquad f_2 = g_2 + h_2,$$

donde $g_1, g_2 \in \mathcal{M}$ y $h_1, h_2 \in \mathcal{M}^{\perp}$. Además,

$$\lambda_1 f_1 + \lambda_2 f_2 = (\lambda_1 g_1 + \lambda_2 g_2) + (\lambda_1 h_1 + \lambda_2 h_2)$$

donde $\lambda_1 g_1 + \lambda_2 g_2 \in \mathcal{M}$ y $\lambda_1 h_1 + \lambda_2 h_2 \in \mathcal{M}^{\perp}$. Como esta descomposición es única (ver [20], Teorema D, pg.251) tenemos

$$P_{\mathcal{M}}(\lambda_1 f_1 + \lambda_2 f_2) = \lambda_1 g_1 + \lambda_2 g_2 = \lambda_1 P_{\mathcal{M}} f_1 + \lambda_2 P_{\mathcal{M}} f_2$$

y por lo tanto $P_{\mathcal{M}}$ es una transformación lineal en $\mathcal{H}$.

Por otra parte tenemos

$$\|P_{\mathcal{M}} f_1\|^2 = \|g_1\|^2 \leq \|g_1\|^2 + \|h_1\|^2 = \|f_1\|^2$$

y entonces $P_{\mathcal{M}}$ está acotado y $\|P_{\mathcal{M}}\| \leq 1$.

Hemos probado que $P_{\mathcal{M}}$ es un operador en $\mathcal{H}$.

Veamos que $P_{\mathcal{M}}$ es autoadjunto. Tenemos

$$\begin{aligned}\langle P_{\mathcal{M}} f_1, f_2\rangle &= \langle g_1, g_2 + h_2\rangle = \langle g_1, g_2\rangle + \langle g_1, h_2\rangle \\ &= \langle g_1, g_2\rangle + 0 = \langle g_1, g_2\rangle + \langle h_1, g_2\rangle \\ &= \langle g_1 + h_1, g_2\rangle = \langle f_1, P_{\mathcal{M}} f_2\rangle\end{aligned}$$

Luego, $P_{\mathcal{M}} = P_{\mathcal{M}}^*$. De donde $P_{\mathcal{M}}$ es autoadjunto.

Por último, si $f \in \mathcal{M}$, entonces $f = f + 0$ es la descomposición de f y por lo tanto $P_{\mathcal{M}} f = f$. Como rang$P_{\mathcal{M}} = \mathcal{M}$, entonces $P_{\mathcal{M}}^2 = P_{\mathcal{M}}$ y por lo tanto $P_{\mathcal{M}}$ es idempotente.

Hemos demostrado que $P_{\mathcal{M}}$ es una proyección con rango $\mathcal{M}$.

Ahora supongamos que P es una proyección en $\mathcal{H}$ y hagamos $\mathcal{M} = \text{rang}P$. Si $\{Pf_n\}_{n=1}^{\infty}$ es una sucesión de Cauchy en $\mathcal{H}$ que converge a g, entonces

$$g = \lim_{n\to\infty} Pf_n = \lim_{n\to\infty} P^2 f_n = P(\lim_{n\to\infty} Pf_n) = Pg.$$

Así $g \in \mathcal{M}$ y por lo tanto $\mathcal{M}$ es un subespacio cerrado de $\mathcal{H}$. Si $g \in \mathcal{M}^{\perp}$, entonces

$$\|Pg\|^2 = \langle Pg, Pg\rangle = \langle g, P^2 g\rangle = 0,$$

ya que $P^2 g \in \mathcal{M}$; y por lo tanto $Pg = 0$.

Si $f \in \mathcal{H}$, entonces $f = Pg + h$, donde $h \in \mathcal{M}^{\perp}$ y por lo tanto

$$P_{\mathcal{M}} f = Pg = P^2 g = P^2 g + Ph = Pf.$$

Finalmente $P = P_{\mathcal{M}}$. □

El teorema anterior es de gran importancia ya que permite establecer una correspondencia entre las proyecciones y los subespacios cerrados de $\mathcal{H}$. De esta manera, siempre hablamos de P como la proyección en $\mathcal{M}$, es decir, $P_{\mathcal{M}}$, para algún $\mathcal{M}$ subespacio cerrado de $\mathcal{H}$.

2.2.2. Algebras-C*

Definición 2.61 Sea $\mathcal{B}$ un álgebra de Banach con involución. Se dice que $\mathcal{B}$ es un *álgebra-C^** si se satisface que

$$\|f^* f\| = \|f\|^2$$

para todo $f \in \mathcal{B}$.

Observación 2.62 En la primera parte de esta sección vimos que el espacio $L(\mathcal{H})$ es un álgebra de Banach con involución $T \longrightarrow T^*$ (donde T^* es el adjunto de T). Ahora, dada la definición anterior vemos (por proposición 2.44) que $L(\mathcal{H})$ es, de hecho, un álgebra-C^*. También se puede ver que toda subálgebra autoadjunta cerrada de $L(\mathcal{H})$ es un álgebra-C^*. Es más, se puede probar que toda álgebra-C^* es isométricamente isomorfa a una subálgebra de $L(\mathcal{H})$ para algún espacio de Hilbert $\mathcal{H}$.

Todas las clases de operadores cuyas definiciones están basadas en el adjunto (involución) pueden ser extendidas a un álgebra-C^*; por ejemplo, un elemento T en un álgebra-C^* se dice *autoadjunto* si $T^* = T$, *normal* si $TT^* = T^*T$, *unitario* si $T^*T = TT^* = I$ y *proyección* si $T^* = T$ y $T^2 = T$.

Proposición 2.63 *La involución en un álgebra-C^* es una isometría, es decir, $\|T\| = \|T^*\|$.*

Demostración:

Sea $\mathcal{U}$ un álgebra-C^*. Sea $T \in \mathcal{U}$. Tenemos

$$\|T\|^2 = \|T^*T\| \leq \|T^*\| \, \|T\|$$

de donde $\|T\| \leq \|T^*\|$. Por lo tanto $\|T\| = \|T^*\|$, ya que $T = T^{**}$. □

Proposición 2.64 *Si T es un elemento normal en un álgebra-C^*, entonces $\|T^2\| = \|T\|^2$.*

Demostración:

Es obvio que $\|T^2\| \leq \|T\|^2$. Probemos la desigualdad contraria.

$$\begin{aligned}
\|T^*\|^2 \, \|T\|^2 = (\|T^*\| \, \|T\|)^2 = \|T^*T\|^2 &= \|(T^*T)^*T^*T\| \\
&= \|T^*TT^*T\| = \|T^*T^*TT\| = \|(T^*)^2T^2\| \\
&= \|(T^2)^*T^2\| = \|(T^2)^*\| \, \|T^2\| = \|(T^*)^2\| \, \|T^2\| \leq \|T^*\|^2 \, \|T^2\|.
\end{aligned}$$

De donde $\|T\|^2 \leq \|T^2\|$.

□

Teorema 2.65 *Si U es un elemento unitario en un álgebra-C^* $\mathcal{U}$, entonces*

$$\sigma(U) \subseteq \mathbb{T},$$

donde $\mathbb{T}$ es la circunferencia unitaria en el plano complejo.

Demostración:

Como U es unitario, entonces U es invertible y $U^{-1} = U^*$. Además, por definición de álgebra-C^* tenemos que $\|U\|^2 = \|U^*U\| = \|I\| = 1$. Entonces

$$1 = \|U\| = \|U^*\| = \|U^{-1}\|.$$

Por proposición 2.13 se cumple que

$$r(U) \leq \|U\| = 1,$$

entonces si $\lambda \in \sigma(U)$ tenemos que $|\lambda| \leq 1$. Queremos probar que $|\lambda| = 1$.

Si $|\lambda| < 1$ entonces $\|\lambda U^*\| < 1$ y por lo tanto

$$\|I - (I - \lambda U^*)\| < 1.$$

De donde $I - \lambda U^*$ es invertible (por proposición 2.8). Luego $U - \lambda = U(I - \lambda U^*)$ es también es invertible. Por lo tanto $\lambda \notin \sigma(U)$.

Hemos demostrado que si $\lambda \in \sigma(U)$ entonces $|\lambda| = 1$, es decir, $\sigma(U) \subseteq \mathbb{T}$. □

A continuación extenderemos la proposición 2.56 a álgebras-C^*. En la prueba de dicha proposición fue esencial que estábamos tratando con operadores en un espacio de Hilbert.

Teorema 2.66 *En un álgebra-C^* todo elemento autoadjunto tiene espectro real.*

Demostración:

Sea $\mathcal{U}$ un álgebra-C^*. Sea $H \in \mathcal{U}$ un elemento autoadjunto y sea $U = \exp iH$. Entonces, del hecho de que H es autoadjunto y de la definición de la función exponencial en álgebras de Banach (análoga a la definida en $\mathbb{R}$ como serie de potencias), se sigue que

$$U^* = \exp(iH)^* = \exp(-iH).$$

Más aún, tenemos

$$UU^* = \exp(iH)\exp(-iH) = \exp(iH - iH) = I = \exp(-iH)\exp(iH) = U^*U,$$

de donde U es unitario. Y usando teorema anterior(2.65) tenemos que

$$\sigma(U) \subseteq \mathbb{T},$$

donde $\mathbb{T}$ es la circunferencia unitaria en el plano complejo.

Por otro lado, usando un corolario del teorema de Gelfand (2.36) tenemos

$$\sigma(U) = \sigma(\exp(iH)) = \exp(i\sigma(H)).$$

Finalmente, sean $\beta \in \sigma(H) \subset \mathbb{C}$ y $\lambda = \exp(i\beta) \in \mathbb{C}$. Como $\sigma(U) = \exp(i\sigma(H))$ y $\sigma(U) \subseteq \mathbb{T}$ entonces $1 = |\lambda| = |\exp(i\beta)|$, de donde $\exp(\mathrm{Re}(i\beta)) = 1$, lo que implica que $\mathrm{Re}(i\beta) = 0$ y esto ocurre si y sólo si β es real.

Luego, $\sigma(H) \subset \mathbb{R}$. $\square$

Teorema 2.67 *Si $\mathcal{B}$ es un álgebra-C^*, $\mathcal{U}$ es una subálgebra autoadjunta cerrada de $\mathcal{B}$, y T es un elemento de $\mathcal{U}$, entonces $\sigma_{\mathcal{U}}(T) = \sigma_{\mathcal{B}}(T)$.*

Demostración:
Como $\mathcal{U} \subset \mathcal{B}$ entonces $\sigma_{\mathcal{B}}(T) \subseteq \sigma_{\mathcal{U}}(T)$. Luego, es suficiente probar que si $T - \lambda$ es invertible en $\mathcal{B}$, entonces la inversa $(T-\lambda)^{-1}$ está en $\mathcal{U}$.

Podemos suponer, sin pérdida de generalidad, que $\lambda = 0$. Así, T es invertible en $\mathcal{B}$ y por lo tanto T^*T es un elemento autoadjunto (por proposición 2.52) de $\mathcal{U}$ que es invertible en $\mathcal{B}$, ya que $(T^*T)^{-1} = T^{-1}(T^*)^{-1} = T^{-1}(T^{-1})^*$.

Como $\sigma_{\mathcal{U}}(T^*T)$ es real (por el teorema anterior), entonces $\sigma_{\mathcal{U}}(T^*T) = \sigma_{\mathcal{B}}(T^*T)$ por un corolario del teorema de Šilov (Ver [9], pg.54).

Así, T^*T es invertible en $\mathcal{U}$ (porque como T^*T es invertible en $\mathcal{B}$ entonces $0 \notin \sigma_{\mathcal{B}}(T^*T)$, luego $0 \notin \sigma_{\mathcal{U}}(T^*T)$ y entonces T^*T es invertible en $\mathcal{U}$) y por lo tanto

$$T^{-1} = (T^{-1}(T^*)^{-1})T^* = (T^*T)^{-1}T^*$$

está en $\mathcal{U}$. $\square$

El siguiente teorema es una caracterización de las álgebras-C^* conmutativas, que usaremos para obtener una forma del teorema espectral para operadores normales.

Teorema 2.68 (Gelfand-Naimark) *Si $\mathcal{U}$ es un álgebra-C^* conmutativa y $\mathbf{M}$ es el espacio de ideales maximales de $\mathcal{U}$, entonces la transformada de Gelfand es un isomorfismo-* isométrico de $\mathcal{U}$ sobre $C(\mathbf{M})$.*

Demostración:

Sea Γ la transformada de Gelfand. Debemos probar que Γ es una función-*, i.e, $\overline{\Gamma(T)} = \Gamma(T^*)$; y que $\|\Gamma(T)\|_\infty = \|T\|$ para $T \in \mathcal{U}$.

Si $T \in \mathcal{U}$, entonces

$$H = (T + T^*)/2 \qquad \text{y} \qquad K = (T - T^*)/2i$$

son elementos autoadjuntos en $\mathcal{U}$, tales que $T = H + iK$ y $T^* = H - iK$.

Por teorema 2.66, $\sigma(H)$ y $\sigma(K)$ están contenidos en $\mathbb{R}$ y además, por corolario del teorema de Gelfand (2.35) tenemos que $\sigma(H) = \operatorname{rang} \Gamma H$ y $\sigma(K) = \operatorname{rang} \Gamma K$. Luego, $\Gamma(H)$ y $\Gamma(K)$ son funciones reales.

Por lo tanto

$$\overline{\Gamma(T)} = \overline{\Gamma(H + iK)} = \overline{\Gamma(H) + i\Gamma(K)} = \Gamma(H) - i\Gamma(K) = \Gamma(H - iK) = \Gamma(T^*),$$

de donde Γ es una función-*.

Veamos que Γ es una isometría.

Como $\mathcal{U}$ es un álgebra-C^* conmutativa y $T \in \mathcal{U}$, entonces T es normal ya que T^* también está en $\mathcal{U}$ y los operadores en $\mathcal{U}$ conmutan. Luego usando la proposición 2.64 tenemos que $\|T^2\| = \|T\|^2$. Por lo tanto Γ es una isometría en vista del corolario 2.38.

Como Γ es lineal e isométrica tenemos la inyectividad. Y finalmente, el hecho de que Γ es sobreyectiva se sigue del teorema de Stone- Weierstrass (ver [9], pg. 46). $\square$

Teorema 2.69 (Espectral) *Si $\mathcal{H}$ es un espacio de Hilbert y T es un operador normal en $\mathcal{H}$, entonces el álgebra-C^* $\mathcal{A}_T$, generada por T, es conmutativa. Más aún, el espacio de ideales maximales $\mathbf{M}$ de $\mathcal{A}_T$ es homeomorfo a $\sigma(T)$, y por lo tanto la transformada de Gelfand Γ es un isomorfismo-* isométrico de $\mathcal{A}_T$ sobre $C(\sigma(T))$. Así, tenemos que $\Gamma(\Gamma T)(z) = z$ para todo $z \in \sigma(T)$.*

Nota: Esto último se suele expresar como que ΓT definida en $\sigma(T)$ es la identidad, es decir, $\Gamma T(z) = z$ para todo $z \in \sigma(T)$.

Demostración:

Como T y T^* conmutan, entonces la colección de polinomios en T y T^* forman una subálgebra autoadjunta conmutativa de $L(\mathcal{H})$, la cual debe estar contenida en $\mathcal{A}_T$ (el álgebra-C^* generada por T).

No es difícil verificar que la clausura de esta colección de polinomios es un álgebra-C^* conmutativa y por lo tanto debe ser $\mathcal{A}_T$. Luego $\mathcal{A}_T$ es conmutativa.

Para demostrar que el espacio de ideales maximales $\mathbf{M}$ de $\mathcal{A}_T$ es homeomorfo a $\sigma(T)$, definamos $\psi : \mathbf{M} \to \sigma(T)$ por

$$\psi(\varphi) = \Gamma(T)(\varphi).$$

Como el rango de $\Gamma(T)$ es $\sigma(T)$ (por corolario del teorema de Gelfand (2.35)), entonces ψ está bien definida y es sobreyectiva.

Ahora, si $\varphi_1, \varphi_2 \in \mathbf{M}$ son tales que $\psi(\varphi_1) = \psi(\varphi_2)$ entonces

$$\Gamma(T)(\varphi_1) = \Gamma(T)(\varphi_2), \qquad \varphi_1(T) = \varphi_2(T) \qquad \text{y}$$

$$\varphi_1(T^*) = \Gamma(T^*)(\varphi_1) = \overline{\Gamma(T)(\varphi_1)} = \overline{\Gamma(T)(\varphi_2)} = \Gamma(T^*)(\varphi_2) = \varphi_2(T^*).$$

Por lo tanto, como φ_1 y φ_2 son lineales y multiplicativos, tenemos

$$\varphi_1\left(\sum_{j,k} c_{jk} T^j {T^*}^k\right) = \varphi_2\left(\sum_{j,k} c_{jk} T^j {T^*}^k\right),$$

es decir, φ_1 y φ_2 concuerdan en todos los polinomios en T y T^*. Luego, como esta colección de operadores es densa en $\mathcal{A}_T$, tenemos que $\varphi_1 = \varphi_2$. De donde ψ es inyectiva.

Por último, si $\{\varphi_n\}_{n=1}^{\infty}$ es una sucesión en $\mathbf{M}$ tal que $\lim_{n\to\infty} \varphi_n = \varphi$, entonces, como $\Gamma(T) \in C(\mathbf{M})$,

$$\lim_{n\to\infty} \psi(\varphi_n) = \lim_{n\to\infty} \Gamma(T)(\varphi_n) = \Gamma(T)(\varphi) = \psi(\varphi)$$

y por lo tanto ψ es continua.

Luego, como $\mathbf{M}$ y $\sigma(T)$ son espacios Hausdorff compactos, entonces ψ es un homeomorfismo. De donde

$$\mathbf{M} \approx \sigma(T).$$

Entonces por el teorema anterior (Gelfand-Naimark (2.68)) tenemos que la transformada de Gelfand es un isomorfismo-* isométrico de $\mathcal{A}_T$ sobre $C(\sigma(T))$.

Finalmente veamos que $\Gamma T(z) = z$ para todo $z \in \sigma(T)$.

Sabemos que la transformada de Gelfand $\Gamma : \mathcal{A}_T \to C(\mathbf{M})$ manda al operador $S \in \mathcal{A}_T$ en una función continua ΓS, definida en $\mathbf{M}$, a la que definimos anteriormente por ψ.

Como probamos que $\mathbf{M} \approx \sigma(T)$, podemos definir a ψ en $\sigma(T)$. De esta manera, la transformada de Gelfand

$$\Gamma(\Gamma S) : \sigma(T) \to \mathbb{C}$$

se puede definir por

$$\Gamma(\Gamma S)(z) = \Gamma S((\Gamma T)^{-1}(z)).$$

En particular, para T, tenemos

$$\Gamma(\Gamma T)(z) = \Gamma T((\Gamma T)^{-1}(z)) = z,$$

de donde, la transformada de Gelfand definida en $\sigma(T)$ es la identidad. $\square$

2.3. Medida Espectral

El objetivo de esta sección es presentar otra versión del teorema espectral donde se hace uso de la medida espectral. Para establecer la construcción de dicha medida necesitamos algunos conceptos y en especial un teorema con respecto a la extensión de la transformada de Gelfand. Comenzaremos esta sección con una breve noción sobre las medidas de Borel regulares y sobre la extensión de la transformada de Gelfand.

Medidas de Borel Regulares.

Sea $\mathcal{X}$ un espacio de Hausdorff. Consideremos la σ-álgebra $\mathcal{B}(\mathcal{X})$ de todos los subconjuntos de Borel de $\mathcal{X}$, es decir, la σ-álgebra generada por los subconjuntos abiertos de $\mathcal{X}$. Una medida definida en $\mathcal{B}(\mathcal{X})$ se denomina *medida de Borel* en $\mathcal{X}$.

Definición 2.70 Sea $\mathcal{X}$ un espacio de Hausdorff. Una medida de Borel μ en $\mathcal{X}$ se dice *regular* si satisface que

(i) $\mu(C) < \infty$ para todo subconjunto compacto $C \subseteq \mathcal{X}$;

(ii) Para cada boreliano $B \in \mathcal{B}(\mathcal{X})$

$$\mu(B) = \text{ínf}\{\mu(C) : B \subseteq C,\ C \ \text{ abierto}\};$$

(iii) La relación

$$\mu(B) = \sup\{\mu(K)) : K \subseteq B\}$$

vale siempre que B sea abierto o cuando $B \in \mathcal{B}(\mathcal{X})$ y $\mu(B) < \infty$.

En la terminología empleada, una medida se refiere a una medida no negativa. Sin embargo, haciendo uso de las medidas no negativas, es posible ampliar el concepto a medidas reales y más aún a medidas complejas, es decir, a medidas a valores complejos. Es por ello que podemos hablar de las *medidas complejas de Borel* y de *medidas complejas de Borel regulares.*

Extensión de la inversa de la transformada de Gelfand.

Sea $\mathcal{U} \subseteq L(\mathcal{H})$ una subálgebra autoadjunta cerrada y conmutativa. Podemos hablar de $\mathcal{U}$ como un álgebra-C^* conmutativa (Ver observación 2.62).

Sea $\mathbf{M}$ el espacio de ideales maximales de $\mathcal{U}$ y consideremos la transformada de Gelfand $\Gamma : \mathcal{U} \to C(\mathbf{M})$. Por teorema de Gelfand-Naimark (2.68) tenemos que la inversa de la transformada de Gelfand $\Gamma^{-1} : C(\mathbf{M}) \to \mathcal{U}$ existe y podemos denotarla

$$\Gamma^{-1}(\varphi) = T_\varphi, \quad \text{para } \varphi \in C(\mathbf{M}).$$

Si definimos $\psi_{fg} : C(\mathbf{M}) \to \mathbb{C}$ por

$$\psi_{fg}(\varphi) = \langle T_\varphi f, g\rangle,$$

con $f, g \in \mathcal{H}$, se puede probar que $\psi_{fg} \in C(\mathbf{M})^*$ y usando el teorema de representación de Riesz (Ver [19], pg.131) tenemos que existe una única medida compleja de Borel regular, μ_{fg}, en $\mathbf{M}$ tal que

$$\psi_{fg}(\varphi) = \int_{\mathbf{M}} \varphi \, d\mu_{fg}, \qquad \text{para todo } \varphi \in C(\mathbf{M})$$

y $\|\mu_{fg}\| = \|\psi_{fg}\|$. De donde,

$$\int_{\mathbf{M}} \varphi \, d\mu_{fg} = \langle T_\varphi f, g\rangle$$

para todo $\varphi \in C(\mathbf{M})$.

Pero esta integral no sólo tiene sentido para $\varphi \in C(\mathbf{M})$, sino también para toda $\varphi \in \mathbf{B}(\mathbf{M})$, donde $\mathbf{B}(\mathbf{M})$ es el conjunto de todas las funciones medibles Borel, acotadas, definidas en $\mathbf{M}$. Así, la inversa de Γ se puede extender a dicho conjunto.

Lo anterior constituye una idea para la prueba del siguiente teorema. Una prueba completa se puede encontrar en libros de análisis funcional en donde, como hemos visto, se hace uso esencial del teorema de representación de Riesz.

Teorema 2.71 (Extensión de la Inversa de la Transformada de Gelfand) *Sea $\mathcal{U}$ un algebra-C^* conmutativa. Sea $\mathbf{M}$ el espacio de ideales maximales de $\mathcal{U}$ y $\Gamma : \mathcal{U} \to C(\mathbf{M})$ la transformada de Gelfand. Sea $\mathcal{R} = \{$medidas complejas de Borel regulares, en $\mathbf{M}\}$ y $\mathbf{B}(\mathbf{M}) = \{$funciones medibles de Borel, acotadas, definidas en $\mathbf{M}\}$. Entonces:*

(1) Existe una correspondencia de $\mathcal{H} \times \mathcal{H} \longrightarrow \mathcal{R}$ que manda al par f, g en μ_{fg}, y es tal que

$$\langle Tf, g\rangle = \int_{\mathbf{M}} \Gamma T \, d\mu_{fg}$$

para todo $T \in \mathcal{U}$. Y tiene las siguientes propiedades:

a) $\mu_{f_1+\lambda f_2, g} = \mu_{f_1, g} + \lambda\, \mu_{f_2, g}, \quad \mu_{f, g_1+\lambda g_2} = \mu_{f, g_1} + \overline{\lambda}\, \mu_{f, g_2}$;

b) $\|\mu_{fg}\| \leq \|f\| \, \|g\|$;

c) $d\mu_{f, T^* g} = \Gamma T \, d\mu_{fg} = d\mu_{Tf, g}$;

d) $\overline{\mu_{fg}} = \mu_{gf}$ *y* $\mu_{ff} \geq 0$.

(2) Existe $\Upsilon : \mathbf{B}(\mathbf{M}) \to L(\mathcal{H})$ como extensión de Γ^{-1} (la inversa de la transformada de Gelfand), definida por $\Upsilon(\phi) = T_\phi$, tal que

$$\langle T_\phi f, g\rangle = \int_{\mathbf{M}} \phi \, d\mu_{fg}$$

para todo $\phi \in \mathbf{B}(\mathbf{M})$. Y cumple las siguientes propiedades:

a) $T_{\phi_1+\lambda\phi_2} = T_{\phi_1} + \lambda\, T_{\phi_2}$, *para* $\phi_1, \phi_2 \in \mathbf{B}(\mathbf{M})$;

b) $\|T_\phi\| \leq \|\phi\|_\infty$, *para* $\phi \in \mathbf{B}(\mathbf{M})$;

c) Si $\phi \geq 0$ *entonces* $\langle T_\phi f, f\rangle \geq 0$, *para* $\phi \in \mathbf{B}(\mathbf{M})$;

d) $T_{\phi_1\phi_2} = T_{\phi_1} T_{\phi_2}$, *para* $\phi_1, \phi_2 \in \mathbf{B}(\mathbf{M})$;

e) $(T_\phi)^* = T_{\overline{\phi}}$, *para* $\phi \in \mathbf{B}(\mathbf{M})$;

f) $T_{\phi_1} T_{\phi_2} = T_{\phi_2} T_{\phi_1}$, *para* $\phi_1 \in \mathbf{B}(\mathbf{M})$ *y* $\phi_2 \in C(\mathbf{M})$;

g) $\phi \, d\mu_{fg} = d\mu_{f, (T_\phi)^* g} = d\mu_{T_\phi f, g}$, *para* $\phi \in \mathbf{B}(\mathbf{M})$;

h) T_ϕ *es un operador normal, para* $\phi \in \mathbf{B}(\mathbf{M})$.

2.3.1. Construcción de la Medida Espectral en un Álgebra-C^*

La construcción de la medida espectral viene dada por el siguiente teorema.

Teorema 2.72 *Sea $\mathcal{U}$ un álgebra-C^* conmutativa. Sea $\mathbf{M}$ el espacio de ideales maximales de $\mathcal{U}$. Entonces*

(1) A cada conjunto de Borel $\triangle \subseteq \mathbf{M}$ se le asocia un operador $P(\triangle)$ en $\mathcal{H}$, tal que para cada $f, g \in \mathcal{H}$, la igualdad

$$\mu_{fg}(\triangle) = \langle P(\triangle)f, g\rangle$$

define una medida compleja de Borel que satisface

$$\mu_{ff} \geq 0, \qquad \overline{\mu_{fg}} = \mu_{gf}, \qquad \|\mu_{fg}\| \leq \|f\|\,\|g\|.$$

(2) a) $P(\triangle)$ es una proyección.

b) $P(\mathrm{M}) = I$ y $P(\varnothing) = 0$.

(3) a) Si $\triangle_1 \cap \triangle_2 = \varnothing$ entonces $P(\triangle_1 \cup \triangle_2) = P(\triangle_1) + P(\triangle_2)$.

b) Se cumple la independencia: $P(\triangle_1 \cap \triangle_2) = P(\triangle_1)P(\triangle_2)$.

c) Si $\triangle_1 \cap \triangle_2 = \varnothing$ entonces $P(\triangle_1)\mathcal{H} \perp P(\triangle_2)\mathcal{H}$.

Demostración:

(1) Sea $\triangle \subseteq \mathbf{M}$ un conjunto de Borel. Usando la notación del teorema anterior (2.71) definimos

$$P(\triangle) = T_{\mathbf{1}_\triangle},$$

donde $\mathbf{1}_\triangle$ es la función indicatriz (o función característica) y $T_{\mathbf{1}_\triangle} = \Upsilon(\mathbf{1}_\triangle)$. Entonces, por el teorema anterior

$$\langle P(\triangle)f, g\rangle = \langle T_{\mathbf{1}_\triangle} f, g\rangle = \int_{\mathbf{M}} \mathbf{1}_\triangle \, d\mu_{fg} = \mu_{fg}(\triangle)$$

donde μ_{fg} es una medida compleja de Borel regular, definida en $\mathbf{M}$, tal que $\mu_{ff} \geq 0$, $\overline{\mu_{fg}} = \mu_{gf}$ y $\|\mu_{fg}\| \leq \|f\|\,\|g\|$.

(2) a) Sea $\triangle \subseteq \mathbf{M}$ un conjunto de Borel. Por el teorema anterior, parte (2)d), Υ es multiplicativa, entonces

$$P(\triangle)^2 = (T_{\mathbf{1}_\triangle})^2 = T_{\mathbf{1}_\triangle^2} = T_{\mathbf{1}_\triangle} = P(\triangle)$$

de donde, $P(\triangle)$ es idempotente. Además, por parte (2)e) del mismo teorema anterior, tenemos

$$P(\triangle)^* = (T_{\mathbf{1}_\triangle})^* = T_{\overline{\mathbf{1}_\triangle}} = T_{\mathbf{1}_\triangle} = P(\triangle)$$

de donde, $P(\triangle)$ es autoadjunto. Luego, $P(\triangle)$ es una proyección.

b) Como $\mathbf{1}_\mathbf{M} \in C(\mathbf{M})$ y $\mathbf{1}_\varnothing \in C(\mathbf{M})$, usando el teorema anterior tenemos

$$P(\mathbf{M}) = T_{\mathbf{1}_\mathbf{M}} = \Upsilon(\mathbf{1}_\mathbf{M}) = \Gamma^{-1}(\mathbf{1}_\mathbf{M}) = I, \qquad \text{ya que } \Gamma I = \mathbf{1}_\mathbf{M}; \text{ y}$$

$$P(\varnothing) = T_{\mathbf{1}_\varnothing} = \Upsilon(\mathbf{1}_\varnothing) = \Gamma^{-1}(\mathbf{1}_\varnothing) = 0, \qquad \text{ya que } \Gamma 0 = \mathbf{1}_\varnothing = 0.$$

(3) a) Sean $\triangle_1, \triangle_2 \subseteq \mathbf{M}$ conjuntos de Borel tales que $\triangle_1 \cap \triangle_2 = \varnothing$. Como Υ es lineal ((2)a) teorema anterior), tenemos

$$\begin{aligned} P(\triangle_1 \cup \triangle_2) &= T_{\mathbf{1}_{\triangle_1 \cup \triangle_2}} = T_{\mathbf{1}_{\triangle_1} + \mathbf{1}_{\triangle_2}} \\ &= T_{\mathbf{1}_{\triangle_1}} + T_{\mathbf{1}_{\triangle_2}} = P(\triangle_1) + P(\triangle_2). \end{aligned}$$

b) Sean $\triangle_1, \triangle_2 \subseteq \mathbf{M}$ conjuntos de Borel. Como Υ es multiplicativa ((2)d) teorema anterior)

$$\begin{aligned} P(\triangle_1)\, P(\triangle_2) &= T_{\mathbf{1}_{\triangle_1}}\, T_{\mathbf{1}_{\triangle_2}} = T_{\mathbf{1}_{\triangle_1} \mathbf{1}_{\triangle_2}} \\ &= T_{\mathbf{1}_{\triangle_1 \cap \triangle_2}} = P(\triangle_1 \cap \triangle_2). \end{aligned}$$

c) Sean $\triangle_1, \triangle_2 \subseteq \mathbf{M}$ tales que $\triangle_1 \cap \triangle_2 = \varnothing$. Usando partes (2)b) y (3)b) anteriormente probadas, tenemos que

$$P(\triangle_1)P(\triangle_2) = 0.$$

Luego, si $f, g \in \mathcal{H}$ tenemos

$$\langle P(\triangle_1)f, P(\triangle_2)g\rangle = \langle P(\triangle_2)P(\triangle_1)f, g\rangle = 0.$$

Y así, $P(\triangle_1)\mathcal{H} \perp P(\triangle_2)\mathcal{H}$. $\square$

Notación 2.73 (Medida Espectral) Sea $\Upsilon(\phi) = T_\phi$ la extensión de la inversa de la transformada de Gelfand, entonces tenemos que

$$\langle T_\phi f, g\rangle = \int_{\mathbf{M}} \phi \, d\mu_{fg}$$

para $\phi \in \mathbf{B}(\mathbf{M})$.

Usando la notación de la medida espectral tenemos que $\mu_{fg}(\triangle) = \langle P(\triangle)f, g\rangle$ y lo anterior lo podemos escribir como

$$\langle T_\phi f, g\rangle = \int_{\mathbf{M}} \phi \, d\langle P(\,)f, g\rangle.$$

Y esto último se abrevia de la siguiente manera, donde P es la medida espectral,

$$T_\phi = \int_{\mathbf{M}} \phi \, dP$$

y se dice que es una *integral débil*.

Observación 2.74 A la medida espectral P también se le llama *resolución de la identidad* en $\mathbf{M}$.

2.3.2. Otras Versiones del Teorema Espectral

Teorema 2.75 (Teorema Espectral para Álgebras de Operadores) *Existe una correspondecia biunívoca entre el conjunto de álgebras-C^* conmutativas de $L(\mathcal{H})$ y el conjunto de medidas espectrales. Es decir,*

(1) Dada un álgebra-C^* $\mathcal{U} \subseteq L(\mathcal{H})$, existe una única medida espectral P tal que

$$T = \int_{\mathbf{M}} \Gamma T \, dP$$

para todo $T \in \mathcal{U}$.

(2) Dada una medida espectral P existe un álgebra-C^* $\mathcal{U}$ tal que

$$T = \int_{\mathbf{M}} \Gamma T \, dP$$

para todo $T \in \mathcal{U}$.

Demostración:

(1) La existencia de la medida espectral fue probada en el teorema anterior (2.72) al hacer la construcción de la medida espectral. Veamos la unicidad.

Sean P, P' medidas espectrales tales que

$$T = \int_{\mathbf{M}} \Gamma T \, dP, \qquad T = \int_{\mathbf{M}} \Gamma T \, dP'$$

para todo $T \in \mathcal{U}$. Entonces

$$\langle Tf, g\rangle = \int_{\mathbf{M}} \Gamma T \, d\langle P(\,)f, g\rangle = \int_{\mathbf{M}} \Gamma T \, d\langle P'(\,)f, g\rangle$$

para todo $\Gamma T \in C(\mathbf{M})$, y $f, g \in \mathcal{H}$. Luego,

$$\langle P(\triangle)f, g\rangle = \langle P'(\triangle)f, g\rangle$$

para todo $f, g \in \mathcal{H}$ y $\triangle \subseteq \mathbf{M}$. De donde

$$P(\triangle) = P'(\triangle)$$

para todo $\triangle \subseteq \mathbf{M}$. Y así,

$$P = P'.$$

(2) Sea P una medida espectral en $\mathbf{M}$ a valores operadores de $L(\mathcal{H})$.

Consideremos una correspondencia de $\mathbf{B}(\mathbf{M}) \to L(\mathcal{H})$ donde a cada función $\varphi \in \mathbf{B}(\mathbf{M})$, le asociaremos un operador $T_\varphi \in L(\mathcal{H})$. Para ello definimos

$$\langle T_\varphi f, g\rangle = \int_{\mathbf{M}} \varphi \, d\langle P(\,)f, g\rangle.$$

Por teorema de extensión de la inversa de la transformada de Gelfand (2.71) sabemos que lo anterior tiene sentido y que dicha correspondencia tiene, entre otras, las siguientes propiedades:

(i) Es lineal: $T_{\varphi_1+\lambda\varphi_2} = T_{\varphi_1} + \lambda T_{\varphi_2}$, para φ_1, $\varphi_2 \in \mathbf{B}(\mathbf{M})$;

(ii) Es multiplicativa: $T_{\varphi_1\varphi_2} = T_{\varphi_1} T_{\varphi_2}$, para ϕ_1, $\phi_2 \in \mathbf{B}(\mathbf{M})$;

(iii) $(T_\varphi)^* = T_{\overline{\varphi}}$, para $\varphi \in \mathbf{B}(\mathbf{M})$;

(iv) $T_{\phi_1}T_{\phi_2} = T_{\phi_2}T_{\phi_1}$, para $\phi_1 \in \mathbf{B}(\mathbf{M})$ y $\phi_2 \in C(\mathbf{M})$.

Entonces, para la medida espectral P podemos definir $\mathcal{U} = \{T_\varphi : \varphi \in C(\mathbf{M})\}$. Así, para todo $T_\varphi \in \mathcal{U}$ se cumplen las propiedades anteriores, ya que $C(\mathbf{M}) \subset \mathbf{B}(\mathbf{M})$. Veamos que $\mathcal{U}$ es un álgebra autoadjunta cerrada de $L(\mathcal{H})$.

De las propiedades (i) y (ii) (linealidad y multiplicatividad) se desprende el hecho de que $\mathcal{U}$ es un álgebra. Además posee una norma que cumple $\|T_\varphi T_\psi\| \leq \|T_\varphi\| \, \|T_\psi\|$. Y utilizando la parte (2)b) del teorema 2.72, tenemos que la identidad $I = P(\mathbf{M}) = T_{\mathbf{1_M}}$ está en $\mathcal{U}$, ya que $\mathbf{1_M} \in C(\mathbf{M})$.

Sea $T_\varphi \in \mathcal{U}$. Entonces por propiedad (iii) tenemos $T_\varphi^* \in \mathcal{U}$. Luego $\mathcal{U}$ es un álgebra autoadjunta.

Además se puede probar que $\mathcal{U}$ es cerrada. Entonces, en vista de la observación 2.62 tenemos que $\mathcal{U}$ es un álgebra-C^*.

Finalmente, por propiedad (iv) tenemos que $\mathcal{U}$ es conmutativa. Y así $\mathcal{U}$ es el álgebra-C^* conmutativa que buscábamos.

$\square$

El siguiente teorema es muy importante ya que da la razón del nombre: medida espectral. Y, al ser consecuencia directa de las dos versiones del teorema espectral que hemos estudiado, constituye en sí, otra versión de dicho teorema.

Teorema 2.76 (Teorema Espectral para Operadores Normales) *Sea $T \in L(\mathcal{H})$ un operador normal. Sea $\mathcal{A}_T$ el álgebra-C^* generada por T. Sea $\sigma(T)$ el espectro de T como elemento de $\mathcal{A}_T$. Entonces existe una única medida espectral E en $\sigma(T) \subset \mathbb{C}$ tal que*

$$T = \int_{\sigma(T)} z\, dE = \int_{\mathbb{C}} z\, dE$$

es decir, para todo $x, y \in \mathcal{H}$

$$\langle Tx, y\rangle = \int_{\sigma(T)} z\, d\langle E(\,)x, y\rangle.$$

Nota: A veces es conveniente pensar a E definida para todos los subconjuntos de Borel en $\mathbb{C}$. Para lograr esto definimos $E(\triangle) = 0$ si $\triangle \cap \sigma(T) = \varnothing$, donde $\triangle \subset \mathbb{C}$ es un conjunto de Borel.

Demostración:

Por el teorema espectral para álgebras de operadores (2.75) y por el teorema espectral (2.69) tenemos

$$T = \int_{\mathbf{M}} \Gamma T\, dP = \int_{\sigma(T)} I\, dE = \int_{\sigma(T)} z\, dE.$$

□

Observación 2.77 A la medida espectral E definida en $\sigma(T) \subset \mathbb{C}$ también se le llama *descomposición espectral*.

Definición 2.78 (Cálculo Simbólico para Operadores Normales) Sea $T \in L(\mathcal{H})$ un operador normal, $\sigma(T)$ el espectro de T, y E la medida espectral dada por el teorema anterior. Sea $\phi \in \mathbf{B}(\sigma(\mathbf{T}))$ una función medible borel, acotada, definida en $\sigma(T)$. Definimos

$$\phi(T) \equiv \int_{\sigma(T)} \phi(z)\, dE.$$

Observación 2.79 La aplicación $\phi \to \phi(T)$ es un homomorfismo de álgebra de $\mathbf{B}(\sigma(\mathbf{T}))$ en $L(\mathcal{H})$, y

$$\phi(T) = \int_{\sigma(T)} \phi(z)\, dE = \int_{\mathbf{M}} \phi \cdot \Gamma T\, dP = T_{\phi \cdot \Gamma T}.$$

El cálculo simbólico permite definir la raíz cuadrada de un operador (Ver [9], pg 93).

Observación 2.80 (Teorema de Wintner) Sea $U \in L(\mathcal{H})$ un operador unitario. Como U es unitario entonces es normal y aplicando el teorema espectral para operadores normales (2.76) tenemos que existe una única medida espectral E en $\sigma(U) \subset \mathbb{C}$ tal que

$$U = \int_{\sigma(U)} z \, dE. \tag{2.3}$$

Como E es una medida espectral, por teorema 2.72, tenemos que $E(\sigma(U)) = I$ y $E(\varnothing) = 0$.

Ahora, usando la definición 2.3.2, para $\phi(z) = z^n$, tenemos

$$U^n = \int_{\sigma(U)} z^n \, dE. \tag{2.4}$$

Por otro lado, $\sigma(U) \subseteq \mathbb{T}$ (teorema 2.65). Y sabemos que $\mathbb{T} \approx [0, 2\pi]$. Entonces, la ecuación (2.3) también la podemos escribir como

$$U = \int_0^{2\pi} e^{i\theta} \, dE_\theta \tag{2.5}$$

donde $\{E_\theta : 0 \leq \theta \leq 2\pi\}$ es una familia de proyecciones con $E_0 = 0$ y $E_{2\pi} = I$ (medida espectral en $[0, 2\pi]$). Para una prueba se puede ver [2], pg. 492.

Y finalmente, de la ecuación (2.5) y usando un argumento análogo a la definición 2.3.2 (Ver [2], pg. 519) podemos obtener la ecuación

$$U^n = \int_0^{2\pi} e^{in\theta} \, dE_\theta. \tag{2.6}$$

Este resultado también se conoce como teorema de Wintner (otra versión del teorema espectral). (Ver [17], pg. 281).

CAPÍTULO 3

Dilatación Unitaria de una Contracción

3.1. Extensión Unitaria de una Isometría

Definición 3.1 Sea V una isometría definida en un espacio de Hilbert $\mathcal{H}$ y U un operador unitario en un espacio de Hilbert $\mathcal{F}$. Decimos que $(U, \mathcal{F})$ es *extensión unitaria de* V si y sólo si U es una extensión de V tal que $\mathcal{F}$ contiene a $\mathcal{H}$ como subespacio cerrado.

Sea $\bigvee_{n\in\mathbb{Z}} U^n\mathcal{H}$ el espacio vectorial generado por $U^n\mathcal{H}$, con $n \in \mathbb{Z}$. Si además,

$$\mathcal{F} = \bigvee_{n\in\mathbb{Z}} U^n\mathcal{H}$$

decimos que $(U, \mathcal{F})$ es *extensión unitaria minimal de* V.

La definición anterior está basada en el siguiente teorema.

Teorema 3.2 *Si D, R son subespacios cerrados de un espacio de Hilbert $\mathcal{H}$ y $V : D \to R$ es una isometría en $\mathcal{H}$, entonces V puede ser extendida a un operador unitario U en un espacio de Hilbert $\mathcal{F}$ que contiene a $\mathcal{H}$ como subespacio cerrado.*

Demostración:

Sean N y M los complementos ortogonales de D y R respectivamente.

Sea $\mathcal{F} = \mathcal{H} \oplus \mathcal{H} = D \oplus N \oplus R \oplus M$ y definamos el operador U en $\mathcal{F}$ por

$$U(d, n, r, m) = (V^{-1}r, n, Vd, m).$$

Así definida, U es una extensión de V.

Veamos que U es una isometría.

$$\begin{aligned} \|(V^{-1}r, n, Vd, m)\|^2 &= \|V^{-1}r\|^2 + \|n\|^2 + \|Vd\|^2 + \|m\|^2 \\ &= \|r\|^2 + \|n\|^2 + \|d\|^2 + \|m\|^2 \\ &= \|(d, n, r, m)\|^2. \end{aligned}$$

De donde, U es una isometría.

Ahora probemos que U es sobreyectiva. Sea $(d, n, r, m) \in \mathcal{F}$. Tomemos $(V^{-1}r, n, Vd, m) \in \mathcal{F}$. Entonces

$$U(V^{-1}r, n, Vd, m) = (V^{-1}(Vd), n, V(V^{-1}r), m) = (d, n, r, m).$$

Por lo tanto U es sobreyectiva.

Luego, por proposición 2.49 (caracterización de operadores unitarios), tenemos que U es unitario.

$\square$

Definición 3.3 Dadas $(U, \mathcal{F})$ y $(U', \mathcal{F}')$ extensiones unitarias minimales de una isometría V en el espacio de Hilbert $\mathcal{H}$, decimos que son *equivalentes* si y sólo si existe un operador unitario $\varphi \in L(\mathcal{F}, \mathcal{F}')$ tal que

$$\varphi U = U'\varphi \qquad \text{y} \qquad \varphi|_{\mathcal{H}} = I_{\mathcal{H}}.$$

Una Extensión Unitaria Especial de una Isometría

Sea $\mathcal{H}$ un espacio de Hilbert. Sean D, R, N, M subespacios cerrados de $\mathcal{H}$ tales que N y M son los complementos ortogonales de D y R respectivamente. Sea $V : D \to R$ una isometría en $\mathcal{H}$.

Consideremos el siguiente espacio de Hilbert

$$\mathcal{F} = \left\{ (f_j)_{j\in\mathbb{Z}} : f_j \in M \quad \text{si} \quad j < 0, \quad f_0 \in \mathcal{H}, \quad f_j \in N \quad \text{si} \quad j > 0; \quad \sum_{j\in\mathbb{Z}} \|f_j\|^2 < \infty \right\}$$

con el producto escalar

$$\langle f, g\rangle = \sum_{j\in\mathbb{Z}} \langle f_j, g_j\rangle_{\mathcal{H}}.$$

Notación 3.4 Los elementos $f = (f_j)_{j\in\mathbb{Z}}$ de $\mathcal{F}$ los escribiremos de la forma $f = (..., f_{-1}, \overbrace{f_0}, f_1, ...)$, usando la llave para indicar el lugar correspondiente a $j = 0$.

Tomando en cuenta al espacio $\mathcal{F}$ podemos identificar a $\mathcal{H}$ como subespacio de $\mathcal{F}$ mediante el conjunto $\{f \in \mathcal{F} : f_j = 0 \quad \text{si} \quad j \neq 0\}$. Así, $\mathcal{H}$ corresponde a los elementos de la forma $f = (..., 0, \overbrace{h}, 0, ...)$.

Ahora definamos $U : \mathcal{F} \to \mathcal{F}$ por

$$(Uf)_j = \begin{cases} f_{j-1}, & \text{si} \quad j < 0; \\ f_{-1} + VP_D f_0, & \text{si} \quad j = 0; \\ P_N f_0, & \text{si} \quad j = 1; \\ f_{j-1}, & \text{si} \quad j > 1; \end{cases}$$

con $f = (..., f_{-1}, \overbrace{f_0}, f_1, ...)$. Es decir,

$$Uf = (..., f_{-3}, f_{-2}, \overbrace{f_{-1} + VP_D f_0}, P_N f_0, f_1, f_2, ...).$$

Proposición 3.5 *Sean U y V los definidos anteriormente. Entonces $(U, \mathcal{F})$ es extensión unitaria minimal de V y $(VP_D)^n = P_{\mathcal{H}} U^n|_{\mathcal{H}}$ para todo $n \geq 0$.*

Demostración:

Primero veamos que U es una extensión de V. Sea $h \in D$. Tenemos

$$U(..., 0, \overbrace{h}, 0, ...) = (..., 0, \overbrace{VP_D h}, P_N h, 0, ...) = (..., 0, \overbrace{Vh}, 0, ...).$$

De donde, $P_{\mathcal{H}} U|_D = V$.

Haciendo uso de la caracterización de los operadores unitarios (proposición 2.49), probemos que U es un operador unitario.

Veamos que U es isométrico.

$$\begin{aligned}
\|U(..., f_{-1}, \overbrace{h}, f_1, ...)\|^2 &= \|(..., f_{-2}, \overbrace{f_{-1} + VP_D h}, P_N h, f_1, ...)\|^2 \\
&= \sum_{n\leq -2} \|f_n\|^2 + \|f_{-1} + VP_D h\|^2 + \|P_N h\|^2 + \sum_{n\geq -1} \|f_n\|^2 \\
&= \sum_{n\leq -2} \|f_n\|^2 + \|f_{-1}\|^2 + \|VP_D h\|^2 + \|P_N h\|^2 + \sum_{n\geq -1} \|f_n\|^2 \\
&= \sum_{n\leq -1} \|f_n\|^2 + \|P_D h\|^2 + \|P_N h\|^2 + \sum_{n\geq -1} \|f_n\|^2 \\
&= \sum_{n\leq -1} \|f_n\|^2 + \|h\|^2 + \sum_{n\geq -1} \|f_n\|^2 = \|(..., f_{-1}, \overbrace{h}, f_1, ...)\|^2
\end{aligned}$$

Por lo tanto U es isométrico.

Ahora veamos que U es sobreyectivo. Sea $f = (..., f_{-1}, \overbrace{h}, f_1, ...) \in \mathcal{F}$. Tomemos $f' = (..., f'_{-1}, \overbrace{f'_0}, f'_1, ...)$ dada por

$$f'_n = \begin{cases} f_{n+1}, & \text{si} \quad n \leq -2; \\ P_M h, & \text{si} \quad n = -1; \\ f_1 + V^* P_R h, & \text{si} \quad n = 0; \\ f_{n+1}, & \text{si} \quad n \geq 1. \end{cases}$$

Entonces

$$\begin{aligned}
Uf' &= U(..., f'_{-1}, \overbrace{f'_0}, f'_1, ...) = U(..., f_{-1}, P_M h, \overbrace{f_1 + V^* P_R h}, f_2, ...) \\
&= (..., f_{-1}, \overbrace{P_M h + VP_D(f_1 + V^* P_R h)}, P_N(f_1 + V^* P_R h), f_2, ...) \\
&= (..., f_{-1}, \overbrace{P_M h + VP_D f_1 + VP_D V^* P_R h}, P_N f_1 + P_N V^* P_R h, f_2, ...) \qquad (3.1)
\end{aligned}$$

Pero, $VP_D f_1 = 0$ y $P_N f_1 = f_1$, ya que $f_1 \in N$. Además, como N es el complemento ortogonal de D y $V^* : R \to D$ entonces $P_N V^* = 0$. Ahora veamos que $VP_D V^* = VV^* = I_R$, donde I_R es la identidad en R.

La primera igualdad es obvia ya que $V^*: R \to D$. Probemos la segunda usando el hecho que si V es isometría siempre se cumple que $V^*V = I_D$. Sea $g \in D$ y $f = Vg$.

$$\langle f, VV^*f\rangle = \langle V^*f, V^*f\rangle = \langle V^*Vg, V^*Vg\rangle = \langle Vg, VV^*Vg\rangle = \langle Vg, Vg\rangle = \langle f, f\rangle,$$

de donde $VV^*f = f$ para todo $f \in R$, es decir, $VV^* = I_R$.

Luego, siguiendo con las igualdades de (3.1) tenemos

$$Uf' = (..., f_{-1}, \overbrace{P_M h + P_R h}, f_1, f_2, ...) = (..., f_{-1}, h, f_1, ...)$$

De donde, U es sobreyectivo.

A continuación probemos inductivamente que $P_{\mathcal{H}}U^n|_{\mathcal{H}} = (VP_D)^n$ para todo $n \geq 0$.

Veamos que la igualdad se cumple para $n = 1$. Ya probamos que $P_{\mathcal{H}}U|_D = V = VP_D|_D$, es decir, $P_{\mathcal{H}}U|_D = VP_D|_D$. Nos falta probar que $P_{\mathcal{H}}U|_N = VP_D|_N = 0$.

Si $h \in N$,

$$U(..., 0, \overbrace{\ h\ }, 0, ...) = (..., 0, \overbrace{VP_D h}, P_N h, 0, ...)$$

y como $P_D h = 0$, entonces $P_{\mathcal{H}}U|_N = 0$. Además, es obvio que $VP_D|_N = 0$, entonces tenemos

$$P_{\mathcal{H}}U|_N = VP_D|_N = 0.$$

Por lo tanto, $P_{\mathcal{H}}U|_{\mathcal{H}} = VP_D$, es decir, se cumple para $n = 1$.

Supongamos que la igualdad se cumple para $n = k$ y veamos que es válida para $n = k+1$.

Como $P_{\mathcal{H}} + P_{\mathcal{F}\ominus\mathcal{H}} = I$ tenemos

$$\begin{aligned} P_{\mathcal{H}}U^{k+1}|_{\mathcal{H}} &= P_{\mathcal{H}}UP_{\mathcal{H}}U^k|_{\mathcal{H}} + P_{\mathcal{H}}UP_{\mathcal{F}\ominus\mathcal{H}}U^k|_{\mathcal{H}} \\ &= P_{\mathcal{H}}U|_{\mathcal{H}}(VP_D)^k + P_{\mathcal{H}}UP_{\mathcal{F}\ominus\mathcal{H}}U^k|_{\mathcal{H}} \\ &= VP_D(VP_D)^k + P_{\mathcal{H}}UP_{\mathcal{F}\ominus\mathcal{H}}U^k|_{\mathcal{H}} \\ &= (VP_D)^{k+1} + P_{\mathcal{H}}UP_{\mathcal{F}\ominus\mathcal{H}}U^k|_{\mathcal{H}} \end{aligned}$$

Debemos probar que $P_{\mathcal{H}}UP_{\mathcal{F}\ominus\mathcal{H}}U^k|_{\mathcal{H}} = 0$. Por inducción.

Veamos para $k = 1$.

$$\begin{aligned} P_{\mathcal{H}}UP_{\mathcal{F}\ominus\mathcal{H}}U(...,0,\widehat{h},0,...) &= P_{\mathcal{H}}UP_{\mathcal{F}\ominus\mathcal{H}}(...,0,\widehat{VP_Dh},P_Nh,0,...) \\ &= P_{\mathcal{H}}U(...,0,\widehat{0},P_Nh,0,...) \\ &= P_{\mathcal{H}}(...,0,\widehat{0},0,P_Nh,0,...) = 0. \end{aligned}$$

Supongamos que se cumple para k y probemos para $k+1$.

$$\begin{aligned} P_{\mathcal{H}}UP_{\mathcal{F}\ominus\mathcal{H}}U^{k+1}(...,0,h,0,...) &= P_{\mathcal{H}}UP_{\mathcal{F}\ominus\mathcal{H}}U^k(...,0,\widehat{VP_Dh},P_Nh,0,...) \\ &= P_{\mathcal{H}}UP_{\mathcal{F}\ominus\mathcal{H}}U^k(...,0,\widehat{VP_Dh},0,...)+ \\ &+ P_{\mathcal{H}}UP_{\mathcal{F}\ominus\mathcal{H}}U^k(...,0,\widehat{0},P_Nh,0,...) \\ &= P_{\mathcal{H}}UP_{\mathcal{F}\ominus\mathcal{H}}U^k(...,0,\widehat{0},P_Nh,0,...) \\ &= P_{\mathcal{H}}UP_{\mathcal{F}\ominus\mathcal{H}}U^{k-1}(...,0,\widehat{0},0,P_Nh,0,...) \\ &\vdots \\ &= P_{\mathcal{H}}UP_{\mathcal{F}\ominus\mathcal{H}}(...,\widehat{0},0,...,0,P_Nh,0,...) \\ &= P_{\mathcal{H}}U(...,\widehat{0},0,...,0,P_Nh,0,...) \\ &= P_{\mathcal{H}}(...,\widehat{0},0,...,0,P_Nh,0,...) = 0. \end{aligned}$$

Por lo tanto queda demostrado para todo $k \in \mathbb{N}$ que $P_{\mathcal{H}}UP_{\mathcal{F}\ominus\mathcal{H}}U^k|_{\mathcal{H}} = 0$ y con esto, finalmente queda demostrado que $P_{\mathcal{H}}U^n|_{\mathcal{H}} = (VP_D)^n$ para todo $n \in \mathbb{N}$. □

3.2. Teorema de Sz.-Nagy

Definición 3.6 Sea $\mathcal{H}$ un espacio de Hilbert. Se dice que $T \in L(\mathcal{H})$ es una *contracción* si $\|T\| \leq 1$.

Definición 3.7 Se dice que $W \in L(\mathcal{M})$ es una *dilatación isométrica minimal* de una contracción $T \in L(\mathcal{H})$ si $\mathcal{M}$ es un espacio de Hilbert que contiene a $\mathcal{H}$ como subespacio cerrado y W es una isometría tal que

$$T^n = P_{\mathcal{H}}W^n|_{\mathcal{H}}, \text{ para todo } n \geq 0 \qquad \text{y} \qquad \mathcal{M} = \bigvee_{n\geq 0} W^n\mathcal{H}\ .$$

Definición 3.8 Se dice que $U \in L(\mathcal{F})$ es una *dilatación unitaria minimal* de una contracción $T \in L(\mathcal{H})$ si $\mathcal{F}$ es un espacio de Hilbert que contiene a $\mathcal{H}$ como subespacio cerrado y U es un operador unitario tal que

$$T^n = P_{\mathcal{H}} U^n|_{\mathcal{H}}, \text{ para todo } n \geq 0 \qquad \text{y} \qquad \mathcal{F} = \bigvee_{n \in \mathbb{Z}} U^n \mathcal{H} \ .$$

Teorema 3.9 (Sz.-Nagy) *Si $\mathcal{K}$ es un espacio de Hilbert, entonces toda contracción $T \in L(\mathcal{K})$ tiene una dilatación isométrica minimal $W \in L(\mathcal{M})$ y una dilatación unitaria minimal $U \in L(\mathcal{F})$, cada una de las cuales es única por isomorfismos unitarios.*

Demostración:

Sea $T \in L(\mathcal{K})$ una contracción, $D_T = (I - T^*T)^{\frac{1}{2}}$ el operador por defecto de T y $\mathcal{D}_T$ la clausura del rango de D_T.

Definamos la isometría $V : \mathcal{K} \to \mathcal{H}$, donde $\mathcal{H} = \mathcal{K} \oplus \mathcal{D}_T$ por

$$Vh = (Th, D_T h), \qquad \text{para todo } h \in \mathcal{K}.$$

Extendamos a $(V, \mathcal{K})$ al par $(U, \mathcal{F})$ como en la proposición 3.5. Sabemos entonces que U es extensión unitaria minimal de V. Veamos que $T^n = P_{\mathcal{K}} U^n|_{\mathcal{K}}$.

Por la igualdad de la proposición 3.5 tenemos que

$$P_{\mathcal{K} \oplus \mathcal{D}_T} U^n|_{\mathcal{K} \oplus \mathcal{D}_T} = (VP_{\mathcal{K}})^n.$$

Luego, usando que $P_{\mathcal{K}} = P_{\mathcal{K}} P_{\mathcal{K} \oplus \mathcal{D}_T}$ tenemos

$$P_{\mathcal{K}} U^n|_{\mathcal{K}} = P_{\mathcal{K}} P_{\mathcal{K} \oplus \mathcal{D}_T} U^n|_{\mathcal{K}} = P_{\mathcal{K}} (VP_{\mathcal{K}})^n|_{\mathcal{K}}. \tag{3.2}$$

Entonces, probemos por inducción que

$$P_{\mathcal{K}} (VP_{\mathcal{K}})^n|_{\mathcal{K}} = T^n.$$

Veamos para $n = 1$. Sea $h \in \mathcal{K}$, tenemos

$$P_{\mathcal{K}} (VP_{\mathcal{K}})h = P_{\mathcal{K}} Vh = P_{\mathcal{K}} (Th, D_T h) = Th.$$

Supongamos para $n = k$,

$$\begin{aligned} P_{\mathcal{K}}(VP_{\mathcal{K}})^{k+1}|_{\mathcal{K}}(h) &= P_{\mathcal{K}}VP_{\mathcal{K}}(VP_{\mathcal{K}})^{k}|_{\mathcal{K}}(h) = P_{\mathcal{K}}VT^{k}|_{\mathcal{K}}(h) \\ &= P_{\mathcal{K}}VT^{k}(h) = P_{\mathcal{K}}(TT^{k}h, D_{T}T^{k}h) = T^{k+1}h. \end{aligned}$$

Hemos probado que $T^n = P_{\mathcal{K}}U^n|_{\mathcal{K}}$. Luego U es dilatación unitaria minimal de T.

Fijemos $\mathcal{M} = \bigvee_{n\geq 0} U^n\mathcal{K}$ y $W = U|_{\mathcal{M}}$, y comprobemos que $W \in L(\mathcal{M})$ es dilatación isométrica de T.

Como $W = U|_{\mathcal{M}}$ y U es isométrico, entonces W es isométrico. Además $W(\mathcal{M}) = U(\mathcal{M}) \subset \mathcal{M}$, y $W^n = U^n|_{\mathcal{M}}$ para todo $n \geq 0$. Probemos que $T^n = P_{\mathcal{K}}W^n|_{\mathcal{K}}$ para todo $n \geq 0$.

Como ya probamos que U es dilatación, entonces para todo $n \geq 0$ tenemos

$$T^n = P_{\mathcal{K}}U^n|_{\mathcal{K}} = P_{\mathcal{K}}W^n|_{\mathcal{K}}.$$

Luego W es dilatación isométrica minimal de T.

Sólo nos falta probar la unicidad.

Sean $(U, \mathcal{F})$ y $(U', \mathcal{F}')$ dos dilataciones unitarias minimales de T, es decir,

$$\mathcal{F} = \bigvee_{n\in\mathbb{Z}} U^n\mathcal{K} \quad \text{y} \quad \mathcal{F}' = \bigvee_{n\in\mathbb{Z}} U'^{\,n}\mathcal{K}.$$

Sea $\varphi : \mathcal{F} \to \mathcal{F}'$ un operador definido sobre las combinaciones lineales finitas $\sum_{|n|\leq s} \alpha_n U^n h_n$, con $h_n \in \mathcal{K}$ y $\alpha_n \in \mathbb{C}$, como sigue

$$\varphi\left(\sum_{|n|\leq s} \alpha_n U^n h_n\right) = \sum_{|n|\leq s} \alpha_n U'^{\,n} h_n.$$

Comprobemos que φ está bien definido, para ello basta probar que es isométrico.

$$\left\|\varphi\left(\sum_{|n|\leq s} \alpha_n U^n h_n\right)\right\|^2 = \left\|\sum_{|n|\leq s} \alpha_n U'^{\,n} h_n\right\|^2 = \left\langle \sum_{|n|\leq s} \alpha_n U'^{\,n} h_n, \sum_{|n|\leq s} \alpha_n U'^{\,n} h_n \right\rangle$$

$$= \sum_{|n|\leq s} \alpha_n \sum_{|m|\leq s} \overline{\alpha_m}\, \langle U'^{\,n} h_n, U'^{\,m} h_m\rangle$$

Para que se cumpla la isometría debemos demostrar que

$$\langle U'^{\,n} h_n, U'^{\,m} h_m \rangle_{\mathcal{F}'} = \langle U^n h_n, U^m h_m \rangle_{\mathcal{F}},$$

para ello, basta ver que si $k \geq 0$ y $h, g \in \mathcal{K}$, se verifica que

$$\langle U'^{\,k} h, g \rangle_{\mathcal{F}'} = \langle U^k h, g \rangle_{\mathcal{F}},$$

y en efecto, para $k \geq 0$ y $h, g \in \mathcal{K}$ tenemos

$$\begin{aligned}
\langle U'^{\,k} h, g \rangle_{\mathcal{F}'} &= \langle U'^{\,k} h, P_{\mathcal{K}} g \rangle_{\mathcal{F}'} = \langle P_{\mathcal{K}} U'^{\,k}|_{\mathcal{K}} h, g \rangle_{\mathcal{F}'} \\
&= \langle T^k h, g \rangle_{\mathcal{K}} = \langle P_{\mathcal{K}} U^k|_{\mathcal{K}} h, g \rangle_{\mathcal{F}} \\
&= \langle U^{\,k} h, P_{\mathcal{K}} g \rangle_{\mathcal{F}} = \langle U^k h, g \rangle_{\mathcal{F}}.
\end{aligned}$$

Hemos demostrado que φ es isométrico. Y como φ es sobreyectivo, tenemos que es unitario.

Por otro lado, si tomamos una combinación lineal finita $\sum_{|n|\leq s} \alpha_n U^n h_n$, con $h_n \in \mathcal{K}$, $\alpha_0 = 1$ y $\alpha_n = 0$ para todo $n \neq 0$ tenemos que

$$\varphi\left(\sum_{|n|\leq s} \alpha_n U^n h_n\right) = \varphi(h_0) = h_0$$

para todo $h_0 \in \mathcal{K}$. De donde,

$$\varphi|_{\mathcal{K}} = I_{\mathcal{K}}.$$

Por lo tanto U y U' son equivalentes. $\square$

CAPÍTULO 4

Dilataciones Unitarias y Cadenas de Markov

"A mathematician is a person who can find analogies between theorems; a better mathematician is one who can see analogies between proofs and the best mathematician can notice analogies between theories. One can imagine that the ultimate mathematician is one who can see analogies between analogies". **Banach, Stefan.**

4.1. Cadenas de Markov

Consideremos el conjunto $\mathcal{E}= \{1, 2, 3, ...\}$. Sea $\{p_{jk} : j, k \in \mathcal{E}\}$ un conjunto de números reales que satisfacen las relaciones

$$p_{jk} \geq 0, \qquad \sum_k p_{jk} = 1$$

para todo $j, k \in \mathcal{E}$. Así, la matriz $P = \{p_{jk}\}$, es una matriz estocástica o de Markov sobre $\mathcal{E}$, con la cual podemos definir una cadena de Markov $X = \{X_n\}$. Sea p^n_{jk} la entrada (j, k) de la n-ésima potencia de la matriz P. Entonces podemos identificar p^n_{jk} con la probabilidad de que un proceso Markoviano, homogéneo en tiempo, inicialmente en el estado j pase al estado k en la n-ésima transición.

Observación 4.1 Sólo consideraremos cadenas de Markov irreducibles.

4.2. Medida Subinvariante Positiva

Definición 4.2 Decimos que una cadena de Markov irreducible con matriz de transición $P = \{p_{jk}\}$ *admite una medida invariante positiva* $\{m_k : k = 1, 2, ...\}$ cuando existen $m_1, m_2, ...$ números reales positivos tales que

$$\sum_j m_j p_{jk} = m_k, \qquad \text{con} \quad k = 1, 2, ...; \tag{4.1}$$

Observación 4.3 No es requerido que $\sum m_j$ converja. Derman (Ver [7]) probó que el conjunto de soluciones de (4.1) es no vacío y que es generado por un único elemento cuando la cadena es irreducible y recurrente. También probó mediante ejemplos que si la cadena es irreducible y no recurrente, el conjunto de soluciones puede ser vacío o puede contener elementos linealmente dependientes (Ver [8]).

Sólo estamos interesados en una analogía más débil que (4.1), definida a continuación.

Definición 4.4 Decimos que una cadena de Markov irreducible con matriz de transición $P = \{p_{jk}\}$ *admite una medida subinvariante positiva* $\{m_k : k = 1, 2, ...\}$ cuando existen $m_1, m_2, ...$ números reales positivos tales que

$$\sum_j m_j p_{jk} \leq m_k, \qquad \text{con} \quad k = 1, 2, ... \tag{4.2}$$

Observación 4.5 De los resultados de Derman se obtiene que la solución de (4.2) no es necesariamente única.

El resultado principal de esta sección es el siguiente teorema.

Teorema 4.6 (Kendall) *Toda cadena de Markov irreducible admite al menos una medida subinvariante positiva.*

Demostración:

Sea X una cadena de Markov irreducible, con el espacio de estados $\mathcal{E}$, y $P = \{p_{jk}\}$ su matriz de transición. Sea $a \in \mathcal{E}$ un estado arbitrario.

Consideremos $\phi(a, k, a)$ como la probabilidad de que el proceso X, inicialmente en el estado a, entre al estado k por lo menos una vez antes de que regrese al estado a (si acaso regresa). Si $k = a$, entonces $\phi(a, a, a)$ es simplemente la probabilidad de que, comenzando en el estado a, regrese a a (Ver figura 4.1, donde los círculos respresentan a los estados, las flechas a las transiciones de estados, los puntos suspensivos significan que la cadena va o viene de otro(s) estado(s) y X_0 indica el estado inicial).

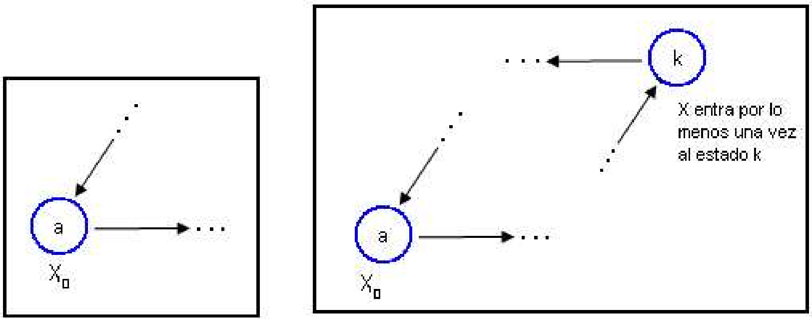

Figura 4.1: Eventos correspondientes a $\phi(a, a, a)$ y $\phi(a, k, a)$.

De manera similar, sea $\phi(k, -a, k)$ la probabilidad de que la cadena inicialmente en el estado k retorne al estado k sin antes pasar por el estado a (Ver figura 4.2).

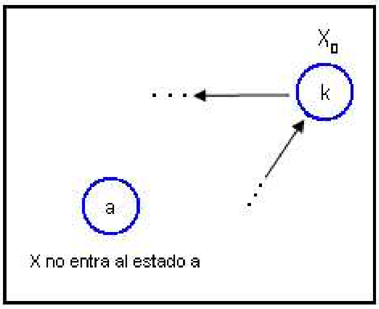

Figura 4.2: Evento correspondiente a $\phi(k, -a, k)$.

Sea $\mathbf{E}(a, k, a)$ la esperanza del número de veces que la cadena, comenzando en a, entre al estado k antes de regresar al estado a (si acaso regresa) (Ver figura 4.3).

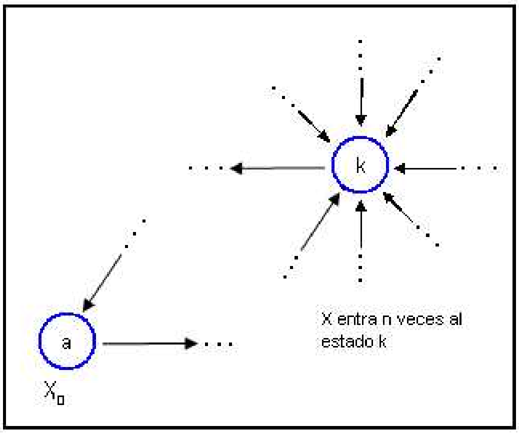

Figura 4.3: Evento correspondiente a $\mathbf{E}(a, k, a)$.

Sea N_{aka} definida para $k \in \mathcal{E}$ de la siguiente manera:

(i) Si $k = a$, entonces $N_{aaa} = \phi(a, a, a)$.

(ii) Si $k \neq a$, entonces $N_{aka} = \mathbf{E}(a, k, a)$

Si $k \neq a$, usando un resultado análogo a una proposición en [6] (pg. 123) se puede ver que

$$\phi(a, k, a)(\phi(k, -a, k))^{n-1}(1 - \phi(k, -a, k))$$

es la probabilidad de que un sistema inicialmente en el estado a entre n veces al estado k antes de regresar al estado a. (Nótese que $1 - \phi(k, -a, k) = \phi(k, a, k)$ es la probabilidad de que la cadena, inicialmente en el estado k, entre al estado a por lo menos una vez antes de regresar al estado k).

Sabiendo esto, podemos hallar una expresión para $N_{aka} = \mathbf{E}(a, k, a)$ en términos de $\phi(a, k, a)$ y $\phi(k, -a, k)$.

Entonces, de la definición de esperanza de una variable aleatoria discreta tenemos

$$
\begin{aligned}
N_{aka} = \mathbf{E}(a,k,a) &= \sum_{n=1}^{\infty} n\,\phi(a,k,a)(\phi(k,-a,k))^{n-1}(1-\phi(k,-a,k)) \\
&= \sum_{n=0}^{\infty} n\,\phi(a,k,a)(\phi(k,-a,k))^{n-1}(1-\phi(k,-a,k)) \\
&= \phi(a,k,a)(1-\phi(k,-a,k)) \sum_{n=0}^{\infty} n\,(\phi(k,-a,k))^{n-1} \\
&= \phi(a,k,a)(1-\phi(k,-a,k)) \sum_{n=0}^{\infty} \frac{d}{d\,\phi(k,-a,k)}(\phi(k,-a,k))^{n} \\
&= \phi(a,k,a)(1-\phi(k,-a,k)) \frac{d}{d\,\phi(k,-a,k)} \sum_{n=0}^{\infty}(\phi(k,-a,k))^{n} \\
&= \phi(a,k,a)(1-\phi(k,-a,k)) \frac{d}{d\,\phi(k,-a,k)} \left(\frac{1}{1-\phi(k,-a,k)}\right) \\
&= \phi(a,k,a)(1-\phi(k,-a,k)) \frac{1}{(1-\phi(k,-a,k))^2} \\
&= \frac{\phi(a,k,a)}{1-\phi(k,-a,k)}.
\end{aligned}
\tag{4.3}
$$

Justifiquemos algunos pasos de la cuenta anterior. Usando proposición 1.10, como la cadena es irreducible tenemos que

$$\phi(a,k,a) \neq 0, \quad \phi(k,-a,k) \neq 0 \quad \text{ y } \quad \phi(k,-a,k) \neq 1.$$

Por lo tanto el primer término de la serie ($n = 0$) es cero y el intercambio de la derivada con la sumatoria es posible porque la serie converge ($0 < \phi(k,-a,k) < 1$). De donde, N_{aka} es finito y positivo para todo $k \in \mathcal{E}$.

Sea $q^n(a,-a,k)$ la probabilidad de que el sistema inicialmente en el estado a entre al estado k en la n-ésima transición, sin haber entrado antes al estado a (Ver figura 4.4).

Obsérvese que si $n = 1$ ocurre que

$$q(a,-a,k) = P\{X_1 = k | X_0 = a\} = p_{ak}. \tag{4.4}$$

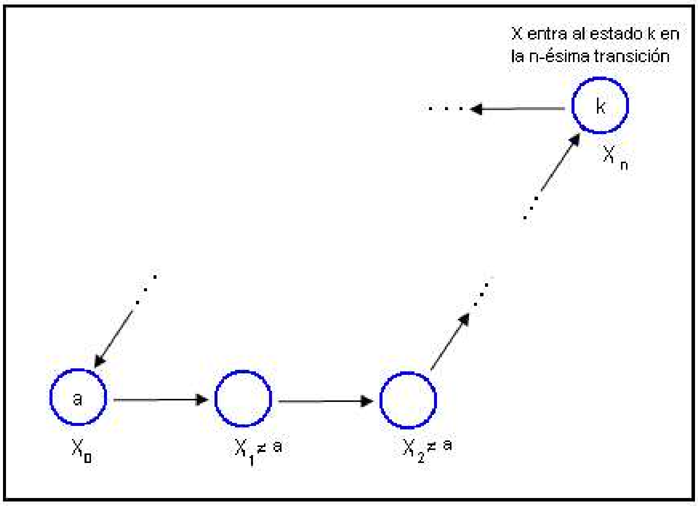

Figura 4.4: Evento correspondiente a $q^n(a, -a, k)$.

Si $n = 2$, usando un procedimiento análogo a una parte de la demostración de la proposición 1.6, y la ecuación anterior (4.4) nos queda

$$\begin{aligned} q^2(a, -a, k) &= P\{X_2 = k, X_1 \neq a | X_0 = a\} \\ &= \sum_{j \neq a} p_{aj}\, p_{jk} = \sum_{j \neq a} q(a, -a, j)\, p_{jk}. \end{aligned}$$

Y en general, obtenemos

$$\begin{aligned} q^{n+1}(a, -a, k) &= P\{X_{n+1} = k, X_i \neq a; i = 1, ..., n | X_0 = a\} \\ &= \sum_{j \neq a} P\{X_n = j, X_i \neq a; i = 1, ..., n | X_0 = a\}\, p_{jk} \\ &= \sum_{j \neq a} q^n(a, -a, j)\, p_{jk}. \end{aligned}$$

Es decir, para todo $n > 1$ tenemos

$$q^{n+1}(a, -a, k) = \sum_{j \neq a} q^n(a, -a, j)\, p_{jk}. \tag{4.5}$$

Por otra parte, tenemos que $N_{aka} = \frac{\phi(a,k,a)}{1-\phi(k,-a,k)}$ (por ecuación (4.3)), pero también podemos obtener la expresión

$$N_{aka} = \sum_{n=1}^{\infty} q^n(a, -a, k) \tag{4.6}$$

Para probarla, denotemos por $\mathbb{E}[\cdot|\cdot]$ la esperanza condicional de una variable aleatoria y por $\mathbf{1}_k(X_n; X_i \neq a, i = 1, ..., n-1)$ la función igual a 1, si $X_n = k$ con $X_i \neq a$, para todo $i = 1, ..., n-1$, y 0, en otro caso. Por teorema de convergencia monótona (intercambio de series de términos positivos), tenemos

$$\begin{aligned}
N_{aka} &= \mathbb{E}\left[\sum_{n=1}^{\infty} \mathbf{1}_k(X_n; X_i \neq a, i = 1, ..., n-1) | X_0 = a\right] \\
&= \sum_{j=1}^{\infty}\sum_{n=1}^{\infty} \mathbf{1}_k(j) P\{X_n = j; X_i \neq a, i = 1, ..., n-1) | X_0 = a\} \\
&= \sum_{j=1}^{\infty} \lim_{m\to\infty} \sum_{n=1}^{m} \mathbf{1}_k(j) P\{X_n = j; X_i \neq a, i = 1, ..., n-1) | X_0 = a\} \\
&= \lim_{m\to\infty} \sum_{n=1}^{m}\sum_{j=1}^{\infty} \mathbf{1}_k(j) P\{X_n = j; X_i \neq a, i = 1, ..., n-1) | X_0 = a\} \\
&= \sum_{n=1}^{\infty} P\{X_n = k; X_i \neq a, i = 1, ..., n-1) | X_0 = a\} = \sum_{n=1}^{\infty} q^n(a, -a, k).
\end{aligned}$$

Finalmente, de las ecuaciones (4.6), (4.5), (4.4), y usando nuevamente el teorema de convergencia monótona (intercambio de series de términos positivos), tenemos

$$\begin{aligned}
\sum_j N_{aja}\, p_{jk} &= N_{aaa}\, p_{ak} + \sum_{j\neq a} N_{aja}\, p_{jk} = \phi(a,a,a)\, p_{ak} + \sum_{j\neq a}\left(\sum_{n=1}^{\infty} q^n(a,-a,j)\right) p_{jk} \\
&= \phi(a,a,a)\, p_{ak} + \sum_{n=1}^{\infty} q^{n+1}(a,-a,k) = \phi(a,a,a)\, p_{ak} + \sum_{m=2}^{\infty} q^m(a,-a,k) \\
&\leq p_{ak} + \sum_{m=2}^{\infty} q^m(a,-a,k) = q(a,-a,k) + \sum_{m=2}^{\infty} q^m(a,-a,k) \\
&= \sum_{m=1}^{\infty} q^m(a,-a,k) = N_{aka}.
\end{aligned}$$

Y tomando $m_j = N_{aja}$, con $j = 1, 2, ...$ y a fijo, tenemos que

$$\sum_j m_j\, p_{jk} \leq m_k,$$

de donde $\{m_j : j = 1, 2, ...\}$ es una medida subinvariante positiva asociada a $P = \{p_{jk}\}$. $\square$

4.3. Dilataciones Unitarias

Sea P una matriz que define una cadena de Markov irreducible. Por el teorema anterior (4.6) existe $\{m_j : j = 1, 2, ...\}$ una medida subinvariante positiva fija asociada a P.

Consideremos el espacio de Hilbert complejo

$$\ell_2 = \left\{ x = (x_j)_{j=1}^{\infty} : x_j \in \mathbb{C} \quad \text{y} \quad \sum_j |x_j|^2 < \infty \right\},$$

con el producto interno

$$\langle x, y \rangle = \sum_j x_j \overline{y_j}.$$

Sea $T : \ell_2 \longrightarrow \ell_2$ la transformación lineal tal que la k-ésima componente de Tx está dada por la serie

$$(Tx)_k = \sum_j x_j (m_j/m_k)^{\frac{1}{2}} p_{jk}.$$

Probemos que esta serie es absolutamente convergente. Usando la desigualdad de Cauchy-Schwarz y la definición de medida subinvariante positiva, tenemos que

$$\begin{aligned}
\left(\sum_j |x_j (m_j/m_k)^{\frac{1}{2}} p_{jk}| \right)^2 &= \left(\sum_j |x_j| (m_j/m_k)^{\frac{1}{2}} p_{jk} \right)^2 \\
&\leq \left(\sum_n |x_n|^2 p_{nk} \right) \left(\sum_i (m_i/m_k) p_{ik} \right) \\
&= \left(\sum_n |x_n|^2 p_{nk} \right) (1/m_k) \left(\sum_i m_i p_{ik} \right) \\
&\leq \left(\sum_n |x_n|^2 p_{nk} \right) (1/m_k)\, m_k \\
&= \sum_n |x_n|^2 p_{nk} \\
&\leq \sum_n |x_n|^2 < \infty
\end{aligned}$$

Además, usando los argumentos anteriores, tenemos

$$\begin{aligned}
\|Tx\|^2 = \langle Tx, Tx \rangle &= \sum_k \left| \sum_j x_j (m_j/m_k)^{\frac{1}{2}} p_{jk} \right|^2 \\
&\leq \sum_k \left(\sum_n |x_n|^2 p_{nk} \right) \left(\sum_i (m_i/m_k) p_{ik} \right) \\
&\leq \sum_n \left(\sum_k |x_n|^2 p_{nk} \right) = \|x\|^2 \qquad (4.7)
\end{aligned}$$

por lo tanto T es acotado. Luego, T es un operador en ℓ_2, es decir, $T \in L(\ell_2)$. Más aún, T es una contracción (i.e,$\|T\| \leq 1$).

Ahora podemos aplicar el teorema de Sz.-Nagy (3.9) a la contracción T que acabamos de definir. Entonces T tiene una dilatación unitaria minimal $U \in L(\mathcal{F})$, donde $\mathcal{F}$ es un espacio de Hilbert que contiene a ℓ_2 como subespacio. Luego, por definición, se cumple que

$$T^n x = P_{\ell_2} U^n x, \quad \text{para todo } x \in \ell_2 \text{ y } n \geq 0,$$

donde P_{ℓ_2} es la proyección de $\mathcal{F}$ sobre ℓ_2.

Además, si $x \in \ell_2$ y $n \geq 0$, tenemos

$$\begin{aligned}
\langle x, (T^*)^n x \rangle &= \langle x, (T^n)^* x \rangle = \langle T^n x, x \rangle \\
&= \langle P_{\ell_2} U^n x, x \rangle = \langle x, P_{\ell_2} (P_{\ell_2} U^n)^* x \rangle \\
&= \langle x, P_{\ell_2} (U^n)^* P_{\ell_2}^* x \rangle = \langle x, P_{\ell_2} (U^*)^n P_{\ell_2} x \rangle \\
&= \langle x, P_{\ell_2} U^{-n} x \rangle
\end{aligned}$$

De donde, $(T^*)^n x = P_{\ell_2} U^{-n} x$ para todo $x \in \ell_2$ y $n \geq 0$.

Sea $e_j \in \ell_2$ el elemento definido por

$$(e_j)_k = \delta_{jk},$$

es decir, $e_j = (0, .., 0, 1, 0, ...)$ donde el 1 se encuentra en la j-ésima componente. Entonces por un lado tenemos que

$$\langle T^n e_j, e_k \rangle = \langle P_{\ell_2} U^n e_j, e_k \rangle = \langle U^n e_j, P_{\ell_2} e_k \rangle = \langle U^n e_j, e_k \rangle$$

de donde

$$\langle T^n e_j, e_k\rangle = \langle U^n e_j, e_k\rangle \tag{4.8}$$

para todo $j, k = 1, 2, ...$ y todo $n \in \mathbb{N}$ y por otro lado tenemos que

$$\langle e_j, T^n e_k\rangle = \langle e_j, P_{\ell_2} U e_k\rangle = \langle U^{-n} P_{\ell_2} e_j, e_k\rangle = \langle U^{-n} e_j, e_k\rangle,$$

de donde

$$\langle e_j, T^n e_k\rangle = \langle U^{-n} e_j, e_k\rangle \tag{4.9}$$

para todo $j, k = 1, 2, ...$ y todo $n \in \mathbb{N}$, donde U es la dilatación unitaria dada por el teorema de Sz.-Nagy.

Ahora, probemos que

$$(T^n x)_k = \sum_j x_j (m_j/m_k)^{\frac{1}{2}} p^n_{jk}, \tag{4.10}$$

por inducción en n. Para $n = 1$ se cumple por definición de T.

Supongamos que se cumple para n.

Sea $y = T^n x$, entonces $y_k = (T^n x)_k = \sum_j x_j (m_j/m_k)^{\frac{1}{2}} p^n_{jk}$. Tenemos

$$\begin{aligned}
(T^{n+1}x)_k &= (TT^n x)_k = (Ty)_k \\
&= \sum_j y_j (m_j/m_k)^{\frac{1}{2}} p_{jk} \\
&= \sum_j \left(\sum_i x_i (m_i/m_j)^{\frac{1}{2}} p^n_{ij} \right) (m_j/m_k)^{\frac{1}{2}} p_{jk} \\
&= \sum_i x_i (m_i/m_k)^{\frac{1}{2}} \sum_j p^n_{ij}\, p_{jk} \\
&= \sum_i x_i (m_i/m_k)^{\frac{1}{2}} p^{n+1}_{ik}
\end{aligned}$$

Hemos probado que

$$(T^n x)_k = \sum_j x_j (m_j/m_k)^{\frac{1}{2}} p^n_{jk}, \qquad \text{para todo } n \geq 1.$$

Por otro lado, como $\{e_j\}_j$ es una base ortonormal de ℓ_2, podemos escribir

$$T^n(x_je_j) = \sum_k \langle T^n(x_je_j), e_k\rangle\, e_k = \sum_k (T^n(x_je_j))_k\, e_k,$$

donde $x_je_j = (0,..,0,x_j,0,...)$ y entonces, usando la ecuación (4.10)

$$\langle T^n(x_je_j), e_k\rangle = (T^n(x_je_j))_k = \sum_i (x_je_j)_i(m_i/m_k)^{\frac{1}{2}}p^n_{ik} = x_j(m_j/m_k)^{\frac{1}{2}}p^n_{jk}.$$

Por lo tanto, $x_j(m_j/m_k)^{\frac{1}{2}}p^n_{jk} = \langle T^n(x_je_j), e_k\rangle$. Sumando en j a ambos lados y por la linealidad de T^n, tenemos

$$(T^nx)_k = \sum_j x_j(m_j/m_k)^{\frac{1}{2}}p^n_{jk} = \sum_j \langle T^n(x_je_j), e_k\rangle = \sum_j x_j\langle T^ne_j, e_k\rangle.$$

De donde

$$(m_j/m_k)^{\frac{1}{2}}p^n_{jk} = \langle T^ne_j, e_k\rangle.$$

Luego, usando (4.8) y despejando, obtenemos

$$p^n_{jk} = (m_k/m_j)^{\frac{1}{2}}\langle U^ne_j, e_k\rangle \qquad n = 0,1,2,... \tag{4.11}$$

Observación 4.7 Usando (4.9) se obtiene la misma igualdad ya que U es unitario y el lado izquierdo de (4.11) es real.

4.4. Representación Integral de las Probabilidades de Transición de Cadenas de Markov Irreducibles

Teorema 4.8 (Kendall) *Si $\{m_j : j = 1,2,...\}$ es una medida subinvariante positiva asociada a una cadena de Markov irreducible entonces las probabilidades de transición de la cadena pueden ser representadas de manera única mediante la forma*

$$p^n_{jk} = (m_k/m_j)^{\frac{1}{2}}\int_0^{2\pi} e^{in\theta}\, d\mu_{jk}(\theta) \qquad n = 0,1,2,...$$

donde las μ_{jk} son medidas complejas de Borel en $[0,2\pi]$ y satisfacen la condición hermítica $\overline{\mu_{jk}} = \mu_{kj}$.

Demostración:

Sea T el operador definido anteriormente y U la dilatación unitaria dada por el teorema de Sz.-Nagy. Por observación 2.80 (teorema de Wintner) tenemos que existe una familia de proyecciones $\{E_\theta : 0 \leq \theta \leq 2\pi\}$ con $E_0 = 0$ y $E_{2\pi} = I$, (medida espectral en $[0, 2\pi]$) tal que

$$U^n = \int_0^{2\pi} e^{in\theta}\, dE_\theta$$

para todo $n \in \mathbb{N}$. Es decir, para $x, y \in \ell_2$, tenemos

$$\langle U^n x, y\rangle = \int_0^{2\pi} e^{in\theta}\, d\langle E_\theta x, y\rangle. \tag{4.12}$$

Sea

$$M_{jk}(\theta) \equiv \langle E_\theta e_j, e_k\rangle, \tag{4.13}$$

por lo tanto, para cada j y k, $M_{jk}(.)$ es una medida compleja de Borel (teorema 2.72) que cumple

$$\overline{M_{jk}(\theta)} = M_{kj}(\theta). \tag{4.14}$$

Entonces, por ecuaciones (4.11), (4.12) y (4.13), tenemos

$$p^n_{jk} = (m_k/m_j)^{\frac{1}{2}} \int_0^{2\pi} e^{in\theta}\, dM_{jk}(\theta) \tag{4.15}$$

para todos los estados j y k y para todo $n \in \mathbb{N}$. (La integral en (4.15) es una integral Riemann-Stieltjes). Este resultado también puede ser escrito en la forma

$$\int_0^{2\pi} e^{in\theta}\, dM_{jk}(\theta) = (m_j/m_k)^{\frac{1}{2}} p^n_{jk} \qquad \text{con } n = 0, 1, 2, ...$$

y por (4.14), tomando conjugado, también tenemos

$$\int_0^{2\pi} e^{-in\theta}\, dM_{jk}(\theta) = (m_k/m_j)^{\frac{1}{2}} p^n_{kj} \qquad \text{con } n = 0, 1, 2, ...$$

Estas fórmulas muestran que *todos* los coeficientes de Fourier-Stieltjes de $M_{jk}(.)$ están determinados por P y la medida subinvariante positiva, y por lo tanto $M_{jk}(.)$ está únicamente determinada. $\square$

Observación 4.9 Es importante notar que, dada la observación 4.5, es posible encontrar sistemas diferentes para representar a las funciones $M_{jk}(.)$.

BIBLIOGRAFÍA

[1] R. Arocena, Unitary extensions of isometries and interpolation problems: dilation and lifting theorems. Publicaciones Matemáticas del Uruguay 22, 1995.

[2] G. Bachman and L. Narici, Functional analysis. Dover Publications. New York, 2000.

[3] A. Bharucha-Reid, Elements of the theory of Markov processes and their aplications. Dover Publications. New York, 1997.

[4] R. Bruzual y M. Domínguez, Espacios de Banach. Publicaciones del Laboratorio de Formas en Grupos. Facultad de Ciencias. UCV, 2003.

[5] R. Bruzual y M. Domínguez, Espacios de Hilbert. Publicaciones del Laboratorio de Formas en Grupos. Facultad de Ciencias. UCV, 2003.

[6] E. Çinlar, Introduction to stochastic processes. Prentice-Hall. New Yersey, 1975.

[7] C. Derman, A solution to a set of fundamental equations in Markov chains. Proc. American Math. Soc. 5. Vol 332-334, 1954.

[8] C. Derman, Some contributions to the theory of denumerable Markov chains. Trans. American Math. Soc. 79. Vol 541-555, 1955.

[9] R. Douglas, Banach algebra techniques in operator theory. Academic Press. New York, 1972.

[10] Y. Farias, Teorema de levantamiento, diversos enfoques y aplicaciones. Trabajo Especial de Grado dirigido por Dra. M. Morán. Escuela de Matemáticas. Facultad de Ciencias. UCV, 1995.

[11] P. Ferrari and A. Galves, Construction of stochastic processes, coupling and regeneration. Escuela Venezolana de Matemáticas. Mérida, 2000.

[12] I. Herstein, Algebra moderna. Trillas. México, 1990.

[13] S. Karlin and H. Taylor, A first course in stochastic processes. Academic Press. New York, 1975.

[14] D. Kendall, Unitary dilations of Markov transition operators and the corresponding integral representations for transition-probability matrices. Probability and Statistic, H. Cramer. Vol 139-161, 1959.

[15] J. Marsden y M. Hoffman, Análisis básico de variable compleja. Trillas. México, 1996.

[16] V. Paulsen, Completely bounded maps and dilations. Pitman Research Notes in Mathematics Series, Vol.146, 1986.

[17] F. Riesz and Sz.-Nagy, Functional analysis. Frederick Ungar Publishing. New York, 1978.

[18] W. Rudin, Functional analysis. Mc Graw-Hill, 1979.

[19] W. Rudin, Real and complex analysis. Mc Graw-Hill, 1966.

[20] G. Simmons, Introduction to topology and modern analysis. Krieger Publishing. Florida, 1963.

ÍNDICE ALFABÉTICO