



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

ANÁLISIS ESPECTRAL DE PROCESOS CUASI DE
NACIMIENTO Y MUERTE

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

LICENCIADO EN MATEMÁTICAS

P R E S E N T A :

JOSÉ RAMÓN TUIRÁN RANGEL

TUTOR

MANUEL DOMÍNGUEZ DE LA IGLESIA



CIUDAD UNIVERSITARIA, Cd. Mx., 2021.



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Agradecimientos

Quisiera aprovechar este espacio para expresar mi profunda gratitud y aprecio a todas las personas e instituciones que contribuyeron con la realización de este trabajo así como al crecimiento de mi persona durante la licenciatura.

Primeramente quiero agradecer a mis padres y a mi hermano por su apoyo incondicional, el aliento y paciencia que me brindaron en mi incursión a las matemáticas a temprana edad. Ciertamente este logro no existiría de no ser por ustedes. Especialmente a mi madre por su dedicación hacia mi persona.

A mis amigos, compañeros y maestros de la Facultad de Ciencias, en particular Ariel, Víctor, Óscar, Jorge, Zeus, Siddharta y Diego. A Diana por su apoyo incondicional.

A mi asesor por su paciencia, su dedicación y su guía en este proyecto.

Finalmente agradezco al Conacyt por otorgarme la beca del proyecto A1-S-16202F y número de registro 30201 para poder realizar esta tesis. A la UNAM por brindarme la oportunidad de realizar mi licenciatura y por todas las experiencias que he vivido gracias a ella.

Índice general

Agradecimientos	II
Introducción	V
1. Polinomios Ortogonales Matriciales	1
1.1. Polinomios Ortogonales Matriciales (POM)	1
1.1.1. Medidas Matriciales	1
1.1.2. Producto interno matricial	3
1.1.3. Expresiones tipo Fourier	5
1.2. POM en la recta real	8
1.2.1. POM Mónicos	10
1.2.2. Relaciones de recurrencia en los POM Mónicos	12
1.2.3. POM Normalizados	13
1.2.4. Matrices de Jacobi por bloques	14
1.2.5. Teorema de Favard	16
1.3. Momentos canónicos en POM	23
1.3.1. Momentos canónicos de una medida matricial	23
1.3.2. POM y Momentos Canónicos	48
2. Procesos cuasi de nacimiento y muerte	55
2.1. Introducción	55
2.2. Existencia de medidas matriciales sobre un proceso cuasi de nacimiento y muerte	57
2.3. Propiedades de recurrencia	78
2.4. Momentos canónicos y su aplicación	80
2.5. Representaciones de la medida invariante	87
3. Ejemplos	90
3.1. Cadenas de nacimiento y muerte en \mathbb{Z}	90
3.2. Un ejemplo proveniente de Teoría de colas	96
3.3. Caminata aleatoria en una cuadrícula	98

3.4. Caminata aleatoria en un árbol	102
3.5. Un ejemplo proveniente de teoría de representación de grupos	105

Introducción

El objetivo principal de esta tesis es estudiar una generalización del análisis espectral de los procesos estocásticos denominados cadenas de nacimiento y muerte a tiempo discreto, los cuales fueron estudiados primeramente en [19]. Esta generalización recibe el nombre de “Quasi-Birth-and-Death-Process (QBD)”, en inglés, cuya traducción más cercana se encuentra en el título del documento, y que hemos decidido usar como “Proceso Cuasi de Nacimiento y Muerte”. La diferencia principal entre ambos procesos es que el espacio de estados pasa de ser \mathbb{N}_0 (para el caso de cadenas de nacimiento y muerte) a ser de la forma $\mathbb{N}_0 \times \{1, 2, \dots, N\}$, con $N \in \mathbb{N}$, en el caso de los procesos QBD. Esto hace que la matriz de transición de probabilidades pase de ser una matriz tridiagonal a ser una matriz tridiagonal por bloques de tamaño $N \times N$.

El estudio de los procesos QBD se remonta a los trabajos de M. F. Neuts en los años 80 (véase [23]) y posteriormente diversos autores como Hajek, Latouche, Ramaswani, Pearce o Taylor, por mencionar algunos, siguieron trabajando en estos procesos (véase [22]). Los procesos QBD poseen significantes aplicaciones en el estudio de sistemas de comunicaciones (véase [24]). En [8] se realiza el primer estudio de análisis espectral de estos procesos, cuya herramienta principal radica en la teoría de polinomios ortogonales matriciales (POM) y medidas matriciales. Es por ello que hemos decidido que el primer capítulo de este texto esté centrado en el estudio de estos objetos.

En el Capítulo 1 definimos pues los POM y estudiaremos varias de sus propiedades principales. En particular nos centraremos en buscar una prueba del Teorema de Favard cuya importancia está en que proporciona una correspondencia directa entre las medidas matriciales asociadas a una sucesión de POM y las matrices tridiagonales de Jacobi por bloques. También en este capítulo estudiaremos los momentos canónicos de una medida matricial, que fueron presentados en su forma escalar en [9], y que posteriormente se definieron para el caso matricial en [10]. Presentamos los principales resultados obtenidos en [8], que nos servirán para poder establecer una relación entre los momentos canónicos y los procesos QBD, así como la sucesión de POM que le corresponde. Un resultado relevante sobre los momentos canónicos, escalares, es la

invarianza que poseen sobre transformaciones lineales a la variable de la medida. Sin embargo, cuando en [10] se definen para el caso matricial, se omite este hecho y a continuación en [8], se da por sentado la veracidad de este resultado sin ofrecer mayor detalle a la prueba. En este trabajo, se probará dicha propiedad mediante la definición de recurrencia, dada por complementos de Schur, de los momentos canónicos en el caso matricial. Para finalizar el capítulo, escribimos los coeficientes de nuestra sucesión de POM en términos de nuestros momentos canónicos.

En el Capítulo 2 de la tesis definimos, usando la matriz de transición por bloques, a los procesos QBD. Este capítulo está basado mayormente en [8]. Estos generan inmediatamente la secuencia de POM, y usamos el Teorema de Favard visto en el Capítulo 1 para probar, con ciertas condiciones suficientes y necesarias, que existe una medida matricial asociada que hace ortogonales a esta secuencia de POM. Continuamos probando el resultado análogo al caso matricial de la representación de Karlin-McGregor, el cual básicamente es la representación del (i, j) -ésimo bloque de la n -ésima potencia de la matriz de transición, con base en el producto de dos integrales que involucran a los POM respecto la medida encontrada previamente. Seguimos el capítulo obteniendo resultados importantes; se demuestra que el soporte de la medida se encuentra en $[-1, 1]$, definimos la transformada de Stieltjes de una medida, la cual nos será útil posteriormente para poder encontrar la medida con base en la matriz de transición, y proporcionamos dos resultados que involucran a la transformada y su relación con los procesos obtenidos de pequeñas modificaciones a la matriz de transición. Después buscamos generalizar el primer resultado sobre la existencia de una medida matricial al relajar las condiciones sobre los coeficientes de la matriz de transición. Concluyendo esta sección del capítulo, nos enfocamos en estudiar las propiedades de recurrencia de nuestro proceso. Para ello obtenemos dos teoremas importantes que nos caracterizan, con base en los coeficientes de la matriz de transición, cuándo un proceso QBD es recurrente y en su caso, este es positivo recurrente si y solo si la medida posee un salto discreto en el punto 1. Seguimos con la siguiente sección en donde buscamos relacionar a los momentos canónicos, definidos en el primer capítulo, con la transformada de Stieltjes de la medida, y usando esto obtener una nueva caracterización de la recurrencia del proceso usando solamente la divergencia de una serie cuyos sumandos dependen únicamente de los coeficientes matriciales de la matriz de transición. Concluimos el capítulo estudiando diversas representaciones de la medida invariante, así como una forma explícita de esta, estudiada en [7].

Por último, el Capítulo 3 busca otorgar al lector una compilación de varios ejemplos de análisis espectral de procesos QBD. Empezamos con el más conocido, el cual es una cadena de nacimiento y muerte en los enteros. Esto difiere del caso escalar

que tradicionalmente se estudia solo en los naturales. Es justo esta diferencia entre usar la cadena en los enteros y los naturales lo que permite transformarlo en un proceso QBD para su estudio, ya que $\mathbb{Z} \simeq \mathbb{N}_0 \times \{1, 2\}$. Obtenemos una expresión para la transformada de Stieltjes, y posteriormente estudiamos un caso particular de este proceso en donde todas las transiciones poseen $1/2$ de probabilidad; damos explícitamente su medida, y estudiamos las propiedades de recurrencia así como su medida invariante. El segundo ejemplo es breve y busca representar una aplicación de estos procesos a la teoría de colas. El tercer ejemplo estudia a las caminatas aleatorias en una cuadrícula, y estudiamos tanto el caso de espacio de estados finito como el infinito, obteniendo la medida matricial, sus propiedades de recurrencia y finalmente la medida invariante. El cuarto ejemplo, estudiado en [16], es una caminata aleatoria definida en un grafo en forma de árbol con l ramas. De este ejemplo obtenemos la medida matricial asociada al proceso, así como nuevamente sus propiedades de recurrencia y su medida invariante. Finalmente, el quinto ejemplo, estudiado en [17], proviene de la teoría de representación de grupos. Estudiaremos la medida matricial del proceso, sus propiedades de recurrencia y finalmente la medida invariante.

Capítulo 1

Polinomios Ortogonales Matriciales

1.1. Polinomios Ortogonales Matriciales (POM)

Empezamos en este capítulo por introducir la teoría de los Polinomios Ortogonales Matriciales o POM. Definimos lo que serán las medidas matriciales en \mathbb{R} , con base en esto propondremos un producto interno con valores matriciales, el cual definirá lo que es una secuencia ortogonal de polinomios matriciales, así como sus propiedades equivalentes al caso escalar. Todo esto será de enorme utilidad para demostrar uno de los teoremas fundamentales: el Teorema de Favard para el caso matricial. Una vez desarrollados los POM, estableceremos los momentos canónicos de una medida matricial y la relación estrecha que mantienen con los POM.

1.1.1. Medidas Matriciales

Definimos primero como $\mathcal{M}_l(\mathbb{C})$ al anillo de todas las matrices de tamaño $l \times l$ y con entradas en los números complejos. Considerando $A \in \mathcal{M}_l(\mathbb{C})$ denotamos a A^* como la matriz hermítica de A . Para la norma euclídea de A , queda definida como: $\|A\| = \inf\{c \geq 0 : \|Av\| \leq c\|v\| \text{ para todo } v \in \mathbb{C}^l\}$, y para el producto interno en \mathbb{C}^l como $\langle \cdot, \cdot \rangle_{\mathbb{C}^l}$. Dado esto, definimos como una medida matricial μ , en la σ -álgebra generada por los conjuntos Borelianos en \mathbb{R} (o en \mathbb{C}), como: $\mu(X) \in \mathcal{M}_l(\mathbb{C})$ y semi-definida positiva, para cualquier conjunto Boreliano X , y con μ contable aditiva (en el sentido matricial). Normalmente normalizaremos la medida en la forma:

$$\mu(\mathbb{R}) = \mathbf{I}, \tag{1.1}$$

(o en otro caso como $\mu(\mathbb{C}) = \mathbf{I}$), donde \mathbf{I} queda definida como la matriz identidad, e igualmente a $\mathbf{0}$ como la matriz con todas las entradas igual a 0. Generalmente pediremos que las medidas tratadas tengan un soporte compacto, aunque en general no es necesario. Note que por la aditividad numerable, cada entrada, $\mu_{i,j}$, de nuestra

medida matricial, es a su vez una medida compleja.

Proposición 1.1. *Cada entrada de la medida matricial es absolutamente continua con respecto a la medida de la traza definida como $\mu_{tr}(X) = \text{Tr}(\mu(X))$ (con X un conjunto Boreliano). Esto implica la existencia de sus derivadas de Radon-Nikodym, las cuales forman una matriz $M'(t) = \left(\frac{d\mu_{i,j}}{d\mu_{tr}}(t) \right)_{i,j}$ (donde claro está que las derivadas son las definidas por Radon-Nikodym), que es semi-definida positiva, salvo un conjunto de medida de la traza igual a 0.*

Demostración. Sea A un conjunto Boreliano de medida de la traza igual a 0. Notemos primero que las entradas diagonales de $\mu(A)$ son reales no negativas. En efecto, son reales pues la matriz $\mu(A)$ es Hermitiana, y son no negativos pues la evaluación $e_j^* \mu(A) e_j = (\mu(A))_{j,j}$ debe ser no negativa para toda j , ya que $\mu(A)$ es semi-definida positiva, con e_j la base canónica (es decir el vector con entradas 0 salvo la entrada j que es 1). Ahora considere la sub-matriz definida por las intersecciones de las columnas y filas i, j :

$$S = \begin{pmatrix} 0 & \mu(A)_{i,j} \\ \mu(A)_{i,j} & 0 \end{pmatrix}.$$

Esta sub-matriz debe ser semi-definida positiva (solo basta considerar en el producto $x^* \mu(A) x$, vectores con entradas no cero solo en i, j), y por ende debe tener determinante no negativo. Pero $\det(S) = -\|\mu(A)_{i,j}\|^2$ lo cual es no negativo si y solo si $\mu(A)_{i,j}$ es 0. Como las i, j fueron arbitrarias esto pasa para todas las entradas, y por ende cada medida de las entradas es absolutamente continua con respecto a la medida de la traza.

Ahora probamos que la matriz $M'(t) = \left(\frac{d\mu_{i,j}}{d\mu_{tr}}(t) \right)_{i,j}$ es semi-definida positiva. Aplicando el teorema de Radon-Nikodym tenemos que:

$$\int_A M'(t) d\mu_{tr}(t) = \mu(A), \quad (1.2)$$

donde la integral es entrada por entrada, y A es un Boreliano, cuya medida de la traza no necesariamente es 0. Ahora viendo a $M'(t)$ y $\mu(A)$ como operadores lineales en \mathbb{C}^l , tenemos que:

$$\int_A x^* M'(t) x d\mu_{tr}(t) = x^* \mu(A) x \geq 0, \quad x \in \mathbb{C}^l,$$

donde la multiplicación matricial de x entra en la integral pues lo hace entrada a entrada y no depende de t al ser un escalar, respecto a la integral. La desigualdad se cumple pues $\mu(A)$ es semi-definida positiva. Como las entradas diagonales de μ_{tr} son

positivas, esta medida es positiva y esto implica que $x^*M'(t)x \geq 0$ salvo un conjunto N_x con $\mu_{tr}(N_x) = 0$. Tomando un conjunto denso y numerable de vectores x (como lo sería los vectores de entradas con parte real e imaginaria racionales), y recordando que el producto $x^*M'(t)x$ es continuo tenemos que $M'(t) \geq 0$ salvo un conjunto N , que es la unión numerable de los N_x , con $\mu_{tr}(N) = 0$. \square

Note de igual forma que como la derivada de Radon-Nikodym es aditiva, y por (1.2):

$$\int_A \text{Tr}(M'(t)) d\mu_{tr}(t) = \text{Tr}(\mu(A)) = \mu_{tr}(A).$$

Pero esto a su vez implica:

$$\int_A (\text{Tr}(M'(t)) - 1) d\mu_{tr}(t) = 0,$$

donde esto ocurre para cualquier A Boreliano. Por ende (por la positividad de μ_{tr}), tenemos que $\text{Tr}(M'(t)) = 1$, para todo t , salvo un conjunto de medida de la traza igual a 0.

A su vez, es fácil ver que cualquier medida escalar positiva en \mathbb{R} tal que $\mu_{tr} = l$ y una función matricial $M'(x)$ define una medida matricial normalizada (ver (1.1)).

1.1.2. Producto interno matricial

Con esta información ya nos es posible definir lo que será una forma sesquilineal, $\langle f, g \rangle_\mu$ una matriz de $l \times l$, con f, g en el conjunto de funciones matriciales de dimensión $l \times l$, de la siguiente forma:

$$\langle f, g \rangle_\mu = \int_{\mathbb{R}} f(t) M'(t) g^*(t) d\mu_{tr}(t),$$

donde la integral es entrada a entrada, es decir la (i, j) entrada es:

$$\sum_{m,n} \int_{\mathbb{R}} f_{i,m}(t) M'_{m,n}(t) \overline{g_{j,n}(t)} d\mu_{tr}(t).$$

Sin embargo normalmente se utiliza la notación siguiente, por convención, que es equivalente a la ecuación anterior:

$$\langle f, g \rangle_\mu = \int_{\mathbb{R}} f(t) d\mu(t) g^*(t). \quad (1.3)$$

Esta definición nos proporciona las siguientes propiedades (con $\alpha \in \mathcal{M}_l(\mathbb{C})$):

$$\langle f, \alpha g \rangle_\mu = \langle f, g \rangle_\mu \alpha^*, \quad \langle \alpha f, g \rangle_\mu = \alpha \langle f, g \rangle_\mu. \quad (1.4)$$

Dada esta definición, inmediatamente obtenemos la siguiente relación:

$$\langle f, g \rangle_\mu^* = \langle g, f \rangle_\mu. \quad (1.5)$$

Las funciones matriciales f, g , pueden ser de cualquier tipo, aunque nos enfocaremos en las polinomiales, siempre y cuando cumplan que las integrales previamente definidas existan y converjan. Para esto solo pediremos que sean continuas (o medibles), y que satisfagan que (tanto f como g):

$$\int_{\mathbb{R}} \text{Tr}(f(t)f^*(t)) d\mu_{tr}(t) < \infty.$$

Esto pues la traza de esa matriz representa:

$$\int_{\mathbb{R}} \text{Tr}(f(t)f^*(t)) d\mu_{tr}(t) = \sum_{i,j} \int_{\mathbb{R}} \|f_{i,j}(t)\|^2 d\mu_{tr}(t) < \infty.$$

Y como la suma es finita, y la medida μ_{tr} es positiva, esto implica que cada sumando converge. Esto es igual para g . Por ende como las funciones $M'_{m,n}(t)$ están acotadas, tenemos que:

$$\begin{aligned} \int_{\mathbb{R}} |f_{m,i}(t) M'_{m,n}(t) \overline{g_{n,j}(t)}| d\mu_{tr}(t) &\leq \int_{\mathbb{R}} |f_{m,i}(t)| |M'_{m,n}(t)| |g_{n,j}(t)| d\mu_{tr}(t) \\ &\leq C \left(\int_{\mathbb{R}} |f_{m,i}(t)|^2 d\mu_{tr}(t) \right)^{1/2} \left(\int_{\mathbb{R}} |g_{n,j}(t)|^2 d\mu_{tr}(t) \right)^{1/2} < \infty, \end{aligned}$$

donde la penúltima desigualdad es por la desigualdad de Cauchy-Schwarz, así las integrales existen.

Regresando, claramente como M' es semi-definida positiva, el producto $\langle f, f \rangle_\mu$ es semi-definido positivo, lo cual nos intuye a definir la siguiente forma sesquilineal:

$$\langle f, g \rangle_{\text{Tr}} = \text{Tr} \langle f, g \rangle_\mu.$$

Esta definición claramente nos dice que es semi-definida positiva, pues cuando $f = g$ tenemos que $\langle f, f \rangle_\mu$ es semi-definida positiva y por ende las entradas de la diagonal son no negativas y así la traza también. Sin embargo que la traza sea 0 no implica

que las funciones lo sean. Por lo tanto esta forma induce la siguiente semi-norma:

$$\|f\|_{\text{Tr}} = (\langle f, f \rangle_{\text{Tr}})^{1/2}. \quad (1.6)$$

A su vez, la definición de esta semi-norma nos otorga la desigualdad de Cauchy-Schwarz:

$$|\text{Tr}\langle f, g \rangle_{\mu}| \leq \|f\|_{\text{Tr}} \|g\|_{\text{Tr}}. \quad (1.7)$$

Con base en esto podemos crear un espacio de Hilbert \mathcal{H} (el cual en efecto es, ver [25, Teorema 3.9]) considerando solo las clases de equivalencia inducidas por la relación: $f \sim g \iff \|f - g\|_{\text{Tr}} = 0$, con $\langle \cdot, \cdot \rangle_{\text{Tr}}$ como el producto interior escalar definido en este espacio.

1.1.3. Expresiones tipo Fourier

Veamos ahora expansiones tipo Fourier sobre este espacio. Definimos a un subconjunto $\{\phi_j\}$ de \mathcal{H} como 'ortonormal', (aunque técnicamente no lo es, pues como se verá en la definición implica ortogonalidad, pero no normalización en $\langle \cdot, \cdot \rangle_{\mu}$), si satisface:

$$\langle \phi_j, \phi_k \rangle_{\mu} = \delta_{j,k} \mathbf{I}. \quad (1.8)$$

Considere $N \in \mathbb{N} \cup \{\infty\}$, y por ahora suponga que es finito. Tomando $a_1, a_2, \dots, a_N \in \mathcal{M}_l(\mathbb{C})$ formamos el siguiente sub-espacio de elementos de la forma $\sum_{j=1}^N a_j \phi_j$, cuidando que la multiplicación sea por la izquierda. Así usando (1.4) y (1.8) tenemos:

$$\left\langle \sum_{j=1}^N a_j \phi_j, \sum_{k=1}^N b_k \phi_k \right\rangle_{\mu} = \sum_{j=1}^N a_j b_j^*. \quad (1.9)$$

Simplemente aprovechamos la 'linealidad' de $\langle \cdot, \cdot \rangle_{\mu}$ y usamos la ortonormalidad de $\{\phi_j\}$. Al espacio de elementos: $\sum_{j=1}^N a_j \phi_j$, lo denotamos por $\mathcal{H}_N^{\{\phi_j\}}$, el cual es un subespacio de \mathcal{H} de dimensión Nl^2 , pues hay N matrices como sumandos y cada una posee l^2 entradas. Ahora para cada $f \in \mathcal{H}$ definimos:

$$\pi_{\{\phi_j\}}(f) = \sum_{j=1}^N \langle f, \phi_j \rangle_{\mu} \phi_j.$$

Note que: $\langle \pi_{\{\phi_j\}}(f), \phi_j \rangle_\mu = \langle f, \phi_j \rangle_\mu$ y por ende $\langle \pi_{\{\phi_j\}}(f), \phi_j \rangle_{\text{Tr}} = \langle f, \phi_j \rangle_{\text{Tr}}$. De igual forma que en el caso escalar obtenemos el teorema de Pitágoras:

$$\begin{aligned} \langle f, f \rangle_\mu &= \langle f - \pi_{\{\phi_j\}}(f), f - \pi_{\{\phi_j\}}(f) \rangle_\mu + \langle f - \pi_{\{\phi_j\}}(f), \pi_{\{\phi_j\}}(f) \rangle_\mu \\ &\quad + \langle \pi_{\{\phi_j\}}(f), f - \pi_{\{\phi_j\}}(f) \rangle_\mu + \langle \pi_{\{\phi_j\}}(f), \pi_{\{\phi_j\}}(f) \rangle_\mu. \end{aligned}$$

Sin embargo, note que:

$$\langle f - \pi_{\{\phi_j\}}(f), \pi_{\{\phi_j\}}(f) \rangle_\mu = \langle f, \pi_{\{\phi_j\}}(f) \rangle_\mu - \langle \pi_{\{\phi_j\}}(f), \pi_{\{\phi_j\}}(f) \rangle_\mu.$$

Así usando (1.8) y la definición (1.9) tenemos que:

$$\langle \pi_{\{\phi_j\}}(f), \pi_{\{\phi_j\}}(f) \rangle_\mu = \sum_{j=1}^N \langle f, \phi_j \rangle_\mu \langle f, \phi_j \rangle_\mu^*,$$

y también:

$$\langle f, \pi_{\{\phi_j\}}(f) \rangle_\mu = \sum_{j=1}^N \langle f, \langle f, \phi_j \rangle_\mu \phi_j \rangle_\mu = \sum_{j=1}^N \langle f, \phi_j \rangle_\mu \langle f, \phi_j \rangle_\mu^*,$$

donde al final usamos la sesquilinealidad de $\langle \cdot, \cdot \rangle_\mu$. Así se obtiene que:

$$\langle f, \pi_{\{\phi_j\}}(f) \rangle_\mu = \langle \pi_{\{\phi_j\}}(f), \pi_{\{\phi_j\}}(f) \rangle_\mu,$$

y por ende:

$$\langle f - \pi_{\{\phi_j\}}(f), \pi_{\{\phi_j\}}(f) \rangle_\mu = \mathbf{0},$$

y de foma análoga al caso anterior obtenemos igualmente:

$$\langle \pi_{\{\phi_j\}}(f), f - \pi_{\{\phi_j\}}(f) \rangle_\mu = \mathbf{0}.$$

Por ende, el teorema de Pitágoras se reduce a:

$$\langle f, f \rangle_\mu = \langle f - \pi_{\{\phi_j\}}(f), f - \pi_{\{\phi_j\}}(f) \rangle_\mu + \sum_{j=1}^N \langle f, \phi_j \rangle_\mu \langle f, \phi_j \rangle_\mu^*. \quad (1.10)$$

Así como las matrices en la expresión son semi-definidas positivas, esto implica que:

$$\sum_{j=1}^N \langle f, \phi_j \rangle_\mu \langle f, \phi_j \rangle_\mu^* \leq \langle f, f \rangle_\mu, \quad (1.11)$$

para toda N finita. Esta desigualdad es la equivalente a la desigualdad de Bessel. En el siguiente teorema se prueba el caso cuando $N = \infty$.

Teorema 1.2. *La serie $\sum_{j=1}^{\infty} \langle f, \phi_j \rangle_{\mu} \langle f, \phi_j \rangle_{\mu}^*$, converge entrada por entrada a una matriz S en $\mathcal{M}_l(\mathbb{C})$.*

Demostración. Definamos primero a $S_N = \sum_{j=1}^N \langle f, \phi_j \rangle_{\mu} \langle f, \phi_j \rangle_{\mu}^*$ como la N -ésima suma parcial. Igualmente note que los elementos matriciales de la serie son semi-definidos positivos, y por ende por la prueba de la Proposición 1.1, sus elementos diagonales son reales no negativos. Además por (1.11) la desigualdad de Bessel, tenemos que, considerando los elementos diagonales de todas las sumas parciales, están acotados por sus respectivos elementos diagonales de $\langle f, f \rangle_{\mu}$, es decir: $(S_N)_{i,i} \leq (\langle f, f \rangle_{\mu})_{i,i}$. Por ende al ser una serie de reales positivos y acotada, tenemos que la sucesión $\{(S_N)_{i,i}\}_N$ converge. Ahora probamos que todas las demás entradas convergen. Por definición, los elementos de las matrices son números complejos, así pues es suficiente demostrar que las entradas forman una sucesión de Cauchy. Sin embargo, por lo anterior, sabemos que las entradas diagonales, de S_N , sí forman una sucesión de Cauchy. Por lo tanto, por definición, para todo ϵ existe un N_i , entero positivo, tal que para todo $m, n > N_i$ enteros positivos, se satisface que:

$$|(S_m - S_n)_{i,i}| < \epsilon. \quad (1.12)$$

Ahora consideremos la siguiente sub-matriz formada por la intersección de las columnas y filas i, j :

$$S'_{i,j} = \begin{pmatrix} (S_m - S_n)_{i,i} & (S_m - S_n)_{i,j} \\ (S_m - S_n)_{i,j} & (S_m - S_n)_{j,j} \end{pmatrix}. \quad (1.13)$$

Tomando $N' = \max(N_i, N_j)$ la ecuación (1.12) también es válida para j , por ende, por la misma observación de la Proposición 1.1, como $S_m - S_n$ es definida positiva, tenemos que el determinante de la matriz $S'_{i,j}$ en (1.13) debe ser positivo:

$$(S_m - S_n)_{i,i} \cdot (S_m - S_n)_{j,j} - |(S_m - S_n)_{i,j}|^2 > 0.$$

Pero por (1.12) ambos $(S_m - S_n)_{i,i}, (S_m - S_n)_{j,j} < \epsilon$ y por ende:

$$\begin{aligned} \epsilon^2 &> |(S_m - S_n)_{i,j}|^2 \\ \implies \epsilon &> |(S_m - S_n)_{i,j}|, \end{aligned}$$

y así $\{(S_N)_{i,j}\}_N$, igual es una sucesión de Cauchy para toda i, j . □

Teorema 1.3. *La serie $\pi_{\{\phi_j\}}(f) = \sum_{j=1}^{\infty} \langle f, \phi_j \rangle_{\mu} \phi_j$, converge entrada por entrada a una función matricial A en \mathcal{H} , en la norma $\|\cdot\|_{\text{Tr}}$.*

Demostración. Definimos a $A_N = \sum_{j=1}^N \langle f, \phi_j \rangle_{\mu} \phi_j$ como la N -ésima suma parcial. Considere, abusando de notación, y usando el teorema anterior así como (1.8), un entero positivo N tal que para todo $m > n > N$:

$$\left| \langle A_m - A_n, A_m - A_n \rangle_{i,i} \right| = \left| \left(\sum_{j=n+1}^m \langle f, \phi_j \rangle_{\mu} \langle f, \phi_j \rangle_{\mu}^* \right)_{i,i} \right| < \epsilon,$$

pues la convergencia de la Proposición 1.1 es entrada por entrada. Así, aplicando la traza y tomando ϵ^2/l en lugar de ϵ , obtenemos:

$$\begin{aligned} \|A_m - A_n\|_{\text{Tr}}^2 &= \text{Tr} \langle A_m - A_n, A_m - A_n \rangle_{\mu} < l \cdot \frac{\epsilon^2}{l} = \epsilon^2 \\ \implies \|A_m - A_n\|_{\text{Tr}} &< \epsilon. \end{aligned}$$

Por ende $\{A_N\}_{N=1}^{\infty}$ es una sucesión de Cauchy en el espacio de Hilbert \mathcal{H} bajo la norma definida, por lo que tiene que converger, al ser completo, a una función matricial A . \square

1.2. POM en la recta real

Empezamos a centrar nuestra atención en las funciones de \mathcal{H} que son polinomios. Así, nombramos como \mathcal{P} al conjunto de polinomios con coeficientes matriciales en $\mathcal{M}_l(\mathbb{C})$, y como \mathcal{V} al conjunto de polinomios con coeficientes en \mathbb{C}^l . Así, como se vio anteriormente, $\|\cdot\|_{\text{Tr}}$ es una semi-norma en \mathcal{P} . El conjunto $\{f : \|f\|_{\text{Tr}} = 0\}$ es un espacio lineal por la desigualdad triangular. Así podemos definir a \mathcal{P}_{Tr} como la completación de $\mathcal{P}/\{f : \|f\|_{\text{Tr}} = 0\}$, visto como un módulo izquierdo en $\mathcal{M}_l(\mathbb{C})$, con respecto a la norma $\|\cdot\|_{\text{Tr}}$.

Note, igualmente, que \mathcal{V} es un espacio lineal y podemos definir una semi-norma asociada al producto interior:

$$\langle f, g \rangle_{\mathcal{V}} = \int_{\mathbb{R}} f(t) M'(t) g^*(t) d\mu_{\text{tr}}(t),$$

con f, g en \mathcal{V} , cuyo valor es un escalar definido explícitamente como:

$$\sum_{i,j} \int_{\mathbb{R}} f_i(t) M'_{i,j}(t) \overline{g_j(t)} d\mu_{tr}(t),$$

el cual será usado con la siguiente notación:

$$\int_{\mathbb{R}} d\langle f(t), \mu(t)g(t) \rangle_{\mathbb{C}^l}.$$

Esto, a su vez, define una nueva semi-norma:

$$|f|_{\mathcal{V}} = \left(\int_{\mathbb{R}} d\langle f(t), \mu(t)f(t) \rangle_{\mathbb{C}^l} \right)^{1/2}.$$

De igual forma que anteriormente definimos como \mathcal{V}_0 al conjunto de polinomios con valores vectoriales tales que $|f|_{\mathcal{V}} = 0$ y llamamos \mathcal{V}_{∞} a la completación del espacio cociente $\mathcal{V}/\mathcal{V}_0$ con respecto a la norma $|\cdot|_{\mathcal{V}}$.

Con base en esto definimos ahora lo que será una norma no-trivial, como en el caso escalar.

Lema 1.4. *Las siguientes afirmaciones son equivalentes:*

- $\|f\|_{\text{Tr}} > 0$ para todo $f \in \mathcal{P}$.
- Para toda n , la dimensión en \mathcal{P}_{Tr} del conjunto de polinomios de grado a lo más n es $(n+1)l^2$.
- Para todo $v \in \mathcal{V}$, distinto de cero, tenemos $|v|_{\mathcal{V}} \neq 0$.
- Para toda n , la dimensión en \mathcal{V}_{∞} de los polinomios con coeficientes vectoriales de grado a lo más n es $(n+1)l$.

Si μ satisface alguna de estas condiciones equivalentes diremos que la medida es no-trivial.

Demostración. Probamos primero que las dos primeras afirmaciones son equivalentes. En efecto, sea el conjunto: $B = \{E_{i,j}t^k \mid 1 \leq i, j \leq l, 0 \leq k \leq n\}$ donde $E_{i,j}$ es la matriz con 0 en todas sus entradas salvo la (i, j) en donde es 1. Este conjunto, de cardinalidad $(n+1)l^2$ claramente genera a todos los polinomios de grado a lo más n , en \mathcal{P} . Así considerando sus respectivas clases de equivalencia en \mathcal{P}_{Tr} igual van a generarlo. Por lo tanto, si tenemos que $\|f\|_{\text{Tr}} > 0$, entonces las clases de equivalencia de los elementos de B son linealmente independientes en \mathcal{P}_{Tr} , pues si no lo fueran

existiría $g \in \mathcal{P}$ como una combinación lineal no trivial de estas clases de equivalencia que pertenece a $\{f : \|f\|_{\text{Tr}} = 0\}$. Pero esto es imposible, pues implicaría que B no es base del conjunto de polinomios de grado a lo más n , pero sí lo genera, y por ende tendría dimensión menor a $(n+1)l^2$. De forma análoga se prueba la equivalencia entre la tercera y cuarta afirmación. Solo resta probar la equivalencia entre la primera y tercera afirmación. Asumiendo la primera, y tomando $v \in \mathcal{V}$, distinto de 0, así como a $f \in \mathcal{M}_l(\mathbb{C})$ como la matriz que tiene a v como su primera columna, y ceros en el resto de las entradas, se cumple claramente que: $0 < \|f\|_{\text{Tr}}^2 = \text{Tr}\langle f, f \rangle_\mu = |v|_{\mathcal{V}}^2$, implicándola. Ahora, suponiendo la tercera afirmación, y asumiendo por contradicción que la primera no se cumple, tenemos que considerando $f \in \mathcal{P}$ distinto de cero, tal que $\|f\|_{\text{Tr}} = 0$. Así pues, alguna de las columnas de f es distinta de cero. Sin pérdida de generalidad podemos suponer que es la primera y es v . Entonces tenemos que:

$$\begin{aligned} 0 = \|f\|_{\text{Tr}} &\implies \text{Tr}\langle f, f \rangle_\mu = 0 \implies \langle f, f \rangle_\mu = 0 \\ \implies E_{1,1}\langle f, f \rangle_\mu E_{1,1}^* &= 0 \implies 0 = \text{Tr}(E_{1,1}\langle f, f \rangle_\mu E_{1,1}^*) = |v|_{\mathcal{V}}^2. \end{aligned}$$

La segunda implicación se da de lo visto en la Proposición 1.1 pues al ser $\langle f, f \rangle_\mu$ semi-definida positiva tiene sus elementos diagonales mayores o iguales a 0, así si la traza es 0, es porque todos estos son 0, y como se probó tiene que ser la matriz $\mathbf{0}$, concluyendo el lema. \square

1.2.1. POM Mónicos

Ahora consideramos los polinomios matriciales ortogonales, y mónicos, esto bajo una medida no-trivial, como fue previamente definida por el Lema 1.3. Llamamos \mathcal{P}_n al conjunto de polinomios de grado a lo más n . Presentamos a continuación un nuevo lema sobre polinomios ortogonales mónicos que generaliza el caso escalar.

Lema 1.5. *Sea μ una medida no-trivial. Tenemos:*

- *Existe un único polinomio mónico \underline{P}_n de grado n , tal que minimiza la norma $\|\underline{P}_n\|_{\text{Tr}}$.*
- *El polinomio \underline{P}_n puede ser definido equivalentemente como el polinomio mónico de grado n que satisface:*

$$\langle f, \underline{P}_n \rangle_\mu = \mathbf{0}, \quad \forall f \in \mathcal{P}, \quad \deg f < n.$$

Demostración. Recordando lo visto en la Sección 1.1.3 tenemos que el sub-espacio \mathcal{P}_n es de dimensión finita, por ende, retomando la proyección vista en aquella sub-

sección, π_n , es la proyección de \mathcal{P} a \mathcal{P}_n . Así tomando la proyección de x^n en \mathcal{P}_{n-1} definimos:

$$\underline{P}_n(t) = t^n - \pi_n(t),$$

y este es nuestro polinomio buscado, pues al ser \mathcal{P}_{n-1} de dimensión finita, por un resultado clásico de análisis, es completo, por lo tanto podemos aplicar el teorema de la proyección y concluir que $\pi_n(t)$ es el polinomio de grado $n - 1$ que minimiza la norma de $t^n - Q$ para todo Q en \mathcal{P}_{n-1} . De igual forma por este mismo teorema tenemos la equivalencia de la ortogonalidad en \mathcal{P}_{n-1} :

$$\langle f, \underline{P}_n \rangle_\mu = \mathbf{0}, \implies \text{Tr}(\langle f, \underline{P}_n \rangle_\mu) = 0, \forall f \in \mathcal{P}_{n-1}.$$

Por lo tanto tomando $\alpha \in \mathcal{M}_l(\mathbb{C})$ tenemos que:

$$\text{Tr}(\alpha \langle f, \underline{P}_n \rangle_\mu) = 0 \implies \text{Tr}(\langle \alpha f, \underline{P}_n \rangle_\mu) = 0, \quad (1.14)$$

cumpliendo la segunda parte del lema. \square

Lema 1.6. *Sea μ una medida no-trivial, y \underline{P} un POM en \mathcal{P} , entonces $\det \langle \underline{P}, \underline{P} \rangle_\mu \neq 0$.*

Demostración. Supongamos por contradicción que la matriz $\langle \underline{P}, \underline{P} \rangle_\mu$ es no-invertible. Entonces podemos tomar una matriz α tal que $\alpha \langle \underline{P}, \underline{P} \rangle_\mu \alpha^* = \mathbf{0}$, solo basta considerar un vector v que anule a $\langle \underline{P}, \underline{P} \rangle_\mu$ y tomar a α como la matriz cuya primera columna es v y el resto se compone de ceros. Lo anterior implica que $\|\alpha \underline{P}\|_{\text{Tr}} = 0$. Sin embargo, al ser \underline{P} mónico, tenemos que $\alpha \underline{P}$ es distinto de cero, pero con norma cero, lo cual contradice el hecho de que nuestra medida sea no-trivial. \square

El siguiente teorema caracteriza a nuestros POM como una base del módulo derecho \mathcal{P}_n .

Teorema 1.7. *Sea μ una medida no-trivial.*

- *Tenemos que:*

$$\langle \underline{P}_m, \underline{P}_n \rangle_\mu = \gamma_n \delta_{n,m}, \quad (1.15)$$

para matrices invertibles γ_j .

- *Los elementos de $f \in \mathcal{P}_n$ tienen una representación única de la siguiente forma:*

$$f = \sum_{j=0}^n f_j \underline{P}_j. \quad (1.16)$$

Además por la ortogonalidad tenemos, esencialmente:

$$f_j = \langle f, \underline{P}_j \rangle_\mu \gamma_j^{-1}. \quad (1.17)$$

Demostración. Si $n < m$, por el Lema 1.5, tenemos que el producto es $\mathbf{0}$. Si $n > m$, concluimos lo mismo por las ecuaciones de (1.5), las γ_j son invertibles por el Lema 1.6. Ahora para probar la segunda afirmación consideramos el siguiente mapeo \mathcal{X} que manda a los elementos de $\mathcal{M}_l(\mathbb{C})^{n+1}$ en \mathcal{P}_n definido como:

$$\mathcal{X} : (\alpha_0, \alpha_1, \dots, \alpha_n) \rightarrow \sum_{j=0}^n \alpha_j \underline{P}_j. \quad (1.18)$$

Multiplicando a $\mathcal{X}(\alpha_0, \alpha_1, \dots, \alpha_n)$ por \underline{P}_i a la derecha, tenemos, por (1.18):

$$\begin{aligned} \langle \mathcal{X}(\alpha_0, \alpha_1, \dots, \alpha_n), \underline{P}_i \rangle_\mu &= \alpha_i \gamma_i \\ \implies \alpha_i &= \langle \mathcal{X}(\alpha_0, \alpha_1, \dots, \alpha_n), \underline{P}_i \rangle_\mu \gamma_i^{-1}. \end{aligned}$$

Esto pues las γ_i son invertibles. De esto es fácil concluir que \mathcal{X} es inyectiva, pues las α_i quedan definidas por la imagen, y por ende tienen que ser únicas para cada imagen. Ahora, resta ver la suprayectividad. Pero esto se sigue de que tanto el dominio como la imagen poseen la misma dimensión y de que \mathcal{X} es una transformación lineal. \square

1.2.2. Relaciones de recurrencia en los POM Mónicos

Veamos ahora las relaciones de recurrencia en este tipo de polinomios inspiradas en el caso escalar. Para ello denotamos como ζ_n , al coeficiente de t^{n-1} en $\underline{P}_n(t)$, (recuerde que al ser polinomios mónicos el coeficiente de x^n es \mathbf{I}). A continuación, usando los γ_n , definidos en el Teorema 1.7, y los ζ_n , es posible definir relaciones de recurrencia para $\underline{P}_n(t)$.

Lema 1.8. *Se cumplen las siguientes relaciones de recurrencia en los polinomios:*

$$t\underline{P}_n(t) = \underline{P}_{n+1}(t) + (\zeta_n - \zeta_{n+1})\underline{P}_n(t) + \gamma_n \gamma_{n-1}^{-1} \underline{P}_{n-1}(t).$$

Demostración. Usando el Teorema 1.7, tenemos que el polinomio $t\underline{P}_n(t)$ tiene una representación única de la forma :

$$t\underline{P}_n(t) = C_{n,n+1}\underline{P}_{n+1}(t) + C_{n,n}\underline{P}_n(t) + C_{n,n-1}\underline{P}_{n-1}(t) + \dots + C_{n,0}\underline{P}_0(t),$$

para algunas matrices complejas $C_{n,n+1}, \dots, C_{n,0}$. Claramente al comparar los coeficientes de t^{n+1} y t^n entre ambos polinomios, y dado que la representación es única, tenemos que $C_{n,n+1} = \mathbf{I}$ y $C_{n,n} = \zeta_n - \zeta_{n+1}$. Al aplicar el producto por la derecha entre $\underline{P}_k(t)$, con k entre 0 y $n-2$, y $t\underline{P}_n(t)$, por la ortogonalidad del Teorema 1.7, todos los términos del lado derecho, salvo el k -ésimo, se eliminan, ya que el lado izquierdo se anula, y por ende obtenemos que $C_{n,k} \langle \underline{P}_k, \underline{P}_k \rangle_\mu = \mathbf{0}$. Pero por el Lema 1.6 esto implica que $C_{n,k} = \mathbf{0}$. Solo resta calcular $C_{n,n-1}$, pero esto es sencillo pues:

$$\gamma_n = \langle \underline{P}_n, \underline{P}_n \rangle_\mu = \langle \underline{P}_n, t^n \rangle_\mu = \langle \underline{P}_n, t\underline{P}_{n-1} \rangle_\mu = \langle t\underline{P}_n, \underline{P}_{n-1} \rangle_\mu.$$

En la igualdad final usamos que $C_{n-1,n} = \mathbf{I}$. Desarrollando la parte izquierda del producto interno de esta última igualdad:

$$\langle \underline{P}_{n+1} + C_{n,n}\underline{P}_n + C_{n,n-1}\underline{P}_{n-1}, \underline{P}_{n-1} \rangle_\mu = C_{n,n-1}\gamma_{n-1},$$

donde lo último se da por la ortogonalidad y definición de γ_{n-1} . Así, obtenemos: $C_{n,n-1} = \gamma_n \gamma_{n-1}^{-1}$, concluyendo. \square

1.2.3. POM Normalizados

Siguiendo con la misma dirección que en el caso escalar, "normalizaremos" nuestra sucesión de polinomios previamente definida, en el sentido de que el producto interior de sí mismo será la identidad. En efecto, llamamos a $p_n \in \mathcal{P}$ como un polinomio ortonormal por la derecha si su grado es a lo más n y:

$$\langle f, p_n \rangle_\mu = \mathbf{0}, \quad \forall f \in \mathcal{P} \text{ con } \deg f < n, \quad \langle p_n, p_n \rangle_\mu = \mathbf{I}. \quad (1.19)$$

Lema 1.9. *Todo polinomio ortonormal posee la siguiente representación:*

$$p_n(t) = \sigma_n \langle \underline{P}_n, \underline{P}_n \rangle_\mu^{-1/2} \underline{P}_n(t), \quad (1.20)$$

donde $\sigma_n \in \mathcal{M}_l(\mathbb{C})$ son matrices unitarias.

Demostración. Considere K_n el coeficiente de t^n en $p_n(t)$, y a $q_n(t) = K_n \underline{P}_n(t) - p_n(t)$, donde justamente $\underline{P}_n(t)$ es el POM mónico definido en la sub-sección anterior. Entonces, claramente $\deg q_n < n$, y por lo tanto del Lema 1.5 y de (1.19) tenemos que

$\langle q_n(t), q_n(t) \rangle_\mu = 0$, lo cual implica necesariamente que $q_n(t) \equiv \mathbf{0}$. Y por ende:

$$\langle p_n, p_n \rangle_\mu = K_n \langle \underline{P}_n, \underline{P}_n \rangle_\mu K_n^* = \mathbf{I}.$$

Así las K_n son invertibles, y justo por esto último obtenemos que $\langle \underline{P}_n, \underline{P}_n \rangle_\mu = K_n^{-1}(K_n^*)^{-1} \implies K_n^* K_n = \langle \underline{P}_n, \underline{P}_n \rangle_\mu^{-1}$. Ahora probamos un resultado sobre esta representación de una matriz positiva-definida (que lo es pues $\langle \underline{P}_n, \underline{P}_n \rangle_\mu^{-1}$ existe por el Lema 1.6). Es bien conocido el hecho de que toda matriz definida positiva tiene una única raíz cuadrada que es definida positiva. Sea B la raíz cuadrada de $\langle \underline{P}_n, \underline{P}_n \rangle_\mu^{-1} = B^2$, por ende $K_n^* K_n = B^* B = B^2$ pues $B = B^*$, por ser definida positiva. Así considerando $U = K_n B^{-1}$, tenemos que $K_n = UB$, así si probamos que U es unitaria, concluimos. En efecto, note que $U^* U = (B^*)^{-1} K_n^* K_n B^{-1} = (B^*)^{-1} \langle \underline{P}_n, \underline{P}_n \rangle_\mu^{-1} B^{-1} = \mathbf{I}$, por ende tomando $\sigma_n = U$, tenemos la representación: $K_n = \sigma_n \langle \underline{P}_n, \underline{P}_n \rangle_\mu^{-1/2}$, y al ser $q(x) \equiv \mathbf{0}$, concluimos que: $p_n(t) = \sigma_n \langle \underline{P}_n, \underline{P}_n \rangle_\mu^{-1/2} \underline{P}_n(t)$. \square

Por el Teorema 1.7 tenemos que todo elemento $f \in \mathcal{P}_{\text{Tr}}$ tiene una representación única de la forma:

$$f(t) = \sum_{m=0}^{\infty} f_m p_m, \quad f_m = \langle f, p_m \rangle_\mu, \quad (1.21)$$

y que además, satisface la identidad de Parseval:

$$\sum_{m=0}^{\infty} \text{Tr}(f_m f_m^*) = \|f\|_{\text{Tr}}^2,$$

en donde claramente hay una cantidad finita de términos distintos de 0 en todo momento pues los f son polinomios.

1.2.4. Matrices de Jacobi por bloques

Veremos ahora, como en el caso escalar, la matriz de Jacobi por bloques generada por nuestros polinomios ortonormales. Considere una sucesión de σ_n unitarias, definidas en el Lema 1.9, tales que $\sigma_0 = \mathbf{I}$ y los polinomios ortonormales p_n igualmente tratados anteriormente. Como se probó, esta sucesión de polinomios es una base para \mathcal{P}_{Tr} . Ahora, considerando el homomorfismo derecho en \mathcal{P} dado por $f(t) \rightarrow tf(t)$, podemos representarlo como una matriz $(J)_{i,j}$, con respecto a la base de nuestros polinomios ortonormales, como:

$$J_{n,m} = \langle p_{n-1}(t), x p_{m-1}(t) \rangle_\mu.$$

Note que la matriz está numerada de $1, 2, \dots$ pero nuestros polinomios de $0, 1, 2, \dots$, es por eso que la anterior definición sigue esa notación (pues el producto tiene $n - 1$ y $m - 1$). Como en el caso escalar y por la ortogonalidad de los p_n (Lema 1.7), tenemos que $J_{n,m} = \mathbf{0}$ si y solo si $|n - m| > 1$. Así podemos definir:

$$B_n = J_{n,n} = \langle p_{n-1}(t), tp_{n-1}(t) \rangle_\mu,$$

y,

$$A_n = J_{n,n+1} = J_{n+1,n}^* = \langle p_{n-1}(t), tp_n(t) \rangle_\mu.$$

Así $(J)_{i,j}$ queda representado como:

$$J = \begin{pmatrix} B_1 & A_1 & \mathbf{0} & \mathbf{0} & \dots \\ A_1^* & B_2 & A_2 & \mathbf{0} & \dots \\ \mathbf{0} & A_2^* & B_3 & A_3 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \quad (1.22)$$

Aplicando J a p_n , obtenemos la relación de recurrencia:

$$tp_n(t) = A_{n+1}p_{n+1}(t) + B_{n+1}p_n(t) + A_n^*p_{n-1}(t). \quad (1.23)$$

Donde lo anterior se da para toda $n > 0$ y podemos incluso definirlo para $n = 0$, si tomamos $p_{-1} = \mathbf{0}$ y $A_0 = \mathbf{I}$. Esta fórmula de recursión es idéntica a la conseguida en el Lema 1.8. Así igualando término a término, y usando el Teorema 1.9 obtenemos:

$$A_n = \sigma_{n-1}\gamma_{n-1}^{-1/2}\gamma_n^{1/2}\sigma_n^*, \quad y \quad B_n = \sigma_{n-1}\gamma_{n-1}^{-1/2}(\zeta_{n-1} - \zeta_n)\gamma_{n-1}^{1/2}\sigma_{n-1}^*, \quad (1.24)$$

y en particular, de aquí obtenemos que todas las A_n 's son invertibles. Como σ_n es unitaria su determinante es 1, y la ecuación anterior implica que $\det(\gamma_n^{1/2}) = \det(\gamma_{n-1}^{1/2})|\det(A_n)|$, y por inducción:

$$\det(\langle \underline{P}_n, \underline{P}_n \rangle_\mu) = \det(A_1 A_2 \dots A_n)^2.$$

Toda matriz tridiagonal por bloques de la forma (1.22) con las B_n hermíticas y las A_n invertibles será llamada *Matriz de Jacobi*.

1.2.5. Teorema de Favard

En esta subsección demostraremos uno de los teoremas fundamentales para el estudio de lo que es nuestro tema, el teorema de Favard, el cual es una generalización, como todo lo anterior, al caso escalar. Empezamos con una matriz de Jacobi con coeficientes A_n , invertibles, y B_n hermíticas, para toda n . Si consideramos el espacio de secuencias, que a su vez es un espacio de Hilbert, $\mathcal{H}_v = \ell^2(\mathbb{Z}_+, \mathbb{C}^l)$, (donde \mathbb{Z}_+ es el conjunto de enteros positivos), con producto interior:

$$\langle f, g \rangle_{\mathcal{H}_v} = \sum_{n=1}^{\infty} \langle f_n, g_n \rangle,$$

donde el producto interno del lado derecho es el clásico en \mathbb{C}^l , y los f_n y g_n son los vectores componentes de f y g . Por ende, dada esta definición, podemos definir la siguiente base ortonormal en este espacio: $\{e_{k,j}\}_{k \in \mathbb{Z}_+, 1 \leq j \leq l}$, donde el vector $e_{k,j}$ es aquel cuyos vectores componentes son todos el vector 0, salvo el k -ésimo componente, en el cual es el vector canónico e_j . Podemos definir, igualmente, como J , nuestra matriz de Jacobi, actúa en \mathcal{H}_v , simplemente realizamos la multiplicación “matricial” de J a un vector de \mathcal{H}_v , es decir:

$$(Jf)_n = A_n^* f_{n-1} + B_{n+1} f_n + A_{n+1} f_{n+1}, \quad f \in \mathcal{H}_v.$$

(en este caso tomamos $f_0 = 0$) el cual claramente, por la simetría de B_n , es un operador simétrico. Usando la invertibilidad de las A_n 's, e inducción en n , es fácil probar que:

$$\text{span}\{e_{k,j} : 1 \leq k \leq n, 1 \leq j \leq l\} = \text{span}\{J^{k-1} e_{1,j} : 1 \leq k \leq n, 1 \leq j \leq l\}, \quad (1.25)$$

para todo n , pues sólo basta con ver el vector n -ésimo componente de $J^{n-1} e_{1,j}$, ya que por inducción las componentes $\leq n-1$ sí producen el mismo generado, el cual es: $A_1^* A_2^* \cdots A_n^* v_j$. Pero al ser las A_n 's invertibles, tenemos que $A_1^* A_2^* \cdots A_n^*$ igual lo es, y por ende:

$$\text{span}\{e_{n,j} : 1 \leq j \leq l\} = \text{span}\{J^{n-1} e_{1,j} : 1 \leq j \leq l\}.$$

En lo que sigue en adelante consideraremos solo operadores de bloques de Jacobi acotados, es decir aquellos que cumplan:

$$\sup_{n \geq 1} \text{Tr}(A_n A_n^* + B_n B_n^*) < \infty.$$

Al igual que en el caso escalar, esto es equivalente a ver que nuestra medida μ tiene soporte compacto. En efecto, esto básicamente se sigue del hecho de que al ser J acotado, su espectro igual lo está, pero como en el caso escalar, el espectro de J se puede identificar como los ceros de los polinomios matriciales (en este caso los ceros del polinomio matricial son los ceros del polinomio definido por su determinante) y nuestros ceros definen el soporte de la medida de ortogonalidad (para más información ver [11]).

Ahora veamos una relación de equivalencia entre matrices de Jacobi. Diremos que dos matrices J y \tilde{J} están relacionadas, o mejor dicho son “equivalentes”, si y solo si existe una sucesión de matrices unitarias; $\{u_k\}_{k=1}^{\infty}$ con $u_1 = \mathbf{I}$ y $\tilde{J}_{n,m} = u_n J_{n,m} u_m^*$, para toda $n, m \geq 1$. Así, usando el Lema 1.9, note que si tenemos dos secuencias de polinomios matriciales ortonormales p_n y \tilde{p}_n respecto a la misma medida (pero con posiblemente distinta normalización), usando σ_n para p_n y $\tilde{\sigma}_n$, tenemos que las matrices son equivalentes pues:

$$\tilde{J}_{n,m} = \langle \tilde{p}_{n-1}, t\tilde{p}_{m-1} \rangle_{\mu} = \langle u_n p_{n-1}, t u_m p_{m-1} \rangle_{\mu} = u_n J_{n,m} u_m^*,$$

con $u_k = \sigma_{k-1} \tilde{\sigma}_{k-1}^*$, y por ende obtenemos, de (1.24):

$$\tilde{B}_n = u_n B_n u_n^*, \quad \tilde{A}_n = u_n A_n u_{n+1}^*. \quad (1.26)$$

De esto podemos crear el siguiente mapeo entre todas las medidas Hermitianas positivas semi-definidas no triviales y con soporte compacto, y las clases de equivalencia de las matrices de Jacobi por bloques y acotadas:

$$\Phi : \mu \rightarrow \{[J] : J \text{ es una matriz de Jacobi, para una normalización de } \mu\}. \quad (1.27)$$

Ahora probaremos la existencia de matrices de Jacobi, con determinadas propiedades que serán definidas a continuación, que son representantes de la clase de equivalencia a la que pertenecen. Para esto considere a J una matriz de Jacobi con parámetros A_n y B_n . Diremos que J es del:

- tipo 1 si A_n es definida positiva para toda n .
- tipo 2 si $A_1 A_2 \cdots A_n$ es definida positiva para toda n .
- tipo 3 si $A_n \in \mathcal{L}$, donde \mathcal{L} es el conjunto de matrices triangulares inferiores con elementos en la diagonal positivos.

Con base en estas definiciones probamos de forma concreta lo mencionado inicialmente.

Lema 1.10.

- Cada clase de equivalencia de una matriz de Jacobi posee exactamente un elemento del tipo 1, tipo 2 y tipo 3.
- Sea J una matriz de Jacobi correspondiente a la secuencia $\{p_n(t)\}$ de polinomios matriciales ortonormales. Entonces J es del tipo 2 si y solo si $\sigma_n = \mathbf{I}$ para toda n .

Demostración. Usamos dos resultados sobre la descomposición matricial polar y de la forma QR para matrices invertibles:

- Sea $M \in \mathcal{M}_l(\mathbb{C})$, invertible, entonces existe una matriz unitaria U tal que MU es positiva definida.
- Sea $M \in \mathcal{M}_l(\mathbb{C})$, invertible, entonces existe una matriz unitaria U tal que $MU \in \mathcal{L}$.

El primer hecho viene de la descomposición polar de $M = PU^*$, con P definida positiva y U^* unitaria, así tomamos simplemente esta U , que es unitaria igualmente, y así $MU = P$ es definida positiva. Además como M es invertible la factorización es única, pues note que $\sqrt{MM^*} = \sqrt{PP^*} = P$, donde el radical simboliza la única raíz cuadrada que una matriz definida positiva tiene. Por ende, P queda definida de forma única y $U = M^{-1}P$ queda definido de forma única. Ahora para el segundo hecho, como M es invertible, sus filas, m_1, m_2, \dots, m_e forman una base para \mathbb{C}^e . Por lo tanto podemos aplicar el proceso de Gram-Schmidt, empezando en la última fila, para obtener al conjunto ortonormal, pues es fácil normalizar, u_1, u_2, \dots, u_e , estos vectores definen una matriz Q , que claramente es unitaria, con sus columnas, y por las ecuaciones obtenidas de Gram-Schmidt al multiplicar a M por Q^* por la derecha, obtenemos una matriz triangular inferior con los elementos de las diagonales representando a las normas de nuestros vectores de las filas de M , y por ende positivas.

Regresando a la demostración, probamos primero que en nuestra clase de equivalencia de J existe un único elemento del tipo 1. Para esto considere la secuencia de matrices A_n , para la cual vamos a definir una sucesión de matrices unitarias $u_1 = \mathbf{I}, u_2, \dots$ tales que $u_n^* A_n u_{n+1} > \mathbf{0}$. Para esto definimos a u_2 como la matriz del primer hecho tal que $A_1 u_2 > \mathbf{0}$, luego a u_3 tal que $u_2^* A_2 u_3 > \mathbf{0}$, y así sucesivamente. Esto junto con (1.26) prueba la equivalencia a una matriz de Jacobi de tipo 1. Nos resta ver la unicidad. Pero esto es simple pues si $A_n > \mathbf{0}$ para toda n , y tomando la sucesión de matrices unitarias previamente definida tenemos que $u_n^* A_n u^{n+1} > \mathbf{0}$ pero $A_1 > \mathbf{0}$ implica que como $A_1 u_2 > \mathbf{0}$ tenemos que $u_2 = \mathbf{I}$, lo que a su vez implica que $A_2 u_3 > \mathbf{0}$ pero $A_2 > \mathbf{0}$ lo que implica que $u_3 = \mathbf{I}$, e inductivamente tenemos que cada $u_n = \mathbf{I}$, por ende la matriz de Jacobi sería única.

Para el tipo 3 simplemente aplicamos los mismos pasos anteriores, pues solo basta usar el segundo hecho del inicio en lugar del primero. Para el tipo 2 el razonamiento es similar pues consideramos u_2 , que existe por el primer hecho, tal que $A_1 u_2 > \mathbf{0}$, para u_3 lo tomamos tal que $(A_1 u_2)(u_2^* A_2 u_3) = A_1 A_2 u_3 > \mathbf{0}$, y definimos el resto de forma inductiva. La unicidad se prueba igual, si $A_1 > \mathbf{0}$ y $A_1 u_2 > \mathbf{0}$, entonces $u_2 = \mathbf{I}$, e inductivamente si $A_1 A_2 \cdots A_n u_{n+1} > \mathbf{0}$, entonces como $A_1 A_2 \cdots A_n > \mathbf{0}$ tenemos que $u_{n+1} = \mathbf{I}$.

Para la segunda proposición del teorema recordamos la ecuación (1.24) que nos da $A_1 A_2 \cdots A_n = \gamma_n^{1/2} \sigma_n$, por ende si J fuera de tipo 2, como las matrices γ_n son definidas positivas, por la unicidad del segundo hecho esto implicaría que $\sigma_n = \mathbf{I}$ para toda n . Ahora si cada $\sigma_n = \mathbf{I}$, entonces $A_1 A_2 \cdots A_n = \gamma_n^{1/2} > \mathbf{0}$, concluyendo. \square

Tenemos ya las herramientas para poder probar el Teorema de Favard, el cual básicamente es probar que la función definida en (1.27) es biyectiva. Probamos primero la inyectividad en el siguiente resultado.

Teorema 1.11. *El mapeo Φ , definido en (1.27) es inyectivo.*

Demostración. Considere μ y $\tilde{\mu}$ dos medidas semi-definidas positivas, no triviales y con soporte compacto. Supongamos que inducen mediante Φ a la misma clase de equivalencia de su matriz de Jacobi, es decir $\Phi(\mu) = \Phi(\tilde{\mu})$.

Considere las secuencias $\{p_n\}$ y $\{\tilde{p}_n\}$ de polinomios matriciales ortonormales correspondientes a μ y $\tilde{\mu}$, respectivamente. Suponga, sin pérdida de generalidad, que la normalización es tal que tanto para p_n como para \tilde{p}_n hemos elegido a $\sigma_n = \tilde{\sigma}_n = \mathbf{I}$, es decir por el lema anterior de tipo 2, (recordar que en el Lema 1.9 cualquier matriz unitaria σ_n cumple, entonces podemos elegirla de tal forma que sean la identidad). Como por el Lema 1.9, las matrices de tipo 2 son únicas en su clase de equivalencia, y por hipótesis $\Phi(\mu) = \Phi(\tilde{\mu})$, es decir están en la misma clase de equivalencia, esto implica que las matrices de Jacobi deben coincidir, es decir $\langle p_n(t), t p_m(t) \rangle_\mu = \langle \tilde{p}_n(t), t \tilde{p}_m(t) \rangle_\mu$, para toda $m, n \geq 0$. Sabiendo esto y recordando que los polinomios ortonormales son definidos por la relación de recurrencia (1.23), tenemos que forzosamente deben de ser los mismos polinomios: $p_n = \tilde{p}_n$, para toda $n \geq 0$.

Así, para toda $n \geq 0$, por (1.21), podemos representar a x^n como:

$$t^n = \sum_{k=0}^n C_k^{(n)} p_k(t) = \sum_{k=0}^n \tilde{C}_k^{(n)} \tilde{p}_k(t).$$

En donde claramente los coeficientes $C_k^{(n)}$ y $\tilde{C}_k^{(n)}$ están determinados en función de los coeficientes de los polinomios p_n y \tilde{p}_n , respectivamente, y al ser estos iguales se

obtiene que $C_k^{(n)} = \tilde{C}_k^{(n)}$. Para los momentos de las medidas, simplemente sustituimos la representación anterior en la integral dada:

$$\int_{\mathbb{R}} t^n d\mu(t) = \langle t^n, \mathbf{I} \rangle_{\mu} = \sum_{k=0}^n \langle C_k^{(n)} p_k(t), \mathbf{I} \rangle_{\mu} = C_0^{(n)} \langle \mathbf{I}, \mathbf{I} \rangle_{\mu} = C_0^{(n)},$$

asumiendo que $\mu(\mathbb{R}) = \mathbf{I}$. Por ende, este mismo cálculo es válido para $\tilde{\mu}$, y así:

$$\int_{\mathbb{R}} t^n d\mu(t) = \int_{\mathbb{R}} t^n d\tilde{\mu}(t), \quad n \geq 0.$$

Por lo tanto, es válido lo siguiente:

$$\int_{\mathbb{R}} f(t) d\mu(t) g^*(t) = \int_{\mathbb{R}} f(t) d\tilde{\mu}(t) g^*(t),$$

para todo polinomio matricial f, g , implicando que las medidas coinciden, como queríamos. \square

Para terminar de probar la biyección nos resta probar la supreyectividad. Sin embargo, procederemos por probar esto de una forma indirecta. Construiremos un mapeo inverso para Φ . Tomemos una matriz de Jacobi J . Usando una versión particular del teorema espectral para operadores auto-adjuntos con multiplicidad finita, la cual puede ser encontrada en [1, Sección 72], existe una medida matricial $d\mu$ con:

$$\langle f(J)e_{1,k}, e_{1,j} \rangle_{\mathcal{H}_v} = \int f(t) d\mu_{j,k}(t),$$

y una isometría:

$$R : \mathcal{H}_v \rightarrow L^2(\mathbb{R}, d\mu; \mathbb{C}^l).$$

Es decir de nuestro espacio \mathcal{H}_v previamente definido, al espacio de funciones L^2 , con dominio en los reales, bajo la medida $d\mu$, que induce a la semi-norma $|\cdot|_{\mathcal{V}}$, y con valores vectoriales en \mathbb{C}^l . Esta isometría satisface que (recordando que e_j es la base canónica de \mathbb{C}^l):

$$[Re_{1,j}](x) = e_j, \quad 1 \leq j \leq l, \quad (1.28)$$

y, para cada $g \in \mathcal{H}_v$, tenemos:

$$(RJg)(t) = t(Rg)(t). \quad (1.29)$$

Si J y \tilde{J} son equivalentes, entonces por las ecuaciones de (1.24), tenemos que $\tilde{J} = UJU^*$ para algún $U = \bigoplus_{n=1}^{\infty} u_n$, con $u_1 = \mathbf{I}$, donde U es una matriz diagonal por bloques con el j, k -ésimo bloque igual a $\delta_{j,k}u_j$, y las u_i son unitarias siguiendo las

ecuaciones de (1.24). Por ende:

$$\langle f(\tilde{J})e_{1,k}, e_{1,j} \rangle_{\mathcal{H}_v} = \langle Uf(J)U^*e_{1,k}, e_{1,j} \rangle_{\mathcal{H}_v} = \langle f(J)U^*e_{1,k}, U^*e_{1,j} \rangle_{\mathcal{H}_v} = \langle f(J)e_{1,k}, e_{1,j} \rangle_{\mathcal{H}_v},$$

donde la primera igualdad se da del hecho de que $\tilde{J}^n = U^*J^nU$, pues claramente $UU^* = \mathbf{I}$, al ser las u_i unitarias y las funciones f polinomios. La segunda igualdad se da de la identidad:

$$\langle f, U^*g \rangle_{\mathcal{H}_v} = \sum_{n=1}^{\infty} \langle f_n, U^*g_n \rangle = \sum_{n=1}^{\infty} \langle Uf_n, g_n \rangle = \langle Uf, g \rangle_{\mathcal{H}_v},$$

y finalmente la tercera se da pues $u_1 = \mathbf{I}$. Por lo tanto las medidas en J y \tilde{J} coinciden, lo cual nos da el siguiente mapeo en las clases de equivalencia:

$$\Psi : \{ \tilde{J} : \tilde{J} \text{ es equivalente a } J \} \rightarrow \mu.$$

Teorema 1.12.

- *Todas las medidas en la imagen de Ψ son no-triviales.*
- $\Phi \circ \Psi = Id_{\{ \tilde{J} : \tilde{J} \text{ es equivalente a } J \}}.$
- $\Psi \circ \Phi = Id_{\{ \mu : \mu \text{ no trivial} \}}.$

Demostración. Empezamos con la primera proposición. Para probar que estas medidas son no-triviales hay que ver primero que la semi-norma definida en la Subsección 1.2, $|\cdot|_{\mathcal{V}}$ en \mathcal{V} es de hecho una norma. Más aún, veremos que $|p|_{\mathcal{V}} > 0$ para todo polinomio p en \mathcal{V} , y por el Lema 1.3 la medida sería no-trivial.

Sea $p \in \mathcal{V}$ un polinomio distinto de cero y con grado n . Retomando el resultado del teorema espectral de (1.28) y (1.29), tenemos que:

$$[RJ^k e_{1,j}](t) = t[RJ^{k-1} e_{1,j}](t) = \dots = t^k v_j, \quad (1.30)$$

para todo k no negativo y $1 \leq j \leq l$. Sin embargo, como R es una isometría, dado k entre 0 y n , mediante una combinación lineal de elementos de la forma $J^k e_{1,j}$ podemos obtener, al aplicar R el coeficiente vectorial de p en su sumando t^k , es decir, existen f_0, \dots, f_n vectores en \mathcal{H}_v , que son combinación lineal de elementos de la forma $e_{1,j}$, tales que si $g = \sum_{k=0}^n J^k f_k$, se satisface que $p(t) = [Rg](t)$. (Note que al ser los f_i generados por los elementos $e_{1,j}$, las únicas componentes distintas de 0 que poseen se encuentran en el primer vector en \mathcal{H}_v).

Como R es una isometría, tenemos que $|p|_{\mathcal{V}} = \|g\|_{\mathcal{H}_v}$, por ende, probar que $|p|_{\mathcal{V}} > 0$ es

igual probar que $\|g\|_{\mathcal{H}_v} > 0$. Procedemos por contradicción suponiendo que $\|g\|_{\mathcal{H}_v} = 0$. Recordando que $\det(A_k) \neq 0$, y fijándonos en la n -ésima componente de g , que es $(J^n f_n)_n$, por la tridiagonalidad de J , si $g = 0 \implies (J^n f_n)_n = 0 \implies f_n = 0$, lo cual es imposible pues al tener p grado n , y $[RJ^n f_n](t) = t^n u_n \neq 0$, donde u_n es el coeficiente de p en t^n , lo cual es una contradicción, concluyendo la no-trivialidad.

Para la segunda proposición del teorema, consideramos nuestra clase de equivalencia de matrices de Jacobi J , y la medida obtenida, μ por la función Ψ . Tomando los elementos $Re_{n,k} \in L^2(\mathbb{R}, d\mu; \mathbb{C}^l)$, note que por (1.25), $e_{n,k}$ es generado por elementos de la forma $J^{i-1}e_{1,j}$, con i entre 0 y $k-1$ y aplicando (1.30) tenemos que $Re_{n,k}$ es un polinomio de grado a lo más $n-1$. Siguiendo la naturaleza de isometría de R , que preserva el producto interior, obtenemos que:

$$\langle Re_{n,k}, Re_{m,j} \rangle_{L^2(\mathbb{R}, d\mu; \mathbb{C}^l)} = \langle e_{n,k}, e_{m,j} \rangle_{\mathcal{H}_v} = \delta_{m,n} \delta_{k,j}. \quad (1.31)$$

Con base en esta información, construimos nuestros candidatos a polinomios matriciales ortonormales $q_n(t)$, usando a $Re_{n,1}, Re_{n,2}, \dots, Re_{n,l}$, (que son vectores), como columnas de $q_{n-1}(t)$:

$$[q_{n-1}(x)]_{j,k} = [Re_{n,k}(t)]_j.$$

Tenemos que $\deg q_n \leq n$ y que por (1.31) $\langle q_m, q_n \rangle_\mu = \delta_{m,n} \mathbf{I}$, lo cual afirma que nuestros candidatos son de hecho polinomios ortonormales con respecto a $d\mu$. Así:

$$\begin{aligned} J_{n,m} &= [\langle e_{n,j}, Je_{m,k} \rangle_{\mathcal{H}_v}]_{1 \leq j,k \leq l} \\ &= [\langle Re_{n,j}, RJe_{m,k} \rangle_{L^2(\mathbb{R}, d\mu; \mathbb{C}^l)}]_{1 \leq j,k \leq l} \\ &= [\langle Re_{n,j}, xRe_{m,k} \rangle_{L^2(\mathbb{R}, d\mu; \mathbb{C}^l)}]_{1 \leq j,k \leq l} \\ &= [\langle [q_{n-1}(t)]_{\cdot,j}, t[q_{m-1}(t)]_{\cdot,k} \rangle_{L^2(\mathbb{R}, d\mu; \mathbb{C}^l)}]_{1 \leq j,k \leq l} \\ &= \langle q_{n-1}, tq_{m-1} \rangle_\mu. \end{aligned}$$

La primera igualdad viene del hecho de que los elementos $e_{i,j}$ son una base ortonormal. La segunda de la propiedad isométrica de R . La tercera de la ecuación (1.29). La cuarta de la definición de los polinomios q_n , y finalmente la quinta viene de la definición inicial de nuestro producto interior. De esto concluimos que $\Phi \circ \Psi$ es la identidad, pues pasamos de J a la medida definida por el teorema espectral y regresamos de nuevo a J .

La última proposición del teorema es inmediata de lo anterior y del hecho de que Φ es inyectiva, como fue visto en el Teorema 1.10.

1.3. Momentos canónicos en POM

1.3.1. Momentos canónicos de una medida matricial

Seguiremos ahora generalizando la teoría desarrollada en los polinomios ortogonales escalares, y obtener su forma equivalente en una secuencia de POM. Nos basaremos principalmente en lo desarrollado en [10]. Veremos bajo qué condiciones dada una secuencia finita de matrices cuadradas, S_0, S_1, \dots, S_n de tamaño $l \times l$, existe una medida matricial $\mu_{[a,b]}$, cuyo soporte se encuentra contenido en el intervalo $[a, b]$, con $a < b$, para la cual su k -ésimo momento matricial queda definido por:

$$S_k = \int_{[a,b]} t^k d\mu_{[a,b]}(t) = \langle t^k \mathbf{I}, \mathbf{I} \rangle_{\mu_{[a,b]}}. \quad (1.32)$$

Después definiremos los momentos canónicos, como en el caso escalar, para esta medida matricial $\mu_{[a,b]}$, y su relación con los POM correspondientes a la medida $\mu_{[a,b]}$.

Sea $M_{n+1}^{[a,b]}$ el $(n+1)$ -ésimo espacio de momentos definido como el conjunto:

$$M_{n+1}^{[a,b]} = \{(S_0, S_1, \dots, S_n) \mid \mu_{[a,b]} \text{ medida matricial}\}, \quad (1.33)$$

es decir el conjunto de vectores matriciales de los primeros $n+1$ momentos matriciales definidos en la ecuación (1.32). Note que este conjunto puede ser visto como un espacio euclideo lineal de dimensión $(n+1)l(l+1)$, o de dimensión $(n+1)l(l+1)/2$, en el caso real, por la simetría de las matrices S_k . Los “ hiperplanos ” de este espacio quedarán definidos de la forma:

$$\sum_{i=0}^n \text{Tr}(A_i S_i) = c,$$

donde $c \in \mathbb{R}$, y las matrices A_0, A_1, \dots, A_n son hermíticas de tamaño $l \times l$. Buscaremos encontrar una forma equivalente de expresar el conjunto $M_{n+1}^{[a,b]}$. Para esto considere el conjunto:

$$C_{n+1}^{[a,b]} = \{(vv^*, tvv^*, \dots, t^n vv^*) \mid a \leq t \leq b, v \in \mathbb{C}^l\},$$

y definimos como $\mathcal{C}(C_{n+1}^{[a,b]})$ al cono convexo generado por $C_{n+1}^{[a,b]}$, es decir la entrada k -ésima de un elemento de $\mathcal{C}(C_{n+1}^{[a,b]})$ puede ser visto de la forma:

$$\sum_i \sum_j t_i^k d_{i,j} v_{i,j} v_{i,j}^*, \quad (1.34)$$

donde los números $d_{i,j}$ son reales positivos y $v_{i,j} \in \mathbb{C}^l$. Note que existe cierta redundancia pues es posible normalizar los vectores $v_{i,j}$ para que tengan norma 1, o equivalentemente, prescindir de los escalares $d_{i,j}$. Probamos a continuación la equivalencia, previamente mencionada, entre $M_{n+1}^{[a,b]}$ y $\mathcal{C}(C_{n+1}^{[a,b]})$.

Lema 1.13. *Los conjuntos $M_{n+1}^{[a,b]}$, definido en la ecuación (1.33) y $\mathcal{C}(C_{n+1}^{[a,b]})$ son iguales.*

Demostración. Considerando una medida matricial, definida en la primera sección de este capítulo, $\mu_{[a,b]}$, tal que $\mu_{[a,b]}(\{s\}) = vv^*$, con $v \in \mathbb{C}^l$ y $a \leq s \leq b$, y $\mu_{[a,b]}(C) = \mathbf{0}$ para todo $C \subset [a, b]$ medible. Tenemos que:

$$S_k = \int_{[a,b]} t^k d\mu_{[a,b]}(t) = s^k vv^* \quad \forall k \in \mathbb{N}_0.$$

Por lo tanto $C_{n+1}^{[a,b]} \subset M_{n+1}^{[a,b]}$, y por la linealidad en las medidas matriciales obtenemos que $\mathcal{C}(C_{n+1}^{[a,b]}) \subset M_{n+1}^{[a,b]}$. Probamos ahora la contención inversa, demostrando que todos los elementos de $M_{n+1}^{[a,b]}$ son puntos límite de elementos en $\mathcal{C}(C_{n+1}^{[a,b]})$. Esto pues el cono convexo es un conjunto cerrado en nuestro espacio en cuestión. Supongamos, por ahora, que nuestra medida matricial $\mu_{[a,b]}$ está formada, entrada a entrada, por densidades diferenciables con respecto a la medida de Lebesgue, es decir, existe una función matricial y diferenciable, f , entrada a entrada, respecto a la medida de Lebesgue, con dominio en el intervalo $[a, b]$. Por lo tanto, las integrales pueden ser vistas como integrales de Riemann, y por ende podemos aproximar la integral $\int t^k f(t) dt$, el k -ésimo momento de $\mu_{[a,b]}$, por una suma de la forma:

$$\sum_i c_i t_i^k f(t_i), \tag{1.35}$$

donde el número $c_i > 0$ y $f(t_i) \in \mathcal{M}_l(\mathbb{C})$ es una matriz definida positiva. En este caso, por la propiedad de $f(t_i)$ de ser definida positiva, esta puede ser escrita como:

$$f(t_i) = \sum_j d_{i,j} v_{i,j} v_{i,j}^*. \tag{1.36}$$

Esta representación se sigue de escribir a $f(t_i) = Q^* D Q$ con D diagonal, y Q unitaria, por el Teorema Espectral. Al ser D diagonal, podemos escribirla como la siguiente suma: $D = \sum_j \lambda_j e_j e_j^*$, donde e_j es el j -ésimo vector canónico de \mathbb{C}^l y los valores λ_j son los auto-valores de $f(t_i)$, que sabemos que son reales por el Teorema Espectral y positivos pues $f(t_i)$ es definida positiva. Así $f(t_i) = \sum_j \lambda_j Q e_j (Q e_j)^*$, simplemente tomamos $Q e_j = v_{i,j}$ y $\lambda_j = d_{i,j}$ y obtenemos (1.36). Sustituyendo (1.36) en (1.35)

obtenemos una aproximación de la forma (1.34), por lo que, para este caso el resultado se concluye. Para el caso general en donde $\mu_{[a,b]}$ no está compuesta de medidas diferenciables respecto la medida de Lebesgue, simplemente usamos el resultado famoso sobre el problema de los momentos truncados para un intervalo finito, ver [20, págs. 62-63], y concluimos que los momentos de cada entrada de la medida matricial $\mu_{[a,b]}$ pueden ser aproximados por una medida diferenciable como las tratadas anteriormente. \square

Continuamos ahora con el siguiente corolario.

Corolario 1.14. *Sean A_0, A_1, \dots, A_n matrices hermíticas de tamaño $l \times l$. Entonces se satisfacen las propiedades:*

1.

$$\sum_{k=0}^n t^k A_k \geq \mathbf{0}, \quad \forall t \in [a, b]$$

$$\iff \sum_{k=0}^n \text{Tr}(A_k S_k) \geq 0, \quad \forall S = (S_0, \dots, S_n) \in M_{n+1}^{[a,b]}.$$

2. *Cada punto $S = (S_0, \dots, S_n)$ en $M_{n+1}^{[a,b]}$ tiene una representación finita de la forma:*

$$S_k = \sum_{i=1}^q d_i a_i a_i^* t_i^k, \quad d_i \in \mathbb{R}^+, \quad a_i \in \mathbb{C}^l, \quad y \quad k = 0, \dots, n,$$

donde $q \leq (n+1)l(l+1)$ o en el caso real $q \leq (n+1)l(l+1)/2$.

3. *Considere la función $f : [a, b] \times \mathbb{C}^l \rightarrow \mathcal{M}_l(\mathbb{C})^n$ definida de la forma:*

$$f(t, v) = (vv^*, tvv^*, \dots, t^n vv^*) \quad \forall (t, v), \quad \text{con } t \in [a, b] \text{ y } v \in \mathbb{C}^l. \quad (1.37)$$

Entonces si $S = (S_0, \dots, S_n)$ está en el interior de $M_{n+1}^{[a,b]}$, S es una combinación convexa, entre $f(t, v)$ y un punto en la frontera de $M_{n+1}^{[a,b]}$, donde esto pasa para todo $(t, v) \in [a, b] \times \mathbb{C}^l$.

Demostración. Empezamos con la primera proposición. Note que:

$$\begin{aligned}
 & \sum_{k=0}^n t^k A_k \geq \mathbf{0}, \quad \forall t \in [a, b], \\
 & \iff \sum_{k=0}^n t^k v^* A_k v \geq 0, \quad \forall t \in [a, b] \text{ y } \forall v \in \mathbb{C}^l, \\
 & \iff \sum_{k=0}^n \text{Tr}(A_k t^k v v^*) \geq 0, \quad \forall t \in [a, b] \text{ y } \forall v \in \mathbb{C}^l, \\
 & \iff \sum_{k=0}^n \text{Tr}(A_k S_k) \geq 0, \quad \forall S = (S_0, \dots, S_n) \in M_{n+1}^{[a,b]},
 \end{aligned}$$

en donde la segunda equivalencia es resultado de la propiedad matricial $\text{Tr}(BC) = \text{Tr}(CB)$ para cualesquiera matrices B, C . La tercera equivalencia es resultado del Lema 1.13.

La segunda proposición se obtiene de aplicar el Teorema de Carathéodory sobre las envolventes convexas, y del hecho que nuestro espacio posee una dimensión de tamaño $(n+1)l(l+1)$ ó $(n+1)l(l+1)/2$ cuando estamos en el caso real.

Finalmente la tercera proposición se sigue de extender la semi recta entre $f(t, v)$ y S , y considerar su punto de intersección con la frontera de $M_{n+1}^{[a,b]}$, que al S pertenecer al interior de $M_{n+1}^{[a,b]}$ este punto de intersección es distinto de S , por ende S puede ser escrito como una combinación lineal entre $f(t, v)$ y el punto correspondiente en la frontera. \square

Lema 1.15. *Sea $S = (S_0, \dots, S_n)$, un vector matricial. Entonces:*

1. $S \in M_{n+1}^{[a,b]}$ si y solo si:

$$\sum_{k=0}^n t^k A_k \geq \mathbf{0}, \quad \forall t \in [a, b] \implies \sum_{k=0}^n \text{Tr}(A_k S_k) \geq 0. \quad (1.38)$$

2. $S \in \text{int}(M_{n+1}^{[a,b]})$ si y solo si:

$$\begin{aligned}
 P(t) = \sum_{k=0}^n t^k A_k \geq \mathbf{0}, \quad \forall t \in [a, b] \text{ y } P(t) \neq \mathbf{0} \quad \forall t \in [a, b], \\
 \implies \sum_{k=0}^n \text{Tr}(A_k S_k) > 0.
 \end{aligned} \quad (1.39)$$

Demostración. Note que por la primera parte del Corolario 1.14, si se cumple que

$S \in M_{n+1}^{[a,b]}$, entonces se sigue la primera implicación de la ecuación (1.38). Ahora, nos enfocamos en la segunda parte de la equivalencia. Por contradicción asumimos que $C \notin M_{n+1}^{[a,b]}$ implica la proposición contrapuesta de (1.38). Explícitamente $C = (C_0, C_1, \dots, C_n)$. Veamos más a detalle cuál es el valor de $\text{Tr}(A_k C_k)$, para una k dada:

$$\text{Tr}(A_k C_k) = \sum_{i=1}^l a_{i,i}^k c_{i,i}^k + \sum_{1 \leq i < j \leq l} \left[a_{i,j}^k \overline{c_{i,j}^k} + \overline{a_{i,j}^k} c_{i,j}^k \right],$$

donde los valores $a_{i,j}^k, c_{i,j}^k$ representan las (i, j) -ésimas entradas de las matrices A_k y C_k , respectivamente. Note que por la misma simetría de A_k y C_k las entradas diagonales $a_{i,i}^k$ y $c_{i,i}^k$, de ambas son números reales. Ahora asumiendo primero que el resto de las entradas son complejas, escribimos a $a_{i,j}^k = x_{i,j}^k + iy_{i,j}^k$ y $c_{i,j}^k = p_{i,j}^k + iq_{i,j}^k$, con $x_{i,j}^k, y_{i,j}^k, p_{i,j}^k, q_{i,j}^k \in \mathbb{R}$. Por ende, la última ecuación se puede reescribir como:

$$\begin{aligned} \text{Tr}(A_k C_k) &= \sum_{i=1}^l a_{i,i}^k c_{i,i}^k + \sum_{1 \leq i < j \leq l} \left[2\text{Re}(a_{i,j}^k \overline{c_{i,j}^k}) \right] \\ &= \sum_{i=1}^l a_{i,i}^k c_{i,i}^k + \sum_{1 \leq i < j \leq l} \left[2(p_{i,j}^k x_{i,j}^k + q_{i,j}^k y_{i,j}^k) \right] \\ &= \sum_{i=1}^l a_{i,i}^k c_{i,i}^k + 2 \sum_{1 \leq i < j \leq l} (p_{i,j}^k x_{i,j}^k + q_{i,j}^k y_{i,j}^k). \end{aligned}$$

Lo anterior es válido para toda k por lo que podemos sustituir en la suma de estas trazas y obtener:

$$\sum_{k=0}^n \text{Tr}(A_k C_k) = \sum_{k=0}^n \left[\sum_{i=1}^l a_{i,i}^k c_{i,i}^k + 2 \sum_{1 \leq i < j \leq l} (p_{i,j}^k x_{i,j}^k + q_{i,j}^k y_{i,j}^k) \right]. \quad (1.40)$$

Note que todo hiperplano en nuestro espacio tiene una ecuación de la forma:

$$\sum_{k=1}^n \text{Tr}(A_k X_k) = r, \quad (1.41)$$

por nuestra representación en (1.40), donde los valores fijos de las entradas de A_k representan a las componentes del vector normal del hiperplano correspondiente, y las matrices X_k son las variables del hiperplano. Por lo tanto, como $M_{n+1}^{[a,b]}$ es un conjunto convexo y cerrado (ver Lema 1.13), y claramente el punto C simboliza un conjunto compacto en nuestro espacio de dimensión $(n+1)l(l+1)$, entonces podemos aplicar el Teorema de Separación del Hiperplano [4, Sección 2.5], y encontrar un hiperplano que separe ambos conjuntos. Es decir, podemos afirmar que existen matri-

ces A_1, A_2, \dots, A_n hermíticas, tales que (1.41) se satisfice, y separa ambos conjuntos. Además como C_0 es invertible, podemos considerar A_0 tal que $\text{Tr}(A_0 C_0) = -r$, y por ende reescribir (1.41) tal que:

$$\sum_{k=0}^n \text{Tr}(A_k X_k) = 0.$$

Por ende, al separar ambos conjuntos se debe cumplir que

$$\sum_{k=0}^n \text{Tr}(A_k C_k) < 0,$$

y que

$$\sum_{k=0}^n \text{Tr}(A_k S_k) \geq 0,$$

para todo $S = (S_0, \dots, S_n) \in M_{n+1}^{[a,b]}$. Sin embargo por el Corolario 1.14, la ecuación anterior es equivalente a que el polinomio:

$$P(t) = \sum_{k=0}^n A_k t^k \geq \mathbf{0}, \quad \forall t \in [a, b],$$

y esto es justamente la contradicción a (1.38) que se buscaba.

Para la segunda afirmación, supongamos primero que $S = (S_0, \dots, S_n)$ está en el interior de $M_{n+1}^{[a,b]}$, y que justamente el polinomio $P(t)$ no es idénticamente la matriz $\mathbf{0}$ para todo t . Por lo tanto existe un vector $v \in \mathbb{C}^l$ tal que $v^* P(t) v > 0$. Así, considerando la representación de S provista en el Corolario 1.14:

$$S = \alpha f(t, v) + (1 - \alpha)V, \quad 0 < \alpha \leq 1,$$

donde $V = (V_0, \dots, V_n)$ es un punto en la frontera de $M_{n+1}^{[a,b]}$, se tiene que

$$\begin{aligned} \sum_{k=0}^n \text{Tr}(A_k S_k) &= \sum_{k=0}^n \text{Tr}(A_k (\alpha t^k v v^* + (1 - \alpha)V_k)) \\ &= \alpha \sum_{k=0}^n \text{Tr}(A_k t^k v v^*) + (1 - \alpha) \sum_{k=0}^n \text{Tr}(A_k V_k) \\ &> \alpha \sum_{k=0}^n t^k v^* A_k v > v^* P(t) v > 0, \end{aligned}$$

en donde nuevamente usamos que $\text{Tr}(BC) = \text{Tr}(CB)$, y que $\alpha > 0$. Ahora para la implicación de conversa, suponemos que S no se encuentra en el interior de $M_{n+1}^{[a,b]}$. En

este caso procedemos exactamente igual a la primera afirmación. Si S no se encuentra en $M_{n+1}^{[a,b]}$ podemos encontrar matrices hermíticas A_0, \dots, A_n tales que:

$$\sum_{k=0}^n \text{Tr}(A_k S_k) < 0, \text{ y } \sum_{k=0}^n A_k t^k \geq \mathbf{0}.$$

Ahora si S se encuentra en la frontera de $M_{n+1}^{[a,b]}$, entonces existen nuevamente matrices hermíticas A_0, \dots, A_n tales que:

$$\sum_{k=0}^n \text{Tr}(A_k S_k) = 0, \text{ y } \sum_{k=0}^n A_k t^k \geq \mathbf{0},$$

concluyendo el caso complejo. El caso real es exactamente igual, simplemente que los coeficientes de las partes imaginarias correspondientes se anulan, lo que nos otorga el mismo procedimiento pero ahora en un espacio de dimensión $(n+1)l(l+1)/2$. \square

Enfocamos ahora nuestra atención en poder otorgar una representación, que nos será útil después, para los polinomios matriciales P_n , que satisfacen:

$$P_n(t) = \sum_{k=0}^n t^k A_k \geq \mathbf{0} \quad \forall t \in [a, b]. \quad (1.42)$$

Para ello enunciamos el siguiente teorema, cuya prueba se omitirá, pero se invita al lector a consultarla en [10, Teorema 2.5].

Teorema 1.16. *Asuma que el polinomio matricial $P_n(t)$ es definido nonegativo en el intervalo $[a, b]$. Si $n = 2m$, entonces existen polinomios matriciales $B_m(t) = \sum_{i=0}^m t^i B_i$, $C_{m-1}(t) = \sum_{i=0}^{m-1} t^i C_i$ tales que:*

$$P_{2m}(t) = B_m(t)B_m^*(t) + (t-a)(b-t)C_{m-1}(t)C_{m-1}^*(t). \quad (1.43)$$

Análogamente si $n = 2m + 1$, entonces existen polinomios matriciales $B_m(t) = \sum_{i=0}^m t^i B_i$, $C_m(t) = \sum_{i=0}^m t^i C_i$ tales que:

$$P_{2m+1}(t) = (t-a)B_m(t)B_m^*(t) + (b-t)C_m(t)C_m^*(t). \quad (1.44)$$

En [10] se trabaja con el caso $a = 0$ y $b = 1$. Para poder obtener el resultado anterior simplemente basta con elegir un simple cambio de variable lineal y obtener el teorema. Teniendo esto, probamos una equivalencia útil sobre cuando un punto se encuentra en

$M_{n+1}^{[a,b]}$ o en su interior. Para esto definimos antes a las siguientes matrices de ‘‘Hankel’’:

$$\underline{H}_{2m}^{[a,b]} = \begin{pmatrix} S_0 & \dots & S_m \\ \vdots & \ddots & \vdots \\ S_m & \dots & S_{2m} \end{pmatrix}, \quad (1.45)$$

$$\overline{H}_{2m}^{[a,b]} = \begin{pmatrix} -abS_0 + (a+b)S_1 - S_2 & \dots & -abS_{m-1} + (a+b)S_m - S_{m+1} \\ \vdots & \ddots & \vdots \\ -abS_{m-1} + (a+b)S_m - S_{m+1} & \dots & -abS_{2m-2} + (a+b)S_{2m-1} - S_{2m} \end{pmatrix}, \quad (1.46)$$

para los pares y para los impares:

$$\underline{H}_{2m+1}^{[a,b]} = \begin{pmatrix} -aS_0 + S_1 & \dots & -aS_m + S_{m+1} \\ \vdots & \ddots & \vdots \\ -aS_m + S_{m+1} & \dots & -aS_{2m} + S_{2m+1} \end{pmatrix}, \quad (1.47)$$

$$\overline{H}_{2m+1}^{[a,b]} = \begin{pmatrix} bS_0 - S_1 & \dots & bS_m - S_{m+1} \\ \vdots & \ddots & \vdots \\ bS_m - S_{m+1} & \dots & bS_{2m} - S_{2m+1} \end{pmatrix}. \quad (1.48)$$

Teorema 1.17. *Sea $S = (S_0, \dots, S_n)$, entonces:*

1. *S está en el espacio de momentos $M_{n+1}^{[a,b]}$ si y solo si $\underline{H}_n^{[a,b]}$ y $\overline{H}_n^{[a,b]}$ son matrices semi-definidas positivas.*
2. *S está en el interior del espacio de momentos $M_{n+1}^{[a,b]}$ si y solo si $\underline{H}_n^{[a,b]}$ y $\overline{H}_n^{[a,b]}$ son matrices definidas positivas.*

Demostración. Consideramos primero el caso par pues el caso impar es totalmente análogo a este. Entonces sin pérdida de generalidad podemos suponer que $n = 2m$. Supongamos primero que $S \in M_{n+1}^{[a,b]}$, y considere el siguiente polinomio:

$$F_{2m}(t) = \left(\sum_{k=0}^m t^k v_k \right) \left(\sum_{k=0}^m t^k v_k \right)^* \geq \mathbf{0}, \quad (1.49)$$

donde los términos v_k son vectores en \mathbb{C}^l . Por ende el anterior polinomio matricial es definido no negativo siempre pues es un producto de la forma yy^* con $y = \sum_{k=0}^m t^k v_k$. Por lo tanto podemos aplicar el Lema 1.14 y obtener que:

$$\sum_{k=0}^{2m} \text{Tr} \left(\left[\sum_{i+j=k} (v_i v_j^* + v_j v_i^*) \right] S_k \right) \geq 0.$$

Ahora veamos que la anterior suma puede reescribirse como:

$$\begin{aligned} \sum_{k=0}^{2m} \operatorname{Tr} \left(\left[\sum_{i+j=k} (v_i v_j^* + v_j v_i^*) \right] S_k \right) &= \sum_{k=0}^{2m} \operatorname{Tr} \left(\left[\sum_{i+j=k} (v_i v_j^* S_k + v_j v_i^* S_k) \right] \right) \\ &= \sum_{k=0}^{2m} \sum_{i+j=k} (v_j^* S_k v_i + v_i^* S_k v_j) \\ &= v^* \underline{H}_{2m}^{[a,b]} v \geq 0, \end{aligned}$$

donde la desigualdad final es por (1.45), y $v^* = (v_0^*, \dots, v_m^*)$ es un vector en $\mathbb{C}^{l(m+1)}$ compuesto por los vectores v_i . Esto último claramente implica que $\underline{H}_{2m}^{[a,b]}$ es definido nonegativo pues el vector v es arbitrario. Para comprobar que $\overline{H}_{2m}^{[a,b]}$ es nonegativa definida procedemos de una forma similar pero ahora consideramos al polinomio:

$$G_{2m}(t) = (t-a)(b-t)F_{2m-2}(t) \geq \mathbf{0},$$

que claramente es definido nonegativo, pues $t \in [a, b]$. Llamando G_k , $k = 0, \dots, 2m$ a los coeficientes matriciales de $G_{2m}(t)$ y aplicando nuevamente el Lema 1.15, se tiene que

$$\sum_{k=0}^{2m} \operatorname{Tr}(G_k S_k) = \sum_{k=0}^{2m-2} \operatorname{Tr} \left[(-abS_k + (a+b)S_{k+1} - S_{k+2}) \sum_{i+j=k} (v_i v_j^* + v_j v_i^*) \right] \geq 0.$$

Aplicando nuevamente la identidad $\operatorname{Tr}(BC) = \operatorname{Tr}(CB)$, y recordando la definición de u_k , obtenemos que la suma anterior se expresa como:

$$\begin{aligned} \sum_{k=0}^{2m-2} \operatorname{Tr} \left[(-abS_k + (a+b)S_{k+1} - S_{k+2}) \sum_{i+j=k} (v_i v_j^* + v_j v_i^*) \right] \\ &= \sum_{k=0}^{2m-2} \sum_{i+j=k} [v_j^* (-abS_k + (a+b)S_{k+1} - S_{k+2}) v_i \\ &\quad + v_i^* (-abS_k + (a+b)S_{k+1} - S_{k+2}) v_j] \\ &= v^* \overline{H}_{2m}^{[a,b]} v \geq 0. \end{aligned}$$

Nuevamente como los vectores v_i fueron arbitrarios esto prueba que $\overline{H}_{2m}^{[a,b]}$ es definida nonegativa.

Ahora probamos la segunda implicación de nuestra equivalencia, es decir si $\underline{H}_{2m}^{[a,b]}$ y $\overline{H}_{2m}^{[a,b]}$ son semi-definidas positivas, entonces tenemos que ver que S está en $M_{n+1}^{[a,b]}$. Para esto consideramos el polinomio $P_{2m}(t)$ definido en (1.42), suponemos que es semi-definido positivo en el intervalo $[a, b]$ y usando el Teorema 1.16 sabemos que

tiene la representación de la forma (1.43). Ahora note que si llamamos B_m^k al k -ésimo coeficiente matricial del polinomio $B_m(t)B_m^*(t)$, y de igual forma observe que $B_m^k = \sum_{i+j=k} B_i B_j^* + B_j B_i^*$, entonces:

$$\begin{aligned} \sum_{k=0}^{2m} \text{Tr}(B_m^k S_k) &= \sum_{k=0}^{2m} \text{Tr} \left(\left[\sum_{i+j=k} (B_i B_j^* + B_j B_i^*) \right] S_k \right) \\ &= \sum_{k=0}^{2m} \text{Tr} \left(\sum_{i+j=k} (B_i^* S_k B_j + B_j^* S_k B_i) \right) \\ &= B^* \underline{H}_{2m}^{[a,b]} B \geq \mathbf{0}. \end{aligned}$$

En donde $B^* = (B_0^*, B_1^*, \dots, B_m^*)$, y la última igualdad se da pues $\underline{H}_{2m}^{[a,b]}$ es semi-definida positiva. Análogamente al anterior procedimiento, pero con el polinomio $(b-t)(t-a)C_{m-1}(t)C_{m-1}^*(t)$, si C_{m-1}^k es el k -ésimo coeficiente matricial, obtenemos:

$$\begin{aligned} \sum_{k=0}^{2m} \text{Tr}(C_m^k S_k) &= \sum_{k=0}^{2m-2} \text{Tr} \left(\left[\sum_{i+j=k} (C_i C_j^* + C_j C_i^*) \right] (-abS_k + (a+b)S_{k+1} - S_{k+2}) \right) \\ &= \sum_{k=0}^{2m-2} \text{Tr} \left(\sum_{i+j=k} [C_i^* (-abS_k + (a+b)S_{k+1} - S_{k+2}) C_j \right. \\ &\quad \left. + C_j^* (-abS_k + (a+b)S_{k+1} - S_{k+2}) C_i] \right) \\ &= C^{*} \overline{H}_{2m}^{[a,b]} C \geq \mathbf{0}, \end{aligned}$$

donde $C^* = (C_0^*, C_1^*, \dots, C_m^*)$, y nuevamente lo último es debido a que $\overline{H}_{2m}^{[a,b]}$ es definida nonegativa. Esto prueba la equivalencia de la primera proposición del Teorema, para el caso par. El caso impar es análogo y se da al considerar los polinomios:

$$H_{2m+1}(t) = (t-a)F_{2m}(t), \text{ y } K_{2m+1}(t) = (b-t)F_{2m}(t).$$

para la primera implicación. El regreso, para este caso, es igualmente análogo usando, nuevamente, el Teorema 1.16, como se usó para el caso par.

La segunda equivalencia del Teorema es resultado de aplicar la segunda equivalencia en lugar de la primera del Lema 1.15, en las ecuaciones correspondientes al caso par y proseguir de forma análoga con el impar, en donde simplemente se cambia la nonegatividad por positividad estricta. \square

Estamos ya en posición de poder definir formalmente los momentos canónicos pa-

ra nuestra medida matricial en $[a, b]$. Para este fin, necesitamos definir los siguientes vectores matriciales:

$$\underline{h}_{2m}^T = (-aS_m + S_{m+1}, \dots, -aS_{2m-1} + S_{2m}), \quad (1.50)$$

$$\underline{h}_{2m-1}^T = (S_m, \dots, S_{2m-1}), \quad (1.51)$$

$$\bar{h}_{2m}^T = (bS_m - S_{m+1}, \dots, bS_{2m-1} - S_{2m}), \quad (1.52)$$

$$\bar{h}_{2m-1}^T = (-abS_{m-1} + (a+b)S_m - S_{m+1}, \dots, -abS_{2m-3} + (a+b)S_{2m-2} - S_{2m-1}), \quad (1.53)$$

que dan lugar a las cotas inferiores y superiores del n -ésimo momento denotadas como S_n^- y S_n^+ , respectivamente. Así para $m \geq 1$, definimos:

$$S_{2m}^- = \underline{h}_{2m-1}^T (\underline{H}_{2m-2}^{[a,b]})^{-1} \underline{h}_{2m-1}, \quad (1.54)$$

$$S_{2m+1}^- = \underline{h}_{2m}^T (\underline{H}_{2m-1}^{[a,b]})^{-1} \underline{h}_{2m} + aS_{2m}, \quad (1.55)$$

$$S_{2m}^+ = -abS_{2m-2} + (a+b)S_{2m-1} - \bar{h}_{2m-1}^T (\bar{H}_{2m-2}^{[a,b]})^{-1} \bar{h}_{2m-1}, \quad (1.56)$$

$$S_{2m+1}^+ = bS_{2m} - \bar{h}_{2m}^T (\bar{H}_{2m-1}^{[a,b]})^{-1} \bar{h}_{2m}, \quad (1.57)$$

y donde definimos como $S_1^- = aS_0$, $S_1^+ = bS_0$ y $S_2^+ = -abS_0 + (a+b)S_1$. Note que S_n^- y S_n^+ dependen únicamente de (S_0, \dots, S_{n-1}) . Antes de continuar note que, para toda n :

$$\underline{H}_n^{[a,b]} = \begin{pmatrix} \underline{H}_{n-2}^{[a,b]} & \underline{h}_{n-1} \\ \underline{h}_{n-1}^T & * \end{pmatrix} \quad \text{y} \quad \bar{H}_n^{[a,b]} = \begin{pmatrix} \bar{H}_{n-2}^{[a,b]} & \bar{h}_{n-1} \\ \bar{h}_{n-1}^T & * \end{pmatrix},$$

donde la matriz en $*$ varía dependiendo de n . Por ende usando el resultado de que una matriz es semi-definida positiva si y solo si el complemento de Schur y su subbloque principal, que en este caso esta determinado por $*$, lo son, por las definiciones anteriores y por (1.55), tenemos que el vector (S_0, \dots, S_{n-1}) está en $M_{n+1}^{[a,b]}$ si y solo si $S_n^- < S_n^+$. Es más, para todo $(S_0, \dots, S_{n-1}) \in M_{n+1}^{[a,b]}$ se tiene que:

$$S_n^- \leq S_n \leq S_n^+.$$

Definición. Si $(S_0, \dots, S_{n-1}) \in \text{int}M_{n+1}^{[a,b]}$, entonces, definimos como el k -ésimo momento cónico a la matriz:

$$U_k^{[a,b]} = (D_k^{[a,b]})^{-1} (S_k - S_k^-), \quad 1 \leq k \leq n. \quad (1.58)$$

donde

$$D_k^{[a,b]} = S_k^+ - S_k^-. \quad (1.59)$$

Estas definiciones son las análogas al caso escalar [9, Sección 1.3]. De igual forma

usaremos las matrices:

$$V_k^{[a,b]} = \mathbf{I} - U_k^{[a,b]} = (D_k^{[a,b]})^{-1}(S_k^+ - S_k), \quad 1 \leq k \leq n. \quad (1.60)$$

Los primeros momentos canónicos son fáciles de calcular. Tenemos que $S_1^+ = bS_0$, $S_1^- = aS_0$, y por ende $D_1^{[a,b]} = (b-a)S_0$ y así $U_1^{[a,b]} = \frac{1}{(b-a)}(S_0^{-1}S_1 - a\mathbf{I})$ y $V_1^{[a,b]} = \frac{1}{(b-a)}(b\mathbf{I} - S_0^{-1}S_1)$. Similarmente, y por las anteriores definiciones de las cotas, tenemos que $S_2^+ = -abS_0 + (a+b)S_1$ y $S_2^- = S_1S_0^{-1}S_1$. Con lo que obtenemos:

$$\begin{aligned} D_2^{[a,b]} &= -abS_0 + (a+b)S_1 - S_1S_0^{-1}S_1 = S_0(S_0^{-1}S_1 - a\mathbf{I})(b\mathbf{I} - S_0^{-1}S_1) \\ &= (b-a)^2 S_0 U_1^{[a,b]} V_1^{[a,b]}, \end{aligned} \quad (1.61)$$

y así:

$$U_2^{[a,b]} = \frac{1}{(b-a)^2} (U_1^{[a,b]} V_1^{[a,b]})^{-1} S_0^{-1} (S_2 - S_1 S_0^{-1} S_1). \quad (1.62)$$

Nuestro siguiente resultado es sobre la generalización, del caso escalar, sobre la propiedad de invarianza de los momentos canónicos respecto la transformación lineal de una medida. El caso escalar es tratado en [9, Sección 1.3], sin embargo, cuando en [10, págs. 177-178] se definen los momentos canónicos matriciales se omitió la prueba de la propiedad de invarianza, que ocurre en el caso escalar. De hecho, simplemente se menciona en [8, pág. 135] sin dar algún detalle de la prueba, la cual se mostrará en las siguientes páginas de este texto. Para este fin, consideramos la medida $\mu(x)$ con soporte en el intervalo $[0, 1]$, y consideramos la transformación lineal $y = (b-a)x + a$, que mapea $x \in [0, 1]$ en $y \in [a, b]$, y por ende manda a la medida $\mu(x)$ en la medida $\mu_{[a,b]}(y)$. Con base en esto, podemos encontrar una relación de recurrencia directa entre los momentos de ambas medidas. Para esto, note que los momentos de $\mu_{[a,b]}$ son:

$$T_k = \int_{[a,b]} y^k d\mu_{[a,b]}(y), \quad k = 1, 2, \dots \quad (1.63)$$

Realizando la sustitución $y = (b-a)x + a$ tenemos:

$$\begin{aligned} T_k &= \int_{[a,b]} y^k d\mu_{[a,b]}(y) = \int_{[0,1]} ((b-a)x + a)^k d\mu(x) \\ &= \int_{[0,1]} \sum_{j=0}^k \binom{k}{j} (b-a)^j a^{k-j} x^j d\mu(x) \\ &= \sum_{j=0}^k \binom{k}{j} (b-a)^j a^{k-j} S_j, \end{aligned} \quad (1.64)$$

donde S_j es el j -ésimo momento de μ . Análogamente, pero utilizando ahora el cambio

de variable $\frac{y-a}{b-a}$, obtenemos las relaciones:

$$S_k = \frac{1}{(b-a)^k} \sum_{j=0}^k \binom{k}{j} (-a)^{k-j} T_j. \quad (1.65)$$

Con base en esto definimos a la siguiente matriz por bloques de tamaño $l \times l$:

$$Q_{n,[a,b]} = \begin{pmatrix} \mathbf{I} & \mathbf{0} & \dots & \mathbf{0} \\ a\mathbf{I} & (b-a)\mathbf{I} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ a^n\mathbf{I} & \binom{n}{1}a^{n-1}(b-a)\mathbf{I} & \dots & (b-a)^n\mathbf{I} \end{pmatrix}, \quad (1.66)$$

donde, explícitamente y siguiendo una numeración de columnas y filas empezando por el 0, la i, j -ésima entrada es:

$$(Q_{n,[a,b]})_{i,j} = \binom{i}{j} (b-a)^j a^{i-j} \mathbf{I}.$$

Esta matriz genera la recurrencia de (1.64). En efecto, si definimos a los vectores matriciales: $S^T = (S_0, S_1, \dots, S_n)$ y $T^T = (T_0, T_1, \dots, T_n)$, entonces tenemos que $Q_{n,[a,b]}S = T$. Igualmente si definimos al vector matricial $V_n^T(x) = (\mathbf{I}, x\mathbf{I}, \dots, x^n\mathbf{I})$, tenemos que:

$$Q_{n,[a,b]}V_n(x) = V_n((b-a)x + a) = V_n(y), \quad (1.67)$$

y que:

$$\int_{[a,b]} V_n(y)V_n^T(y) d\mu_{[a,b]}(y) = \underline{H}_{2n}^{[a,b]}, \quad (1.68)$$

pues

$$V_n(y)V_n^T(y) = \begin{pmatrix} 1 & \dots & y^n \\ \vdots & \ddots & \vdots \\ y^n & \dots & y^{2n} \end{pmatrix},$$

y al integrar, entrada a entrada, obtenemos (1.68). Por ende, usando (1.67) en (1.68),

tenemos la relación:

$$\begin{aligned}
 \underline{H}_{2n}^{[a,b]} &= \int_{[a,b]} V_n(y) V_n^T(y) d\mu_{[a,b]}(y) \\
 &= \int_{[0,1]} Q_{n,[a,b]} V_n(x) V_n^T(x) Q_{n,[a,b]}^T d\mu(x) \\
 &= Q_{n,[a,b]} \left(\int_{[0,1]} V_n(x) V_n^T(x) d\mu(x) \right) Q_{n,[a,b]}^T \\
 &= Q_{n,[a,b]} \underline{H}_{2n}^{[0,1]} Q_{n,[a,b]}^T.
 \end{aligned} \tag{1.69}$$

Análogamente si integramos $V_n(y) V_n^T(y)$ respecto las medidas $(y-a)(b-y)d\mu_{[a,b]}(y)$, $(y-a)d\mu_{[a,b]}(y)$ y $(b-y)d\mu_{[a,b]}(y)$, obtenemos que:

$$\overline{H}_{2n}^{[a,b]} = \int_{[a,b]} V_{n-1}(y) V_{n-1}^T(y) (y-a)(b-y) d\mu_{[a,b]}(y), \tag{1.70}$$

$$\underline{H}_{2n+1}^{[a,b]} = \int_{[a,b]} V_n(y) V_n^T(y) (y-a) d\mu_{[a,b]}(y), \tag{1.71}$$

$$\overline{H}_{2n+1}^{[a,b]} = \int_{[a,b]} V_n(y) V_n^T(y) (b-y) d\mu_{[a,b]}(y). \tag{1.72}$$

Si realizamos el cambio de variable $y = (b-a)x + a$, obtenemos que $(y-a)(b-y)d\mu_{[a,b]}(y) = (b-a)^2 x(1-x)d\mu(x)$, $(y-a)d\mu_{[a,b]}(y) = (b-a)x d\mu(x)$ y, finalmente, $(b-y)d\mu_{[a,b]}(y) = (b-a)(1-x)d\mu(x)$. Así, replicando lo ocurrido en (1.69) tenemos:

$$\overline{H}_{2n}^{[a,b]} = (b-a)^2 Q_{n-1,[a,b]} \overline{H}_{2n}^{[0,1]} Q_{n-1,[a,b]}^T, \tag{1.73}$$

$$\underline{H}_{2n+1}^{[a,b]} = (b-a) Q_{n,[a,b]} \underline{H}_{2n+1}^{[0,1]} Q_{n,[a,b]}^T, \tag{1.74}$$

$$\overline{H}_{2n+1}^{[a,b]} = (b-a) Q_{n,[a,b]} \overline{H}_{2n+1}^{[0,1]} Q_{n,[a,b]}^T. \tag{1.75}$$

Antes de continuar, note que la matriz inversa de $Q_{n,[a,b]}$ está dada por:

$$Q_{n,[a,b]}^{-1} = \begin{pmatrix} \mathbf{I} & \mathbf{0} & \dots & \mathbf{0} \\ \frac{-a}{b-a} \mathbf{I} & \frac{1}{b-a} \mathbf{I} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{(-a)^n}{(b-a)^n} \mathbf{I} & \frac{\binom{n}{1} (-a)^{n-1}}{(b-a)^n} \mathbf{I} & \dots & \frac{1}{(b-a)^n} \mathbf{I} \end{pmatrix}. \tag{1.76}$$

Donde la i, j -ésima entrada, nuevamente empezando a enumerar las filas y columnas desde 0, está dada por:

$$(Q_{n,[a,b]}^{-1})_{i,j} = \frac{1}{(b-a)^n} \binom{i}{j} (-a)^{k-j} \mathbf{I}. \tag{1.77}$$

Para obtener (1.76), primero, al ser $Q_{n,[a,b]}$ triangular inferior, su inverso igual lo es, e inductivamente podemos calcular la última fila de forma sencilla. Note, de hecho, que la matriz inversa es debido a (1.65), así como (1.66) se debe a (1.64), pues la recursión de (1.65) es la recursión inversa de (1.64), es decir, se satisface que $Q_{n,[a,b]}^{-1}T = S$. Teniendo esto enunciamos el siguiente Lema que nos será de utilidad a continuación.

Lema 1.18. *Considere los vectores matriciales $T_{[m,k]}^T = (T_m, T_{m+1}, \dots, T_k)$, con $m < k$ y en \mathbb{N} , y $S_{[p,l]}^T = (S_p, S_{p+1}, \dots, S_l)$, nuevamente con $p < l$, naturales. Entonces se satisface:*

$$Q_{(k-m),[a,b]}^{-1}T_{[m,k]} = \sum_{j=0}^m \binom{m}{j} (b-a)^j a^{m-j} S_{[j,k-m+j]}.$$

Demostración. Consideremos la i -ésima fila del vector $Q_{(k-m),[a,b]}^{-1}T_{[m,k]}$, y reescribimos los momentos en su forma integral:

$$\begin{aligned} \frac{1}{(b-a)^i} \sum_{j=0}^i \binom{i}{j} (-a)^{i-j} T_{j+m} &= \frac{1}{(b-a)^i} \int_{[a,b]} \sum_{j=0}^i \binom{i}{j} (-a)^{i-j} y^{j+m} d\mu_{[a,b]}(y) \\ &= \int_{[a,b]} y^m \left(\frac{y-a}{b-a} \right)^i d\mu_{[a,b]}(y) \\ &= \int_{[0,1]} ((b-a)x + a)^m x^i d\mu(x) \\ &= \int_{[0,1]} \sum_{j=0}^m \binom{m}{j} (b-a)^j a^{m-j} x^{i+j} d\mu(x) \\ &= \sum_{j=0}^m \binom{m}{j} (b-a)^j a^{m-j} S_{i+j}. \end{aligned}$$

Como i fue arbitrario, esto ocurre para todos los componentes de $Q_{(k-m),[a,b]}^{-1}T_{[m,k]}$, así fijando j y variando i , tenemos la ecuación deseada. \square

Estamos ya en condiciones de probar la invarianza de los momentos canónicos, respecto a la transformación lineal $y = (b-a)x + a$. Para ello probamos el siguiente Teorema que es equivalente.

Teorema 1.19. *Sea μ una medida matricial cuyo soporte se encuentra en $[0, 1]$, y considere la transformación lineal $y = (b-a)x + a$, que manda a $x \in [0, 1]$ a $y \in [a, b]$, y la medida μ a $\mu_{[a,b]}$. Entonces, dadas las definiciones anteriores, donde los momentos de la medida μ están dados por S_0, S_1, \dots y los momentos de $\mu_{[a,b]}$ por*

T_0, T_1, \dots , se satisface para toda $n \in \mathbb{N}$:

$$T_n^- = (b-a)^n S_n^- + \sum_{j=0}^{n-1} \binom{n}{j} (b-a)^j a^{n-j} S_j, \quad (1.78)$$

$$T_n^+ = (b-a)^n S_n^+ + \sum_{j=0}^{n-1} \binom{n}{j} (b-a)^j a^{n-j} S_j. \quad (1.79)$$

Demostración. Empezamos probando primero la ecuación (1.78). Procedemos primero probando cuando n es par, y realizaremos un procedimiento análogo al caso impar. Así sea $n = 2m$. Recordando la ecuación (1.51), y renombrando $\underline{g}_{2m-1}^T = (T_m, \dots, T_{2m-1})$, tenemos que $T_{2m}^- = \underline{g}_{2m-1}^T (\underline{H}_{2m-2}^{[a,b]})^{-1} \underline{g}_{2m-1}$. Por otro lado, nombramos $\underline{h}_{2m-1}^T = (S_m, \dots, S_{2m-1})$ y nuevamente $S_{2m}^- = \underline{h}_{2m-1}^T (\underline{H}_{2m-2}^{[0,1]})^{-1} \underline{h}_{2m-1}$. Usando (1.69), y el Lema 1.18 tenemos que:

$$\begin{aligned} T_{2m}^- &= \underline{g}_{2m-1}^T (\underline{H}_{2m-2}^{[a,b]})^{-1} \underline{g}_{2m-1} \\ &= \underline{g}_{2m-1}^T (Q_{m-1,[a,b]} \underline{H}_{2m-2}^{[0,1]} Q_{m-1,[a,b]}^T)^{-1} \underline{g}_{2m-1} \\ &= \underline{g}_{2m-1}^T (Q_{m-1,[a,b]}^{-1})^T (\underline{H}_{2m-2}^{[0,1]})^{-1} Q_{m-1,[a,b]}^{-1} \underline{g}_{2m-1} \\ &= (Q_{m-1,[a,b]}^{-1} \underline{g}_{2m-1})^T (\underline{H}_{2m-2}^{[0,1]})^{-1} Q_{m-1,[a,b]}^{-1} \underline{g}_{2m-1} \\ &= \left(\sum_{j=0}^m \binom{m}{j} (b-a)^j a^{m-j} S_{[j,m-1+j]} \right)^T (\underline{H}_{2m-2}^{[0,1]})^{-1} \left(\sum_{i=0}^m \binom{m}{i} (b-a)^i a^{m-i} S_{[i,m-1+i]} \right) \\ &= \sum_{r=0}^{2m} \sum_{\substack{i+j=r \\ 0 \leq i, j \leq m}} \binom{m}{i} \binom{m}{j} (b-a)^r a^{2m-r} S_{[j,m-1+j]}^T (\underline{H}_{2m-2}^{[0,1]})^{-1} S_{[i,m-1+i]}. \end{aligned} \quad (1.80)$$

Para continuar, note que $S_{[j,m-1+j]}^T (\underline{H}_{2m-2}^{[0,1]})^{-1} = (\mathbf{0}, \dots, \mathbf{I}, \mathbf{0}, \dots, \mathbf{0}) = E_{m,j}^T$, que es el vector matricial, de m entradas matriciales, cuya j -ésima componente es \mathbf{I} , y con $j < m$. Esto último es cierto, pues $S_{[j,m-1+j]}^T$ es la j -ésima fila y columna de $\underline{H}_{2m-2}^{[0,1]}$, asumiendo que la numeración de filas y columnas empieza en 0. Análogamente para $i < m$, $(\underline{H}_{2m-2}^{[0,1]})^{-1} S_{[i,m-1+i]} = (\mathbf{0}, \dots, \mathbf{I}, \mathbf{0}, \dots, \mathbf{0})^T = E_{m,i}$, se satisface por la misma razón. Para $i = j = m$ note que $S_{[m,2m-1]}^T (\underline{H}_{2m-2}^{[0,1]})^{-1} S_{[m,2m-1]} = \underline{h}_{2m-1}^T (\underline{H}_{2m-2}^{[0,1]})^{-1} \underline{h}_{2m-1} = S_{2m}^-$. Teniendo estas últimas identidades en cuenta, la cadena de igualdades de (1.80) sigue:

$$= (b-a)^{2m} S_{2m}^- + \sum_{r=0}^{2m-1} \sum_{\substack{i+j=r \\ 0 \leq i, j \leq m}} \binom{m}{i} \binom{m}{j} (b-a)^r a^{2m-r} S_r, \quad (1.81)$$

donde se usó que $S_{[j,m-1+j]}^T (\underline{H}_{2m-2}^{[0,1]})^{-1} S_{[i,m-1+i]} = E_{m,j} S_{[i,m-1+i]} = S_{i+j} = S_r$. Ahora usamos la identidad:

$$\sum_{\substack{i+j=r \\ 0 \leq i,j \leq m}} \binom{m}{i} \binom{m}{j} = \binom{2m}{r}, \quad (1.82)$$

que se obtiene de comparar los coeficientes entre los polinomios $(x+y)^m(x+y)^m = (x+y)^{2m}$, por lo que sustituyendo en (1.81):

$$T_{2m}^- = (b-a)^{2m} S_{2m}^- + \sum_{r=0}^{2m-1} \binom{2m}{r} (b-a)^r a^{2m-r} S_r,$$

concluyendo. Para el caso impar, asumimos que $n = 2m + 1$, y consideramos ahora a los vectores matriciales $\underline{g}_{2m}^T = (-aT_m + T_{m+1}, \dots, -aT_{2m-1} + T_{2m})$ y $\underline{h}_{2m}^T = (S_{m+1}, \dots, S_{2m})$. Usamos (1.50), sustituimos en (1.74) y obtenemos:

$$\begin{aligned} T_{2m+1}^- &= \underline{g}_{2m}^T (\underline{H}_{2m-1}^{[a,b]})^{-1} \underline{g}_{2m} + aT_{2m} \\ &= \frac{1}{b-a} (Q_{m-1,[a,b]}^{-1} \underline{g}_{2m})^T (\underline{H}_{2m-1}^{[0,1]})^{-1} Q_{m-1,[a,b]}^{-1} \underline{g}_{2m} + aT_{2m}. \end{aligned} \quad (1.83)$$

Usando la notación del Lema 1.18, tenemos que $\underline{g}_{2m} = -aT_{[m,2m-1]} + T_{[m+1,2m]}$, así, por el Lema 1.18:

$$\begin{aligned} Q_{m-1,[a,b]}^{-1} \underline{g}_{2m} &= Q_{m-1,[a,b]}^{-1} (-aT_{[m,2m-1]} + T_{[m+1,2m]}) \\ &= -aQ_{m-1,[a,b]}^{-1} T_{[m,2m-1]} + Q_{m-1,[a,b]}^{-1} T_{[m+1,2m]} \\ &= -\sum_{j=0}^m \binom{m}{j} (b-a)^j a^{m-j+1} S_{[j,m-1+j]} + \sum_{j=0}^{m+1} \binom{m+1}{j} (b-a)^j a^{m-j+1} S_{[j,m-1+j]} \\ &= (b-a)^{m+1} S_{[m+1,2m]} + \sum_{j=1}^m \binom{m}{j-1} (b-a)^j a^{m-j+1} S_{[j,m-1+j]} \\ &= (b-a)^{m+1} S_{[m+1,2m]} + \sum_{j=0}^{m-1} \binom{m}{j} (b-a)^{j+1} a^{m-j} S_{[j+1,m+j]} \\ &= (b-a) \sum_{j=0}^m \binom{m}{j} (b-a)^j a^{m-j} S_{[j+1,m+j]}. \end{aligned} \quad (1.84)$$

De igual forma que con el caso par note que para todo $0 \leq i, j < m$ se tiene que $S_{[j+1,m+j]}^T (\underline{H}_{2m-1}^{[0,1]})^{-1} = E_{m,j}^T$, solo basta con sustituir $a = 0$ y $b = 1$ en (1.47) y notar que $S_{[j+1,m+j]}$ es la j -ésima columna y fila de $\underline{H}_{2m-1}^{[0,1]}$. Por las mismas razones $(\underline{H}_{2m-1}^{[0,1]})^{-1} S_{[i+1,m+i]} = E_{m,i}$. Para $i = j = m$ tenemos que $S_{[m+1,2m]}^T (\underline{H}_{2m-1}^{[0,1]})^{-1} S_{[m+1,2m]} =$

$\underline{h}_{2m}^T (\underline{H}_{2m-1}^{[0,1]})^{-1} \underline{h}_{2m} = S_{2m+1}^-$. Así tomando en cuenta esto, y sustituyendo (1.84) en (1.83), así como considerando nuevamente (1.82):

$$\begin{aligned}
&= \frac{1}{b-a} \left((b-a) \sum_{j=0}^m \binom{m}{j} (b-a)^j a^{m-j} S_{[j+1, m+j]}^T \right)^T (\underline{H}_{2m-1}^{[0,1]})^{-1} \\
&\quad \times \left((b-a) \sum_{i=0}^m \binom{m}{i} (b-a)^i a^{m-i} S_{[i+1, m+i]} \right) + aT_{2m} \\
&= (b-a)^{2m+1} S_{2m+1}^- + \sum_{r=0}^{2m-1} \sum_{\substack{i+j=r \\ 0 \leq i, j \leq m}} \binom{m}{i} \binom{m}{j} (b-a)^{r+1} a^{2m-r} S_{r+1} + aT_{2m} \\
&= (b-a)^{2m+1} S_{2m+1}^- + \sum_{r=0}^{2m-1} \binom{2m}{r} (b-a)^{r+1} a^{2m-r} S_{r+1} + aT_{2m},
\end{aligned}$$

recordando la ecuación recursiva de (1.64), y sustituyendo:

$$= (b-a)^{2m+1} S_{2m+1}^- + \sum_{r=0}^{2m-1} \binom{2m}{r} (b-a)^{r+1} a^{2m-r} S_{r+1} + \sum_{j=0}^{2m} \binom{2m}{j} (b-a)^j a^{2m-j+1} S_j.$$

Excluyendo $a^{2m+1} S_0$ del primer término de la última suma, y realizando el cambio de índices $j = i + 1$, tenemos:

$$\sum_{j=0}^{2m} \binom{2m}{j} (b-a)^j a^{2m-j+1} S_j = a^{2m+1} S_0 + \sum_{i=0}^{2m-1} \binom{2m}{i+1} (b-a)^{i+1} a^{2m-i} S_{i+1}.$$

Sustituyendo, nuevamente:

$$\begin{aligned}
&= (b-a)^{2m+1} S_{2m+1}^- + a^{2m+1} S_0 + \sum_{i=0}^{2m-1} \left(\binom{2m}{i+1} + \binom{2m}{i} \right) (b-a)^{i+1} a^{2m-i} S_{i+1} \\
&= (b-a)^{2m+1} S_{2m+1}^- + a^{2m+1} S_0 + \sum_{i=0}^{2m-1} \binom{2m+1}{i+1} (b-a)^{i+1} a^{2m-i} S_{i+1} \\
&= (b-a)^{2m+1} S_{2m+1}^- + \sum_{i=0}^{2m} \binom{2m+1}{i} (b-a)^i a^{2m+1-i} S_i,
\end{aligned}$$

concluyendo (1.78).

Resta ahora probar la ecuación (1.79). Para ello, nuevamente, empezamos por probar el caso par con $n = 2m$, y a continuación el impar. Consideramos, ahora, a los vectores matriciales $\bar{g}_{2m-1}^T = (-abT_{m-1} + (a+b)T_m - T_{m+1}, \dots, -abT_{2m-3} + (a+b)T_{2m-2} - T_{2m-1})$

y $\bar{h}_{2m-1}^T = (S_m - S_{m+1}, \dots, S_{2m-2} - S_{2m-1})$, y sustituimos (1.73) en (1.56) para obtener:

$$\begin{aligned} T_{2m}^+ &= -\bar{g}_{2m-1}^T (\bar{H}_{2m-2}^{[a,b]})^{-1} \bar{g}_{2m-1} - abT_{2m-2} + (a+b)T_{2m-1} \\ &= \frac{-1}{(b-a)^2} (Q_{m-2,[a,b]}^{-1} \bar{g}_{2m-1})^T (\bar{H}_{2m-2}^{[0,1]})^{-1} Q_{m-2,[a,b]}^{-1} \bar{g}_{2m-1} - abT_{2m-2} + (a+b)T_{2m-1}. \end{aligned} \quad (1.85)$$

Como en el caso anterior, usamos la notación del Lema 1.18 y tenemos que $\bar{g}_{2m-1} = -abT_{[m-1,2m-3]} + (a+b)T_{[m,2m-2]} - T_{[m+1,2m-1]}$. Usando este mismo lema obtenemos:

$$\begin{aligned} Q_{m-2,[a,b]}^{-1} \bar{g}_{2m-1} &= Q_{m-2,[a,b]}^{-1} (-abT_{[m-1,2m-3]} + (a+b)T_{[m,2m-2]} - T_{[m+1,2m-1]}) \\ &= -abQ_{m-2,[a,b]}^{-1} T_{[m-1,2m-3]} + (a+b)Q_{m-2,[a,b]}^{-1} T_{[m,2m-2]} \\ &\quad - Q_{m-2,[a,b]}^{-1} T_{[m+1,2m-1]} \\ &= -ab \sum_{j=0}^{m-1} \binom{m-1}{j} (b-a)^j a^{m-1-j} S_{[j,m-2+j]} \\ &\quad + (a+b) \sum_{j=0}^m \binom{m}{j} (b-a)^j a^{m-j} S_{[j,m-2+j]} \\ &\quad - \sum_{j=0}^{m+1} \binom{m+1}{j} (b-a)^j a^{m+1-j} S_{[j,m-2+j]}. \end{aligned} \quad (1.86)$$

Antes de seguir, note que:

$$\begin{aligned} (a+b) \sum_{j=0}^m \binom{m}{j} (b-a)^j a^{m-j} S_{[j,m-2+j]} &= \sum_{j=0}^m \binom{m}{j} (b-a)^j a^{m+1-j} S_{[j,m-2+j]} \\ &\quad + b \sum_{j=0}^m \binom{m}{j} (b-a)^j a^{m-j} S_{[j,m-2+j]}, \end{aligned}$$

por lo que sustituyendo en las igualdades de (1.86), y reordenando:

$$\begin{aligned}
 &= b \sum_{j=1}^{m-1} \left(\binom{m}{j} - \binom{m-1}{j} \right) (b-a)^j a^{m-j} S_{[j, m-2+j]} \\
 &\quad - \sum_{j=1}^m \left(\binom{m+1}{j} - \binom{m}{j} \right) (b-a)^j a^{m+1-j} S_{[j, m-2+j]} \\
 &\quad + b(b-a)^m S_{[m, 2m-2]} - (b-a)^{m+1} S_{[m+1, 2m-1]} \\
 &= b \sum_{j=1}^{m-1} \binom{m-1}{j-1} (b-a)^j a^{m-j} S_{[j, m-2+j]} - \sum_{j=1}^m \binom{m}{j-1} (b-a)^j a^{m+1-j} S_{[j, m-2+j]} \\
 &\quad + b(b-a)^m S_{[m, 2m-2]} - (b-a)^{m+1} S_{[m+1, 2m-1]} \\
 &= b(b-a) \sum_{i=0}^{m-1} \binom{m-1}{i} (b-a)^i a^{m-1-i} S_{[i+1, m-1+i]} \\
 &\quad - (b-a) \sum_{i=0}^m \binom{m}{i} (b-a)^i a^{m-i} S_{[i+1, m-1+i]} \\
 &= (b-a) \sum_{i=1}^{m-1} \left(b \binom{m-1}{i} - a \binom{m}{i} \right) (b-a)^i a^{m-1-i} S_{[i+1, m-1+i]} \\
 &\quad - (b-a)^{m+1} S_{[m+1, 2m-1]} + (b-a)^2 a^{m-1} S_{[1, m-1]} \\
 &= (b-a) \sum_{i=1}^{m-1} \left((b-a) \binom{m-1}{i} - a \binom{m-1}{i-1} \right) (b-a)^i a^{m-1-i} S_{[i+1, m-1+i]} \\
 &\quad - (b-a)^{m+1} S_{[m+1, 2m-1]} + (b-a)^2 a^{m-1} S_{[1, m-1]} \\
 &= (b-a)^2 \sum_{i=0}^{m-1} \binom{m-1}{i} (b-a)^i a^{m-1-i} S_{[i+1, m-1+i]} \\
 &\quad - (b-a) \sum_{i=1}^m \binom{m-1}{i-1} (b-a)^i a^{m-i} S_{[i+1, m-1+i]} \\
 &= (b-a)^2 \sum_{i=0}^{m-1} \binom{m-1}{i} (b-a)^i a^{m-1-i} S_{[i+1, m-1+i]} \\
 &\quad - (b-a)^2 \sum_{j=0}^{m-1} \binom{m-1}{j} (b-a)^j a^{m-1-j} S_{[j+2, m+j]} \\
 &= (b-a)^2 \sum_{i=0}^{m-1} \binom{m-1}{i} (b-a)^i a^{m-1-i} [S_{[i+1, m-1+i]} - S_{[i+2, m+i]}]. \tag{1.87}
 \end{aligned}$$

Prosiguiendo como en los anteriores casos, notamos que para todo $0 \leq i, j < m-1$ tenemos que $(S_{[j+1, m-1+j]} - S_{[j+2, m+j]})^T (\overline{H}_{2m-2}^{[0,1]})^{-1} = E_{m-1, j}^T$, que se sigue, nuevamente,

de sustituir $a = 0$ y $b = 1$ en la ecuación (1.46) y notar que $S_{[j+1,m-1+j]} - S_{[j+2,m+j]}$ corresponde a la j -ésima columna y fila de $\overline{H}_{2m-2}^{[0,1]}$. De igual forma $(\overline{H}_{2m-2}^{[0,1]})^{-1}(S_{[i+1,m-1+i]} - S_{[i+2,m+i]}) = E_{m-1,i}$. El caso $i = j = m - 1$, es especial y analizándolo tenemos $(S_{[m+1,2m-1]} - S_{[m+2,2m]})^T (\overline{H}_{2m-2}^{[0,1]})^{-1} (S_{[m+1,2m-1]} - S_{[m+2,2m]}) = \overline{h}_{2m-1}^T (\overline{H}_{2m-2}^{[0,1]})^{-1} \overline{h}_{2m-1} = S_{2m-1} - S_{2m}^+$. Con base en estos resultados podemos sustituir (1.87) en (1.85) y continuar con:

$$\begin{aligned}
 &= \frac{-1}{(b-a)^2} \left((b-a)^2 \sum_{i=0}^{m-1} \binom{m-1}{i} (b-a)^i a^{m-1-i} [S_{[i+1,m-1+i]} - S_{[i+2,m+i]}] \right)^T (\overline{H}_{2m-2}^{[0,1]})^{-1} \\
 &\quad \times \left((b-a)^2 \sum_{j=0}^{m-1} \binom{m-1}{j} (b-a)^j a^{m-1-j} [S_{[j+1,m-1+j]} - S_{[j+2,m+j]}] \right) \\
 &\quad - abT_{2m-2} + (a+b)T_{2m-1} \\
 &= (b-a)^{2m} S_{2m}^+ - \sum_{r=0}^{2m-3} \sum_{\substack{i+j=r \\ 0 \leq i, j \leq m}} \binom{m-1}{i} \binom{m-1}{j} (b-a)^{r+2} a^{2m-2-r} [S_{r+1} - S_{r+2}] \\
 &\quad - abT_{2m-2} + (a+b)T_{2m-1} - (b-a)^{2m} S_{2m-1} \\
 &= (b-a)^{2m} S_{2m}^+ + \sum_{r=0}^{2m-3} \binom{2m-2}{r} (b-a)^{r+2} a^{2m-2-r} [S_{r+2} - S_{r+1}] - abT_{2m-2} + \\
 &\quad (a+b)T_{2m-1} - (b-a)^{2m} S_{2m-1}, \tag{1.88}
 \end{aligned}$$

recordando la recursión de (1.64) y extrapolando los sumandos $-abT_{2m-2} + (a+b)T_{2m-1}$:

$$\begin{aligned}
 &-abT_{2m-2} + (a+b)T_{2m-1} - (b-a)^{2m} S_{2m-1} \\
 &= -b \sum_{i=0}^{2m-2} \binom{2m-2}{i} (b-a)^i a^{2m-1-i} S_i + \sum_{i=0}^{2m-1} \binom{2m-1}{i} (b-a)^i a^{2m-i} S_i \\
 &\quad + b \sum_{i=0}^{2m-1} \binom{2m-1}{i} (b-a)^i a^{2m-1-i} S_i \\
 &= b \sum_{j=0}^{2m-2} \binom{2m-2}{j} (b-a)^{j+1} a^{2m-2-j} S_{j+1} + \sum_{i=0}^{2m-1} \binom{2m-1}{i} (b-a)^i a^{2m-i} S_i.
 \end{aligned}$$

Con esto en mente y tomando $2 \leq i \leq 2m - 2$ arbitraria, note que el coeficiente de

S_i en la suma es:

$$\begin{aligned}
& \binom{2m-2}{i-2} (b-a)^i a^{2m-i} - \binom{2m-2}{i-1} (b-a)^{i+1} a^{2m-1-i} + \binom{2m-2}{i-1} b (b-a)^i a^{2m-1-i} \\
& \quad + \binom{2m-2}{i-2} (b-a)^i a^{2m-i} \\
& = (b-a)^i a^{2m-1-i} \left(\binom{2m-2}{i-2} a - \binom{2m-2}{i-1} (b-a) + \binom{2m-2}{i-1} b + \binom{2m-1}{i-2} a \right) \\
& = (b-a)^i a^{2m-1-i} \left(\binom{2m-2}{i-2} a + \binom{2m-2}{i-1} a + \binom{2m-1}{i-2} a \right) \\
& = \binom{2m}{i} (b-a)^i a^{2m-i},
\end{aligned}$$

igualmente es fácil ver que el coeficiente de S_0 es a^{2m} , el de S_1 es $2m(b-a)a^{2m-1}$ y finalmente el de S_{2m-1} es $2m(b-a)^{2m-1}a$, por lo que la suma final de (1.88) es:

$$T_{2m}^+ = (b-a)^{2m} S_{2m}^+ + \sum_{i=0}^{2m-1} \binom{2m}{i} (b-a)^i a^{2m-i} S_i,$$

que es lo que queríamos probar.

Para terminar la prueba del Teorema, nos enfocamos ahora en el caso impar de la ecuación (1.79). Análogamente a los anteriores casos, consideramos los vectores matriciales: $\bar{g}_{2m}^T = (bT_m - T_{m+1}, \dots, bT_{2m-1} - T_{2m})$ y $\bar{h}_{2m}^T = (S_m - S_{m+1}, \dots, S_{2m-1} - S_{2m})$. Ahora usando (1.75) y (1.57), tenemos que:

$$\begin{aligned}
T_{2m+1}^+ & = bT_{2m} - \bar{g}_{2m}^T (\bar{H}_{2m-1}^{[a,b]})^{-1} \bar{g}_{2m} \\
& = bT_{2m} - \frac{1}{b-a} (Q_{n,[a,b]}^{-1} \bar{g}_{2m})^T (\bar{H}_{2m-1}^{[0,1]})^{-1} Q_{n,[a,b]}^{-1} \bar{g}_{2m}.
\end{aligned} \tag{1.89}$$

Como en el caso par, note que $\bar{g}_{2m} = bT_{[m,2m-1]} - T_{[m+1,2m]}$, y usando el Lema 1.18:

$$\begin{aligned}
& Q_{n,[a,b]}^{-1} \bar{g}_{2m} \\
&= bQ_{n,[a,b]}^{-1} T_{[m,2m-1]} - Q_{n,[a,b]}^{-1} T_{[m+1,2m]} \\
&= b \sum_{i=0}^m \binom{m}{i} (b-a)^i a^{m-i} S_{[i,m-1+i]} - \sum_{j=0}^{m+1} \binom{m+1}{j} (b-a)^j a^{m+1-j} S_{[j,m-1+j]} \\
&= (b-a)^m S_{[0,m-1]} + \sum_{i=1}^m \left(\binom{m}{i} b (b-a)^i a^{m-i} - \binom{m+1}{i} (b-a)^i a^{m+1-i} \right) S_{[i,m-1+i]} \\
&\quad - (b-a)^{m+1} S_{[m+1,2m]} \\
&= (b-a)^m S_{[0,m-1]} + \sum_{i=1}^m \left(\binom{m}{i} (b-a)^{i+1} a^{m-i} - \binom{m}{i-1} (b-a)^i a^{m+1-i} \right) S_{[i,m-1+i]} \\
&\quad - (b-a)^{m+1} S_{[m+1,2m]} \\
&= (b-a) \sum_{i=0}^m \binom{m}{i} (b-a)^i a^{m-i} S_{[i,m-1+i]} - \sum_{j=1}^{m+1} \binom{m}{j-1} (b-a)^j a^{m+1-j} S_{[j,m-1+j]} \\
&= (b-a) \sum_{i=0}^m \binom{m}{i} (b-a)^i a^{m-i} S_{[i,m-1+i]} - (b-a) \sum_{r=0}^m \binom{m}{r} (b-a)^r a^{m-r} S_{[r+1,m+r]} \\
&= (b-a) \sum_{i=0}^m \binom{m}{i} (b-a)^i a^{m-i} [S_{[i,m-1+i]} - S_{[i+1,m+i]}]. \tag{1.90}
\end{aligned}$$

Análogamente a los anteriores casos, notamos, para simplificar la expresión de (1.89), que para todo $0 \leq i, j < m$ se cumple que $(S_{[j,m-1+j]} - S_{[j+1,m+j]})^T (\bar{H}_{2m-1}^{[0,1]})^{-1} = E_{m,j}^T$, cuya veracidad radica en notar que $S_{[j,m-1+j]} - S_{[j+1,m+j]}$ es la j -ésima columna y fila de la matriz $\bar{H}_{2m-1}^{[0,1]}$. Igualmente se obtiene que $(\bar{H}_{2m-1}^{[0,1]})^{-1} (S_{[i,m-1+i]} - S_{[i+1,m+i]}) = E_{m,i}$. El caso $i = j = m$, se trata diferente, pero basta con observar que $(S_{[m,2m-1]} - S_{[m+1,2m]})^T (\bar{H}_{2m-1}^{[0,1]})^{-1} (S_{[m,2m-1]} - S_{[m+1,2m]}) = \bar{h}_{2m}^T (\bar{H}_{2m-1}^{[0,1]})^{-1} \bar{h}_{2m} = S_{2m} - S_{2m+1}^+$. Con esta información, la expresión de (1.89) es equivalente a:

$$\begin{aligned}
&= \frac{-1}{b-a} \left((b-a) \sum_{i=0}^m \binom{m}{i} (b-a)^i a^{m-i} [S_{[i,m-1+i]} - S_{[i+1,m+i]}] \right)^T (\overline{H}_{2m-1}^{[0,1]})^{-1} \\
&\quad \times \left((b-a) \sum_{i=0}^m \binom{m}{i} (b-a)^i a^{m-i} [S_{[i,m-1+i]} - S_{[i+1,m+i]}] \right) + bT_{2m} \\
&= (b-a)^{2m+1} S_{2m+1}^+ - \sum_{r=0}^{2m-1} \sum_{\substack{i+j=r \\ 0 \leq i, j \leq m}} \binom{m}{i} \binom{m}{j} (b-a)^{r+1} a^{2m-r} [S_r - S_{r+1}] + bT_{2m} \\
&\quad - (b-a)^{2m+1} S_{2m} \\
&= (b-a)^{2m+1} S_{2m+1}^+ + \sum_{r=0}^{2m-1} \binom{2m}{r} (b-a)^{r+1} a^{2m-r} [S_{r+1} - S_r] \\
&\quad + b \sum_{i=0}^{2m} \binom{2m}{i} (b-a)^i a^{2m-i} S_i - (b-a)^{2m+1} S_{2m} \\
&= (b-a)^{2m+1} S_{2m+1}^+ + \sum_{r=0}^{2m-1} \binom{2m}{r} (b-a)^{r+1} a^{2m-r} S_{r+1} + \sum_{j=0}^{2m-1} \binom{2m}{j} (b-a)^j a^{2m+1-j} S_j \\
&\quad + b(b-a)^{2m} S_{2m} - (b-a)^{2m+1} S_{2m} \\
&= (b-a)^{2m+1} S_{2m+1}^+ + \sum_{i=1}^{2m} \binom{2m}{i-1} (b-a)^i a^{2m+1-i} S_i + \sum_{j=0}^{2m} \binom{2m}{j} (b-a)^j a^{2m+1-j} S_j \\
&= (b-a)^{2m+1} S_{2m+1}^+ + \sum_{i=0}^{2m} \binom{2m+1}{i} (b-a)^i a^{2m+1-i} S_i,
\end{aligned}$$

como se quería, completando todos los casos del Teorema. \square

Para completar la invarianza respecto al intervalo donde se encuentra el soporte de nuestra medida, note que el Teorema anterior y la ecuación (1.64), mediante una simple resta nos dice que:

$$T_n - T_n^- = (b-a)^n (S_n - S_n^-), \quad (1.91)$$

$$T_n - T_n^+ = (b-a)^n (S_n - S_n^+), \quad (1.92)$$

$$T_n^+ - T_n^- = (b-a)^n (S_n^+ - S_n^-), \quad (1.93)$$

lo cual claramente implica la invarianza. Abusando de la notación, nos referiremos de ahora en adelante, salvo que se indique lo contrario, a S_i como el i -ésimo momento de la medida matricial $\mu_{[a,b]}$, y a U_n , D_n y V_n como las matrices definidas en (1.58), (1.59) y (1.60), respectivamente (note que en el Teorema anterior usamos a T_i como

el i -ésimo momento de $\mu_{[a,b]}$).

El siguiente Teorema da una representación de D_{n+1} en términos de los momentos canónicos U_k y V_k .

Teorema 1.20. *Sea (S_0, \dots, S_n) en el interior de M_{n+1} . Entonces:*

$$D_{n+1} = (b - a)^{n+1} S_0 U_1 V_1 \dots U_n V_n,$$

y además

$$U_k V_k = V_k U_k, \quad \forall k, \text{ tal que } 1 \leq k \leq n.$$

Demostración. La segunda expresión es un simple resultado de la definición (1.60). Para la primera proposición, usamos la siguiente identidad cuya prueba puede ser localizada en [10, Teorema 2.7] que es puramente una identidad de álgebra lineal, en donde la única diferencia entre la identidad encontrada en [10, (2.20)] , y la nuestra, radica en que estamos trabajando con medidas matriciales en $[a, b]$ y en la referencia se utiliza en $[0, 1]$. Sin embargo, gracias al Teorema 1.19 es que sabemos que dicha identidad de [10, (2.20)] igual es válida sin importar el intervalo en cuestión.

$$D_{n+1} = (S_n^+ - S_n)(S_{n-1}^+ - S_{n-1})^{-1}(S_{n-1}^+ - S_{n-1}^-)(S_{n-1} - S_{n-1}^-)^{-1}(S_n - S_n^-). \quad (1.94)$$

Asumiendo la anterior expresión, obtenemos la prueba mediante un argumento inductivo. Note que $D_1 = (b - a)S_0$ y por (1.61) que $D_2 = (b - a)^2 S_0 U_1 V_1$. Por ende, usando (1.94) tenemos que:

$$\begin{aligned} D_{n+1} &= D_n V_n (D_{n-1} V_{n-1})^{-1} D_{n-1} (D_{n-1} U_{n-1})^{-1} D_n U_n \\ &= D_n V_n V_{n-1}^{-1} U_{n-1}^{-1} D_{n-1}^{-1} D_n U_n \\ &= (b - a) D_n V_n U_n = (b - a)^{n+1} S_0 U_1 V_1 \dots U_n V_n. \end{aligned}$$

Donde usamos que $D_n = (b - a) D_{n-1} U_{n-1} V_{n-1}$ de (1.94). □

Ahora consideramos las definiciones de las siguientes cantidades, asumiendo que (S_0, \dots, S_n) está en el interior de M_{n+1} :

$$\zeta_n := V_{n-1} U_n = (\mathbf{I} - U_{n-1}) U_n, \quad (1.95)$$

$$\gamma_n := U_{n-1} V_n = (\mathbf{I} - V_{n-1}) V_n. \quad (1.96)$$

Por el anterior teorema tenemos:

$$\begin{aligned}
 \mathbf{0} < S_n - S_n^- &= D_n U_n \\
 &= (b-a)^n S_0 U_1 V_1 \dots U_{n-1} V_{n-1} U_n \\
 &= (b-a)^n \left(\frac{1}{(b-a)^{n-1}} D_{n-1} \right) U_{n-1} V_{n-1} U_n \\
 &= (b-a)(S_{n-1} - S_{n-1}^-) \zeta_n \\
 \implies \zeta_n &= \frac{1}{b-a} (S_{n-1} - S_{n-1}^-)^{-1} (S_n - S_n^-).
 \end{aligned}$$

Note que la invertibilidad de $S_{n-1} - S_{n-1}^-$ es producto de que (S_0, \dots, S_n) esté en el interior de M_{n+1} . Análogamente tenemos que:

$$\gamma_n = \frac{1}{b-a} (S_{n-1}^+ - S_{n-1})^{-1} (S_n^+ - S_n) > \mathbf{0}.$$

1.3.2. POM y Momentos Canónicos

Centramos ahora nuestra atención en la relación entre los POM y los momentos canónicos definidos. Para ello recordamos la definición (1.3) del producto interior $\langle \cdot, \cdot \rangle_{\mu_{[a,b]}}$. Como se vio en la Subsección 1.2, nuestra sucesión de polinomios ortogonales se obtiene de ortogonalizar la secuencia $\mathbf{I}, \mathbf{I}t, \mathbf{I}t^2, \dots$, respecto al producto interno definido. Nuestro objetivo será, nuevamente, generalizar el caso escalar, y encontrar una relación entre los coeficientes de los polinomios ortogonales y los momentos canónicos de su medida asociada. Para ello consideramos el determinante de la matriz:

$$\underline{D}_{2m} = \begin{vmatrix} S_0 & \dots & S_m \\ \vdots & \ddots & \vdots \\ S_m & \dots & S_{2m} \end{vmatrix}. \quad (1.97)$$

En el caso escalar, la última fila es reemplazada por el vector $(1, t, \dots, t^m)$ para definir los polinomios ortogonales. En el caso matricial se procede de otra forma. Para obtener la (i, j) entrada del POM $P_m(t)$ escribimos el determinante de arriba con una modificación en la última fila de bloques (S_m, \dots, S_{2m}) . En forma explícita:

$$P_m(t) = \underline{H}_{2m-1}(t) = (H_{i,j}(t))_{i,j}, \quad (1.98)$$

donde los polinomios $H_{i,j}(t)$ son de la forma:

$$H_{i,j}(t) = \begin{vmatrix} S_0 & S_1 & \dots & S_m \\ \vdots & & & \vdots \\ S_{m-1} & S_m & \dots & S_{2m-1} \\ S_m^{i,j}(t) & S_{m+1}^{i,j}(t) & \dots & S_{2m}^{i,j}(t) \end{vmatrix}, \quad i, j = 1, \dots, l. \quad (1.99)$$

Las matrices $S_{m+k}^{i,j}(t)$ son obtenidas al reemplazar la j -ésima fila de S_{m+k} por $e_i^T t^k$, donde $e_i \in \mathbb{C}^l$ es el i -ésimo vector canónico. Un ejemplo de como lucen estos polinomios matriciales puede ser visto en la página 185 de [10]. Para ejemplificar la definición de $P_1(t) = \underline{H}_1(t)$, consideramos el caso en el que $m = 1$ y $l = 2$, y denotamos por $s_{k,(i,j)}$ a la (i, j) -ésima entrada de S_k , el k -ésimo momento matricial. En este caso el polinomio lineal está dado por la matriz de 2×2 de determinantes:

$$P_1(t) = \underline{H}_1(t) = \begin{pmatrix} \begin{vmatrix} \begin{pmatrix} s_{0,1,1} & s_{0,1,2} \\ s_{0,2,1} & s_{0,2,2} \end{pmatrix} & \begin{pmatrix} s_{1,1,1} & s_{1,1,2} \\ s_{1,2,1} & s_{1,2,2} \end{pmatrix} \\ \begin{pmatrix} 1 & 0 \\ s_{1,2,1} & s_{1,2,2} \end{pmatrix} & \begin{pmatrix} t & 0 \\ s_{2,2,1} & s_{2,2,2} \end{pmatrix} \end{vmatrix} & \begin{vmatrix} \begin{pmatrix} s_{0,1,1} & s_{0,1,2} \\ s_{0,2,1} & s_{0,2,2} \end{pmatrix} & \begin{pmatrix} s_{1,1,1} & s_{1,1,2} \\ s_{1,2,1} & s_{1,2,2} \end{pmatrix} \\ \begin{pmatrix} s_{1,1,1} & s_{1,1,2} \\ 1 & 0 \end{pmatrix} & \begin{pmatrix} s_{2,1,1} & s_{2,1,2} \\ t & 0 \end{pmatrix} \end{vmatrix} \\ \begin{vmatrix} \begin{pmatrix} s_{0,1,1} & s_{0,1,2} \\ s_{0,2,1} & s_{0,2,2} \end{pmatrix} & \begin{pmatrix} s_{1,1,1} & s_{1,1,2} \\ s_{1,2,1} & s_{1,2,2} \end{pmatrix} \\ \begin{pmatrix} 0 & 1 \\ s_{1,2,1} & s_{1,2,2} \end{pmatrix} & \begin{pmatrix} 0 & t \\ s_{2,2,1} & s_{2,2,2} \end{pmatrix} \end{vmatrix} & \begin{vmatrix} \begin{pmatrix} s_{0,1,1} & s_{0,1,2} \\ s_{0,2,1} & s_{0,2,2} \end{pmatrix} & \begin{pmatrix} s_{1,1,1} & s_{1,1,2} \\ s_{1,2,1} & s_{1,2,2} \end{pmatrix} \\ \begin{pmatrix} s_{1,1,1} & s_{1,1,2} \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} s_{2,1,1} & s_{2,1,2} \\ 0 & t \end{pmatrix} \end{vmatrix} \end{pmatrix}.$$

Nuestro siguiente resultado prueba que en efecto estos polinomios son ortogonales respecto al producto interior definido, e identifica el coeficiente principal de $P_m(t)$.

Teorema 1.21. *El polinomio $P_m(t) = \underline{H}_{2m-1}^*(t)$ es un m -ésimo POM respecto al producto interior $\langle \cdot, \cdot \rangle_{\mu_{[a,b]}}$ y su coeficiente principal es:*

$$L_m = [\underline{D}_{2m}(S_{2m} - S_{2m}^-)^{-1}]^*,$$

donde \underline{D}_{2m} es el determinante definido en (1.97).

Demostración. Es suficiente probar que para todo $0 \leq k \leq m - 1$ entero arbitra-

rio se tiene que:

$$B = \int_{[a,b]} t^k d\mu_{[a,b]}(t) \underline{H}_{2m-1}(t) = \mathbf{0} \in \mathbb{C}^{l \times l}.$$

Enfocándonos en la (i, j) entrada de B :

$$B_{i,j} = \sum_{q=1}^l \int_{[a,b]} H_{q,j}(t) t^k d\mu_{i,q,[a,b]}(t),$$

donde $\mu_{i,q,[a,b]}$ es la medida de la entrada (i, q) de $\mu_{[a,b]}$. Como solo la $(mp + j)$ -ésima fila de $H_{q,j}(t)$ depende de la variable de integración t , por la forma desarrollada del determinante, la integral entra al determinante en cada una de las entradas de esta fila. Luego por la linealidad del determinante, y al correr la suma de 1 a l , es fácil observar que la expresión $B_{i,j}$ es equivalente a calcular \underline{D}_{2m} , pero reemplazando la $mp + j$ -ésima fila por:

$$\int_{[a,b]} t^k (d\mu_{i,1,[a,b]}, \dots, d\mu_{i,l,[a,b]}), \dots, \int_{[a,b]} t^{m+k} (d\mu_{i,1,[a,b]}, \dots, d\mu_{i,l,[a,b]}),$$

donde $\int_{[a,b]} t^{r+k} (d\mu_{i,1,[a,b]}, \dots, d\mu_{i,l,[a,b]}) = (\int_{[a,b]} t^{r+k} d\mu_{i,1,[a,b]}, \dots, \int_{[a,b]} t^{r+k} d\mu_{i,l,[a,b]})$. Esto último es la i -ésima fila del vector matricial (S_k, \dots, S_{m+k}) , y es por esto que $B_{i,j}$ es 0, pues esta i -ésima fila se repite dos veces, dado que $k < m$. Ahora bien si $k = m$, se aplica el mismo argumento siempre y cuando $i \neq j$ (la fila de bloques (S_m, \dots, S_{2m}) tendrá dos filas iguales), y en el caso $i = j$, $B_{i,i} = \underline{D}_{2m}$, para todo $i = 1, \dots, m$.

Si llamamos L_m al coeficiente principal de $P_m(t)$, tenemos por (1.4) y los cálculos anteriores que:

$$\langle P_m, P_m \rangle_{\mu_{[a,b]}} = L_m \underline{D}_{2m} \mathbf{I} = L_m \underline{\Delta}_{2m}, \quad (1.100)$$

donde $\underline{\Delta}_{2m} = \underline{D}_{2m} \mathbf{I}$. Para encontrar L_m , nos enfocamos en el coeficiente principal de $H_{i,j}$, que es la entrada (i, j) de L_m^* . Note que al ser la $(mp + j)$ -ésima fila del determinante la única que depende de la variable t , el coeficiente principal de $H_{i,j}$ se obtiene al calcular el determinante menor dado por eliminar la fila y columna donde se encuentra la entrada t^m , es decir la $(mp + j)$ -ésima fila y la $(mp + i)$ -ésima columna, y multiplicar por $(-1)^{i+j}$. Note que justamente este valor es el mismo al de eliminar la misma fila y columna de $\underline{H}_{2m}^{[a,b]}$, definida en (1.45), y calcular el correspondiente menor, el cual llamaremos $\underline{D}_{2m}^{mp+j, mp+i}$, con $1 \leq i, j \leq l$. Así, observando que la matriz:

$$((-1)^{i+j} \underline{D}_{2m}^{mp+j, mp+i})_{i,j}$$

es proporcional al bloque en la posición $(m+1, m+1)$ de $(\underline{H}_{2m}^{[a,b]})^{-1}$, por la regla de Cramer. Además recordando que:

$$\underline{H}_{2m}^{[a,b]} = \begin{pmatrix} \underline{H}_{2m-2}^{[a,b]} & h_{2m-1} \\ h_{2m-1}^T & S_{2m} \end{pmatrix},$$

usamos el hecho de que el bloque inferior derecho de la inversa de una matriz de 4 bloques, como la de arriba, es justamente la inversa del complemento de Schur, que es $S_{2m} - S_{2m}^-$, por nuestra definición de S_{2m}^- en (1.54). Es decir, tenemos que:

$$L_m^* = ((-1)^{i+j} \underline{D}_{2m}^{mp+j, mp+i})_{i,j} = \underline{D}_{2m} (S_{2m} - S_{2m}^-)^{-1},$$

concluyendo la proposición restante del Teorema. \square

De este mismo Teorema se sigue que los POM mónicos están dados por:

$$\underline{P}_m(t) = L_m^{-1} P_m(t) = \frac{1}{\underline{D}_{2m}} (S_{2m} - S_{2m}^-)^* P_m(t),$$

lo último se da pues \underline{D}_{2m} es real ya que $\underline{H}_{2m}^{[a,b]}$ es hermítica. Se multiplica a L_m^{-1} por la izquierda por el producto interior que estamos usando. De manera similar es posible encontrar la sucesión de POM con respecto a la medida $(t-a)d\mu_{[a,b]}(t)$, como se ve a continuación en el siguiente Teorema.

Teorema 1.22. *Los polinomios matriciales $Q_m(t)$ definidos por*

$$Q_m(t) = \underline{H}_{2m}(t) = (K_{i,j}(t))_{i,j},$$

donde

$$K_{i,j}(t) = \begin{vmatrix} -aS_0 + S_1 & -aS_1 + S_2 & \dots & -aS_m + S_{m+1} \\ \vdots & & & \vdots \\ -aS_{m-1} + S_m & -aS_m + S_{m+1} & \dots & -aS_{2m-1} + S_{2m} \\ -aS_m^{i,j}(t) + S_{m+1}^{i,j}(t) & -aS_{m+1}^{i,j}(t) + S_{m+2}^{i,j}(t) & \dots & -aS_{2m}^{i,j}(t) + S_{2m+1}^{i,j}(t) \end{vmatrix},$$

son POM respecto la medida $d\lambda(t) = (t-a)d\mu_{[a,b]}(t)$, con $i, j = 1, \dots, l$, y donde $-aS_{k+m}^{i,j} + S_{k+m+1}^{i,j}$ se obtiene, como en el Teorema 1.20, cambiando la j -ésima fila del bloque $(-aS_m + S_{m+1}, \dots, -a_{2m} + S_{2m+1})$ de la última fila de $\underline{D}_{2m+1} = |\underline{H}_{2m+1}^{[a,b]}|$ por el vector $e_i^T t^k$. Más aún, el coeficiente principal de $Q_m(t)$, K_m , está dado por:

$$K_m = [((-1)^{i+j} \underline{D}_{2m+1}^{mp+j, mp+i})_{i,j}]^* = [\underline{D}_{2m+1} (S_{2m+1} - S_{2m+1}^-)^{-1}]^*,$$

donde, análogamente al Teorema anterior, $\underline{D}_{2m+1}^{mp+j, mp+i}$ es el menor obtenido de eliminar la $(mp+j)$ -ésima fila y la $(mp+i)$ -ésima columna de $\underline{H}_{2m+1}^{[a,b]}$.

La prueba del Teorema es totalmente análoga a la del Teorema 1.20, con las ligeras variaciones que produce integrar respecto a $d\lambda(t)$, así como el complemento de Schur obtenido al calcular K_m . De este mismo procedimiento se obtiene que:

$$\langle Q_m, Q_m \rangle_\lambda = K_m \underline{D}_{2m+1} \mathbf{I} = K_m \underline{\Delta}_{2m}. \quad (1.101)$$

Los polinomios mónicos respecto de λ quedan dados por:

$$\underline{Q}_m(t) = K_m^{-1} Q_m(t) = \frac{1}{\underline{D}_{2m+1}} (S_{2m+1} - S_{2m+1}^-) Q_m(t).$$

Finalizamos la sección con el resultado final que expresa los coeficientes de nuestros POM asociados a $d\mu_{[a,b]}(t)$, en términos de las definiciones de (1.95) y (1.96).

Teorema 1.23. *La secuencia $\{\underline{P}_k(t)\}_{k \geq 0}$ de POM mónicos con respecto a la medida matricial $\mu_{[a,b]}$, satisface la relación de recurrencia $\underline{P}_0(t) = \mathbf{I}$, $\underline{P}_{-1}(t) = \mathbf{0}$, y para $m \geq 0$:*

$$(t-a)\underline{P}_m(t) = \underline{P}_{m+1}(t) + (b-a)(\zeta_{2m+1}^* + \zeta_{2m}^*)\underline{P}_m(t) + (b-a)^2 \zeta_{2m-1}^* \zeta_{2m}^* \underline{P}_{m-1}(t),$$

donde las cantidades $\zeta_j \in \mathbb{C}^{l \times l}$ quedan definidas como $\zeta_0 = \mathbf{0}$, $\zeta_1 = U_1$ y $\zeta_j = V_{j-1} U_j$ para $j \geq 2$, y las secuencias $\{U_j\}$, $\{V_j\}$, quedan definidas como en (1.58) y (1.60). Similarmente, la secuencia $\{\underline{Q}_k(t)\}_{k \geq 0}$ ortogonal con respecto $d\lambda(t) = (t-a)d\mu_{[a,b]}(t)$, satisface la recurrencia $\underline{Q}_0(t) = \mathbf{I}$, $\underline{Q}_{-1}(t) = \mathbf{0}$, y para $m \geq 0$:

$$(t-a)\underline{Q}_m(t) = \underline{Q}_{m+1}(t) + (b-a)(\zeta_{2m+1}^* + \zeta_{2m+2}^*)\underline{Q}_m(t) + (b-a)^2 \zeta_{2m+1}^* \zeta_{2m}^* \underline{Q}_{m-1}(t).$$

Demostración. Considere la expresión $(t-a)\underline{Q}_m(t) - A_m \underline{P}_m(t)$, con A_m una matriz de tamaño $l \times l$ por definir, de tal forma que la expresión anterior sea igual a $\underline{P}_{m+1}(t)$. Note que:

$$\langle (t-a)\underline{Q}_m(t), t^k \rangle_{\mu_{[a,b]}} - A_m \langle \underline{P}_m(t), t^k \rangle_{\mu_{[a,b]}} = \mathbf{0}, \quad k = 1, \dots, m.$$

Claramente, por ortogonalidad de los polinomios $(t-a)\underline{Q}_m(t)$ y $\underline{P}_m(t)$ ambos sumandos son $\mathbf{0}$ para $k < m$. Si $k = m$, entonces:

$$\langle (t-a)\underline{Q}_m(t), x^m \rangle_{\mu_{[a,b]}} = A_m \langle \underline{P}_m(t), t^m \rangle_{\mu_{[a,b]}}. \quad (1.102)$$

Para poder seguir calculamos $\langle (t-a)\underline{Q}_m(t), t^m \rangle_{\mu_{[a,b]}}$ y $\langle \underline{P}_m(t), x^m \rangle_{\mu_{[a,b]}}$, independientemente.

Por la ortogonalidad de $\underline{P}_m(t)$ y siguiendo la expresión (1.100) tenemos:

$$\begin{aligned} \langle t^m, \underline{P}_m(t) \rangle_{\mu_{[a,b]}} &= \langle \underline{P}_m(t), \underline{P}_m(t) \rangle_{\mu_{[a,b]}} \\ &= \langle L_m^{-1} P_m(t), L_m^{-1} P_m(t) \rangle_{\mu_{[a,b]}} \\ &= L_m^{-1} \langle P_m(t), P_m(t) \rangle_{\mu_{[a,b]}} (L_m^{-1})^* \\ &= L_m^{-1} L_m \underline{\Delta}_{2m} (L_m^{-1})^* \\ &= S_{2m} - S_{2m}^-. \end{aligned}$$

Pero usando la identidad (1.5), tenemos que:

$$\langle \underline{P}_m(t), x^m \rangle_{\mu_{[a,b]}} = (S_{2m} - S_{2m}^-)^*.$$

Análogamente:

$$\begin{aligned} \langle t^m, (t-a)\underline{Q}_m(t) \rangle_{\mu_{[a,b]}} &= \langle \underline{Q}_m(t), (t-a)\underline{Q}_m(t) \rangle_{\mu_{[a,b]}} \\ &= \langle K_m^{-1} Q_m(t), K_m^{-1} Q_m(t) \rangle_{\lambda} \\ &= K_m^{-1} \langle Q_m(t), Q_m(t) \rangle_{\lambda} (K_m^{-1})^* \\ &= \underline{\Delta}_{2m+1} (K_m^{-1})^* \\ &= S_{2m+1} - S_{2m+1}^-, \end{aligned}$$

y como en el cálculo anterior:

$$\langle (t-a)\underline{Q}_m(t), x^m \rangle_{\mu_{[a,b]}} = (S_{2m+1} - S_{2m+1}^-)^*,$$

sustituyendo en (1.102) tenemos que:

$$A_m = (S_{2m+1} - S_{2m+1}^-)^* ((S_{2m} - S_{2m}^-)^*)^{-1} = (b-a) \zeta_{2m+1}^*.$$

Análogamente si consideramos la expresión $\underline{P}_{m+1}(t) - B_m \underline{Q}_m(t)$, con B_m por determinar para que la expresión sea idénticamente $\underline{Q}_{m+1}(t)$, y la multiplicamos por t^k , resulta $\mathbf{0}$ para $k < m$ y como en el caso anterior para $k = m$ obtenemos que, usando (1.101):

$$B_m = (S_{2m+2} - S_{2m+2}^-)^* ((S_{2m+1} - S_{2m+1}^-)^*)^{-1} = (b-a) \zeta_{2m+2}^*.$$

Así las expresiones quedan como:

$$\begin{aligned} \underline{P}_{m+1}(t) &= (t-a)\underline{Q}_m(t) - (b-a)\zeta_{2m+1}^* \underline{P}_m(t), \\ \underline{Q}_{m+1}(t) &= \underline{P}_{m+1}(t) - (b-a)\zeta_{2m+2}^* \underline{Q}_m(t). \end{aligned}$$

De las últimas ecuaciones tenemos que:

$$\underline{Q}_m(t) = \frac{1}{t-a} (\underline{P}_{m+1}(t) + (b-a)\zeta_{2m+1}^* \underline{P}_m(t)),$$

sustituyendo en $\underline{P}_m(t) = \underline{Q}_m(t) + (b-a)\zeta_{2m}^* \underline{Q}_{m-1}(t)$:

$$\begin{aligned} \underline{P}_m(t) &= \frac{1}{t-a} (\underline{P}_{m+1}(t) + (b-a)\zeta_{2m+1}^* \underline{P}_m(t)) \\ &\quad + \frac{1}{t-a} (b-a)\zeta_{2m}^* (\underline{P}_m(t) + (b-a)\zeta_{2m-1}^* \underline{P}_{m-1}(t)), \end{aligned}$$

obteniendo,

$$(t-a)\underline{P}_m(t) = \underline{P}_{m+1}(t) + (b-a)(\zeta_{2m+1}^* + \zeta_{2m}^*)\underline{P}_m(t) + (b-a)^2\zeta_{2m-1}^*\zeta_{2m}^*\underline{P}_{m-1}(t).$$

De forma análoga, usando un argumento similar, se obtiene la segunda parte del Lema. \square

Vale la pena notar que estas ecuaciones de recurrencia son equivalentes a las presentadas en la subsección 1.2.2.

Capítulo 2

Procesos cuasi de nacimiento y muerte

Durante este capítulo definiremos los procesos cuasi de nacimiento y muerte, los cuales básicamente son una generalización matricial de los procesos de nacimiento y muerte. Con base en los resultados del Capítulo 1, específicamente el Teorema de Favard, podremos obtener condiciones de equivalencia para la existencia de una medida matricial μ tal que la recursión de polinomios asociadas al proceso sea ortogonal respecto a esta medida. A continuación, obtendremos propiedades de recurrencia del proceso, y haremos uso de los momentos canónicos para relacionar estas expresiones y la Transformada de Stieltjes de la medida.

2.1. Introducción

Empezamos con la parte central de nuestro escrito. Estudiaremos las principales propiedades de este tipo de procesos, los cuales quedan caracterizados como procesos de Markov homogéneos tales que su matriz de transición es tridiagonal por bloques. De forma más explícita, nuestra cadena de Markov tiene por espacio de estados a:

$$\mathcal{C}_l = \{(i, j) \in \mathbb{N}_0 \times \mathbb{N} \mid 1 \leq j \leq l\}, \quad (2.1)$$

y con matriz de transición tridiagonal por bloques:

$$P = \begin{pmatrix} B_0 & A_0 & \mathbf{0} & \mathbf{0} & \dots \\ C_1^T & B_1 & A_1 & \mathbf{0} & \dots \\ \mathbf{0} & C_2^T & B_2 & A_2 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad (2.2)$$

donde l es un entero finito y $\{A_k\}_{k=0}^\infty$, $\{B_k\}_{k=0}^\infty$ y $\{C_k\}_{k=1}^\infty$ son matrices de dimensión $l \times l$ que contienen en sí mismas las probabilidades de transición en un paso, (usaremos el super índice de T para indicar la matriz transpuesta). Así si representamos a nuestra matriz de transición tridiagonal por bloques como

$$P = (P_{i,i'})_{i,i'=0,1,\dots,}$$

donde las entradas $P_{i,i'}$ son las matrices cuadradas de tamaño $l \times l$, la probabilidad de ir de (i, j) a (i', j') queda almacenada en (j, j') -ésima entrada de $P_{i,i'}$. En el estado (i, j) la entrada i es usualmente referida como el nivel del estado, y la entrada j como la fase. De forma análoga al caso escalar definiremos a $P_{i,j}^n$ como el (i, j) bloque matricial de la matriz P^n , y por las ecuaciones de Chapman-Kolmogorov la probabilidad de ir de (i, j) a (i', j') en n pasos está almacenada en la (j, j') -ésima entrada de $P_{i,i'}$. Normalmente, como lo enuncia el título del capítulo, a este tipo de procesos de Markov se le nombrará como cuasi de nacimiento y muerte ('Quasi Birth and Death' en inglés, o QBD por sus siglas).

Este tipo de procesos esta íntimamente relacionado con la sucesión de polinomios con valores matriciales $Q_n(x)$ definidos por la recurrencia:

$$xQ_n(x) = A_nQ_{n+1}(x) + B_nQ_n(x) + C_n^TQ_{n-1}(x), \quad n \geq 1. \quad (2.3)$$

Con polinomios iniciales $Q_0(x) = \mathbf{I}$ y $Q_{-1}(x) = \mathbf{0}$. El lector podrá notar que la matriz de transición de nuestro proceso, así como la ecuación de recurrencia previa, es familiar a la definida en la ecuación (1.23) de la sección anterior. La diferencia radica, sin considerar que la numeración de las matrices por bloques definidas en (2.2) empieza en 0, en que al representar probabilidades las entradas ahora son particularmente reales, así como la simetría en B_k no es necesaria, ni que se satisfaga que $A_n = C_{n+1}$. Sin embargo, por el Teorema de Favard, si se cumplen estas propiedades, esto implica la existencia de una medida matricial hermítica definida positiva sobre la recta real, introducida en la subsección 1.1.1 de nuestro capítulo anterior, $\mu = (\mu_{i,j})$ con $1 \leq i, j \leq l$, para la cual la sucesión de polinomios definidos en (2.3) es ortogonal con respecto a la forma sesquilineal de la subsección 1.1.2:

$$\langle Q_i, Q_j \rangle_\mu = \int_{\mathbb{R}} Q_i(x) d\mu(x) Q_j^T(x) = \delta_{i,j} F_i,$$

pues al tener los Q_n entradas reales tenemos que $Q_n^T(x) = Q_n^*(x)$.

2.2. Existencia de medidas matriciales sobre un proceso cuasi de nacimiento y muerte

En esta sección nuestro principal objetivo es encontrar, en determinadas condiciones para nuestra matriz de transición P , una medida matricial μ , para la cual la sucesión de polinomios definida por la recurrencia (2.3) es un conjunto de POM, y además cuya medida está soportada en el intervalo $[-1, 1]$. Es decir, buscamos una forma de generalización al Teorema de Favard visto previamente en el capítulo anterior.

Para una medida matricial μ definimos a su k -ésimo momento matricial como el definido en (1.32), donde solo consideramos a las medidas μ cuyos momentos existen y son finitos. Por lo tratado en la sección 1.3, una sucesión de matrices $\{S_k\}_{k \geq 0}$ que satisface que las matrices de Hankel por bloques

$$\underline{H}_{2m} = \begin{pmatrix} S_0 & S_1 & \cdots & S_m \\ S_1 & S_2 & \cdots & S_{m+1} \\ \vdots & \vdots & \ddots & \vdots \\ S_m & S_{m+1} & \cdots & S_{2m} \end{pmatrix},$$

son definidas positivas, entonces, implicará que existe una medida matricial μ que posee a S_k como k -ésimo momento matricial, y como se vio en el capítulo anterior a esta medida se le asocia una sucesión de POM.

Tomando como $\{Q_j(x)\}_{j \geq 0}$ la sucesión de polinomios definida por la recurrencia (2.3), consideramos que las matrices C_i , $i \in \mathbb{N}$ y A_j , $j \in \mathbb{N}_0$ son invertibles. El siguiente Teorema otorgará una caracterización sobre la existencia de una medida matricial para la cual la sucesión $\{Q_j(x)\}_{j \geq 0}$ es ortogonal. Note que esto es una generalización al resultado del capítulo anterior al otorgar más libertad a las matrices A_j y C_i .

Teorema 2.1. *Asumiendo que las matrices C_i , $i \in \mathbb{N}$ y A_j , $j \in \mathbb{N}_0$, definidas en la matriz de transición por bloques y tridiagonal, son invertibles, existe una medida matricial μ soportada en la recta real, cuyas matrices por bloques de Hankel \underline{H}_{2m} son definidas positivas, tal que la sucesión de polinomios $\{Q_j(x)\}_{j \geq 0}$ definida por (2.3) es ortogonal con respecto $\langle \cdot, \cdot \rangle_\mu$ si y solo si existe una sucesión de matrices $\{R_n\}_{n \geq 0}$ invertibles, para las cuales se cumplen las siguientes ecuaciones:*

$$R_n B_n R_n^{-1}, \text{ es simétrica } \forall n \in \mathbb{N}_0, \quad (2.4)$$

$$R_n^T R_n = C_n^{-1} \cdots C_1^{-1} (R_0^T R_0) A_0 \cdots A_{n-1} \forall n \in \mathbb{N}. \quad (2.5)$$

Demostración. Asumimos primero que la secuencia $\{Q_j(x)\}_{j \geq 0}$ de polinomios matriciales es ortogonal con respecto a la medida $d\mu$, cuyas matrices de Hankel \underline{H}_{2m} son definidas positivas. Esto último implica que:

$$\int_{\mathbb{R}} Q_i(x) d\mu(x) Q_i^T(x) = F_i \geq \mathbf{0}, \quad i \in \mathbb{N}_0.$$

En el capítulo 1 vimos que estas F_i son semi-definidas positivas. Para ver que en efecto son definidas positivas basta escribir a $Q_i(x) = \sum_{j=0}^i M_j x^j$, y sustituir esto en la ecuación anterior. Al desarrollar obtenemos:

$$\begin{aligned} F_i &= \int_{\mathbb{R}} \left(\sum_{j=0}^i M_j x^j \right) d\mu(x) \left(\sum_{r=0}^i M_r^T x^r \right) = \int_{\mathbb{R}} \sum_{k=0}^{2i} \sum_{i+j=k} x^k M_i d\mu(x) M_j^T \\ &= \sum_{k=0}^{2i} \sum_{i+j=k} \int_{\mathbb{R}} x^k M_i d\mu(x) M_j^T = \sum_{k=0}^{2i} \sum_{i+j=k} M_i S_k M_j^T = M \underline{H}_{2i} M^T > \mathbf{0}, \end{aligned}$$

donde la matriz M queda definida como el vector compuesto de matrices:

$$M = (M_0, M_1, \dots, M_i).$$

Con base en esto claramente F_i es definida positiva pues \underline{H}_{2i} lo es.

Ahora siguiendo con la prueba, definimos a $R_n = F_n^{-1/2}$ y a $\tilde{Q}_n(x) = R_n Q_n(x)$ con $n \geq 0$. Como en el Lema 1.9, estos polinomios son ortonormales. Por ende podemos aplicar el Teorema de Favard del capítulo anterior para encontrar matrices $\{D_n\}$ invertibles y $\{E_n\}$ simétricas tales que:

$$x\tilde{Q}_n(x) = D_{n+1}\tilde{Q}_{n+1}(x) + E_n\tilde{Q}_n(x) + D_n^T\tilde{Q}_{n-1}(x),$$

se satisface para todo entero positivo n , y además $\tilde{Q}_{-1}(x) = \mathbf{0}$ y $\tilde{Q}_0(x) = R_0$, pues $Q_0(x) = \mathbf{I}$. Por otro lado como $\{Q_j(x)\}_{j \geq 0}$ satisface (2.3) sustituyendo tenemos:

$$x\tilde{Q}_n(x) = R_n A_n R_{n+1}^{-1} \tilde{Q}_{n+1}(x) + R_n B_n R_n^{-1} \tilde{Q}_n(x) + R_n C_n R_{n-1}^{-1} \tilde{Q}_{n-1}(x). \quad (2.6)$$

Como la recurrencia es única, los coeficientes respectivos deben ser iguales, es decir comparándolos en las últimas dos ecuaciones, implican:

$$D_{n+1} = R_n A_n R_{n+1}^{-1}, \quad E_n = R_n B_n R_n^{-1}, \quad D_n^T = R_n C_n R_{n-1}^{-1}.$$

Por ende tomando la simetría:

$$R_n A_n R_{n+1}^{-1} = (R_{n+1} C_{n+1} R_n^{-1})^T = (R_n^{-1})^T C_{n+1} R_{n+1}^T.$$

Multiplicando lo anterior por R_n a la derecha, y por $C_{n+1}^{-1}R_n^T$ a la izquierda, se obtiene:

$$R_{n+1}^T R_{n+1} = C_{n+1}^{-1} (R_n^T R_n) A_n.$$

De aquí es fácil ver que podemos proceder de forma inductiva para lograr la ecuación pedida:

$$R_n^T R_n = C_n^{-1} \cdots C_1^{-1} (R_0^T R_0) A_0 \cdots A_{n-1} \quad \forall n \in \mathbb{N},$$

concluyendo la primera parte del Teorema.

Probamos ahora la otra implicación. Asumiendo que las ecuaciones (2.4) y (2.5) se cumplen, y consideramos nuevamente los polinomios $\tilde{Q}_n(x) = R_n Q_n(x)$. Al igual que en la primera parte de la prueba, esta sustitución nos otorga la misma relación de recurrencia que en (2.6), e igualmente comparando coeficientes y asumiendo (2.4), tenemos que:

$$E_n = R_n B_n R_n^{-1}, \quad (2.7)$$

es simétrica. De igual forma:

$$D_{n+1} = R_n A_n R_{n+1}^{-1} = (R_{n+1} C_{n+1} R_n^{-1})^T, \quad (2.8)$$

esto se obtiene de asumir (2.5), y de proceder al revés en la cadena de igualdades en las últimas tres ecuaciones de la prueba de la primera parte. Por ende, la relación de recurrencia es de la forma (2.6), y así nuevamente estamos en posición de poder usar el Teorema de Favard y concluir que esta sucesión de polinomios es ortonormal con respecto a una medida matricial μ . De igual forma esto implica la ortogonalidad de los polinomios $Q_n(x) = R_n^{-1} \tilde{Q}_n(x)$ con respecto a μ .

Nos resta probar que en efecto las matrices de Hankel generadas por esta medida son positivas definidas. Note que usando (2.6) para $n = 0$ y despejando $\tilde{Q}_1(x)$ obtenemos que el coeficiente principal de este polinomio es $D_1^{-1} R_0$, por lo que procediendo de forma inductiva, y usando:

$$\tilde{Q}_{n+1}(x) = x D_{n+1}^{-1} \tilde{Q}_n(x) - D_{n+1}^{-1} E_n \tilde{Q}_n(x) - D_{n+1}^{-1} D_n^T \tilde{Q}_{n-1}(x), \quad (2.9)$$

si el coeficiente principal de $\tilde{Q}_n(x)$ es $D_n^{-1} D_{n-1}^{-1} \cdots D_1^{-1} R_0$, entonces el de $\tilde{Q}_{n+1}(x)$ es, por (2.9), $D_{n+1}^{-1} D_n^{-1} \cdots D_1^{-1} R_0$. Por ende si definimos a $\underline{Q}_n(x) = R_0^{-1} D_1 \cdots D_n \tilde{Q}_n(x)$ estos tendrán como coeficiente principal a la identidad. Por lo tanto obtenemos que:

$$\langle \underline{Q}_n, \underline{Q}_n \rangle_\mu = \int \underline{Q}_n(x) \mu(x) \underline{Q}_n^T(x) = R_0^{-1} D_1 \cdots D_n D_n^T \cdots D_1^T (R_0^T)^{-1},$$

que es invertible y por ende definida positiva. Por otro lado, se sigue que, por los resultados obtenidos en el Capítulo 1, en la sección 1.3, que el lado izquierdo de la ecuación anterior es $(S_{2n} - S_{2n}^-)^T$. De igual forma, por el Teorema 1.17, \underline{H}_{2n} es definida positiva si y solo si \underline{H}_{2n-2} y su complemento de Schur, es decir $\langle \underline{Q}_n, \underline{Q}_n \rangle_\mu$, es definida positiva. Procediendo de forma inductiva, es claro que todas las matrices de Hankel generadas por esta medida son definidas positivas. \square

Como los POM $\{\tilde{Q}_n\}_{n \geq 0}$ son ortonormales con respecto a la medida μ se sigue que:

$$\mathbf{I} = \langle \tilde{Q}_0, \tilde{Q}_0 \rangle_\mu = \int \tilde{Q}_0(x) \mu(x) \tilde{Q}_0^T(x) = R_0 S_0 R_0^T,$$

o lo que es equivalente:

$$S_0 = R_0^{-1} ((R_0^T)^{-1}) = (R_0^T R_0)^{-1}, \quad (2.10)$$

y como sabemos, S_0 es el 0-ésimo momento de la medida matricial μ . Por último note que, por el Lema 1.9, las matrices $\{R_k\}_{k \geq 0}$ no son únicas, pues de hecho, por este mismo resultado, cualquier secuencia $\{\sigma_k R_k\}_{k \geq 0}$ con σ_k unitaria para toda $k \in \mathbb{N}$, satisface las condiciones deseadas.

Ahora, obtendremos el análogo de la representación de Karlin-McGregor en este tipo de procesos.

Teorema 2.2. *Considerando que las suposiciones del Teorema 2.1 se satisfacen, el bloque $P_{i,j}^n$ de la posición (i, j) de la matriz de transición P , dada en (2.2) tiene la siguiente representación:*

$$P_{i,j}^n = \left(\int_{\mathbb{R}} x^n Q_i(x) d\mu(x) Q_j^T(x) \right) \left(\int_{\mathbb{R}} Q_j(x) d\mu(x) Q_j^T(x) \right)^{-1}, \quad (2.11)$$

donde μ es la medida matricial correspondiente a P .

Demostración. Para esto definamos al vector cuyas entradas son polinomios matriciales $Q(x) = (Q_0^T(x), Q_1^T(x), \dots)^T$, generados por la recurrencia (2.3). Esta representación nos permite obtener la equivalencia de (2.3) con:

$$xQ(x) = PQ(x),$$

donde P es la matriz de transición previamente definida en (2.2). Aplicando de forma iterativa esta igualdad tenemos:

$$x^n Q(x) = P^n Q(x).$$

Por ende, si multiplicamos por $d\mu(x)Q^T(x)$ por la derecha ambos lados e integramos entrada a entrada, de acuerdo al producto punto, obtenemos (abusando la notación):

$$\int_{\mathbb{R}} x^n Q(x) d\mu(x) Q_j^T(x) = P^n \int_{\mathbb{R}} Q(x) d\mu(x) Q_j^T(x). \quad (2.12)$$

En el lado derecho de la anterior igualdad, note que al ser $\{Q_j(x)\}_{j \geq 0}$ una sucesión de polinomios ortogonales todas las entradas, salvo la j -ésima, se anulan. Por lo tanto considerando solo la i -ésima entrada del vector matricial del lado izquierda, tenemos que es $\int x^n Q_i(x) d\mu(x) Q_j^T(x)$, y por otro lado la i -ésima entrada del lado derecho, por lo anterior, al ser solo la j -ésima entrada de $\int Q(x) d\mu(x) Q_j^T(x)$ la que no se anula, en la multiplicación de P^n y $\int Q(x) d\mu(x) Q_j^T(x)$ solo permanece $P_{i,j}^n \int Q_j(x) d\mu(x) Q_j^T(x)$, y así:

$$P_{i,j}^n \int_{\mathbb{R}} Q_j(x) d\mu(x) Q_j^T(x) = \int_{\mathbb{R}} x^n Q_i(x) d\mu(x) Q_j^T(x),$$

o equivalentemente:

$$P_{i,j}^n = \left(\int_{\mathbb{R}} x^n Q_i(x) d\mu(x) Q_j^T(x) \right) \left(\int_{\mathbb{R}} Q_j(x) d\mu(x) Q_j^T(x) \right)^{-1},$$

la cual es la ecuación análoga a la representación de Karlin-McGregor para el caso matricial. \square

Note, antes de continuar con el siguiente Teorema, que la medida correspondiente a P , no necesariamente está determinada de forma única por P . Sin embargo, sustituyendo $i = j = 0$ en (2.11), tenemos la representación de los momentos:

$$S_n = P_{0,0}^n S_0,$$

para todo natural n . Por ende, todos los momentos de la medida matricial quedan únicamente definidos por $P_{0,0}$ y S_0 . En el siguiente resultado encontraremos condiciones suficientes para que la medida correspondiente μ , si es que existe, quede con su respectivo soporte contenido en $[-1, 1]$, y en este caso quedará determinada por sus momentos.

Teorema 2.3. *Asumiendo nuevamente las condiciones del Teorema 2.1 y definiendo*

la matriz diagonal por bloques:

$$R = \begin{pmatrix} R_0 & \mathbf{0} & \mathbf{0} & \dots \\ \mathbf{0} & R_1 & \mathbf{0} & \dots \\ \mathbf{0} & \mathbf{0} & R_2 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad (2.13)$$

si R es simétrica, y la matriz:

$$\tilde{P} = R^T P R^{-1}, \quad (2.14)$$

tiene entradas no-negativas, entonces la medida matricial correspondiente a los polinomios de (2.3) está soportada en $[-1, 1]$, dicho de otra forma, se cumple que:

$$\text{supp}(\mu_{i,j}) \subset [-1, 1] \quad \forall i, j = 1, \dots, l.$$

Demostración. Antes de continuar probamos una versión de un resultado de análisis, conocido como la Prueba de Schur, para poder encontrar una cota a la norma del operador P .

Prueba de Schur. Sean X, Y espacios medibles, y sea T un operador integral en $L^2 \rightarrow L^2$ con representación de la forma:

$$Tf(x) = \int_Y K(x, y)f(y)dy, \quad (2.15)$$

con $K(x, y)$ no negativo, $x \in X$ and $y \in Y$. Si existen funciones $p(x) > 0$, $q(y) > 0$ y números $\alpha, \beta > 0$ tales que:

$$\int_Y K(x, y)q(y)dy \leq \alpha p(x), \quad (2.16)$$

para casi todo x, y

$$\int_X p(x)K(x, y)dx \leq \beta q(y), \quad (2.17)$$

para casi todo y , entonces T está acotada en norma por:

$$\|T\|_2^2 \leq \alpha\beta.$$

Demostración. Aplicamos la desigualdad de Cauchy-Schwarz en la ecuación (2.15) de la siguiente forma:

$$\begin{aligned} |Tf(x)|^2 &= \left| \int_Y K(x, y) f(y) dy \right|^2 \leq \left(\int_Y K(x, y) q(y) dy \right) \left(\int_Y \frac{K(x, y) f(y)^2}{q(y)} dy \right) \\ &\leq \alpha p(x) \left(\int_Y \frac{K(x, y) f(y)^2}{q(y)} dy \right), \end{aligned}$$

de aquí integramos en todo X , aplicamos el Teorema de Fubini a la resultante integral y finalmente aplicamos la desigualdad (2.16):

$$\begin{aligned} \|Tf\|_2^2 &\leq \alpha \int_Y \left(\int_X p(x) K(x, y) dx \right) \frac{f(y)^2}{q(y)} dy \\ &\leq \alpha \beta \int_Y f(y)^2 dy \\ &= \alpha \beta \|f\|_2^2, \end{aligned}$$

lo cual claramente implica que $\|T\|_2^2 \leq \alpha \beta$, concluyendo. \square

Regresando al problema, tomando $X = \mathbb{N}_0$ y $Y = \mathbb{N}$, y a $T = \tilde{P}$ como el operador integral, el cual es, pues tomamos a $K(i, j)$ como la (i, j) entrada de \tilde{P} , que sabemos es no negativa, y por ende si f es un vector, en donde $f(i)$ es su i -ésima entrada, claramente el producto matricial $\tilde{P}f$ tiene la representación integral (2.15), pues en este caso nuestra integral es discreta. Por ende si tomamos $\alpha = \beta = 1$, y encontramos vectores u, w tales que:

$$\tilde{P}v \leq w \text{ y } \tilde{P}w \leq v,$$

(donde como en este caso $v = p$, y como dijimos anteriormente $v(i)$ es la i -ésima entrada de v , y así las anteriores desigualdades son entrada a entrada para satisfacer (2.16) y (2.17)) se cumplirán las condiciones de la prueba de Schur, y podremos concluir que:

$$\|\tilde{P}\|_2 \leq 1. \quad (2.18)$$

Por lo tanto, tomando $v = w = R\mathbf{1}$, donde $\mathbf{1}$ es el vector infinito con todas sus entradas igual a 1. Por ende, aplicando la prueba de Schur, implica:

$$\tilde{P}v = \tilde{P}R\mathbf{1} = R^T P\mathbf{1} = RP\mathbf{1} \leq R\mathbf{1} = w, \quad (2.19)$$

donde la última igualdad se da por ser R simétrica, y la desigualdad por ser P una matriz de transición y por ende sub-estocástica. Ahora, consideremos:

$$\Pi_j = C_j^{-1} \cdots C_1^{-1} R_0^T R_0 A_0 \cdots A_{j-1} = R_j^T R_j, \quad (2.20)$$

y el siguiente producto interior (el cual es fácil ver que en efecto lo es):

$$\langle x, y \rangle_{\Pi} = \sum_{j=0}^{\infty} x_j^T \Pi_j y_j, \quad (2.21)$$

en donde x, y son vectores infinitos cuyas entradas son a su vez vectores de dimensión l , es decir en el espacio $\mathcal{H}_v = \ell^2(\mathbb{Z}_+, \mathbb{R}^l)$ visto en la subsección 1.2.5, pero con distinto producto interior. Este producto interior define, a su vez, una norma $\|\cdot\|_{\Pi}$. Así definimos a:

$$\ell_{\Pi}^2(\mathbb{N}, \mathbb{C}^l) = \{x = (x_0^T, x_1^T, \dots) \mid x_j \in \mathbb{R}^l \text{ y } \|x\|_{\Pi}^2 < \infty\}.$$

Por la definición de $P_{i,j}$, Π_i y Π_j tenemos que hay 3 casos para $P_{i,j} \neq \mathbf{0}$:

1. Que $P_{i,i+1} = A_i$ y $P_{i+1,i} = C_{i+1}^T$ y así:

$$R_i^T R_i A_i = C_i^{-1} \cdots C_1^{-1} R_0^T R_0 A_0 \cdots A_i = C_{i+1} R_{i+1}^T R_{i+1},$$

por definición de Π_i y Π_{i+1} . A su vez es equivalente a:

$$\Pi_i P_{i,i+1} = P_{i+1,i}^T \Pi_{i+1}.$$

2. Que $j = i - 1$ y entonces $P_{i,i-1} = C_i^T$ y $P_{i-1,i}^T = A_{i-1}$ y así, tomando la traspuesta a la ecuación anterior obtenemos:

$$A_{i-1}^T R_{i-1}^T R_{i-1} = R_i^T R_i C_i^T,$$

y otra vez esto nuevamente es equivalente a:

$$\Pi_i P_{i,i-1} = P_{i-1,i}^T \Pi_{i-1}.$$

3. Finalmente si $i = j$, $P_{i,i} = B_i$ y como por hipótesis se satisfacen las condiciones del Teorema 2.1, por la simetría de (2.4):

$$R_i B_i R_i^{-1} = (R_i B_i R_i^{-1})^T = (R_i^T)^{-1} B_i R_i^T,$$

y esto último se puede escribir reordenando:

$$R_i^T R_i B_i = B_i R_i^T R_i,$$

lo cual es equivalente a:

$$\Pi_i P_{i,i} = P_{i,i}^T \Pi_i.$$

Finalmente si $P_{i,j} = \mathbf{0}$, claramente se satisface que $\Pi_i P_{i,j} = P_{j,i}^T \Pi_j = \mathbf{0}$, por lo que esta ecuación es válida para todo $i, j \in \mathbb{N}_0$. Esto último implica que P es auto-adjunto con respecto este producto interior ya que:

$$\langle Px, y \rangle_{\Pi} = \sum_{(i,j) \in \mathbb{N}_0 \times \mathbb{N}_0} x_i^T P_{j,i}^T \Pi_j y_j = \sum_{(i,j) \in \mathbb{N}_0 \times \mathbb{N}_0} x_i^T \Pi_i P_{i,j} y_j = \langle x, Py \rangle_{\Pi},$$

donde la segunda igualdad se da $\Pi_i P_{i,j} = P_{j,i}^T \Pi_j$ que recién probamos. Así, conociendo esto, note que para todo vector infinito x , y definiendo $\Pi = R^T R$:

$$\begin{aligned} \|Px\|_{\Pi} &= x^T P^T \Pi P x = x^T P^T R^T R P x = x^T R^T \tilde{P}^T \tilde{P} R x = \|\tilde{P} R x\|_2 \\ &\leq \|\tilde{P}\|_2 \|R x\|_2 \leq x^T R^T R x = x^T \Pi x = \|x\|_{\Pi}, \end{aligned}$$

en donde usamos (2.19), y que $\|R x\|_2 = x^T R^T R x = x^T \Pi x = \|x\|_{\Pi}$. Así por la invertibilidad de R , esto implica que la norma de operador de P es menor que 1, tanto en $\|\cdot\|_2$ como en $\|\cdot\|_{\Pi}$. Para concluir usamos el resultado de [12, Sección 3], que establece que al ser P un operador acotado, auto-adjunto con norma menor que 1 su espectro es real, compacto y está contenido en $[-1, 1]$. Después, usando [12, Sección 2], notamos que las raíces de los polinomios ortonormales definidos por la recurrencia (2.3), forman parte del espectro de P , y nuevamente por [12, Sección 3], el soporte de μ queda contenido en $[-1, 1]$. \square

Ahora, enfocamos nuestra atención en buscar una relación entre dos medidas matriciales, diga μ y $\tilde{\mu}$, correspondientes a un proceso cuasi de nacimiento y muerte, mediante la transformada de Stieltjes, cuando sus respectivas matrices de transición, sean P y \tilde{P} , difieren únicamente en el bloque $(1, 1)$ el cual es B_0 y \tilde{B}_0 , respectivamente.

Teorema 2.4. *Considere la matriz de transición P , dada en (2.2), correspondien-*

te a un proceso cuasi de nacimiento y muerte, y la nueva matriz:

$$\tilde{P} = \begin{pmatrix} \tilde{B}_0 & A_0 & \mathbf{0} & \mathbf{0} & \dots \\ C_1^T & B_1 & A_1 & \mathbf{0} & \dots \\ \mathbf{0} & C_2^T & B_2 & A_2 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

y asuma que existe una medida matricial μ correspondiente a la matriz de transición P , es decir que se cumplen las condiciones del Teorema 2.1. Además suponga que R_0 es una matriz obtenida en el Teorema 2.1 que satisface (2.4) y (2.5), y que $R_0 \tilde{B}_0 R_0^{-1}$ es simétrica. Entonces existe igualmente una medida matricial $\tilde{\mu}$ correspondiente a la matriz de transición \tilde{P} , y además si μ y $\tilde{\mu}$ son determinadas, están relacionadas mediante la fórmula:

$$\int_{-1}^1 \frac{d\mu(t)}{z-t} = \left[\left(\int_{-1}^1 \frac{d\tilde{\mu}(t)}{z-t} \right)^{-1} - S_0^{-1}(B_0 - \tilde{B}_0) \right]^{-1}. \quad (2.22)$$

La función $\Phi(z, \mu) = \int_{-1}^1 \frac{d\mu(t)}{z-t}$ se le conoce como la **transformada de Stieltjes** de la medida μ .

Demostración. Como la matriz $R_0 B_0 R_0^{-1}$ es simétrica, las matrices P y \tilde{P} difieren únicamente en el bloque de la entrada $(0, 0)$, y por hipótesis la secuencia de matrices R_0, R_1, \dots puede ser usada para satisfacer las ecuaciones del Teorema 2.1. Podemos usar justamente este resultado para concluir que existe una medida matricial correspondiente a la matriz de transición \tilde{P} . Al igual que antes nombremos a $\{Q_j(x)\}_{j \geq 0}$ a la sucesión de polinomios ortogonales generados por la medida matricial μ correspondiente a P , y a $\{\tilde{Q}_j(x)\}_{j \geq 0}$ a la sucesión de polinomios ortogonales, pero generada por la medida matricial $\tilde{\mu}$ correspondiente a \tilde{P} . Como ambos polinomios están generados por la misma recursión (2.3), pero con \tilde{B}_0 en lugar de B_0 para $\{\tilde{Q}_j(x)\}_{j \geq 0}$, restando ambas recursiones obtenemos que los polinomios:

$$R_j(x) = \tilde{Q}_j(x) - Q_j(x),$$

también satisfacen la recursión (2.3), para $n \geq 2$, y con condiciones iniciales $R_0(x) = \mathbf{0}$ y $R_1(x) = A_0^{-1}(B_0 - \tilde{B}_0)$. Estas condiciones iniciales nos hacen pensar que los polinomios $\{R_k(x)\}_{k \geq 0}$ son “proporcionales” a los POM asociados de primera especie, definidos por:

$$Q_n^{(1)}(x) = \int_{-1}^1 \frac{Q_n(x) - Q_n(t)}{x-t} d\mu(t). \quad (2.23)$$

Vale la pena notar que estos nuevos polinomios satisfacen nuevamente la relación de recurrencia definida en (2.3). En efecto, como $Q_0(x) = \mathbf{I}$, claramente $Q_0^{(1)} = \mathbf{0}$, entonces por inducción:

$$\begin{aligned} xQ_n^{(1)}(x) - C_n^T Q_{n-1}^{(1)}(x) - B_n Q_n^{(1)}(x) = \\ x \int_{-1}^1 \frac{Q_n(x) - Q_n(t)}{x-t} d\mu(t) - C_n^T \int_{-1}^1 \frac{Q_{n-1}(x) - Q_{n-1}(t)}{x-t} d\mu(t) \\ - B_n \int_{-1}^1 \frac{Q_n(x) - Q_n(t)}{x-t} d\mu(t), \end{aligned}$$

en donde reordenando y recordando que $\{Q_j(x)\}_{j \geq 0}$ satisfacen (2.3) esto es igual a:

$$\begin{aligned} &= \int_{-1}^1 \frac{A_n Q_{n+1}(x) + tQ_n(t) - A_n Q_{n+1}(t) - xQ_n(t)}{x-t} d\mu(t) \\ &= A_n \int_{-1}^1 \frac{Q_{n+1}(x) - Q_{n+1}(t)}{x-t} d\mu(t) - \int_{-1}^1 \frac{Q_n(t)(x-t)}{x-t} d\mu(t) \\ &= A_n Q_{n+1}^{(1)}(x) - \int_{-1}^1 Q_n(t) d\mu(t) \\ &= A_n Q_{n+1}^{(1)}(x), \end{aligned}$$

donde la última igualdad es por la ortogonalidad de la secuencia $\{Q_j(x)\}_{j \geq 0}$. Y así esta nueva secuencia igual satisface (2.3). Note que los primeros 2 polinomios de esta secuencia son $Q_0^{(1)}(x) = \mathbf{0}$ y $Q_1^{(1)}(x) = A_0^{-1} S_0 = A_0^{-1} (R_0^T R_0)^{-1}$, por ende:

$$R_n(x) = Q_n^{(1)}(x) R_0^T R_0 (B_0 - \tilde{B}_0),$$

se satisface inicialmente para los valores $n = 0, 1$, y para el resto de los naturales igualmente se cumple, pues ambas secuencias $\{Q_j^{(1)}(x)\}_{j \geq 0}$ y $\{R_k(x)\}_{k \geq 0}$ satisfacen (2.3).

Recordando la prueba del Teorema 2.1 sabemos que las secuencias de polinomios $\{R_k Q_k(x)\}_{k \geq 0}$ y $\{R_k \tilde{Q}_k(x)\}_{k \geq 0}$ son ortonormales con respecto las medidas μ y $\tilde{\mu}$, y por ende las secuencias $\{R_k Q_k(x) R_0^{-1}\}_{k \geq 0}$ y $\{R_k \tilde{Q}_k(x) R_0^{-1}\}_{k \geq 0}$ son nuevamente ortonormales pero ahora con respecto las medidas $dm(x) = R_0 d\mu(x) R_0^T$ y $d\tilde{m}(x) = R_0 d\tilde{\mu}(x) R_0^T$, respectivamente. Por lo tanto, aplicando el Teorema de Markov para POM enunciado y demostrado en [13] en la secuencia $\{R_k \tilde{Q}_k(x) R_0^{-1}\}_{k \geq 0}$:

$$\begin{aligned} \int_{-1}^1 \frac{d\tilde{\mu}(t)}{x-t} &= R_0^{-1} \int_{-1}^1 \frac{d\tilde{m}(t)}{x-t} (R_0^T)^{-1} \\ &= \lim_{n \rightarrow \infty} R_0^{-1} (R_n \tilde{Q}_n(x) R_0^{-1})^{-1} (R_n \tilde{Q}_n^{(1)}(x) R_0^T) (R_0^T)^{-1}, \end{aligned}$$

donde esto último se da pues el POM asociado de primera especie a la secuencia $\{R_k \tilde{Q}_k(x) R_0^{-1}\}_{k \geq 0}$ es $R_n \tilde{Q}_n^{(1)}(x) R_0^T$. Siguiendo con la última igualdad, tenemos:

$$\begin{aligned}
 &= \lim_{n \rightarrow \infty} (\tilde{Q}_n(x))^{-1} \tilde{Q}_n^{(1)}(x) \\
 &= \lim_{n \rightarrow \infty} [Q_n(x) + Q_n^{(1)}(x) R_0^T R_0 (B_0 - \tilde{B}_0)]^{-1} Q_n^{(1)}(x) \\
 &= \lim_{n \rightarrow \infty} [Q_n^{(1)}(x) [(Q_n(x))^{-1} Q_n^{(1)}(x)]^{-1} + Q_n^{(1)}(x) R_0^T R_0 (B_0 - \tilde{B}_0)]^{-1} Q_n^{(1)}(x) \\
 &= \lim_{n \rightarrow \infty} [Q_n^{(1)}(x) [(Q_n(x))^{-1} Q_n^{(1)}(x)]^{-1} + R_0^T R_0 (B_0 - \tilde{B}_0)]^{-1} Q_n^{(1)}(x) \\
 &= \lim_{n \rightarrow \infty} [[(Q_n(x))^{-1} Q_n^{(1)}(x)]^{-1} + R_0^T R_0 (B_0 - \tilde{B}_0)]^{-1} \\
 &= \lim_{n \rightarrow \infty} [R_0^T [(R_n Q_n(x) R_0^{-1})^{-1} R_n Q_n^{(1)}(x) R_0^T]^{-1} R_0 + R_0^T R_0 (B_0 - \tilde{B}_0)]^{-1},
 \end{aligned}$$

donde $\tilde{Q}_n^{(1)}(x)$ es el POM asociado de primera especie, obtenido de la sucesión $\{\tilde{Q}_k(x)\}_{k \geq 0}$, y usamos que $\tilde{Q}_n^{(1)}(x) = Q_n^{(1)}(x) \forall n$, pues sabemos que las secuencias $\{Q_k(x)\}_{k \geq 0}$ y $\{\tilde{Q}_k(x)\}_{k \geq 0}$ cumplen (2.3), para $n \geq 1$ y como solo se diferencian por B_0 y \tilde{B}_0 , en los polinomios esto solo se ve reflejado sólomente en el término constante, es decir, $Q_n(x)$ y $\tilde{Q}_n(x)$ poseen los mismos coeficientes matriciales, salvo el constante, en donde aparece B_0 en el primero y \tilde{B}_0 en el segundo. De esto es fácil concluir que: $Q_n(x) - Q_n(t) = \tilde{Q}_n(x) - \tilde{Q}_n(t)$ y por la definición de $R_n(x)$ obtenemos $\tilde{Q}_n^{(1)}(x) = Q_n^{(1)}(x)$. Ahora, regresando a la anterior cadena de igualdades, aplicando nuevamente el Teorema de Markov al primer término de la última igualdad se llega a que:

$$\begin{aligned}
 &= \left[R_0^T \left(\int_{-1}^1 \frac{dm(t)}{x-t} \right)^{-1} R_0 + R_0^T R_0 (B_0 - \tilde{B}_0) \right]^{-1} \\
 &= \left[\left(\int_{-1}^1 \frac{d\mu(t)}{x-t} \right)^{-1} + R_0^T R_0 (B_0 - \tilde{B}_0) \right]^{-1} \\
 &= \left[\left(\int_{-1}^1 \frac{d\mu(t)}{x-t} \right)^{-1} + S_0^{-1} (B_0 - \tilde{B}_0) \right]^{-1},
 \end{aligned}$$

y la última igualdad se da usando $S_0^{-1} = R_0^T R_0$. \square

Es posible extender el resultado anterior para buscar una relación entre la transformada de Stieltjes de μ y la de la medida asociada al proceso \hat{X} , el cual se obtiene al considerar como su matriz de transición a la matriz resultante de descartar la primera fila y columna de bloques de P . Explícitamente, si \hat{P} es esta matriz, tenemos

que:

$$\widehat{P} = \begin{pmatrix} B_1 & A_1 & \mathbf{0} & \mathbf{0} & \dots \\ C_2^T & B_2 & A_2 & \mathbf{0} & \dots \\ \mathbf{0} & C_3^T & B_3 & A_3 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \quad (2.24)$$

Teorema 2.5. *Sea P la matriz de transición P , dada en (2.2) asociada a un proceso cuasi de nacimiento y muerte, y la matriz \widehat{P} dada en (2.24). Asumiendo que se cumplen las condiciones del Teorema 2.1, existe una medida matricial μ correspondiente a la matriz P . Análogamente, como se probará, existe una medida matricial $\widehat{\mu}$ correspondiente a la matriz de transición \widehat{P} . Entonces, se satisface que:*

$$\Phi(z)\Pi_\mu = [z\mathbf{I}_1 - B_0 - A_0\widehat{\Phi}(z)\Pi_{\widehat{\mu}}C_1^T]^{-1}. \quad (2.25)$$

Demostración. Por las mismas razones que en la prueba del teorema anterior, la secuencia de matrices $\{R_k\}_{k \geq 1}$ satisface las condiciones del Teorema 2.1, y por ende, existe una medida matricial $\widehat{\mu}$ correspondiente a la matriz de transición \widehat{P} . Como antes, $\{Q_j(x)\}_{j \geq 0}$ es la sucesión de POM generada por μ , y definimos como $\{\widehat{Q}_j(x)\}_{j \geq 0}$ a la sucesión de POM generada por $\widehat{\mu}$. Análogamente nombramos a $\Phi(z)$ y $\widehat{\Phi}(z)$ como las transformadas de Stieltjes correspondientes a las medidas μ y $\widehat{\mu}$.

Consideramos como en la subsección 1.2.5 al espacio $\mathcal{H}_v = \ell^2(\mathbb{Z}_+, \mathbb{R}^l)$, con su producto interior definido como ahí, así como el operador que define P en este espacio. A este espacio lo escribimos como la suma ortogonal:

$$\ell^2(\mathbb{Z}_+, \mathbb{R}^l) = V_1 \oplus V_2,$$

donde $V_1 = \text{span}\{e_{1,1}, \dots, e_{1,l}\}$ y $V_2 = \text{span}\{e_{2,1}, \dots, e_{2,l}, e_{3,1}, \dots, e_{3,l}, \dots\}$, es decir V_1 es el sub-espacio de \mathcal{H}_v cuyas primeras l entradas pueden diferir de 0 y el resto son 0, y V_2 el sub-espacio de \mathcal{H}_v cuyas primeras l entradas son 0 y el resto pueden diferir de 0 (note, de hecho, que V_2 es isomorfo a \mathcal{H}_v). Así, viendo a P como operador en \mathcal{H}_v , de la misma forma que en la subsección 1.2.5, tenemos que P se descompone como:

$$P = \begin{pmatrix} B_0 & A \\ C & \widehat{P} \end{pmatrix},$$

donde $B_0 : V_1 \rightarrow V_1$, $\widehat{P} : V_2 \rightarrow V_2$,

$$A = \begin{pmatrix} A_0 & \mathbf{0} & \mathbf{0} & \dots \end{pmatrix} : V_2 \rightarrow V_1,$$

y finalmente

$$C = \begin{pmatrix} C_1^T \\ \mathbf{0} \\ \mathbf{0} \\ \vdots \end{pmatrix} : V_1 \rightarrow V_2.$$

Sean, con el fin de diferenciar, \mathbf{I} , \mathbf{I}_1 y \mathbf{I}_2 los operadores identidad en los espacios \mathcal{H}_v , V_1 y V_2 , respectivamente. Entonces, para z adecuada los operadores $(z\mathbf{I} - P)$, y $(z\mathbf{I}_2 - \widehat{P})$ son invertibles, y usando la teoría de complementos de Schur, ver [5], $(z\mathbf{I} - P)^{-1}$ se descompone como:

$$(z\mathbf{I} - P)^{-1} = \begin{pmatrix} [z\mathbf{I}_1 - B_0 - A(z\mathbf{I}_2 - \widehat{P})^{-1}C]^{-1} & * \\ * & * \end{pmatrix}. \quad (2.26)$$

Ahora, por el Teorema de Convergencia Dominada de Lebesgue tenemos que:

$$\Phi(z) = z^{-1} \sum_{n=0}^{\infty} z^{-n} \int x^n d\mu(x) = z^{-1} \sum_{n=0}^{\infty} z^{-n} P_{0,0}^n \Pi_{\mu}^{-1} = (z\mathbf{I} - P)_{0,0}^{-1} \Pi_{\mu}^{-1},$$

donde usamos (2.11) con $i = j = 0$ y $\Pi_{\mu} = (\int d\mu(x))^{-1} = R_0^T R_0$, por (2.10). Análogamente obtenemos que:

$$\widehat{\Phi}(z) = (z\mathbf{I}_2 - \widehat{P})_{0,0}^{-1} \Pi_{\widehat{\mu}}^{-1},$$

donde $\Pi_{\widehat{\mu}} = (\int d\widehat{\mu}(x))^{-1}$. De estas dos últimas ecuaciones, así como de (2.26), tenemos el resultado:

$$\Phi(z)\Pi_{\mu} = [z\mathbf{I}_1 - B_0 - A_0\widehat{\Phi}(z)\Pi_{\widehat{\mu}}C_1^T]^{-1}.$$

□

Llamaremos a un proceso *auto-similar* si ocurre que $X = \widehat{X}$. En este caso, la fórmula anterior nos otorga una ecuación cuadrática en términos de la transformada. Es más, si un proceso no es por sí mismo similar, puede ocurrir que \widehat{X} lo sea, y nuevamente la fórmula puede usarse para encontrar $\Phi(z)$.

En ocasiones, y en consecuencia de la ecuación anterior, es posible obtener, con base en la matriz de transición de un proceso cuasi de nacimiento y muerte, la transformada de Stieltjes asociada a la medida. Dado esto, es importante buscar una relación inversa para obtener la medida partiendo de la transformada. Dicha relación se conoce como *Fórmula de inversión de Perron-Stieltjes*. Presentamos la prueba del caso escalar.

Fórmula de inversión de Perron-Stieltjes. Sea ψ una medida escalar en \mathbb{R} , con

momentos finitos y sea $\Phi(z, \psi)$ su transformada de Stieltjes. Entonces:

$$\int_a^b d\psi(x) + \frac{1}{2}\psi(\{a\}) + \frac{1}{2}\psi(\{b\}) = -\frac{1}{\pi} \lim_{\epsilon \downarrow 0} \int_a^b \Im(\Phi(x + i\epsilon, \psi)) dx, \quad (2.27)$$

donde \Im denota a la parte imaginaria, y $\psi(\{a\}), \psi(\{b\}) \geq 0$ son la magnitud de los pesos en a y b , respectivamente.

Demostración. Primeramente observe que:

$$\begin{aligned} 2i\Im(\Phi(z, \psi)) &= \Phi(z, \psi) - \overline{\Phi(z, \psi)} = \Phi(z, \psi) - \Phi(\bar{z}, \psi) = \int_{\mathbb{R}} \left[\frac{1}{z-x} - \frac{1}{\bar{z}-x} \right] d\psi(x) \\ &= - \int_{\mathbb{R}} \frac{z-\bar{z}}{|x-z|^2} d\psi(x) = -2i \int_{\mathbb{R}} \frac{\Im(z)}{|x-z|^2} d\psi(x). \end{aligned}$$

Por ende sustituyendo $z = x + i\epsilon$, con $x \in \mathbb{R}$, tenemos:

$$\Im(\Phi(x + i\epsilon, \psi)) = - \int_{\mathbb{R}} \frac{\epsilon}{|s - (x + i\epsilon)|^2} d\psi(s) = - \int_{\mathbb{R}} \frac{\epsilon}{(s-x)^2 + \epsilon^2} d\psi(s).$$

Integrando esto último en el intervalo $[a, b]$, e intercambiando las integrales (lo cual es posible pues el integrando es positivo y podemos aplicar el Teorema de Fubini), obtenemos

$$\int_a^b -\Im(\Phi(x + i\epsilon, \psi)) dx = \int_{\mathbb{R}} \left[\int_a^b \frac{\epsilon}{(s-x)^2 + \epsilon^2} dx \right] d\psi(s).$$

La integral entre los corchetes se puede calcular usando el cambio de variable $y = (x-s)/\epsilon$:

$$\chi_{\epsilon}(s) = \int_a^b \frac{\epsilon}{(s-x)^2 + \epsilon^2} dx = \int_{(a-s)/\epsilon}^{(b-s)/\epsilon} \frac{1}{1+y^2} dy = \arctan y \Big|_{y=(a-s)/\epsilon}^{y=(b-s)/\epsilon}.$$

Tenemos que $0 \leq \chi_{\epsilon}(s) \leq \pi$ y cuando tomamos el límite (lo cual es válido por el Teorema de Convergencia Dominada, y $\chi_{\epsilon}(s)$ es acotado y positivo) obtenemos:

$$\lim_{\epsilon \downarrow 0} \chi_{\epsilon}(s) = \begin{cases} \pi, & \text{si } a < s < b, \\ \frac{\pi}{2}, & \text{si } s = a \text{ ó } s = b. \end{cases}$$

La fórmula es obtenida simplemente al integrar $\chi_{\epsilon}(s)$. □

Como consecuencia de esta misma fórmula obtenemos la identidad:

$$\int_a^b d\psi(x) = -\frac{1}{\pi} \lim_{\epsilon \downarrow 0} \lim_{\eta \downarrow 0} \int_{a+\eta}^{b-\nu} \Im(\Phi(x+i\epsilon, \psi)) dx.$$

Si la medida es absolutamente continua respecto la medida de Lebesgue, es decir $d\psi(x) = \psi(x)dx$, abusando de la notación, tenemos:

$$\psi(x) = \frac{1}{\pi} \lim_{\epsilon \downarrow 0} \Im(\Phi(x+i\epsilon, \psi)) = \lim_{\epsilon \downarrow 0} \frac{-\Phi(x+i\epsilon, \psi) + \Phi(x-i\epsilon, \psi)}{2\pi i}.$$

Finalmente, para medidas que se componen de una parte absolutamente continua y otra discreta, existe una forma directa de calcular el tamaño del salto. Para esto, asuma que $\psi = \widehat{\psi} + \psi(\{a\})\delta_a$, donde $\delta_a(x) = \delta(x-a)$ es la distribución delta de Dirac, definida como $\int_{\mathbb{R}} f(x)\delta(x-a)dx = f(a)$. Entonces, como la transformada de Stieltjes es lineal, obtenemos:

$$\Phi(z, \psi) = \Phi(z, \widehat{\psi}) + \frac{\psi(\{a\})}{z-a}.$$

Evaluando en $z = a + i\epsilon$ y tomando partes imaginarias

$$\Im(\Phi(a+i\epsilon, \psi)) = \Im(\Phi(a+i\epsilon, \widehat{\psi})) + \Im\left(\frac{\psi(\{a\})}{i\epsilon}\right) = \Im(\Phi(a+i\epsilon, \widehat{\psi})) - \frac{\psi(\{a\})}{\epsilon}.$$

Por ende, obtenemos

$$\psi(\{a\}) = -\epsilon \Im(\Phi(a+i\epsilon, \psi)) + \epsilon \Im(\Phi(a+i\epsilon, \widehat{\psi})).$$

Tomando límites $\epsilon \downarrow 0$ observamos que $\Phi(a+i\epsilon, \widehat{\psi})$ está acotado por ser $\widehat{\psi}$ absolutamente continua. Por lo tanto los únicos puntos aislados relevantes, es decir $\psi(\{a\}) > 0$, son aquellos que satisfacen:

$$\lim_{\epsilon \downarrow 0} \Im(\Phi(a+i\epsilon, \psi)) = -\infty,$$

mientras que el tamaño del salto en $x = a$ está dado por

$$\psi(\{a\}) = -\epsilon \lim_{\epsilon \downarrow 0} \Im(\Phi(a+i\epsilon, \psi)) \geq 0.$$

La fórmula de inversión, así como la discusión posterior, se generalizan al caso matricial simplemente notando que todas las ecuaciones anteriores son válidas cuando se tiene una medida matricial, pues en este caso la integración es entrada a entrada.

Regresando a los resultados anteriores, cabe la pena recalcar que el Teorema 2.1 y varias de sus consecuencias vistas, se derivan de asumir que tanto las matrices A_n como C_n son invertibles. Veremos ahora que es posible encontrar una forma de relajar esta condición de invertibilidad para las matrices C_n . Para este objetivo, reescribamos las condiciones del Teorema 2.1 como se sigue:

$$C_{n+1}R_{n+1}^T R_{n+1} = R_n^T R_n A_n \quad \forall n \in \mathbb{N}, \quad (2.28)$$

y

$$R_n B_n = E_n R_n \quad \forall n \in \mathbb{N}, \quad (2.29)$$

donde tenemos que asumir que la secuencia de matrices $\{E_k\}_{k \geq 0}$ es simétrica. De hecho, se puede observar que estas condiciones son equivalentes a las encontradas en las ecuaciones (2.5) y (2.4) en el Teorema 2.1. Ahora probaremos que estas condiciones son suficientes para demostrar la existencia de una medida matricial asociada a la matriz de transición P .

Teorema 2.6. *Sea P la matriz de transición de un proceso cuasi de nacimiento y muerte, definida en (2.2), y $\{Q_k(x)\}_{k \geq 0}$ la sucesión de polinomios matriciales asociada a esta, por la relación de recurrencia (2.3). Si además existe una sucesión de matrices de tamaño $l \times l$ dada por $\{R_k\}_{k \geq 0}$ que se satisface las ecuaciones (2.28) y (2.29), entonces existe una medida matricial μ para la cual los polinomios $\{Q_k(x)\}_{k \geq 0}$ forman un conjunto ortogonal.*

Demostración. Al igual que en (2.20) definimos:

$$\Pi_j = R_j^T R_j, \quad j \in \mathbb{N}_0, \quad (2.30)$$

y definimos el siguiente pseudo producto interior con valor matricial por:

$$\langle\langle X, Y \rangle\rangle_{\Pi} = \sum_{j=0}^{\infty} X_j^T \Pi_j Y_j, \quad (2.31)$$

donde $X = (X_0^T, X_1^T, \dots)^T$ y $Y = (Y_0^T, Y_1^T, \dots)^T$ son dos vectores infinitos con valores matriciales, cada $X_i, Y_i \in \mathbb{R}^{l \times l}$ y además X, Y se encuentran en el siguiente espacio:

$$\ell^2(\mathbb{R}^{l \times l}) := \{X = (X_0^T, X_1^T, \dots)^T \mid X_i \in \mathbb{R}^{l \times l}, \langle\langle X, X \rangle\rangle_{\Pi} < \infty\},$$

donde que el producto $\langle\langle X, X \rangle\rangle_{\Pi}$ sea menor que infinito quiere decir que la serie definida converge a una matriz bajo la norma euclídeana. A este mismo espacio lo

podemos equipar con un producto interior escalar definido por:

$$\langle X, Y \rangle_{\Pi, \text{Tr}} = \text{Tr}(\langle\langle X, Y \rangle\rangle_{\Pi}),$$

el cual es un espacio de Hilbert. Note que la matriz P de (2.2) define un operador lineal actuando en el espacio $\ell^2(\mathbb{R}^{l \times l})$, y análogamente define un operador lineal en $\ell^2(\mathbb{R}^l)$, denotados, respectivamente y con la finalidad de poder diferenciarlos, por P y J . Es decir:

$$(PX)_n = A_n X_{n+1} + B_n X_n + C_n^T X_{n-1}, \quad n \in \mathbb{N}, X_{-1} = \mathbf{0}, \quad (2.32)$$

$$(Jx)_n = A_n x_{n+1} + B_n x_n + C_n^T x_{n-1}, \quad n \in \mathbb{N}, x_{-1} = \mathbf{0}, \quad (2.33)$$

son la n -ésima entrada matricial y vectorial, respectivamente. Análogamente al Teorema 2.3 es fácil concluir las siguientes relaciones de simetría:

$$P_{i,j}^T \Pi_i = \Pi_j P_{j,i}. \quad (2.34)$$

De hecho llegar a estas ecuaciones es más sencillo por las relaciones (2.28) y (2.29). Recordando nuevamente la prueba del Teorema 2.3, sabemos que el operador J que actúa en $\ell^2(\mathbb{R}^l)$ es auto-adjunto. Por lo tanto estamos en condiciones de aplicar el Teorema Espectral para Operadores Auto-adjuntos [26, Capítulo 12], y encontrar la única resolución de la identidad asociada a J , y denotada por $(E_\lambda)_\lambda$, y como usualmente se denota $E_\lambda = E((-\infty, \lambda))$, de la forma:

$$J = \int_{\sigma(J)} \lambda dE(\lambda),$$

donde lo anterior es realmente una abreviación a la expresión:

$$\langle Jx, y \rangle_{\Pi} = \int_{\sigma(J)} \lambda dE_{x,y}(\lambda).$$

De esta resolución de la identidad podemos construir un nuevo operador, llamémoslo $(\mathcal{E}_\lambda)_\lambda$, correspondiente al operador P en $\ell^2(\mathbb{R}^{l \times l})$, de la siguiente manera:

$$(\mathcal{E}_\lambda U)x := E_\lambda(Ux), \quad U \in \ell^2(\mathbb{R}^{l \times l}), \quad x \in \mathbb{R}^l.$$

Lo anterior está bien definido pues E_λ es un operador actuando en $\ell^2(\mathbb{R}^l)$, Ux pertenece igualmente a $\ell^2(\mathbb{R}^l)$, y así $(\mathcal{E}_\lambda U)x$ manda un vector x de \mathbb{R}^l a un vector infinito en $\ell^2(\mathbb{R}^l)$, por lo cual $(\mathcal{E}_\lambda U)x$ puede ser visto como un elemento de $\ell^2(\mathbb{R}^{l \times l})$.

Ahora consideremos a los elementos $E^{(j)} = (\mathbf{0}, \dots, \mathbf{0}, \mathbf{I}, \mathbf{0}, \dots)^T \in \ell^2(\mathbb{R}^{l \times l})$, los cuales nombramos como una pseudo base canónica del espacio $\ell^2(\mathbb{R}^{l \times l})$. De hecho bajo el

pseudo producto interior usual es fácil ver que este conjunto es ortonormal en el mismo sentido de la subsección 1.1.3 [3]. Con base en este conjunto de vectores matriciales construimos nuestro candidato a la medida matricial asociada a P , y comprobamos que en efecto lo es. Para este objetivo considere la definición:

$$\Sigma(\lambda) = \langle\langle E^{(0)}, \mathcal{E}_\lambda E^{(0)} \rangle\rangle_{\Pi} \in \mathbb{R}^{l \times l}. \quad (2.35)$$

Para verificar que es la medida matricial asociada a P , seguimos los mismos pasos a los usados en [3]. Por eso, con este objetivo en mente, enunciaremos los dos siguientes lemas tratados en esta misma referencia. Aunque omitiremos las respectivas pruebas invitamos al lector a poder consultarlas. Ambos lemas son el análogo al caso escalar.

Lema 2.7. *Para $j, k = 0, 1, \dots$ fijos, el operador:*

$$\langle\langle E^{(j)}, \mathcal{E}_\lambda E^{(k)} \rangle\rangle_{\Pi}, \quad \lambda \in (-\infty, \infty),$$

es de variación acotada y todas las integrales de la forma:

$$\int_{-\infty}^{\infty} |\lambda|^m |dE_{x,y}(\lambda)|, \quad x, y \in \mathbb{R}^l; m \in \mathbb{N}_0,$$

convergen para este operador, y para el caso $j = k$ este operador es semi-definido positivo.

Lema 2.8. *Suponga que $F(\lambda) = (F_0(\lambda), F_1(\lambda), \dots)$ es una secuencia de operadores con variación acotada tales que:*

$$\int_{-\infty}^{\infty} |\lambda|^m |dE_{x,y}(\lambda)| < \infty \quad (x, y \in \mathbb{R}^l; m \in \mathbb{N}_0), \quad (2.36)$$

$$(PF(\lambda))_i = \int_{-\infty}^{\lambda} t dF_i(t), \quad i = 0, 1, \dots, \quad (2.37)$$

donde en la ecuación (2.37) la integral es matricial y definida análogamente a la sección 1.1. Entonces se cumple la siguiente representación:

$$F_i(\lambda) = \int_{-\infty}^{\lambda} Q_i(t) dF_0(t), \quad i = 0, 1, \dots, \quad (2.38)$$

donde $Q_i(t)$ es el i -ésimo polinomio matricial definido por la recurrencia (2.3).

Regresando al problema, para una k fija consideramos a la siguiente sucesión de funciones $F_j(\lambda) = \langle\langle E^{(j)}, \mathcal{E}_\lambda E^{(k)} \rangle\rangle_{\Pi}$, con $j \in \mathbb{N}$. Por el Lema 2.7 estas funciones satisfacen la condición (2.36) del lema anterior. Comprobamos que igual satisfacen la

condición (2.37). Considere $x, y \in \mathbb{R}^l$, tenemos que:

$$\begin{aligned}
& \langle (PF(\lambda))_j x, y \rangle \\
&= \langle A_j \langle \langle E^{(j+1)}, \mathcal{E}_\lambda E^{(k)} \rangle \rangle_{\Pi} x, y \rangle + \langle B_j \langle \langle E^{(j)}, \mathcal{E}_\lambda E^{(k)} \rangle \rangle_{\Pi} x, y \rangle + \langle C_j^T \langle \langle E^{(j-1)}, \mathcal{E}_\lambda E^{(k)} \rangle \rangle_{\Pi} x, y \rangle \\
&= \langle \langle \langle E^{(j+1)}, \mathcal{E}_\lambda E^{(k)} \rangle \rangle_{\Pi} x, A_j^T y \rangle + \langle \langle \langle E^{(j)}, \mathcal{E}_\lambda E^{(k)} \rangle \rangle_{\Pi} x, B_j y \rangle + \langle \langle \langle E^{(j-1)}, \mathcal{E}_\lambda E^{(k)} \rangle \rangle_{\Pi} x, C_j y \rangle \\
&= \langle (\mathcal{E}_\lambda E^{(k)}) x, E^{(j+1)} A_j^T y \rangle_{\Pi} + \langle (\mathcal{E}_\lambda E^{(k)}) x, E^{(j)} B_j y \rangle_{\Pi} + \langle (\mathcal{E}_\lambda E^{(k)}) x, E^{(j-1)} C_j y \rangle_{\Pi} \\
&= \langle E_\lambda(E^{(k)} x), E^{(j+1)} A_j^T y \rangle_{\Pi} + \langle (E_\lambda(E^{(k)} x), E^{(j)} B_j y) \rangle_{\Pi} + \langle E_\lambda(E^{(k)} x), E^{(j-1)} C_j y \rangle_{\Pi} \\
&= \langle (E_\lambda(E^{(k)} x))_{j+1}, A_j^T y \rangle + \langle (E_\lambda(E^{(k)} x))_j, B_j y \rangle + \langle (E_\lambda(E^{(k)} x))_{j-1}, C_j y \rangle \\
&= \langle A_j (E_\lambda(E^{(k)} x))_{j+1}, y \rangle + \langle B_j (E_\lambda(E^{(k)} x))_j, y \rangle + \langle C_j^T (E_\lambda(E^{(k)} x))_{j-1}, y \rangle \\
&= \langle (P[E_\lambda(E^{(k)} x)])_j, y \rangle = \langle (PE_\lambda(E^{(k)} x))_j, y \rangle = \left\langle \left(\int_{-\infty}^{\lambda} tdE_t(E^{(k)} x) \right)_j, y \right\rangle \\
&= \int_{-\infty}^{\lambda} td \langle (E_t(E^{(k)} x))_j, y \rangle = \int_{-\infty}^{\lambda} td \langle (\mathcal{E}_t E^{(k)}) x, E^{(j)} y \rangle_{\Pi} \\
&= \int_{-\infty}^{\lambda} td \langle \langle \langle E^{(j)}, \mathcal{E}_t E^{(k)} \rangle \rangle_{\Pi} x, y \rangle = \left\langle \left(\int_{-\infty}^{\lambda} tdF_j(t) \right) x, y \right\rangle \\
&\implies (PF(\lambda))_j = \int_{-\infty}^{\lambda} tdF_j(t), \tag{2.39}
\end{aligned}$$

donde $\langle \cdot, \cdot \rangle$ es el producto interior usual en \mathbb{R}^l . Todas las anteriores igualdades emanan esencialmente de (2.35), (2.32), de la identidad, sencilla de demostrar:

$$\langle \langle \langle X, Y \rangle \rangle_{\Pi} x, y \rangle = \langle Y x, X y \rangle_{\Pi},$$

en donde $X, Y \in \ell^2(\mathbb{R}^{l \times l})$ y $x, y \in \mathbb{R}^l$, así como del Teorema Espectral, y del resultado de que el producto escalar no afecta la integral. Por ende, se satisface la condición (2.37).

Ahora podemos aplicar el Lema 2.8 para obtener:

$$\langle \langle E^{(j)}, \mathcal{E}_\lambda E^{(k)} \rangle \rangle_{\Pi} = \int_{-\infty}^{\lambda} Q_j(t) d \langle \langle E^{(0)}, \mathcal{E}_t E^{(k)} \rangle \rangle_{\Pi} \quad \forall j, k \in \mathbb{N}_0.$$

Ahora usando que los operadores \mathcal{E}_λ son auto-adjuntos pues al ser E_λ una resolución de la identidad lo tiene que ser, y por (2.35), \mathcal{E}_λ también lo es y obtenemos:

$$\langle \langle E^{(j)}, \mathcal{E}_\lambda E^{(k)} \rangle \rangle_{\Pi} = \langle \langle \mathcal{E}_\lambda E^{(j)}, E^{(k)} \rangle \rangle_{\Pi}^T = \langle \langle E^{(k)}, \mathcal{E}_\lambda E^{(j)} \rangle \rangle_{\Pi}^T. \tag{2.40}$$

Aplicamos (2.40) a esta última igualdad:

$$\begin{aligned}
 &= \left(\int_{-\infty}^{\lambda} Q_k(t) d\langle\langle E^{(0)}, \mathcal{E}_t E^{(j)} \rangle\rangle_{\Pi} \right)^T = \int_{-\infty}^{\lambda} d\langle\langle E^{(0)}, \mathcal{E}_t E^{(j)} \rangle\rangle_{\Pi}^T Q_k^T(t) \\
 &= \int_{-\infty}^{\lambda} d\langle\langle E^{(j)}, \mathcal{E}_t E^{(0)} \rangle\rangle_{\Pi} Q_k^T(t) \\
 &\implies \langle\langle E^{(j)}, \mathcal{E}_{\lambda} E^{(k)} \rangle\rangle_{\Pi} = \int_{-\infty}^{\lambda} d\langle\langle E^{(j)}, \mathcal{E}_t E^{(0)} \rangle\rangle_{\Pi} Q_k^T(t),
 \end{aligned}$$

lo anterior se cumple para todo $j, k \in \mathbb{N}$. Ahora usando esto último lo sustituimos en (2.40), de la siguiente forma:

$$\begin{aligned}
 \langle\langle E^{(j)}, \mathcal{E}_{\lambda} E^{(k)} \rangle\rangle_{\Pi} &= \int_{-\infty}^{\lambda} Q_j(t) dt \left(\int_{-\infty}^t d\langle\langle E^{(0)}, \mathcal{E}_v E^{(0)} \rangle\rangle_{\Pi} Q_k^T(v) \right) \\
 &= \int_{-\infty}^{\lambda} Q_j(t) d\langle\langle E^{(0)}, \mathcal{E}_t E^{(0)} \rangle\rangle_{\Pi} Q_k^T(t),
 \end{aligned}$$

así considerando la resolución evaluada en toda la recta real obtenemos:

$$\langle\langle E^{(j)}, \mathcal{E}(\mathbb{R}) E^{(k)} \rangle\rangle_{\Pi} = \int_{\mathbb{R}} Q_j(t) d\langle\langle E^{(0)}, \mathcal{E}_t E^{(0)} \rangle\rangle_{\Pi} Q_k^T(t).$$

Recordando que $E(\mathbb{R}) = \mathbf{I}$ y esto a su vez implica $\mathcal{E}(\mathbb{R}) = \mathbf{I}$, por (2.35), y que los elementos $\{E^{(k)}\}_{k \geq 0}$ son un conjunto ortonormal como fue definido en la primera sección obtenemos la identidad buscada:

$$\int_{\mathbb{R}} Q_j(t) d\langle\langle E^{(0)}, \mathcal{E}_t E^{(0)} \rangle\rangle_{\Pi} Q_k^T(t) = \delta_{j,k} \Pi_j.$$

Que usando la notación de (2.74) queda en su forma explícita:

$$\int_{\mathbb{R}} Q_j(t) d\Sigma(t) Q_k^T(t) = \delta_{j,k} \Pi_j, \tag{2.41}$$

concluyendo. □

Nuevamente aplicamos el mismo razonamiento del Teorema 2.2 y encontramos, para estas condiciones, la representación de Karlin-McGregor:

$$\begin{aligned}
 P_{i,j}^n &= \left(\int_{\mathbb{R}} t^n Q_i(t) d\Sigma(t) Q_j^T(t) \right) (\Pi_j)^{-1} \\
 \implies P_{i,j}^n \Pi_j &= \int_{\mathbb{R}} t^n Q_i(t) d\Sigma(t) Q_j^T(t).
 \end{aligned} \tag{2.42}$$

2.3. Propiedades de recurrencia

En esta subsección desarrollaremos equivalencias sobre estados recurrentes basándonos, principalmente, en los resultados obtenidos anteriormente. En particular el Teorema 2.2 será de gran ayuda. Durante el resto de esta subsección asumiremos de forma implícita que se satisfacen las condiciones del Teorema 2.1, y que la medida matricial correspondiente a este proceso tiene su soporte en el intervalo $[-1, 1]$.

Empezamos definiendo la función matricial generadora del bloque (i, j) de nuestra matriz de transición P como:

$$H_{i,j}(z) = \sum_{n=0}^{\infty} (P_{i,j}^n) z^n.$$

Usando el Teorema 2.2, por la representación de Karlin-McGregor, esta última ecuación se puede reescribir como:

$$H_{i,j}(z) = \sum_{n=0}^{\infty} z^n \left(\int_{-1}^1 t^n Q_i(t) d\mu(t) Q_j^T(t) \right) \left(\int_{-1}^1 Q_j(t) d\mu(t) Q_j^T(t) \right)^{-1}.$$

Así reagrupando y factorizando $(\int Q_j(x) d\mu(x) Q_j^T(x))^{-1}$, que no depende de n , obtenemos:

$$H_{i,j}(z) = \left(\sum_{n=0}^{\infty} \int_{-1}^1 (tz)^n Q_i(t) d\mu(t) Q_j^T(t) \right) \left(\int_{-1}^1 Q_j(t) d\mu(t) Q_j^T(t) \right)^{-1}.$$

Sin embargo, aplicando el teorema de Fubini, entrada a entrada, podemos intercambiar la suma y la integral para obtener:

$$\begin{aligned} H_{i,j}(z) &= \left(\int_{-1}^1 \sum_{n=0}^{\infty} (tz)^n Q_i(t) d\mu(t) Q_j^T(t) \right) \left(\int_{-1}^1 Q_j(t) d\mu(t) Q_j^T(t) \right)^{-1} \\ &= \left(\int_{-1}^1 \frac{Q_i(t) d\mu(t) Q_j^T(t)}{1 - tz} \right) \left(\int_{-1}^1 Q_j(t) d\mu(t) Q_j^T(t) \right)^{-1}. \end{aligned}$$

Recordando el resultado conocido sobre las cadenas de Markov, un estado j es recurrente si y solo si la serie $\sum_{n=0}^{\infty} P_{j,j}^n = \infty$, tenemos que fijándonos en los bloques matriciales de la diagonal de la matriz P , y un estado $(i, j) \in \mathcal{C}_i$, este último es

recurrente si y solo si:

$$\begin{aligned} \sum_{n=0}^{\infty} e_j^T P_{i,i}^n e_j &= \lim_{z \rightarrow 1} e_j^T H_{i,i}(z) e_j \\ &= e_j^T \left(\int_{-1}^1 \frac{Q_i(t) d\mu(t) Q_i^T(t)}{1-t} \right) \left(\int_{-1}^1 Q_i(t) d\mu(t) Q_i^T(t) \right)^{-1} e_j = \infty, \end{aligned} \quad (2.43)$$

donde $e_j^T = (0, \dots, 0, 1, 0, \dots, 0)$ es el j -ésimo vector canónico en \mathbb{R}^l . Note incluso que si nuestro proceso es irreducible, basta con fijarnos en la anterior relación para $i = 0$ y obtener que el proceso es recurrente si y solo si:

$$e_j^T \left(\int_{-1}^1 \frac{d\mu(t)}{1-t} S_0^{-1} \right) e_j = \infty, \quad (2.44)$$

para algún $j \in \{1, 2, \dots, l\}$, y en caso de ocurrir, por la irreducibilidad, de hecho se satisface para cualquier $j \in \{1, 2, \dots, l\}$. Lo anterior se resume en el siguiente corolario:

Corolario 2.9. *Suponiendo que se satisfacen las condiciones del Teorema 2.1 para la matriz de transición P definida en (2.2), correspondiente a nuestro proceso cuasi de nacimiento y muerte, cuyos estados radican en \mathcal{C}_l y con una medida matricial correspondiente μ con soporte contenido en el intervalo $[-1, 1]$, entonces un estado $(i, j) \in \mathcal{C}_l$ es recurrente si y solo si la condición (2.43) se cumple, y además bajo lo expuesto anteriormente, si el proceso es irreducible, este es también recurrente si y solo si se satisface (2.44).*

Ahora, el siguiente resultado será una caracterización para el caso de que nuestro proceso, asumiendo que es irreducible, sea positivo recurrente.

Teorema 2.10. *Suponiendo que se satisfacen las condiciones del Teorema 2.1 para la matriz de transición P definida en (2.2), correspondiente a nuestro proceso cuasi de nacimiento y muerte, cuyos estados radican en \mathcal{C}_l y con una medida matricial correspondiente μ con soporte contenido en el intervalo $[-1, 1]$ y si además suponemos que el proceso es irreducible, entonces este será positivo recurrente si y solo si alguna de las medidas matriciales $d\tau_j(t) = e_j^T d\mu(t) S_0^{-1} e_j$, con $j \in \{1, 2, \dots, l\}$, tiene un salto en el punto 1. Análogamente al corolario anterior, si este es el caso entonces todas las medidas poseen un salto en el punto 1.*

Demostración. Como se presentó en el enunciado del teorema, consideramos a las medidas $d\tau_j(t) = e_j^T d\mu(t) S_0^{-1} e_j$. Es fácil ver que en efecto son medidas pues justo como se vio en la prueba de la Proposición 1.1, los elementos diagonales de una matriz

definida positiva son positivos. Anuado a esto, la matriz S_0^{-1} es definida positiva, y por ende $d\mu(t)S_0^{-1}$ es definida positiva y así sus elementos diagonales forman medidas reales. Ahora por la irreducibilidad de nuestro proceso basta solo enfocarse en los estados de la forma $(0, j)$. Por ende, la probabilidad de retorno del estado $(0, j)$ al estado $(0, j)$ en k pasos está dada por:

$$\alpha_k = e_j^T (P_{0,0}^k) e_j = e_j^T \int_{-1}^1 t^k d\mu(t) S_0^{-1} e_j = \int_{-1}^1 t^k d\tau_r(t).$$

Así el proceso es positivo recurrente si y solamente si se cumple que $\alpha = \lim_{k \rightarrow \infty} \alpha_k$ existe y es positivo. Ahora note que si definimos las funciones $f_k(t) = t^k \forall k \in \mathbb{N}$ con dominio en el intervalo $[-1, 1]$, es fácil ver que la subsecuencia de funciones definida, generada cuando $k = 2n$ es par, converge puntualmente a la función $f(t) = \mathbf{1}_{\{t=-1,1\}}(t)$, es decir la función indicadora del conjunto $\{-1, 1\}$. Análogamente, si $k = 2n + 1$ es impar la subsecuencia generada converge a la función $g(t) = \mathbf{1}_{\{t=1\}}(t) - \mathbf{1}_{\{t=-1\}}(t)$, es decir a la función cuyo valor en 1 es 1, en -1 es -1 y es 0 en el restante. Como las integrales de estas funciones claramente están acotadas, por el teorema de Lebesgue de Convergencia Dominada obtenemos:

$$\begin{aligned} \lim_{n \rightarrow \infty} \int_{-1}^1 t^{2n} d\tau_j(t) &= \int_{-1}^1 f(t) d\tau_j(t) = \tau_j(-1) + \tau_j(1) \\ \lim_{n \rightarrow \infty} \int_{-1}^1 t^{2n+1} d\tau_j(t) &= \int_{-1}^1 g(t) d\tau_j(t) = \tau_j(1) - \tau_j(-1). \end{aligned}$$

Por lo tanto si α existe ambos límites deben ser iguales lo que implica claramente que $\tau_j(-1) = 0$, y por ende $\alpha = \tau_j(1) > 0$, y así τ_j posee un salto en $t = 1$. \square

Note de hecho que en el caso de que $S_0 = \mathbf{I}$ las propiedades anteriores están caracterizadas únicamente por los elementos diagonales de nuestra medida matricial μ .

2.4. Momentos canónicos y su aplicación

Representaremos la transformada de Stieltjes de una medida matricial μ asociada de un proceso cuasi de nacimiento y muerte, con soporte en el intervalo $[-1, 1]$, en términos de sus momentos canónicos, como los presentamos en la sección 1.3 del capítulo primero.

Teorema 2.11. *La transformada de Stieltjes de una medida μ asociada a un proceso cuasi de nacimiento y muerte, con soporte en el intervalo $[-1, 1]$, tiene la siguiente*

expansión en términos de fracciones continuas matriciales:

$$\begin{aligned}
 \int_{-1}^1 \frac{d\mu(t)}{z-t} &= \lim_{n \rightarrow \infty} S_0^{1/2} \left\{ z\mathbf{I} + \mathbf{I} - 2\zeta_1^T - \left\{ z\mathbf{I} + \mathbf{I} - 2\zeta_2^T - 2\zeta_3^T - \left\{ z\mathbf{I} + \mathbf{I} - 2\zeta_4^T \right. \right. \right. \\
 &\quad \left. \left. \left. - 2\zeta_5^T - \dots - \left\{ z\mathbf{I} + \mathbf{I} - 2\zeta_{2n}^T - 2\zeta_{2n+1}^T \right\}^{-1} 4\zeta_{2n}^T \zeta_{2n-1}^T \right\}^{-1} \dots \right. \\
 &\quad \left. 4\zeta_4^T \zeta_3^T \right\}^{-1} 4\zeta_2^T \zeta_1^T \left. \right\}^{-1} S_0^{1/2}, \\
 &= \lim_{n \rightarrow \infty} S_0^{1/2} \left\{ (z+1)\mathbf{I} - \left\{ \mathbf{I} - \left\{ (z+1)\mathbf{I} - \dots - \left\{ (z+1)\mathbf{I} - 2\zeta_{2n+1}^T \right\}^{-1} \right. \right. \right. \right. \\
 &\quad \left. \left. \left. \times 2\zeta_{2n}^T \right\}^{-1} \dots 2\zeta_2^T \right\}^{-1} 2\zeta_1^T \right. \left. \right\}^{-1} S_0^{1/2},
 \end{aligned}$$

donde las cantidades ζ_j quedan definidas como en el Teorema 1.23, con base en los momentos canónicos de la subsección 1.3.1. La convergencia anterior es uniforme en subconjuntos compactos de \mathbb{C} con distancia positiva de $[-1, 1]$. En particular, se satisface la siguiente representación:

$$\int_{-1}^1 \frac{d\mu(t)}{1-t} = \frac{1}{2} S_0^{1/2} \left[\mathbf{I} + \sum_{i=0}^{\infty} (V_1^T)^{-1} \dots (V_i^T)^{-1} U_i^T \dots U_1^T \right] S_0^{1/2},$$

donde las secuencias de matrices $\{U_k\}$ y $\{V_k\}$, son los momentos canónicos de μ .

Demostración. Como en la subsección 1.3.2 denotamos al n -ésimo POM mónico, respecto μ como $\underline{P}_n(t)$, y recordamos el Lema 1.22, con $a = 1$ y $b = -1$, para obtener que:

$$\underline{P}_{n+1}(t) = ((t+1)\mathbf{I} - 2\zeta_{2n+1}^T - 2\zeta_{2n}^T) \underline{P}_n(t) - 4\zeta_{2n}^T \zeta_{2n-1}^T \underline{P}_{n-1}(t), \quad (2.45)$$

y $\underline{P}_0(t) = \mathbf{0}$, $\underline{P}_{-1}(t) = \mathbf{I}$, y en este caso usamos $\zeta_j^T = \zeta_j^*$ del Teorema 1.23, por tratarse de una medida real. Ahora consideramos a las cantidades, que son producto de los cálculos del Teorema 1.23 y el Teorema 1.21:

$$\Delta_{2n} := \langle \underline{P}_n, \underline{P}_n \rangle_\mu = (S_{2n} - S_{2n}^-)^T = 2^{2n} (S_0 \zeta_1 \dots \zeta_{2n})^T,$$

que son definidas positivas, ver la discusión posterior a las definiciones 1.95 y 1.96, por lo que por el Lema 1.9, los polinomios:

$$P_n(t) = \Delta_{2n}^{-1/2} \underline{P}_n(t),$$

son ortonormales con respecto a μ . Sustituyendo esta última ecuación en (2.45), tenemos que los polinomios $\{P_n(t)\}$ satisfacen la recurrencia:

$$tP_n(t) = A_{n+1}P_{n+1}(t) + B_nP_n(t) + A_n^T P_{n-1}(t), \quad n \in \mathbb{N}_0, \quad (2.46)$$

cuyas condiciones iniciales son $P_{-1}(x) = \mathbf{0}$ y $P_0(x) = \mathbf{I}$. Además sus coeficientes están dados por:

$$A_{n+1} = \Delta_{2n}^{-1/2} \Delta_{2n+2}^{1/2}, \quad (2.47)$$

$$B_n = -\Delta_{2n}^{-1/2} (\mathbf{I} - 2\zeta_{2n}^T - 2\zeta_{2n+1}^T) \Delta_{2n}^{1/2}, \quad (2.48)$$

$$A_n^T = 4\Delta_{2n}^{-1/2} \zeta_{2n}^T \zeta_{2n-1}^T \Delta_{2n-2}^{1/2}, \quad (2.49)$$

(note que $\Delta_{2n} = 4\Delta_{2n-2}\zeta_{2n-1}\zeta_{2n}$, por lo que las representaciones (2.47) y (2.49) son equivalentes). Si denotamos a $P_n^{(1)}(t)$ al POM asociado de primera especie de $P_n(t)$, y usamos el Teorema 1.2 de [27], obtenemos la representación:

$$\begin{aligned} F_n(z) &:= (P_{n+1}(z))^{-1} P_{n+1}^{(1)}(z) \\ &= S_0 \left\{ z\mathbf{I} - B_0 - A_1 \left\{ z\mathbf{I} - B_1 - A_2 \left\{ z\mathbf{I} - B_1 - \dots \right. \right. \right. \\ &\quad \left. \left. \left. - A_n \left\{ z\mathbf{I} - B_n \right\}^{-1} A_n^T \right\}^{-1} \dots A_2^T \right\}^{-1} A_1^T \right\}^{-1}, \end{aligned}$$

Donde $F_n(z)$ denota al n -ésimo aproximante. Sustituyendo (2.47) – (2.49) en la ecuación anterior se tiene

$$\begin{aligned} F_n(z) &= S_0^{1/2} \left\{ z\mathbf{I} + \mathbf{I} - 2\zeta_1^T - \left\{ z\mathbf{I} + \mathbf{I} - 2\zeta_2^T - 2\zeta_3^T - \left\{ z\mathbf{I} + \mathbf{I} - 2\zeta_4^T \right. \right. \right. \\ &\quad \left. \left. \left. - 2\zeta_5^T - \dots - \left\{ z\mathbf{I} + \mathbf{I} - 2\zeta_{2n}^T - 2\zeta_{2n+1}^T \right\}^{-1} 4\zeta_{2n}^T \zeta_{2n-1}^T \right\}^{-1} \dots \right. \\ &\quad \left. \left. 4\zeta_4^T \zeta_3^T \right\}^{-1} 4\zeta_2^T \zeta_1^T \right\}^{-1} S_0^{-1/2}, \end{aligned}$$

si aplicamos iterativamente la identidad matricial:

$$\mathbf{I} + A^{-1}B = (\mathbf{I} - (B + A)^{-1}B)^{-1},$$

tomando en este caso $A = (z + 1)\mathbf{I} - 2\zeta_{2k}^T - 2\zeta_{2k+1}^T$ y $B = 2\zeta_{2k}^T$, obtenemos:

$$\begin{aligned} F_n(z) &= S_0^{1/2} \left\{ (z + 1)\mathbf{I} - \left\{ \mathbf{I} - \left\{ (z + 1)\mathbf{I} - \dots - \left\{ (z + 1)\mathbf{I} - 2\zeta_{2n+1}^T \right\}^{-1} \right. \right. \right. \\ &\quad \left. \left. \left. \times 2\zeta_{2n}^T \right\}^{-1} \dots 2\zeta_2^T \right\}^{-1} 2\zeta_1^T \right\}^{-1} S_0^{-1/2}, \end{aligned}$$

y aplicando el Teorema de Markov [13, Teorema 1.1] para el caso matricial:

$$\begin{aligned} \int_{-1}^1 \frac{d\mu(x)}{z-t} &= \lim_{n \rightarrow \infty} F_n(z) \\ &= \lim_{n \rightarrow \infty} S_0^{1/2} \left\{ (z+1)\mathbf{I} - \left\{ \mathbf{I} - \left\{ (z+1)\mathbf{I} - \cdots - \left\{ (z+1)\mathbf{I} - 2\zeta_{2n+1}^T \right\}^{-1} \right. \right. \right. \\ &\quad \left. \left. \left. \times 2\zeta_{2n}^T \right\}^{-1} \cdots 2\zeta_2^T \right\}^{-1} 2\zeta_1^T \right\}^{-1} S_0^{-1/2}. \end{aligned}$$

Esto prueba la primera parte del Teorema. Para la segunda parte, sustituimos $z = 1$ en la ecuación anterior, usamos el resultado (1.3) de [15], con las sustituciones adecuadas:

$$\begin{aligned} \int_{-1}^1 \frac{d\mu(x)}{1-x} &= \lim_{n \rightarrow \infty} \frac{1}{2} S_0^{1/2} \left\{ \mathbf{I} - \left\{ \mathbf{I} - \left\{ \mathbf{I} - \cdots - \left\{ \mathbf{I} - \zeta_{2n+1}^T \right\}^{-1} \right. \right. \right. \\ &\quad \left. \left. \left. \times \zeta_{2n}^T \right\}^{-1} \cdots \zeta_2^T \right\}^{-1} \zeta_1^T \right\}^{-1} S_0^{-1/2} \\ &= \lim_{n \rightarrow \infty} \frac{1}{2} S_0^{1/2} \sum_{j=0}^{n+1} X_{j+1}^{-1} \zeta_j^T X_{j-1} X_j^{-1} \zeta_{j-1}^T X_{j-2} X_{j-1}^{-1} \cdots X_1 X_2^{-1} \zeta_1^T S_0^{-1/2}, \end{aligned} \tag{2.50}$$

donde $X_0 = \mathbf{I}$, $X_1 = \mathbf{I}$ y:

$$X_{n+1} = X_n - \zeta_n^T X_{n-1}, \quad (n \geq 1).$$

Note que $X_2 = \mathbf{I} - \zeta_1^T = \mathbf{I} - U_1^T = V_1^T$, e inductivamente si suponemos que $X_n = (V_1 V_2 \cdots V_{n-1})^T$, entonces:

$$\begin{aligned} X_{n+1} &= (V_1 V_2 \cdots V_{n-1})^T - \zeta_n^T (V_1 V_2 \cdots V_{n-2})^T \\ &= (V_1 V_2 \cdots V_{n-1})^T - (V_{n-1} U_n)^T (V_1 V_2 \cdots V_{n-2})^T \\ &= (\mathbf{I} - U_n^T) (V_1 V_2 \cdots V_{n-1})^T \\ &= (V_1 V_2 \cdots V_n)^T. \end{aligned} \tag{2.51}$$

Por inducción lo anterior es válido para toda n , lo cual al sustituir en (2.50) obtenemos la segunda expresión del Teorema. \square

El siguiente resultado es una generalización sobre una famosa recurrencia en un proceso de nacimiento y muerte al caso matricial.

Teorema 2.12. *Asuma, nuevamente, que se satisfacen las condiciones del Teorema 2.1 para la matriz de transición tridiagonal por bloques asociada a un proceso cuasi de nacimiento y muerte, y cuya medida correspondiente tiene su soporte ubicado en*

el intervalo $[-1, 1]$. El estado $(0, p)$ es recurrente si y solo si:

$$e_p^T S_0^{1/2} \sum_{i=0}^{\infty} T_{i+1}^{-1} \zeta_j^T A_i^{-1} C_i^T T_{i-1} T_{i-1}^{-1} A_{i-1}^{-1} C_{i-1}^T T_{i-2} T_{i-1}^{-1} \cdots T_1 T_2^{-1} A_1^{-1} C_1^T T_0 T_1^{-1} A_0^{-1} T_0 S_0^{-1/2} e_p = \infty,$$

donde $T_i = Q_i(1)$, con $i \in \mathbb{N}$, $T_2 = T_1 = \mathbf{I}$, y $Q_i(x)$ denota al i -ésimo polinomio asociado al proceso, generados por la ecuación de recurrencia (2.3). En particular, un proceso irreducible en nuestro espacio de estados \mathcal{C}_1 , es recurrente si y solo si, alguno de los elementos de la diagonal de la matriz:

$$S_0^{1/2} \sum_{i=0}^{\infty} T_{i+1}^{-1} A_i^{-1} C_i^T T_{i-1} T_{i-1}^{-1} A_{i-1}^{-1} C_{i-1}^T T_{i-2} T_{i-1}^{-1} \cdots T_1 T_2^{-1} A_1^{-1} C_1^T T_0 T_1^{-1} A_0^{-1} T_0 S^{-1/2},$$

es infinito (si este es el caso, todos los elementos de la diagonal cumplen con esta propiedad).

Demostración. Utilizando el Corolario 2.9 y el Teorema 2.11 tenemos que un estado $(0, p)$ es recurrente si y solo si:

$$r_p = \frac{1}{2} e_p^T S_0^{1/2} \left[\mathbf{I} + \sum_{i=0}^{\infty} (V_1^T)^{-1} \cdots (V_i^T)^{-1} U_i^T \cdots U_1^T \right] S_0^{-1/2} e_p = \infty, \quad (2.52)$$

donde U_k y V_k son los momentos canónicos de μ . Ahora considerando la recursión (2.3), notamos, mediante una simple inducción, que el coeficiente principal de $Q_n(t)$ está dado por $A_{n-1}^{-1} \cdots A_0^{-1}$ (ver la discusión posterior a la ecuación (2.9)). Por lo tanto el n -ésimo POM mónico es: $\underline{Q}_n(t) = A_0 \cdots A_{n-1} Q_n(t)$ y satisface la recursión:

$$\begin{aligned} \underline{Q}_{n+1}(t) &= t \underline{Q}_n(t) - A_0 \cdots A_{n-1} B_n A_{n-1}^{-1} \cdots A_0^{-1} \underline{Q}_n(t) \\ &\quad - A_0 \cdots A_{n-1} C_n^T A_{n-2}^{-1} \cdots A_0^{-1} \underline{Q}_{n-1}(t). \end{aligned}$$

Comparando con los coeficientes de (2.45) tenemos que:

$$A_0 \cdots A_{n-1} B_n A_{n-1}^{-1} \cdots A_0^{-1} = -\mathbf{I} + 2\zeta_{2n+1}^T + 2\zeta_{2n}^T, \quad (2.53)$$

$$A_0 \cdots A_{n-1} C_n^T A_{n-2}^{-1} \cdots A_0^{-1} = 4\zeta_{2n}^T \zeta_{2n-1}^T. \quad (2.54)$$

Note que, mediante unos sencillos cálculos: $T_1 = Q_1(1) = A_0^{-1}(\mathbf{I} - B_0) = 2A_0^{-1}(\mathbf{I} - U_1^T) = 2A_0^{-1}V_1^T$, por lo tanto si de forma inductiva suponemos que

$$T_n = Q_n(1) = 2^n A_{n-1}^{-1} \cdots A_0^{-1} V_{2n-1}^T \cdots V_1^T,$$

o equivalentemente:

$$A_0 \cdots A_{n-1} T_n = A_0 \cdots A_{n-1} Q_n(1) = \underline{Q}_n(1) V_{2n-1}^T \cdots V_1^T,$$

entonces de (2.53), por inducción, por el Teorema 1.20, (1.95), y de la identidad $U_k = \mathbf{I} - V_k$, tenemos que;

$$\begin{aligned} \underline{Q}_{n+1}(1) &= \underline{Q}_n(1) - (-\mathbf{I} + 2\zeta_{2n+1}^T + 2\zeta_{2n}^T) \underline{Q}_n(1) - 4\zeta_{2n}^T \zeta_{2n-1}^T \underline{Q}_{n-1}(1) \\ &= 2^{n+1} V_{2n-1}^T \cdots V_1^T - 2^{n+1} U_{2n}^T V_{2n-1}^T V_{2n-1}^T \cdots V_1^T - 2^{n+1} U_{2n+1}^T V_{2n}^T \cdots V_1^T \\ &\quad - 2^{n+1} U_{2n}^T U_{2n-1}^T V_{2n-1}^T \cdots V_1^T \\ &= 2^{n+1} (\mathbf{I} - U_{2n}^T V_{2n-1}^T - U_{2n+1}^T V_{2n}^T - U_{2n}^T U_{2n-1}^T) V_{2n-1}^T \cdots V_1^T \\ &= 2^{n+1} (\mathbf{I} - U_{2n+1}^T V_{2n}^T - U_{2n}^T) V_{2n}^T \cdots V_1^T \\ &= 2^{n+1} (\mathbf{I} - U_{2n+1}^T (\mathbf{I} - U_{2n}^T) - U_{2n}^T) V_{2n-1}^T \cdots V_1^T \\ &= 2^{n+1} (\mathbf{I} - U_{2n+1}^T + U_{2n+1}^T U_{2n}^T - U_{2n}^T) V_{2n-1}^T \cdots V_1^T \\ &= 2^{n+1} (V_{2n+1}^T - (\mathbf{I} - U_{2n+1}^T) U_{2n}^T) V_{2n-1}^T \cdots V_1^T \\ &= 2^{n+1} (V_{2n+1}^T - V_{2n+1}^T U_{2n}^T) V_{2n-1}^T \cdots V_1^T \\ &= 2^{n+1} V_{2n+1}^T V_{2n}^T \cdots V_1^T, \end{aligned}$$

concluyendo. Por lo que podemos definir $\widehat{Q}_n(x) = T_n^{-1} Q_n(x)$, y es fácil observar que satisface la siguiente relación de recurrencia:

$$x \widehat{Q}_n(x) = \widehat{A}_n \widehat{Q}_{n+1}(x) + \widehat{B}_n \widehat{Q}_n(x) + \widehat{C}_n^T \widehat{Q}_{n-1}(x), \quad (2.55)$$

donde los coeficientes quedan definidos como:

$$\widehat{A}_n = T_n^{-1} A_n T_{n+1}, \quad \widehat{B}_n = T_n^{-1} B_n T_n, \quad \widehat{C}_n^T = T_n^{-1} C_n^T T_{n-1}. \quad (2.56)$$

(Note que sustituyendo $x = 1$ en (2.55) obtenemos que $\widehat{A}_n + \widehat{B}_n + \widehat{C}_n^T = \mathbf{I}$). Sustituyendo (2.56) en (2.53) – (2.54) obtenemos que:

$$\widehat{A}_0 \cdots \widehat{A}_{n-1} \widehat{B}_n \widehat{A}_{n-1}^{-1} \cdots \widehat{A}_0^{-1} = -\mathbf{I} + 2\zeta_{2n+1}^T + 2\zeta_{2n}^T, \quad (2.57)$$

$$\widehat{A}_0 \cdots \widehat{A}_{n-1} \widehat{C}_n^T \widehat{A}_{n-2}^{-1} \cdots \widehat{A}_0^{-1} = 4\zeta_{2n}^T \zeta_{2n-1}^T. \quad (2.58)$$

Note que usando la identidad $\widehat{A}_n + \widehat{B}_n + \widehat{C}_n^T = \mathbf{I}$:

$$\begin{aligned} \widehat{A}_0 \cdots \widehat{A}_{n-1} \widehat{B}_n \widehat{A}_{n-1}^{-1} \cdots \widehat{A}_0^{-1} &= \mathbf{I} - \widehat{A}_0 \cdots \widehat{A}_{n-1} \widehat{A}_n \widehat{A}_{n-1}^{-1} \cdots \widehat{A}_0^{-1} \\ &\quad - \widehat{A}_0 \cdots \widehat{A}_{n-1} \widehat{C}_n^T \widehat{A}_{n-1}^{-1} \cdots \widehat{A}_0^{-1}, \end{aligned}$$

pero:

$$\begin{aligned} -\mathbf{I} + 2\zeta_{2n+1}^T + 2\zeta_{2n}^T &= -\mathbf{I} + 2U_{2n}^T V_{2n-1}^T + 2U_{2n+1}^T V_{2n}^T \\ &= \mathbf{I} - 2U_{2n}^T U_{2n-1}^T - 2V_{2n+1}^T V_{2n}^T, \end{aligned}$$

por lo que, igualando estas últimas dos ecuaciones:

$$\begin{aligned} 2U_{2n}^T U_{2n-1}^T + 2V_{2n+1}^T V_{2n}^T &= \widehat{A}_0 \cdots \widehat{A}_{n-1} \widehat{A}_n \widehat{A}_{n-1}^{-1} \cdots \widehat{A}_0^{-1} \\ &\quad + \widehat{A}_0 \cdots \widehat{A}_{n-1} \widehat{C}_n^T \widehat{A}_{n-1}^{-1} \cdots \widehat{A}_0^{-1}. \end{aligned} \quad (2.59)$$

Ahora, por un argumento inductivo podemos suponer que $\widehat{A}_0 \cdots \widehat{A}_{n-2} \widehat{A}_{n-1} \widehat{A}_{n-2}^{-1} \cdots \widehat{A}_0^{-1} = 2V_{2n-1}^T V_{2n-2}^T$, por lo que de (2.58):

$$\widehat{A}_0 \cdots \widehat{A}_{n-1} \widehat{C}_n^T \widehat{A}_{n-2}^{-1} \cdots \widehat{A}_0^{-1} = 4U_{2n}^T U_{2n-1}^T V_{2n-1}^T V_{2n-2}^T,$$

por lo que:

$$\widehat{A}_0 \cdots \widehat{A}_{n-1} \widehat{C}_n^T \widehat{A}_{n-1}^{-1} \cdots \widehat{A}_0^{-1} = 2U_{2n}^T U_{2n-1}^T, \quad (2.60)$$

y sustituyendo esto último en (2.59):

$$\widehat{A}_0 \cdots \widehat{A}_{n-1} \widehat{A}_n \widehat{A}_{n-1}^{-1} \cdots \widehat{A}_0^{-1} = 2V_{2n+1}^T V_{2n}^T, \quad (2.61)$$

completando la inducción y haciendo válidas las últimas identidades (2.60) y (2.61) para toda $n \in \mathbb{N}$. Así, usando estos valores en el lado izquierdo de (2.52):

$$\begin{aligned} r_p &= \frac{1}{2} e_p^T S_0^{1/2} \sum_{j=0}^{\infty} \left[(V_1^T)^{-1} \cdots (V_{2j}^T)^{-1} U_{2j}^T \cdots U_1^T \right. \\ &\quad \left. + (V_1^T)^{-1} \cdots (V_{2j+1}^T)^{-1} U_{2j+1}^T \cdots U_1^T \right] S_0^{-1/2} e_p \\ &= \frac{1}{2} e_p^T S_0^{1/2} \sum_{j=0}^{\infty} \widehat{A}_j^{-1} \widehat{C}_j^T \widehat{A}_{j-1}^{-1} \cdots \widehat{A}_1^{-1} \widehat{C}_1^T \widehat{A}_0^{-1} S_0^{-1/2} e_p \\ &= e_p^T S_0^{1/2} \sum_{i=0}^{\infty} T_{i+1}^{-1} A_i^{-1} C_i^T T_{i-1}^{-1} T_{i-1}^{-1} A_{i-1}^{-1} C_{i-1}^T T_{i-2}^{-1} T_{i-1}^{-1} \\ &\quad \cdots T_1 T_2^{-1} A_1^{-1} C_1^T T_0 T_1^{-1} A_0^{-1} T_0 S_0^{-1/2} e_p = \infty, \end{aligned}$$

lo que concluye la demostración. \square

En el caso en el que las matrices T_i , A_i , C_i conmutan, la anterior identidad se re-

duce a:

$$e_p^T S_0^{1/2} \sum_{i=0}^{\infty} T_{i+1}^{-1} T_i^{-1} (C_1 \cdots C_i)^T (A_0 \cdots A_i)^{-1} S_0^{-1/2} e_p = \infty.$$

2.5. Representaciones de la medida invariante

En esta sección nos enfocaremos en obtener diversas expresiones de la medida invariante de nuestro proceso, donde una medida invariante queda definida como en el caso escalar, es decir, es un vector real infinito $\pi = (\pi_0, \pi_1, \dots)$ tal que $\pi P = \pi$. Por los resultados de [21], sabemos que todo proceso cuasi de nacimiento y muerte, irreducible, posee siempre una medida invariante en forma de un producto matricial. En particular, si el proceso es positivo recurrente, la medida invariante coincide con la distribución estacionaria, definida de igual forma al caso escalar, $x = (x_0^T, x_1^T, \dots)$, que puede ser representada como:

$$x_k^T = x_0^T \prod_{i=0}^{k-1} \tilde{R}_i, \quad (2.62)$$

donde el conjunto $\{\tilde{R}_i\}_{i \geq 0}$ es la solución minimal no negativa de las ecuaciones:

$$R_k = A_k + R_k B_{k+1} + R_k R_{k+1} C_{k+2}^T, \quad k = 0, 1, \dots, \quad (2.63)$$

y x_0 satisface:

$$x_0^T (B_0 + \tilde{R}_0 C_1^T) = x_0^T, \quad (2.64)$$

normalizado tal que $x^T \mathbf{1}_\infty = 1$, donde $\mathbf{1}_\infty$ denota al vector infinito cuyas todas entradas son 1. Supondremos ahora que las condiciones del Teorema 2.4 se cumplen para nuestra cadena de Markov, aperiódica e irreducible en \mathcal{C}_l . Por ende, por un resultado clásico, los límites:

$$L'_i = \lim_{n \rightarrow \infty} P_{i,i'}^n,$$

existen y son independientes de la elección de i . Ahora, recordando la ecuación (2.11), y procediendo como en la prueba del Teorema 2.10, tenemos que $x^k \rightarrow \mathbf{1}_{\{x=-1,1\}}(x)$ puntualmente para k par, y $x^k \rightarrow \mathbf{1}_{\{x=1\}}(x) - \mathbf{1}_{\{x=-1\}}(x)$ para k impar. Por ende podemos aplicar el Teorema de convergencia dominada y obtener que:

$$L_{i'} = \lim_{n \rightarrow \infty} [Q_i(1)\mu(1)Q_{i'}^T(1) + (-1)^n Q_i(-1)\mu(-1)Q_{i'}^T(-1)] Z_{i'}^{-1},$$

donde $Z_{i'} = \int_{-1}^1 Q_{i'}(t) d\mu(t) Q_{i'}^T(t)$. Procediendo con la misma idea de la prueba del Teorema 2.10, consideramos las subsecuencias obtenidas para k par e impar y con-

cluimos que $\mu(-1) = \mathbf{0}$. Por lo que esta última ecuación se reduce a:

$$L_{i'} = \lim_{n \rightarrow \infty} P_{i,i'}^n = Q_i(1)\mu(1)Q_{i'}^T(1)Z_{i'}^{-1}. \quad (2.65)$$

Note que el lado izquierdo de esta ecuación es claramente independiente de i , lo que provee de diversas representaciones para la misma cantidad. Por ejemplo, si tomamos $i = 0$ y notamos que el rango de las matrices $L_{i'}$ es 1, obtenemos de la identidad $L_0 = \mu(1)Z_0^{-1}$ que la matriz $\mu(1)$ tiene rango 1. Es más, si el proceso es positivo recurrente, la distribución estacionaria está dada por:

$$x = (x_0^T, x_1^T, \dots) = e_0^T(L_0, L_1, \dots),$$

y de aquí se sigue, igualando entradas, que:

$$x_k^T = e_0^T L_k = e_0^T \mu(1)Q_k^T(1)Z_k^{-1}, \quad k \geq 0,$$

que es una representación alterna a nuestra distribución estacionaria. En particular:

$$x_0^T = e_0^T \mu(1)S_0^{-1}.$$

Si las matrices $Q_k(1)$ son invertibles, mediante un cálculo telescópico, obtenemos que las definiciones siguientes satisfacen (2.62):

$$\tilde{R}_j = Z_j(Q_j^T(1))^{-1}Q_{j+1}^T(1)Z_{j+1}^{-1}.$$

Usando ahora las ecuaciones de (2.5) y (2.4), es fácil observar que la secuencia $\{\tilde{R}_j\}_{j \geq 0}$ es una solución al sistema (2.63) y (2.64). Finalmente, observamos que la matriz $Z_j^{-1} = R_j^T R_j$ puede ser expresada únicamente en términos de los bloques A_j , C_j y S_0 .

Veamos una condición necesaria para la recurrencia positiva. Para ello usamos la identidad de (2.65), para dos valores distintos i, k y obtener la igualdad:

$$Q_i(1)\mu(1)Q_{i'}^T(1) = Q_k(1)\mu(1)Q_{i'}^T(1) = L_{i'},$$

que para cuando $i' = 0$, se reduce a:

$$(Q_i(1) - Q_k(1))\mu(1) = \mathbf{0}, \quad i, k \geq 0.$$

Utilizando el Teorema 2.10, tenemos que el proceso es positivo recurrente si y solamente si las medidas $e_i^T d\mu(x)S_0^{-1}e_i$ poseen un salto en 1. En este caso, se sigue de la última ecuación que las matrices $Q_i(1) - Q_k(1)$ tienen que ser no invertibles para que

el salto exista, lo cual nos otorga el siguiente Teorema:

Teorema 2.13. *Asuma que P definida en (2.2) satisface las condiciones del Teorema 2.4. Si el proceso es positivo recurrente, entonces las matrices:*

$$Q_i(1) - Q_k(1), \quad \forall i, k \in \mathbb{N}_0,$$

son no invertibles.

Por último, en [7], se da explícitamente una expresión para una medida invariante, la cual se presenta en el siguiente y último Teorema de este capítulo.

Teorema 2.14. *Considere a P nuestra matriz por bloques, tridiagonal y de transición que satisface las condiciones del Teorema 2.4. Considere la sucesión de matrices $\{\Pi_j\}_{j \geq 0}$ definidas en (2.20), donde $\Pi_0 = S_0^{-1}$. Considere el siguiente vector fila:*

$$\pi = ((\Pi_0 \mathbf{1}_l)^T, (\Pi_1 \mathbf{1}_l)^T, \dots),$$

donde $\mathbf{1}_l$ es el vector con todas las entradas 1 y de dimensión l . Entonces π es una medida invariante para el proceso P .

Demostración. De (2.20) es fácil ver que las matrices Π_j son semi-definidas positivas, de aquí que las entradas de π sean no negativas. Para probar que $\pi P = \pi$ tenemos que verificar, por la definición de P , que:

$$(\Pi_0 \mathbf{1}_l)^T B_0 + (\Pi_1 \mathbf{1}_l)^T C_1^T = (\Pi_0 \mathbf{1}_l)^T,$$

y que

$$(\Pi_{n+1} \mathbf{1}_l)^T A_{n-1} + (\Pi_n \mathbf{1}_l)^T B_n + (\Pi_{n+1} \mathbf{1}_l)^T C_{n+1}^T = (\Pi_n \mathbf{1}_l)^T, \quad \forall n \geq 1.$$

La primera igualdad es producto de las ecuaciones $\Pi_i P_{i,j} = P_{i,j}^T \Pi_j$, y de que P sea estocástica, ya que:

$$\mathbf{1}_l^T (\Pi_0 B_0 + \Pi_1 C_1^T) = \mathbf{1}_l^T (B_0 \Pi_0 + A_0^T \Pi_0) = [(B_0 + A_0) \mathbf{1}_l]^T \Pi_0 = (\Pi_0 \mathbf{1}_l)^T,$$

pues $(B_0 + A_0) \mathbf{1}_l = \mathbf{1}_l$ por ser P estocástica. Similarmente:

$$\mathbf{1}_l^T (\Pi_{n-1} A_{n-1} + \Pi_n B_n + \Pi_{n+1} C_{n+1}^T) = [(A_n + B_n + C_n) \mathbf{1}_l]^T \Pi_n = (\Pi_n \mathbf{1}_l)^T,$$

concluyendo. □

Capítulo 3

Ejemplos

En este último capítulo presentaremos diversos ejemplos concretos de procesos cuasi de nacimiento y muerte. Empezamos mostrando el ejemplo clásico de este tipo de procesos en los enteros, el cual puede verse como una matriz infinita en ambas direcciones, pero que su matriz de transición puede ser transformada a una matriz de transición por bloques y ser infinita solamente en una dirección, y con esto poder aplicar los resultados vistos. Después presentamos un ejemplo de aplicación en Teoría de Colas, donde encontrar la secuencia matrices correspondientes $\{R_i\}_{i \geq 0}$ es relativamente fácil, pero obtener la medida se complica, incluso para el caso de $l = 2$. Posteriormente, estudiamos dos caminatas aleatorias, la primera en una cuadrícula y la segunda en un árbol infinito y particular conocido en inglés como “Walsh’s Spider” o “Araña de Walsh” en español. Por último, mostramos por encima un ejemplo no trivial proveniente de la teoría de representación de grupos con un mayor número de transiciones entre los estados.

3.1. Cadenas de nacimiento y muerte en \mathbb{Z}

Consideramos una cadena de nacimiento y muerte con espacio de estados en los enteros, donde este queda definido por sus probabilidades de transición: r_i es la probabilidad de permanecer en el estado i , q_i la de avanzar al estado $i + 1$ y p_i la de retroceder al estado $i - 1$, donde $i \in \mathbb{Z}$. Si consideramos que $p_i + q_i + r_i \leq 1$, la desigualdad estricta $p_i + q_i + r_i < 1$ es interpretada como la existencia de un estado i^* en permanente absorción y cuya probabilidad de cualquier estado i de ir a i^* es $1 - p_i - q_i - r_i$. Mediante una biyección entre los enteros y los naturales podemos

encontrar el siguiente mapeo:

$$\psi : \begin{cases} \mathbb{Z} \rightarrow \mathcal{C}_2, \\ i \rightarrow \begin{cases} (i, 1) & \text{si } i \in \mathbb{N}_0, \\ (-i - 1, 2) & \text{otro caso,} \end{cases} \end{cases}$$

el cual tiene su conjunto de estados en \mathcal{C}_2 , y la transición entre la primera y segunda fila solo es posible cuando estamos en el estado $(0, 1)$. La matriz de transición está dada por (2.2) cuyos bloques son de dimensión 2×2 y están definidos como:

$$B_0 = \begin{pmatrix} r_0 & q_0 \\ p_{-1} & r_{-1} \end{pmatrix}, \quad B_n = \begin{pmatrix} r_n & 0 \\ 0 & r_{-n-1} \end{pmatrix}, \\ A_n = \begin{pmatrix} p_n & 0 \\ 0 & q_{-n-1} \end{pmatrix}, \quad C_n^T = \begin{pmatrix} q_n & 0 \\ 0 & p_{-n-1} \end{pmatrix}.$$

Es fácil observar que se satisfacen las condiciones del Teorema 2.1 con la sucesión de matrices $\{R_n\}_{n \geq 0}$ definidas por:

$$R_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{\frac{q_0}{p_{-1}}} \end{pmatrix}, \quad R_n = \begin{pmatrix} \sqrt{\frac{p_0 \cdots p_{n-1}}{q_1 \cdots q_n}} & 0 \\ 0 & \sqrt{\frac{q_0 q_{-1} \cdots q_{-n}}{p_{-1} p_{-2} \cdots p_{-n-1}}} \end{pmatrix} \quad n \geq 1. \quad (3.1)$$

Consecuentemente, por este Teorema, existe una medida matricial μ asociada a este proceso y cuyo soporte está contenido en el intervalo $[-1, 1]$, (véase el Teorema 2.3). Ahora calculamos la transformada de Stieltjes de esta medida usando el Teorema 2.4 y obtenemos:

$$\Phi(z, \mu) = \int_{-1}^1 \frac{d\mu(t)}{z-t} = \left[\tilde{\Phi}(z)^{-1} - R_0^T R_0 (B_0 - \tilde{B}_0) \right]^{-1}.$$

Aquí $\tilde{\Phi}(z)$ es la transformada de Stieltjes asociada a la medida matricial $\tilde{\mu}$ del proceso obtenido de reemplazar B_0 por \tilde{B}_0 en la matriz de transición, (ver (2.48)). \tilde{B}_0 queda definido como:

$$\tilde{B}_0 = \begin{pmatrix} r_0 & 0 \\ 0 & r_{-1} \end{pmatrix}.$$

De esta misma definición obtenemos que $B_0 - \tilde{B}_0$ es:

$$B_0 - \tilde{B}_0 = \begin{pmatrix} 0 & q_0 \\ p_{-1} & 0 \end{pmatrix}.$$

Note que $\tilde{\Phi}(z)$ es una matriz diagonal pues los coeficientes respectivos a esta medida son diagonales, por lo que aplicando el Teorema de Markov, como en la prueba del Teorema 2.4, obtenemos que es una matriz diagonal. Así, si nombramos a $\tilde{\Phi}^+(z)$ y $\tilde{\Phi}^-(z)$ como los componentes diagonales de $\tilde{\Phi}(z)$, respectivamente, de la ecuación (3.1) obtenemos:

$$\begin{aligned} \Phi(z) &= \int_{-1}^1 \frac{d\mu(t)}{z-t} = \begin{pmatrix} \frac{1}{\tilde{\Phi}^+(z)} & -q_0 \\ -q_0 & \frac{1}{\tilde{\Phi}^-(z)} \end{pmatrix}^{-1} \\ &= \frac{1}{1 - q_0^2 \tilde{\Phi}^+(z) \tilde{\Phi}^-(z)} \begin{pmatrix} \tilde{\Phi}^+(z) & q_0 \tilde{\Phi}^+(z) \tilde{\Phi}^-(z) \\ q_0 \tilde{\Phi}^+(z) \tilde{\Phi}^-(z) & \tilde{\Phi}^-(z) \end{pmatrix}. \end{aligned} \quad (3.2)$$

Estudiemos ahora el caso especial en el que $p_i = p$, $q_i = q$ y $r_i = 0$ para todo $i \in \mathbb{Z}$. En [19] se obtiene que:

$$\tilde{\Phi}^+(z) = \frac{z - \sqrt{z^2 - 4pq}}{2pq}, \quad \tilde{\Phi}^-(z) = \frac{p}{q} \tilde{\Phi}^+(z). \quad (3.3)$$

Esta última ecuación igual se puede obtener mediante la fórmula (2.25), pues en este caso el proceso es auto-similar (ver la discusión posterior a (2.25)), y por ende nos otorga:

$$\tilde{\Phi}(z) R_0^T R_0 = [z\mathbf{I} - pq\tilde{\Phi}(z) R_0^T R_0]^{-1}.$$

Multiplicando $z\mathbf{I} - pq\tilde{\Phi}(z) R_0^T R_0$ por la derecha y acomodando obtenemos la ecuación cuadrática:

$$pq(\tilde{\Phi}(z)(R_0^T R_0)^2 - z\tilde{\Phi}(z)R_0^T R_0 + \mathbf{I}) = \mathbf{0}.$$

Usando (3.1) obtenemos las ecuaciones:

$$\begin{aligned} pq(\tilde{\Phi}^+(z))^2 - z\tilde{\Phi}^+(z) + 1 &= 0, \\ pq\left(\frac{q}{p}\tilde{\Phi}^-(z)\right)^2 - z\frac{q}{p}\tilde{\Phi}^-(z) + 1 &= 0. \end{aligned}$$

Lo cual claramente implica (3.3). Sustituyendo (3.3) en (3.2) tenemos que:

$$\Phi(z) = \begin{pmatrix} \frac{1}{\sqrt{z^2 - 4pq}} & \frac{-1}{2q} \left(1 - \frac{z}{\sqrt{z^2 - 4pq}}\right) \\ \frac{-1}{2q} \left(1 - \frac{z}{\sqrt{z^2 - 4pq}}\right) & \frac{p}{q\sqrt{z^2 - 4pq}} \end{pmatrix}. \quad (3.4)$$

De hecho es posible encontrar nuestra medida explícitamente gracias a la fórmula de inversión de Perron-Stieltjes, vista en (2.27), que en el caso matricial es aplicar a cada entrada de la medida matricial μ . Primero aplicamos la inversión a las partes

absolutamente continuas. Para la entrada (1, 1) de $\Phi(z)$ queda como:

$$\begin{aligned}\mu_{1,1}(x) &= -\frac{1}{\pi} \lim_{\epsilon \rightarrow 0} \Im \left(\frac{1}{\sqrt{(x+i\epsilon)^2 - 4pq}} \right) = -\frac{1}{\pi} \lim_{\epsilon \rightarrow 0} \Im \left(\frac{1}{\sqrt{x^2 + 2ix\epsilon - \epsilon^2 - 4pq}} \right) \\ &= -\frac{1}{\pi} \Im \left(\frac{1}{\sqrt{x^2 - 4pq}} \right),\end{aligned}$$

donde \Im indica la parte imaginaria. Para que la última expresión tenga parte imaginaria positiva se requiere que $|x| < 2\sqrt{pq}$, y en este caso tenemos que:

$$d\mu_{1,1}(x) = \frac{1}{\pi\sqrt{4pq - x^2}} dx, \quad |x| < 2\sqrt{pq}.$$

Para la entrada (2, 2) de μ , por (3.3), tenemos que:

$$d\mu_{2,2}(x) = \frac{p}{\pi q\sqrt{4pq - x^2}} dx, \quad |x| < 2\sqrt{pq}.$$

Finalmente para la entrada (1, 2) de μ , que es igual a la entrada (2, 1), aplicando la fórmula de inversión:

$$\begin{aligned}\mu_{1,2}(x) &= -\frac{1}{\pi} \lim_{\epsilon \rightarrow 0} \Im \left(\frac{-1}{2q} \left(1 - \frac{x+i\epsilon}{\sqrt{(x+i\epsilon)^2 - 4pq}} \right) \right) \\ &= \frac{1}{\pi} \lim_{\epsilon \rightarrow 0} \Im \left(\frac{1}{2q} \left(1 - \frac{x+i\epsilon}{\sqrt{x^2 + 2ix\epsilon - \epsilon^2 - 4pq}} \right) \right) \\ &= \frac{1}{\pi} \Im \left(\frac{1}{2q} \left(1 - \frac{x}{\sqrt{x^2 - 4pq}} \right) \right).\end{aligned}$$

Para que la última expresión tenga parte imaginaria positiva se requiere que $|x| < 2\sqrt{pq}$, y tenemos que:

$$d\mu_{1,2}(x) = \frac{x}{2\pi q\sqrt{4pq - x^2}} dx, \quad |x| < 2\sqrt{pq}.$$

Observe que los únicos saltos posibles de nuestras medidas están localizados en $x = \pm 2\sqrt{pq}$. Sin embargo, por la misma fórmula de inversión el tamaño de los saltos es 0. Por ende, la medida matricial está dada por:

$$d\mu(x) = \frac{1}{\pi\sqrt{4pq - x^2}} \begin{pmatrix} 1 & x/2q \\ x/2q & p/q \end{pmatrix} dx, \quad |x| < 2\sqrt{pq}. \quad (3.5)$$

Estudiamos ahora las propiedades de recurrencia de este proceso usando los resultados obtenidos en la subsección 2.3. Para ello, primeramente notamos que el proceso es claramente irreducible, pues es posible llegar de cualquier estado a otro, (pues si estamos en el estado i , podemos avanzar a $i + 1$ o disminuir a $i - 1$ en un paso), en una cantidad finita de pasos. Sabiendo esto y aplicando el Teorema 2.9 con la fórmula explícita de μ (véase (3.5)) tenemos que si $p \neq q$, entonces, por la desigualdad entre la media geométrica y aritmética $2\sqrt{pq} < p + q = 1$. Así, por (3.5), la integral en (2.44) está definida en el intervalo $[-2\sqrt{pq}, 2\sqrt{pq}]$ y es finita, pues este está completamente contenido en $[-1, 1]$. De esta misma observación es fácil ver que cuando $p = q = 1/2$ la integral claramente diverge, por lo que el proceso es transitorio si y solo si $p \neq q$ y recurrente si y solo si $p = q = 1/2$. Como la fórmula de inversión no nos otorgó ningún salto esto implica que el proceso nunca es positivo recurrente por el Teorema 2.10.

Para obtener la medida invariante usamos el Teorema 2.13 de la subsección 2.5. En este caso, por (3.1), tenemos que la expresión buscada de la medida invariante es:

$$\pi = \left(1, \frac{q_0}{p_{-1}}, \frac{p_0}{q_1}, \frac{q_0 q_{-1}}{p_{-1} p_{-2}}, \frac{p_0 p_1}{q_1 q_2}, \frac{q_0 q_{-1} q_{-2}}{p_{-1} p_{-2} p_{-3}}, \dots \right).$$

Veamos ahora cómo obtener los momentos canónicos de la medida μ para el caso en el que $p = q = 1/2$ y veremos una manera alternativa de demostrar que el proceso es recurrente. Para poder encontrar los momentos canónicos de esta medida aplicamos la transformación lineal $x = 2t - 1$ a μ para que el soporte cambie de $[-1, 1]$ a $[0, 1]$ obteniendo la medida $\mu_{[0,1]}$, y multiplicamos por $2\mathbf{I}$ para normalizarla y obtener $2\mu_{[0,1]} = \lambda$. Por (1.91) – (1.93) sabemos que los momentos canónicos de μ y $\mu_{[0,1]}$ son los mismos, y además es fácil comprobar que la multiplicación por escalares, distintos de 0, a la medida preserva los momentos canónicos, por lo que $\mu_{[0,1]}$ y λ también poseen los mismos momentos canónicos. Con base en esto la medida λ queda definida explícitamente como:

$$\lambda(t) = \frac{1}{\pi\sqrt{t(1-t)}} \begin{pmatrix} 1 & 2t-1 \\ 2t-1 & 1 \end{pmatrix}, \quad |t| < 1.$$

Para encontrar los momentos canónicos calculamos primeramente los momentos. Así, usando la notación de la sección 1.3, si S_k representa el k -ésimo momento matricial de λ , entonces

$$(S_k)_{1,1} = \int_0^1 \frac{t^k}{\pi\sqrt{t(1-t)}} dt = \frac{1}{\pi} \int_0^1 \frac{t^{\frac{2k-1}{2}}}{\sqrt{1-t}} dt.$$

Considerando el cambio de variable $t = u^2$ obtenemos que esto último es igual a

$$\frac{2}{\pi} \int_0^1 \frac{u^{2k}}{\sqrt{1-u^2}} du,$$

por lo que realizando la sustitución trigonométrica $u = \sin(v)$:

$$\frac{2}{\pi} \int_0^{\pi/2} \frac{\cos(v) \sin^{2k}(v)}{\sqrt{1-\sin^2(v)}} dv = \frac{2}{\pi} \int_0^{\pi/2} \sin^{2k}(v) dv.$$

Esta última integral se puede obtener integrando por partes, y al evaluar en $v = 0$ y $v = \pi/2$ se obtiene:

$$(S_k)_{1,1} = \binom{2k}{k} \frac{1}{2^{2k}}, \quad k \geq 0.$$

Calculamos $(S_k)_{1,2}$ análogamente:

$$(S_k)_{1,2} = \binom{2k}{k} \frac{k}{2^{2k}(k+1)}, \quad k \geq 0.$$

Por lo que explícitamente el momento k -ésimo viene dado por

$$S_k = \binom{2k}{k} \frac{k}{2^{2k}(k+1)} \begin{pmatrix} k+1 & k \\ k & k+1 \end{pmatrix}, \quad k \geq 0.$$

Con base en esto y por la discusión posterior a (1.60), así como (1.62), tenemos los primeros dos momentos canónicos:

$$U_1 = S_0^{-1} S_1 = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} \end{pmatrix}, \quad U_2 = \begin{pmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{3} \end{pmatrix}.$$

Finalmente por [10] el resto de los momentos canónicos son:

$$U_{2k} = \frac{k}{2k+1} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad U_{2k-1} = \frac{1}{2} \begin{pmatrix} 1 & \frac{1}{2k} \\ \frac{1}{2k} & 1 \end{pmatrix},$$

con $k \in \mathbb{N}$. Ahora note que por (1.60) tenemos que $V_{2k} = \frac{k+1}{2k+1} \mathbf{I}$ y

$$V_{2k+1} = \frac{1}{2} \begin{pmatrix} 1 & -\frac{1}{2k} \\ -\frac{1}{2k} & 1 \end{pmatrix},$$

por lo que sus respectivos inversos son $V_{2k}^{-1} = \frac{2k+1}{k+1} \mathbf{I}$ y $V_{2k+1}^{-1} = \frac{16k^2}{4k^2-1} U_{2k+1}$. Por ende utilizando el Corolario 2.9 y el Teorema 2.11 tenemos que un estado $(0, p)$ es recurrente

si y solo si se satisface (2.52). Calculando los sumandos de la serie de esta ecuación, tenemos que:

$$(V_1^T)^{-1} \dots (V_{2k}^T)^{-1} U_{2k}^T \dots U_1^T = \left(\prod_{j=1}^k \frac{2j+1}{j} \right) \left(\prod_{j=1}^k \frac{16j^2}{4j^2-1} \right) \mathbf{I},$$

para $i = 2k$ en (2.52). Análogamente, tomando $i = 2k + 1$:

$$(V_1^T)^{-1} \dots (V_{2k+1}^T)^{-1} U_{2k+1}^T \dots U_1^T = \left(\prod_{j=1}^k \frac{2j+1}{j} \right) \left(\prod_{j=1}^{k+1} \frac{16j^2}{4j^2-1} \right) \mathbf{I}.$$

Al considerar la serie de (2.52) las entradas diagonales de esta matriz claramente divergen. Así usando el Teorema 2.11 y el Corolario 2.9 tenemos que todos los estados $(0, p)$, de nuestro proceso asociado a μ , son recurrentes.

3.2. Un ejemplo proveniente de Teoría de colas

Veremos ahora un ejemplo en Teoría de Colas, introducido en 2002 en [6]. Para ello consideramos un sistema independiente de colas, donde la primera cola es de la forma $M/M/1$ y la segunda $M/M/1/l-1$. Ambas colas poseen un proceso de Poisson de llegada con parámetros λ_i , $i = 1, 2$ y un servicio de distribución exponencial con parámetros μ_i , $i = 1, 2$. Considerando el proceso cuasi de nacimiento y muerte correspondiente a la longitud de la primera cola, que es no acotado, y la longitud de la segunda cola, que varía entre $0, 1, \dots, l-1$, poseen una matriz de transición por bloques de la forma (2.2), y con bloques A_i , B_i y C_i dados por:

$$B_0 = \begin{pmatrix} 0 & \frac{\lambda_2}{\lambda_1 + \lambda_2} & 0 & \dots & 0 \\ \frac{\mu_2}{\gamma - \mu_1} & 0 & \frac{\lambda_2}{\gamma - \mu_1} & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & \frac{\mu_2}{\gamma - \mu_1} & 0 & \frac{\lambda_2}{\gamma - \mu_1} \\ 0 & \dots & 0 & \frac{\mu_2}{\lambda_1 + \mu_2} & 0 \end{pmatrix}, \quad B_i = \begin{pmatrix} 0 & \frac{\lambda_2}{\gamma - \mu_2} & 0 & \dots & 0 \\ \frac{\mu_2}{\gamma} & 0 & \frac{\lambda_2}{\gamma} & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & \frac{\mu_2}{\gamma} & 0 & \frac{\lambda_2}{\gamma} \\ 0 & \dots & 0 & \frac{\mu_2}{\gamma - \mu_2} & 0 \end{pmatrix}, \quad i \geq 1,$$

$$A_0 = \begin{pmatrix} \frac{\lambda_1}{\lambda_1 + \lambda_2} & 0 & \dots & 0 & 0 \\ 0 & \frac{\lambda_1}{\gamma - \mu_1} & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 0 & \frac{\lambda_1}{\gamma - \mu_1} & 0 \\ 0 & \dots & 0 & 0 & \frac{\lambda_1}{\lambda_1 + \mu_2} \end{pmatrix}, \quad A_i = \begin{pmatrix} \frac{\lambda_1}{\gamma - \mu_2} & 0 & \dots & 0 & 0 \\ 0 & \frac{\lambda_1}{\gamma} & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 0 & \frac{\lambda_1}{\gamma} & 0 \\ 0 & \dots & 0 & 0 & \frac{\lambda_1}{\gamma - \lambda_2} \end{pmatrix}, \quad i \geq 1,$$

y finalmente,

$$C_i = \begin{pmatrix} \frac{\mu_1}{\gamma - \mu_2} & 0 & \cdots & 0 & 0 \\ 0 & \frac{\mu_1}{\gamma} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & \frac{\mu_1}{\gamma} & 0 \\ 0 & \cdots & 0 & 0 & \frac{\mu_1}{\gamma - \lambda_2} \end{pmatrix}, i \geq 1,$$

respectivamente, donde $\gamma = \lambda_1 + \lambda_2 + \mu_1 + \mu_2$ y $\lambda_1 < \mu_2$. Recordando nuevamente el Teorema 2.1, si se definen a las matrices $\{R_i\}_{i \geq 0}$ como:

$$R_0 = \text{diag} \left(\frac{\sqrt{(\lambda_1 + \lambda_2)\mu_2}}{\sqrt{(\gamma - \mu_1)\lambda_2}}, 1, \frac{\sqrt{\lambda_2}}{\sqrt{\mu_2}}, \frac{\lambda_2}{\mu_2}, \dots, \left(\frac{\sqrt{\lambda_2}}{\sqrt{\mu_2}} \right)^{l-3}, \frac{\sqrt{(\lambda_1 + \mu_2)\lambda_2^{l-2}}}{\sqrt{(\gamma - \mu_1)\mu_2^{l-2}}} \right),$$

$$R_1 = \text{diag} \left(\frac{\sqrt{\lambda_1(\gamma - \mu_2)\mu_2}}{\sqrt{\lambda_2(\gamma - \mu_1)\mu_1}}, \frac{\sqrt{\gamma\lambda_1}}{\sqrt{(\gamma - \mu_1)\mu_1}}, \dots, \frac{\sqrt{\gamma\lambda_1\lambda_2^{l-3}}}{\sqrt{(\gamma - \mu_1)\mu_1\mu_2^{l-3}}}, \frac{\sqrt{\lambda_1(\gamma - \lambda_2)\lambda_2^{l-2}}}{\sqrt{\mu_1(\gamma - \mu_1)\mu_2^{l-3}}} \right),$$

y finalmente,

$$R_i = \left(\frac{\sqrt{\lambda_1}}{\sqrt{\mu_1}} \right)^{i-1} R_1, \quad i \geq 2,$$

se satisfacen las condiciones del Teorema 2.3 y la medida correspondiente al proceso tiene su soporte contenido en $[-1, 1]$.

Estudiamos, finalmente, la medida invariante asociada a este proceso. Para ello, nuevamente hacemos uso del Teorema 2.13 de la subsección 2.5. Considerando a la siguiente sucesión de vectores horizontales de dimensión l :

$$v_0 = \left(\frac{(\lambda_1 + \lambda_2)\mu_2}{(\gamma - \mu_1)\lambda_2}, 1, \frac{\lambda_2}{\mu_2}, \frac{\lambda_2^2}{\mu_2^2}, \dots, \left(\frac{\lambda_2}{\mu_2} \right)^{l-3}, \frac{(\lambda_1 + \mu_2)\lambda_2^{l-2}}{(\gamma - \mu_1)\mu_2^{l-2}} \right)$$

$$v_1 = \left(\frac{\lambda_1(\gamma - \mu_2)\mu_2}{\lambda_2(\gamma - \mu_1)\mu_1}, \frac{\gamma\lambda_1}{(\gamma - \mu_1)\mu_1}, \dots, \frac{\gamma\lambda_1\lambda_2^{l-3}}{(\gamma - \mu_1)\mu_1\mu_2^{l-3}}, \frac{\lambda_1(\gamma - \lambda_2)\lambda_2^{l-2}}{\mu_1(\gamma - \mu_1)\mu_2^{l-3}} \right),$$

y para $i \geq 2$ definimos $v_i = \left(\frac{\lambda_1}{\mu_1} \right)^{i-1} v_1$. Con base en esto nuestra medida invariante queda expresada, en términos de estas previas definiciones, de la forma:

$$\pi = (v_0; v_1; v_2; \dots).$$

3.3. Caminata aleatoria en una cuadrícula

Consideramos ahora la medida al proceso cuasi de nacimiento y muerte en \mathcal{C}_l , donde las probabilidades de ir de (i, j) a $(i, j+1)$, $(i, j-1)$, $(i-1, j)$ y $(i+1, j)$ están dadas por los reales positivos u, v, q, r , respectivamente, y $u + v + q + r = 1$. En este caso, se sigue que $A_i = r\mathbf{I}$, para todo $i \geq 0$, $C_i = q\mathbf{I}$, para $i \geq 1$, y:

$$B_i = \begin{pmatrix} 0 & u & 0 & 0 & \dots & 0 \\ v & 0 & u & 0 & \dots & 0 \\ 0 & v & 0 & u & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & v & 0 & u \\ 0 & \dots & 0 & 0 & v & 0 \end{pmatrix}, i \geq 0. \quad (3.6)$$

Nuevamente definiendo a las matrices $\{R_i\}_{i \geq 0}$ como:

$$R_0 = \text{diag} \left(1, \sqrt{\frac{u}{v}}, \sqrt{\frac{u^2}{v^2}}, \dots, \sqrt{\frac{u^{l-1}}{v^{l-1}}} \right), \quad R_i = \left(\sqrt{\frac{r}{q}} \right)^i R_0, \quad i \geq 1,$$

se satisfacen las condiciones del Teorema 2.1. Del Teorema 2.3 obtenemos que la medida matricial correspondiente a este proceso tiene su soporte contenido en el intervalo $[-1, 1]$. De la anterior definición de R_i y de las ecuaciones (2.13) tenemos a los coeficientes constantes $D = D_n = \sqrt{qr}\mathbf{I}$, y:

$$E = E_n = \sqrt{uv} \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 1 & 0 & 1 & 0 & \dots & 0 \\ 0 & 1 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 0 & 1 \\ 0 & \dots & 0 & 0 & 1 & 0 \end{pmatrix}. \quad (3.7)$$

Ahora usando los resultados obtenidos en [14], se sigue que la transformada de Stieltjes de nuestra medida esta dada por:

$$\int_{-1}^1 \frac{d\mu(t)}{z-t} = \frac{1}{2qr} \left[z\mathbf{I} - E - [(z\mathbf{I} - E)^2 - 4qr\mathbf{I}]^{1/2} \right].$$

De este mismo artículo obtenemos que el soporte de la medida viene dado por el conjunto:

$$\text{supp}(\mu) = \{x \in \mathbb{R} \mid x\mathbf{I} - E \text{ tiene un valor propio en } [-2\sqrt{qr}, 2\sqrt{qr}]\}. \quad (3.8)$$

Es bien conocido, ver [2], que los valores propios de la matriz E en (3.7) son:

$$2\sqrt{uv} \cos\left(\frac{j\pi}{l+1}\right), \quad j = 1, \dots, l,$$

y con vectores propios normalizados y correspondientes:

$$x_j^T = \sqrt{\frac{2}{l+1}} \left(\sin\left(\frac{j\pi}{l+1}\right), \sin\left(2\frac{j\pi}{l+1}\right), \dots, \sin\left(l\frac{j\pi}{l+1}\right) \right).$$

Por ende, se sigue de (3.8) y esto último que:

$$\text{supp}(\mu) = \left[-2\sqrt{qr} + 2\sqrt{uv} \cos\left(\frac{l\pi}{l+1}\right), 2\sqrt{qr} + 2\sqrt{uv} \cos\left(\frac{\pi}{l+1}\right) \right], \quad (3.9)$$

(note que $\text{supp}(\mu) \subset [-1, 1]$). Para poder calcular la medida matricial μ , tenemos primero que determinar la descomposición espectral de la matriz:

$$\begin{aligned} -H(x) &= 4\mathbf{I} - D^{-1/2}(x\mathbf{I} - E)D^{-1}(x\mathbf{I} - E)D^{-1/2} \\ &= \frac{1}{qr} [4qr\mathbf{I} - (x\mathbf{I} - E)^2]. \end{aligned}$$

Es fácil ver que al conocer los valores propios de E , los de esta matriz son:

$$\lambda_j(x) = \frac{1}{qr} \left[4qr - \left(x - 2\sqrt{uv} \cos\left(\frac{j\pi}{l+1}\right) \right)^2 \right],$$

y nuevamente por [14], el peso de la medida matricial está dada por:

$$d\mu(x) = \frac{1}{2\pi\sqrt{qr}} U \Lambda(x) U^T dx,$$

donde la matriz $\Lambda(x)$ es definida como:

$$\Lambda(x) = [\text{diag}(\text{máx}(\lambda_1(x), 0), \dots, \text{máx}(\lambda_d(x), 0))]^{1/2},$$

y la (i, j) -ésima entrada de la matriz U está dada por:

$$u_{i,j} = \sqrt{\frac{2}{l+1}} \sin\left(j\frac{i\pi}{l+1}\right).$$

Ahora para estudiar las propiedades de recurrencia, con base en la medida asociada de este proceso, hacemos uso del Teorema 2.9. Siguiendo un razonamiento similar al Ejemplo 3.1, y empleando (3.9), veremos en qué casos ocurre que $\text{supp}(\mu) = [-1, 1]$,

para poder distinguir cuándo el proceso es transitorio y cuándo recurrente. Por ende, usando la desigualdad entre las medias aritmética y geométrica, así como el hecho de que $\cos\left(\frac{l\pi}{l+1}\right) \leq 1$, tenemos que $2\sqrt{qr} + 2\sqrt{uv} \cos\left(\frac{l\pi}{l+1}\right) \leq 2\sqrt{qr} + 2\sqrt{uv} \leq q + r + u + v = 1$, donde la igualdad, asumiendo que todos son positivos, se da cuando $q = r$, $u = v$ y $\cos\left(\frac{l\pi}{l+1}\right) = 1$. Sin embargo, esta última condición no es posible pues l es un entero positivo. Por lo tanto, forzosamente debe ocurrir que $u = v = 0$ y $q = r = 1/2$, y cualquier otra elección de u, v, q, r forzosamente debe satisfacer que $\text{supp}(\mu) \subset (-1, 1)$. Así, esto último implica que el proceso es transitorio en cualquier caso, salvo $u = v = 0$ y $q = r = 1/2$. Falta ver si el proceso es recurrente en esta situación. Para ello analizamos la medida asociada y notamos que en este caso $\lambda_j(x) = 4(1 - x^2)$, así usando el Teorema 2.9 queda claro que el proceso sí es recurrente en este caso. Al igual que en el Ejemplo 3.1, usando el Teorema 2.10, la medida al no poseer ningún salto hace imposible que el proceso sea recurrente positivo.

Para obtener la medida invariante procedemos a usar el Teorema 2.13. Por la definición de las matrices R_i que se da al inicio del Ejemplo, si definimos a la sucesión de vectores horizontales de dimensión l , $\{v_i\}_{i \geq 0}$ de la siguiente forma:

$$v_0 = \left(1, \frac{u}{v}, \frac{u^2}{v^2}, \dots, \frac{u^{l-1}}{v^{l-1}}\right), \quad v_i = \left(\frac{r}{q}\right)^i v_0, \quad i \geq 1,$$

entonces por el Teorema 2.13 nuestra medida invariante está dada por

$$\pi = (v_0; v_1; v_2; \dots).$$

Un caso particular de la anterior caminata es considerar el caso finito. Es fácil ver que los resultados demostrados en las secciones anteriores permanecen válidos para procesos en espacios de estados finitos. Como un ejemplo, consideramos un proceso en:

$$\mathcal{C}_{l,N} = \{(i, j) \in \mathbb{N}_0 \times \mathbb{N} \mid 0 \leq i \leq N - 1, 1 \leq j \leq l\},$$

donde, nuevamente, las probabilidades de ir de (i, j) a $(i, j + 1)$, $(i, j - 1)$, $(i - 1, j)$ y $(i + 1, j)$ están dadas por los reales positivos u, v, q, r , respectivamente, y $u + v + l + r =$

1. Entonces, la matriz de transición por bloques está dada por:

$$P = \begin{pmatrix} B_0 & A_0 & 0 & \dots & 0 \\ C_1^T & B_1 & A_1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & C_{N-2}^T & B_{N-2} & A_{N-2} \\ 0 & \dots & 0 & C_{N-1}^T & B_{N-1} \end{pmatrix},$$

donde $A_i = r\mathbf{I}$, con $0 \leq i \leq N-2$, $C_i = q\mathbf{I}$, con $1 \leq i \leq N-1$, y las matrices $B = B_i$ quedan definidas como en (3.6). Para encontrar los POM asociados a P , empezamos por notar que $Q_1(x) = \frac{1}{r}(x\mathbf{I} - B)$, dado que $Q_{-1}(x) = \mathbf{0}$ y $Q_0(x) = \mathbf{I}$, así si nombramos $A = \frac{1}{r}(x\mathbf{I} - B)$ y usando (2.3) tenemos que:

$$Q_{n+1}(x) = AQ_n(x) - \frac{q}{r}Q_{n-1}(x). \quad (3.10)$$

Antes de seguir, nombramos como $U_n(z)$ al n -ésimo polinomio de Chebyshev de segundo tipo. Es bien conocido que $U_0(z) = \mathbf{I} = Q_0(x)$ y si hacemos la sustitución $z = \frac{1}{2}\sqrt{\frac{r}{q}}A$, tenemos que $\sqrt{\frac{q}{r}}U_1(z) = 2\sqrt{\frac{q}{r}}z = A = Q_1(x)$. Así, si asumimos por inducción, que:

$$Q_k(x) = \left(\sqrt{\frac{q}{r}}\right)^k U_k\left(\frac{1}{2}\sqrt{\frac{r}{q}}A\right), \quad (3.11)$$

se cumple para $0 \leq k \leq n$, entonces para $k = n+1$, tenemos por (3.10), (3.11) y la conocida recurrencia de los polinomios $U_n(z)$:

$$\begin{aligned} Q_{n+1}(x) &= AQ_n(x) - \frac{q}{r}Q_{n-1}(x) \\ &= A \left(\sqrt{\frac{q}{r}}\right)^n U_n\left(\frac{1}{2}\sqrt{\frac{r}{q}}A\right) - \left(\sqrt{\frac{q}{r}}\right)^{n+1} U_{n-1}\left(\frac{1}{2}\sqrt{\frac{r}{q}}A\right) \\ &= \left(\sqrt{\frac{q}{r}}\right)^{n+1} \left[2\frac{1}{2}\sqrt{\frac{r}{q}}AU_n\left(\frac{1}{2}\sqrt{\frac{r}{q}}A\right) - U_{n-1}\left(\frac{1}{2}\sqrt{\frac{r}{q}}A\right)\right] \\ &= \left(\sqrt{\frac{q}{r}}\right)^{n+1} U_{n+1}\left(\frac{1}{2}\sqrt{\frac{r}{q}}A\right), \end{aligned}$$

como se pedía, haciendo (3.11) válido para todo k natural. Recordando las raíces de los polinomios de Chebyshev de segundo tipo y usando que $\det(A) = \det\frac{1}{r}(x\mathbf{I} - B)$, que es el polinomio característico de B , obtenemos:

$$U_N\left(\frac{z}{2}\right) = \prod_{j=1}^N \left(z - 2\cos\left(\frac{j\pi}{N+1}\right)\right), \quad \det A = \left(\frac{\sqrt{uv}}{r}\right)^l U_l\left(\frac{1}{2\sqrt{uv}}x\right). \quad (3.12)$$

Por ende, los ceros del polinomio $Q_N(x)$, que son las raíces escalares de $\det Q_N(x)$, están dados por los ceros de $U_N\left(\frac{1}{2}\sqrt{\frac{r}{q}}A\right)$. Por (3.11), y por lo anterior tenemos que los ceros de este último polinomio son:

$$\det U_N\left(\frac{1}{2}\sqrt{\frac{r}{q}}A\right) = \prod_{j=1}^N \det\left(\sqrt{\frac{r}{q}}A - 2\cos\left(\frac{j\pi}{N+1}\right)\right).$$

Los ceros están determinados por las raíces de cada uno de los elementos del producto, extrapolarlo:

$$\begin{aligned} \det\left(\sqrt{\frac{r}{q}}A - 2\cos\left(\frac{j\pi}{N+1}\right)\right) &= \det\left(\frac{1}{\sqrt{qr}}(x\mathbf{I} - B) - 2\cos\left(\frac{j\pi}{N+1}\right)\right) \\ &= \left(\frac{1}{\sqrt{qr}}\right)^d \det\left(\left(x - 2\sqrt{qr}\cos\left(\frac{j\pi}{N+1}\right)\right)\mathbf{I} - B\right). \end{aligned}$$

Como $\det A = \det(x\mathbf{I} - B)$, el polinomio característico de B , que se calculó en (3.12), tenemos que las raíces de esto último están dadas por:

$$\lambda_{i,j} = 2\left(\sqrt{uv}\cos\left(\frac{i\pi}{l+1}\right) + \sqrt{rl}\cos\left(\frac{j\pi}{N+1}\right)\right),$$

con $1 \leq i \leq l$ y $1 \leq j \leq N$.

3.4. Caminata aleatoria en un árbol

Considere el proceso inducido por la gráfica generada por l rayos infinitos conectados de un punto, llamado el origen, conocida como “araña de Walsh” (“Walsh’s spider”). En cada rayo la probabilidad de alejarnos del origen en distancia 1 es p , y la de acercarnos a este es q , donde $p + q = 1$. Del origen, la probabilidad de ir al i -ésimo rayo es $x_i > 0$, ($i = 1, \dots, l$) (ver la Figura 3.1, en el caso $l = 5$). Es fácil observar que este proceso tiene como estados a \mathcal{C}_l , y con matriz de transición P como en (2.2), donde $B_i = \mathbf{0}$, si $i \geq 1$, $A_0 = \text{diag}(x_1, p, \dots, p)$, $A_i = p\mathbf{I}$ para $i \geq 1$, y

$$B_0 = \begin{pmatrix} 0 & x_2 & \dots & x_l \\ q & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ q & 0 & \dots & 0 \end{pmatrix}, \quad C_0 = \begin{pmatrix} x_1 & 0 & \dots & 0 \\ 0 & p & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & p \end{pmatrix},$$

y $C_i = p\mathbf{I}$, $i \geq 1$, donde $\sum_{i=1}^l x_i = 1$. Claramente este conjunto de matrices satisface las hipótesis del Teorema 2.1 con la secuencia de matrices $\{R_j\}_{j \geq 0}$:

$$R_0 = \text{diag} \left(1, \sqrt{\frac{x_2}{q}}, \dots, \sqrt{\frac{x_l}{q}} \right), \quad R_1 = \text{diag} \left(\sqrt{\frac{x_1}{q}}, \sqrt{\frac{x_2 p}{q^2}}, \dots, \sqrt{\frac{x_l p}{q^2}} \right),$$

y

$$R_i = \left(\sqrt{\frac{p}{q}} \right)^{i-1} R_1, \quad i \geq 2.$$

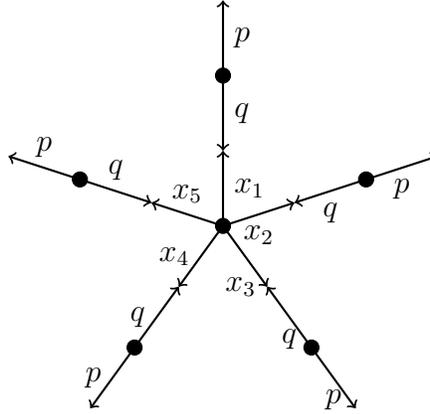


Figura 3.1: El proceso en el caso que $l = 5$.

Para encontrar la medida asociada a nuestro proceso, consideramos a $\tilde{B}_0 = \mathbf{0}$ y al proceso \tilde{X} que este induce con su matriz de transición \tilde{P} , como en el Teorema 2.4. Es fácil ver que este nuevo proceso es auto-similar y por ende podemos usar la fórmula de (2.25) para encontrar la transformada de Stieltjes de la medida $\tilde{\mu}$ asociada a \tilde{X} . Para esto note que la secuencia de matrices $\{R_j\}_{j \geq 0}$ definida como $R_0 = \mathbf{I}$ y $R_j = \left(\sqrt{\frac{p}{q}} \right)^j \mathbf{I}$ satisface las condiciones del Teorema 2.1, por lo que usando esto para aplicar (2.25) en $\tilde{\Phi}$, que es la transformada de $\tilde{\mu}$

$$\tilde{\Phi}(z) = [z\mathbf{I} - pq\tilde{\Phi}(z)]^{-1}.$$

Esta última ecuación es fácilmente reducida a la ecuación cuadrática

$$pq\tilde{\Phi}(z)^2 - z\tilde{\Phi}(z) + \mathbf{I} = \mathbf{0}.$$

Cuya solución es similar a la obtenida en el Ejemplo 3.1 y está dada por:

$$\tilde{\Phi}(z) = \left(\frac{z - \sqrt{z^2 - 4pq}}{2pq} \right) \mathbf{I}.$$

Con base en esto podemos usar el Teorema 2.4 para obtener la transformada de Stieltjes asociada a μ y una vez obtenida la transformada podemos aplicar (2.27) para calcular la medida. Omitiremos los cálculos de la obtención de la medida, pero presentamos explícitamente esta como en [16]. Apoyándonos de las matrices $E_{i,j}$ de tamaño $l \times l$, definidas como en la prueba del Lema 1.4, donde su única entrada no cero está en (i, j) y tiene valor 1, consideramos a la siguiente matriz de tamaño $l \times l$:

$$\begin{aligned} M(x) = & \frac{\sqrt{4pq - x^2}}{1 - x^2} \left\{ \left(E_{1,1} + \sum_{2 \leq i,j \leq l} E_{i,j} \right) + x \sum_{i=2}^l (E_{1,i} + E_{i,1}) \right\} \\ & + \sqrt{4pq - x^2} \frac{x_1 - p}{px_2^2} \left((2x_2 + x_1 - 1)E_{2,2} + \sum_{i=3}^l (E_{2,i} + E_{i,2}) \right) \\ & + \chi_{\{p < 1/2\}}(x) (1 - 2p)\pi \left\{ \left(E_{1,1} - \sum_{i=2}^l (E_{1,i} + E_{i,1}) + \sum_{2 \leq i,j \leq l} E_{i,j} \right) \delta_{-1}(x) \right. \\ & \left. + \left(\sum_{1 \leq i,j \leq l} E_{i,j} \right) \delta_1(x) \right\}. \end{aligned}$$

Definimos a las matrices M_i , con $l \geq i \geq 3$ y $M_{i,j}$, donde nuevamente $l \geq i \neq j \geq 3$, como

$$\begin{aligned} M_i &= \frac{x_i^2}{x_2^2} E_{2,2} - \frac{x_i}{x_2} (E_{2,i} + E_{i,2}) + E_{i,i} \\ M_{i,j} &= 2 \frac{x_i x_j}{x_2^2} E_{2,2} - \frac{x_j}{x_2} (E_{2,i} + E_{i,2}) - \frac{x_i}{x_2} (E_{2,j} + E_{j,2}) + (E_{i,j} + E_{j,i}). \end{aligned}$$

Dado esto la medida buscada tiene la forma:

$$\mu(x) = M(x) + \sqrt{4pq - x^2} \left(\sum_{i=3}^l c_i M_i + \sum_{3 \leq i < j \leq l} c_{i,j} M_{i,j} \right),$$

donde los coeficientes c_i y $c_{i,j}$ son elegidos con base en

$$\begin{aligned} c_i &= \left(\frac{1}{p} - 1 \right) \frac{1}{x_i} - \frac{1}{p}, \quad 3 \leq i \leq l, \\ c_{i,j} &= -\frac{1}{p}, \quad 3 \leq i < j \leq l. \end{aligned}$$

Finalmente analizamos las propiedades de recurrencia de este proceso. Al analizar la definición de $M(x)$ y la ecuación que define la medida asociada, para que esta quede bien definida, es necesario que $|x| \leq 2\sqrt{pq}$, por lo que replicando el argumento del Ejemplo 3.1, el proceso es transitorio si y solo si $p \neq q$ y $p \geq 1/2$, pues sino poseería

un salto y por el Teorema 2.10 no solo sería recurrente sino recurrente positivo. En el caso que $p = q = 1/2$, por la propia definición de la medida es fácil ver que la integral del Teorema 2.9 diverge y por ende, en este caso, el proceso es recurrente. Regresando a la recurrencia positiva del proceso, notemos que para el caso en el que $p < 1/2$ la medida posee saltos en 1 y -1 , por lo que en este caso, usando el Teorema 2.10, el proceso es recurrente positivo.

Para obtener la medida invariante ocupamos el Teorema 2.13. Por la definición de las matrices R_i , si definimos a la siguiente sucesión de vectores horizontales $\{v_i\}_{i \geq 0}$ de dimensión l :

$$v_0 = \left(1, \frac{x_2}{q}, \dots, \frac{x_l}{q}\right), \quad v_1 = \left(\frac{x_1}{q}, \frac{x_2 p}{q^2}, \dots, \frac{x_l p}{q^2}\right),$$

y

$$v_i = \left(\frac{p}{q}\right)^{i-1} v_1, \quad i \geq 2.$$

Entonces la medida invariante, nuevamente queda definida como:

$$\pi = (v_0; v_1; v_2; \dots).$$

De hecho, por lo estudiado en la sección 2.5, en el caso que $p < 1/2$, como el proceso resulta recurrente positivo, la medida invariante coincide con la distribución estacionaria que para obtenerla solo basta normalizar el vector π . Para esto calculamos la suma de las entradas de π , la cual llamamos S

$$S = 1 + \sum_{i=1}^l \frac{x_i}{q} \sum_{j=0}^{\infty} \left(\frac{p}{q}\right)^j = 1 + \sum_{i=1}^l \frac{x_i}{q} \cdot \frac{q}{q-p} = 1 + \frac{1}{q-p} \sum_{i=1}^l x_i = 1 + \frac{1}{q-p} = \frac{2q}{q-p},$$

donde usamos que $\sum_{i=1}^l x_i = 1$, y $p + q = 1$. Por lo que la distribución estacionaria está dada por $\frac{q-p}{2q} \pi$.

3.5. Un ejemplo proveniente de teoría de representación de grupos

Nuestro siguiente y último ejemplo aparece, primeramente, en [17]. Como su nombre lo menciona este proviene de la teoría de representación de grupos. Presentamos a las matrices por bloques que componen a su matriz de transición, así como la medida matricial asociada a esta. Por último, estudiamos sus propiedades de recurrencia y expresamos explícitamente su medida invariante.

Los coeficientes de nuestro proceso en cuestión quedan definidos de la siguiente forma (note que se usa a la matriz $E_{i,j}$ para referirse a la matriz definida en la prueba del Lema 1.4 y que fue usada en el Ejemplo 3.4):

$$A_n = \sum_{i=0}^{l-1} \frac{(k+n)(\beta+n+l)(\alpha+\beta+n+l+i)(\alpha+\beta-k+n+i+1)}{(k+n+l-i-1)(\alpha+\beta-k+n+2i+1)(\alpha+\beta+2n+l+i)_2} E_{i,i} \\ + \sum_{i=0}^{l-2} \frac{(i+1)(k+n)(k+l-i-2)(\beta+n+l)}{(\alpha+\beta+2n+l+i+1)(\alpha+\beta-k+n+2i+3)(k+n+l-i-2)_2} E_{i+1,i},$$

para $n \geq 0$. Las matrices B_n quedan definidas como

$$B_n = \sum_{i=0}^{l-1} \left(1 + \frac{n(k+n-1)(k+n+l-1)(\beta+n+l-1)}{(\alpha+\beta+2n+l+i-1)(k+n+l-i-2)_2} \right. \\ - \frac{(n+1)(k+n)(k+n+l)(\beta+n+l)}{(\alpha+\beta+2n+l+i+1)(k+n+l-i-1)_2} \\ + \frac{i(l-i)(k+l-i-1)(\beta-k+i)}{(\alpha+\beta-k+n+2i)(k+n+l-i-1)_2} \\ \left. - \frac{(i+1)(l-i-1)(k+l-i-2)(\beta-k+i+1)}{(\alpha+\beta-k+n+2i+2)(k+n+l-i-2)_2} \right) E_{i,i} \\ + \sum_{i=0}^{l-2} \frac{(l-i-1)(\beta-k+i+1)(\alpha+\beta-k+n+i+1)(\alpha+\beta+n+l+i)}{(k+n+l-i-1)(\alpha+\beta+2n+l+i)(\alpha+\beta-k+n+2i+1)_2} E_{i,i+1} \\ + \sum_{i=0}^{l-2} \frac{(i+1)(\alpha+n+i+1)(k+l-i-2)(\alpha+\beta-k+n+l+i+1)}{(k+n+l-i-2)(\alpha+\beta+2n+l+i+1)(\alpha+\beta-k+n+2i+2)_2} E_{i+1,i},$$

para $n \geq 0$, y finalmente las matrices C_n^T

$$C_n^T = \sum_{i=0}^{l-1} \frac{n(\alpha+n+i)(k+n+l-1)(\alpha+\beta-k+n+l+i)}{(k+n+l-i-1)(\alpha+\beta-k+n+2i+1)(\alpha+\beta+2n+l+i-1)_2} E_{i,i} \\ + \sum_{i=0}^{l-2} \frac{n(l-i-1)(k+n+l-1)(\beta-k+i+1)}{(\alpha+\beta+2n+l+i)(\alpha+\beta-k+n+2i+1)(k+n+l-i-2)_2} E_{i,i+1},$$

para $n \geq 1$. Donde el subíndice $(a)_r$ denota al símbolo de Pochhammer $(a)_r = a(a+1) \dots (a+r-1)$, y los parámetros $\alpha, \beta > -1$ y $k \in \mathbb{R}$ tal que $0 < k < \beta + 1$. Ahora, asociada a estos coeficientes encontramos que la secuencia de matrices $\{R_n\}_{n \geq 0}$ dada

por

$$R_n = \sum_{i=0}^{l-1} \left((-1)^i \frac{\Gamma(n+1)\Gamma(\beta+l)\Gamma(\alpha+n+i+1)(1-k-l-n)_i}{\binom{l-1}{i}\Gamma(l)\Gamma(\alpha+\beta+l+i+2n+1)} \right. \\ \left. \times \frac{(k+l-i-1)_n(\alpha+\beta+l+i+n)_n(\alpha+\beta-k+i+n+1)_l}{(\alpha+\beta-k+2i+n+1)(k)_n(\beta-k+1)_i(\beta+l)_n} \right)^{-1/2} E_{i,i}, \quad (3.13)$$

con $n \geq 0$. Para definir la medida matricial asociada a dicho proceso es necesario primero definir a las siguientes matrices T y $Z(x)$

$$T = \sum_{i \leq j} (-1)^i \frac{(-j)_i}{(1-l)_i} \frac{(\alpha+\beta-k+j+1)_i}{(\beta-k+1)_i} E_{i,j},$$

y

$$Z(x) = \sum_{i,j=0}^{l-1} \left(\sum_{r=0}^{l-1} \binom{r}{i} \binom{r}{j} \binom{l+k-r-2}{l-r-1} \binom{\beta-k+r}{r} (1-x)^{i+j} x^{l-r-1} \right) E_{i,j}.$$

En base a esto la medida matricial queda dada por

$$d\mu(x) = x^\alpha (1-x)^\beta T^* Z(x) T dx, \quad x \in [0, 1]. \quad (3.14)$$

En base a esta información y por los resultados obtenidos en el capítulo 2, podemos estudiar las propiedades de recurrencia de este proceso. Note, primeramente, que la matriz de transición definida por los anteriores coeficientes es estocástica, y que el proceso de Markov que induce es irreducible y aperiódico. Para verificar esto último solo basta con ver que dados dos estados (i, j) y (i', j') existe un camino finito entre ellos. En efecto, por la forma de las matrices coeficientes definidas al inicio de la sección, si partimos de un estado arbitrario (i, j) siempre podemos movernos a $(i+1, j)$, $(i-1, j)$, $(i, j+1)$ y $(i, j-1)$, así al variar la primera coordenada del estado para pasar de i a i' y posteriormente de j a j' encontramos nuestro camino, esto implica que el proceso es irreducible, el hecho de que sea aperiódico es resultado inmediato de que todo estado puede permanecer en sí mismo después de una transición.

Considerando esta última información podemos hacer uso del Corolario 2.9, en específico de la expresión (2.44), tenemos que el proceso es recurrente si y solo si la expresión (2.44) se satisface para algún j . Poniendo más atención en (3.14), y específicamente en $T^* Z(x) T$, tenemos

$$T^* Z(x) T = \frac{(\beta+1)_{d-1}}{(d-1)!} \sum_{i,j=0}^{d-1} E_{ij}.$$

Por lo tanto, dado el factor $\frac{1}{1-x}$ de la integral en (2.44), entonces la integral diverge si y solo si se satisface que $-1 < \beta \leq 0$. De hecho, gracias al Teorema 2.10, sabemos que el proceso es positivo recurrente si y solo si la medida posee un salto en 1, lo cual claramente no pasa aquí, por lo que el proceso es recurrente nulo.

Finalmente estudiamos la medida invariante, como en los ejemplos anteriores. Para ello, nuevamente, hacemos uso del Teorema 2.13. Si abreviamos como $c_{i,n}$ a el i -ésimo sumando de (3.13), entonces definiendo la siguiente sucesión de vectores horizontales de dimensión l

$$v_n = (c_{0,n}^2, \dots, c_{l-1,n}^2), \quad n \geq 0,$$

obtenemos la medida invariante

$$\pi = (v_0; v_1; v_2; \dots).$$

Recientemente, en [18] se puede encontrar un modelo de urnas asociado a este ejemplo.

Bibliografía

- [1] AKHIEZER N. y GLAZMAN I., *Theory of Linear Operators in Hilbert Space*. Dover Publications, New York, 1993.
- [2] BASILEVSKY A., *Applied Matrix Algebra in the Statistical Sciences*. North-Holland, Amsterdam, 1983.
- [3] BEREZANSKII J. M., *Expansions in Eigenfunctions of Selfadjoint Operators*. Translations of Mathematical Monographs, 17, AMS, Providence, RI, págs. 547-564, 1968.
- [4] BOYD S. y VANDENBERGHE L., *Convex Optimization*. Cambridge University Press, 2004.
- [5] DAMANIK D., PUSHNITSKI, A. y SIMON B., *The Analytic Theory of Matrix Orthogonal Polynomials*. Surveys in Approximation Theory, Volumen 4, 2008, págs. 1-85.
- [6] DAYAR T. y QUESSETTE F., *Quasi-birth-and-death processes with level-geometric distribution*. SIAM Journal Matrix Anal. Appl., Volumen 24, págs. 281-291, 2002.
- [7] DE LA IGLESIA M. D., *A note on the invariant distribution of a quasi-birth-and-death process*. Journal of Physics A: Mathematical and Theoretical, Volumen 44, pág. 4, 2011.
- [8] DETTE H., REUTHER B., STUDDEN W. J. y ZYGMUNT M., *Matrix Measures and Random Walks with a Block Tridiagonal Transition Matrix*. SIAM Journal on Matrix Analysis and Applications, Volumen 29, No. 1, págs. 117-142, 2006.
- [9] DETTE H. y STUDDEN W. J., *The Theory of Canonical Moments with Applications in Statistics, Probability, and Analysis*. Wiley Series in Probability and Statistics: Applied Probability and Statistics. John Wiley & Sons Inc., New York, págs. 1-13, 1997.

- [10] DETTE H. y STUDDEN W. J., *Matrix measures, moment spaces, and Favard's theorem for the interval $[0, 1]$ and $[0, \infty)$* . Linear Algebra Appl., Volumen 345, págs. 163-193, 2001.
- [11] DURÁN A. J., *On Orthogonal Polynomials With Respect to a Positive Definite Matrix of Measures*. Canadian Journal of Mathematics, 47(1), págs. 88-112, 1995.
- [12] DURÁN A. J. y LÓPEZ-RODRÍGUEZ P., *Orthogonal Matrix Polynomials: Zeros and Blumenthal's Theorem* Journal of Approximation Theory, 1996.
- [13] DURÁN A. J., *Markov's Theorem for orthogonal matrix polynomials*. Canadian Journal of Mathematics, Volumen 48, págs. 1180-1195, 1996.
- [14] DURÁN A. J., *Ratio asymptotics for orthogonal matrix polynomials*. Journal Approx. Theory, Volumen 100, págs. 304-344, 1999.
- [15] FAIR W., *Noncommutative continued fractions*. SIAM Journal Math. Anal., Volumen 2, págs. 226-232, 1971.
- [16] GRÜNBAUM A., *A Spectral Weight Matrix for a Discrete Version of Walsh's Spider*. Operator Theory: Advances and Applications, Vol. 202, 253–264, 2010.
- [17] GRÜNBAUM A. y DE LA IGLESIA M. D., *Matrix valued orthogonal polynomials arising from group representation theory and a family of Quasi-Birth-and-Death processes*. SIAM Journal on Matrix Analysis and Applications, Vol. 30, No. 2, págs. 741–761.
- [18] GRÜNBAUM A. y DE LA IGLESIA M. D., *Stochastic Darboux transformations for quasi-birth-and-death processes and urn models*. J. Math. Anal. Appl., Vol. 478, 634–654, 2019.
- [19] KARLIN S. y MCGREGOR J., *Random walks*. Illinois Journal Math., Volumen 3, págs. 66-81, 1959.
- [20] KREIN M. G. y NUDELMAN A. A., *The Markov moment problem and extremal problems*. Translations of Mathematical Monographs., Volumen 50, American Mathematical Society, págs. 62-63, 1977.
- [21] LATOUCHE G., *Matrix Chebyshev polynomials and continued fractions*. Linear Algebra Appl., Volumen 14, págs. 443-460, 1998.
- [22] LATOUCHE G. y RAMASWAMI V., *Introduction to Matrix Analytic Methods in Stochastic Modeling*. Society for Industrial and Applied Mathematics, 1999.

- [23] NEUTS M. F., *Structured Stochastic Matrices of M/G/1 Type and Their Applications*. CRC Press, Primera Edición, 1989.
- [24] OST A., *Performance of Communication Systems*. Springer-Verlag Berlin Heidelberg, Primera Edición, 2001.
- [25] ROSENBERG M., *The square-integrability of matrix-valued functions with respect to a non-negative Hermitian measure*. Duke Mathematical Journal, Volumen 31, págs. 291-298, 1964.
- [26] RUDIN W., *Functional Analysis*. McGraw-Hill, 1973.
- [27] ZYGMUNT J., PEARCE C. E. M. y TAYLOR P. G., *Invariant measures for quasi-birth-and-death process*, Comm. Statist., Volumen 340, págs. 155-158, 1998.