

Centro de Investigación en Matemáticas, A.C.

CIMAT

**Análisis del Espectro Asintótico
de Canales Multiantena vía
Probabilidad Libre**

T E S I S

Que para obtener el grado de

Maestro en Ciencias

con Orientación en

Probabilidad y Estadística

P r e s e n t a

Ing. Mario Alberto Diaz Torres

Director de Tesis:

**Dr. Víctor Manuel Pérez Abreu
Carrión**

Guanajuato, Gto.. Agosto de 2013

Integrantes del Jurado

Presidente: Dr. Francisco Javier Torres Ayala

Secretario: Dr. Octavio Arizmendi Echegaray

Vocal: Dr. Víctor Manuel Pérez Abreu Carrión

Lector Especial: Mtro. Carlos Vargas Obieta

Asesor:

Dr. Víctor Manuel Pérez Abreu Carrión

Sustentante:

Ing. Mario Alberto Diaz Torres

Agradecimientos

Antes que nada, quiero agradecer a mi familia por el apoyo que me ha brindado a lo largo de toda mi vida.

También agradezco a mi director de tesis, el Dr. Víctor Manuel Pérez Abreu Carrión, por todo el apoyo que me ha brindado a lo largo de estos años de conocerlo.

Así mismo, doy las gracias a mis sinodales, el Dr. Octavio Arizmendi Echegaray y el Dr. Francisco Javier Torres Ayala, por sus valiosas sugerencias y comentarios. También agradezco a mi lector especial, el Mtro. Carlos Vargas Obieta, por todo su apoyo antes, durante y después de la elaboración de esta tesis.

Finalmente, agradezco al Centro de Investigación en Matemáticas CIMAT por todas las facilidades ofrecidas. Así mismo, agradezco al Consejo Nacional de Ciencia y Tecnología CONACYT por la beca de maestría que me fue otorgada (CVU 418057).

Es difícil mencionar a todas las personas e instituciones que contribuyeron de alguna manera a la realización de esta tesis, aprovecho estas líneas para brindarles mi más sincero agradecimiento.

Mario Alberto Díaz Torres

Índice general

Introducción	IX
Notación	XIII
1. Motivación general	1
2. Probabilidad libre escalar	7
2.1. Definiciones básicas	7
2.2. Variables aleatorias libres	16
2.3. Teorema del límite central libre	19
2.4. Cumulantes y transformadas	23
2.5. Notas adicionales	28
3. Probabilidad libre y matrices aleatorias	29
3.1. Matrices aleatorias	29
3.2. Convergencia de matrices aleatorias clásicas	33
3.3. Libertad asintótica	36
3.4. Notas adicionales	39
4. Probabilidad libre valuada en operadores	41
4.1. Definiciones básicas	41
4.2. Variables aleatorias semicirculares y el teorema del límite central	49
4.3. Transformadas y ecuaciones de subordinación	54

4.4. Espacios de probabilidad rectangulares	59
4.5. Matrices a bloques	66
4.6. Notas adicionales	76
5. Análisis del espectro asintótico de canales multiantena	77
5.1. Canal multiacceso	78
5.2. Redes retransmisoras con multisalto	81
5.3. Sistema monousuario con interferencia entre símbolos	87
5.4. Sistema monousuario con correlación angular	88
5.4.1. Modelos de correlación a bloques	88
5.4.2. Extensión a elementos no circulares y correlación angular	90
5.4.3. Un caso de estudio: familia circular libre	94
5.4.4. Conclusiones y trabajo futuro	97
5.5. Notas adicionales	98
A. Teoría general de la inversión de Möbius	101

Introducción

La teoría de probabilidad libre fue fundada en la década de los 80 por Dan Virgil Voiculescu. En palabras de Dimitri Shlyakhtenko: “Noncommutative probability theory views elements of a noncommutative von Neumann algebra as analogs of classical random variables. One of Voiculescu’s discoveries is that in this more general noncommutative framework there is room for a new notion of independence, called freeness. [...] Amazingly, as Voiculescu showed, many theorems and concepts in classical probability have very nice free analogs; [...] Free probability theory has now grown into a rich field, with connections and applications to many other areas of mathematics. For example, Voiculescu’s discovery that certain random matrices are asymptotically free as their sizes go to infinity makes possible computations of expected asymptotic spectral density of their eigenvalues. [...]”.

En otras palabras, la teoría de probabilidad libre es un análogo a la teoría de probabilidad cuando las variables aleatorias son no conmutativas. En la década de los 90 Roland Speicher introdujo el concepto de cumulantes libres y relacionó a estos últimos con la teoría combinatoria de las particiones que no se cruzan. En esa misma época Voiculescu demostró que ciertos ensambles de matrices aleatorias son libres asintóticamente, lo que permitió poner a las matrices aleatorias en el contexto de la probabilidad libre. Específicamente, permitió calcular los espectros asintóticos de ciertos ensambles de matrices aleatorias vía las herramientas, tanto analíticas como combinatorias, de la probabilidad libre. Para más información acerca del contexto histórico de la probabilidad libre véase la introducción del libro de Vladislav Kargin [14] del 2008.

En comunicación inalámbrica, particularmente al trabajar con sistemas multiantena, ciertas cantidades de interés dependen del espectro de ciertas matrices aleatorias. Puesto que es posible trabajar con el espectro asintótico de matrices aleatorias vía probabilidad libre, esta última es la base de diversas aplicaciones en comunicación inalámbrica. Por ejemplo véanse los artículos de Mestre, Fonollosa y Pagès-Zamora [18] del 2003; Far, Oraby, Bryc y Speicher [11] del 2008; Speicher, Vargas y Mai [24] del 2011; y Debbah, Fawaz, Gesbert y Zarifi [7] también del 2011.

Un sistema multiantena es un sistema de comunicaciones en el cual se transmite y se recibe con varias antenas que operan simultáneamente. El sistema multiantena ideal supone la ausencia de correlación entre las variables aleatorias que modelan a los canales entre las antenas del sistema. Como consecuencia de la no correlación se puede demostrar, a partir del teorema célebre de Marchenko-Pastur (Teorema 3.5), que dicho sistema ideal es capaz de transmitir cualquier cantidad de información. Específicamente, la cantidad de información crece indefinidamente al incrementar el número de antenas. Sin embargo, cuando el número de antenas en un mismo dispositivo aumenta es muy difícil evitar que estas se correlacionen [23]. Por lo tanto, surge la necesidad de modelar la correlación de tal forma que seamos capaces de estimar la escalabilidad del sistema, es decir, cuantificar la cantidad de información que pasará a través del sistema cuando el número de antenas se vuelva arbitrariamente grande.

En la literatura existen diversos modelos para la correlación de sistemas multiantena. Uno de los modelos más utilizados es el modelo de correlación separable o correlación de Kronecker, véase por ejemplo el artículo de Lozano, Tulino y Verdú [17] del 2005. En dicho artículo se calcula la capacidad de un sistema multiantena con este tipo de correlación mediante una serie de teoremas derivados a partir de un teorema de matrices aleatorias deducido por Girko.

En este contexto, una extensión reciente de la probabilidad libre, llamada probabilidad libre valuada en operadores, nos permite calcular el espectro de dicho canales multiantena correlacionados. Más aún, explotando el concepto de libertad, el análogo en probabilidad libre al de independencia en probabilidad, es posible contruir nuevos

modelos para la correlación de sistemas multiantena. En este sentido, la probabilidad libre valuada en operadores nos proporciona una serie de herramientas de utilidad desde el punto de vista de la modelación.

En esta tesis de maestría proporcionamos un modelo para modelar la correlación en un sistema multiantena el cual, hasta donde sabemos, no ha sido explorado en la literatura. Este modelo surge de adaptar y extender el modelo a bloques utilizado por Far, Oraby, Bryc y Speicher en el artículo [11] del 2008 en el contexto de sistemas monousuario con interferencia intersímbolos. Para calcular la cantidad de información asintótica en el número de antenas transmitida bajo este modelo es necesario calcular el espectro de la matriz de operadores que nos proporciona el modelo. En este trabajo describimos a grandes rasgos una forma de resolver este problema en un caso particular. Sin embargo, la solución del caso general es parte del trabajo futuro que surge de esta tesis.

En el Capítulo 1 discutimos la metodología usual con la cual posteriormente pondremos a las matrices aleatorias en el contexto de la probabilidad libre. Este capítulo posee un carácter introductorio por lo cual procuramos mantener los tecnicismos al mínimo.

En el Capítulo 2 presentamos algunos de los fundamentos de la teoría de probabilidad libre. En particular, definimos espacio de probabilidad no conmutativo, variable aleatoria no conmutativa y objetos relacionados, como por ejemplo la relación de libertad, los cumulantes libres y la transformada de Cauchy.

En el tercer capítulo enunciamos los teoremas de la teoría de matrices aleatorias que nos permiten aplicar el procedimiento discutido en el Capítulo 1.

En el Capítulo 4 presentamos la probabilidad libre valuada en operadores. Después de dar algunos elementos básicos de esta teoría, trabajamos con dos espacios de probabilidad valuados en operadores los cuales nos permitirán operar con matrices rectangulares y matrices a bloque. Este tipo de espacios serán de utilidad en las aplicaciones tratadas en el último capítulo. En particular, las matrices rectangulares nos permiten solucionar el problema ya resuelto en Debbah, Fawaz, Gesbert y Zarifi [7] mediante el uso exclusivo de probabilidad libre (valuada en operadores); mientras que el modelo

que proponemos está basado fuertemente en matrices a bloques.

Como ya se mencionó, en el Capítulo 5 presentamos algunas aplicaciones de la probabilidad libre valuada en operadores en comunicación inalámbrica. Después de revisar algunas aplicaciones existentes, en la Sección 5.4 presentamos nuestra propuesta para modelar la correlación en sistemas multiantena.

Al final de cada capítulo se incluyen notas adicionales de carácter bibliográfico e histórico relacionadas con el capítulo al que pertenecen.

En el Apéndice A damos algunos elementos de la Teoría de Inversión de Möbius, la cual es una parte importante de la teoría combinatoria de los cumulantes libres.

Notación

\mathbb{R}	Campo de los números reales
\mathbb{C}	Campo de los números complejos
$\text{Mat}_d(\mathbb{K})$	Conjunto de las matrices de $d \times d$ con coeficientes en el campo \mathbb{K}
$\text{Mat}_{d \times e}(\mathbb{K})$	Conjunto de las matrices de $d \times e$ con coeficientes en \mathbb{K}
tr	Función traza
tr_d	Función traza normalizada por la dimensión (d)
$\mathbf{1}_{\mathcal{A}}$	Identidad en el álgebra \mathcal{A}
$\mathbf{0}_{\mathcal{A}}$	Elemento neutro en el álgebra \mathcal{A}
$C(X, \mathbb{K})$	Conjunto de las funciones continuas en X con valores en \mathbb{K}
$\mathcal{B}_p(r)$	Bola centrada en p de radio r
$\mathbb{K}\langle X_i \mid i \in I \rangle$	Polinomios con coeficientes en \mathbb{K} con indeterminadas $(X_i)_{i \in I}$
$L^p(\Omega, P)$	Conjunto de las funciones $f : \Omega \rightarrow \mathbb{C}$ tales que $\int f ^p dP < \infty$
$L^\infty(\Omega, P)$	Conjunto de las funciones $f : \Omega \rightarrow \mathbb{C}$ acotadas P c.s.
$L^{\infty-}(\Omega, P)$	Conjunto de las $f : \Omega \rightarrow \mathbb{C}$ tales que $f \in L^p(\Omega, P)$ para toda $p \geq 1$
\mathcal{S}^1	Subconjunto de los números complejos con norma unitaria
$\text{NC}(n)$	Particiones que no se cruzan del conjunto $\{1, \dots, n\}$
$\text{NC}_2(n)$	Emparejamientos que no se cruzan del conjunto $\{1, \dots, n\}$
NC	Conjunto de todas las particiones que no se cruzan
NC_2	Conjunto de todos los emparejamientos que no se cruzan
$\mathfrak{B}(X)$	σ -álgebra de Borel del espacio topológico X
$[A]_{i,j}$	Coficiente i, j de la matriz A

Capítulo 1

Motivación general

Como se dijo en la introducción, el objetivo de este capítulo es ilustrar la metodología usual para operar con el espectro de matrices aleatorias mediante probabilidad libre. Para ello consideremos lo siguiente.

Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad en el cual están definidas variables aleatorias $(X_n, Y_n)_{n \geq 1}$ independientes, tales que las $(X_n)_{n \geq 1}$ son idénticamente distribuidas con media μ_x y las $(Y_n)_{n \geq 1}$ también son idénticamente distribuidas de media μ_y . Supongamos además que las variables aleatorias en cuestión tienen soporte uniformemente acotado, i.e. $X_n, Y_n \in L^\infty(\Omega, \mathbb{P})$ para toda $n \in \mathbb{N}$ y existe $K > 0$ tal que $|X_n|, |Y_n| \leq K$ para toda n . Definamos para cada $N \in \mathbb{N}$

$$S_N^X := \frac{1}{N} \sum_{n=1}^N X_n,$$
$$S_N^Y := \frac{1}{N} \sum_{n=1}^N Y_n,$$
$$\mathbb{E}_N := \mathbb{E},$$

donde \mathbb{E} es el valor esperado asociado a \mathbb{P} . Observemos que \mathbb{E}_N es un funcional lineal $\mathbb{E}_N : L^\infty(\Omega, \mathbb{P}) \rightarrow \mathbb{C}$.

Si por alguna razón únicamente operamos con las variables aleatorias S_N^X y S_N^Y

a través del funcional \mathbb{E}_N , y además únicamente estamos interesados en el valor de $\mathbb{E}_N((S_N^X + S_N^Y)^2)$, entonces un cálculo directo nos muestra que

$$\mathbb{E}_N((S_N^X + S_N^Y)^2) = \frac{\mathbb{E}(X_1^2) + \mathbb{E}(Y_1^2)}{N} + \frac{N-1}{N}(\mu_x^2 + \mu_y^2) + 2\mu_x\mu_y.$$

El cálculo anterior es relativamente sencillo, sin embargo, este tipo de expresiones se puede complicar, e.g. $\mathbb{E}_N((S_N^X S_N^Y + S_N^{Y^2})^3 - S_N^X S_N^Y)$. Una observación importante es que¹ $\mathbb{E}_N((S_N^X + S_N^Y)^2) \rightarrow (\mu_x + \mu_y)^2$ y

$$\mathbb{E}_N((S_N^X S_N^Y + S_N^{Y^2})^3 - S_N^X S_N^Y) \rightarrow (\mu_x\mu_y + \mu_y^2)^3 - \mu_x\mu_y$$

cuando $N \rightarrow \infty$. Por lo tanto, para N suficientemente grande es posible utilizar la cantidad $(\mu_x\mu_y + \mu_y^2)^3 - \mu_x\mu_y$ como una aproximación a $\mathbb{E}_N((S_N^X S_N^Y + S_N^{Y^2})^3 - S_N^X S_N^Y)$.

Observemos que para cada $N \in \mathbb{N}$ únicamente hemos trabajado con las variables aleatorias S_N^X, S_N^Y y el funcional lineal \mathbb{E}_N . Puesto que las variables aleatorias S_N^X y S_N^Y son variables aleatorias de soporte compacto, la frase anterior es equivalente a decir que para cada $N \in \mathbb{N}$ trabajamos con la pareja $(\mathcal{A}_N, \mathbb{E}_N)$, donde $\mathcal{A}_N = L^\infty(\Omega, \mathbb{P})$. Si denotamos por $\text{id} : \mathbb{C} \rightarrow \mathbb{C}$ a la identidad en \mathbb{C} , la aproximación propuesta en el párrafo anterior es equivalente a aproximar a nuestro espacio $(\mathcal{A}_N, \mathbb{E}_N)$ por el espacio $(\mathcal{A} := \mathbb{C}, \varphi := \text{id})$. En general, no es claro como aproximar a una variable aleatoria cualquiera en \mathcal{A}_N por un elemento en \mathcal{A} . Sin embargo, las variables aleatorias S_N^X y S_N^Y pueden ser aproximadas por μ_x y μ_y respectivamente.

Si bien es cierto que la simplificación propuesta nos evita una gran cantidad de cálculos, en esta era en la que el ordenador nos permite realizar millones de operaciones por segundo parece poco útil tal simplificación. Sin embargo, la razón de fondo es que el espacio (\mathcal{A}, φ) contiene la esencia del fenómeno en cuestión. En este sentido, si las parejas $(\mathcal{A}_N, \mathbb{E}_N)$ provienen de cierto fenómeno físico entonces podemos pensar a $(\mathcal{A}_N, \mathbb{E}_N)$ como una versión imperfecta de (\mathcal{A}, φ) , aun que no sea de esta manera.

¹Puesto que $(X_n, Y_n)_{n \geq 1}$ tienen soporte uniformemente acotado, los límites siguientes son una consecuencia directa del Teorema de Convergencia Dominada.

De esta forma optamos por la simplificación de espacio para evitar distraernos con detalles posiblemente irrelevantes. En este sentido, la probabilidad libre nos proporciona un espacio esencial el cual nos permite operar y entender mejor a las matrices aleatorias de dimensión grande. En capítulos subsecuentes presentaremos la noción de libertad asintótica, el concepto análogo a la independencia en el contexto de ensambles de matrices aleatorias. Puesto que algunas aplicaciones suponen libertad asintótica entre algunas matrices, es imperativo desde el punto de vista de modelación entender la relación de libertad asintótica. En particular, este es un ejemplo de los beneficios de entender la esencia detrás de las matrices de dimensión grande.

Para ilustrar la metodología usual al trabajar con matrices aleatorias consideremos lo siguiente. Supóngase que se tiene matrices autoadjuntas² $(A_n)_{n \in \mathbb{N}}$ cuyos coeficientes son variables aleatorias con valores en los números complejos. Denotemos por F^{A_n} a la distribución empírica espectral de A_n , es decir, la distribución uniforme en los eigenvalores de A_n .

Puesto que los coeficientes de las A_n son aleatorios, también las distribuciones F^{A_n} son aleatorias. Sin embargo, existen familias de matrices aleatorias $(A_n)_{n \in \mathbb{N}}$ cuyas distribuciones empíricas espectrales F^{A_n} convergen en distribución casi seguramente a una función de distribución F^A no aleatoria. En el Capítulo 3 damos ejemplos de este tipo de familias de matrices aleatorias.

En el ánimo de las aproximaciones hechas en párrafos anteriores, en este contexto utilizaremos a F^A en lugar de F^{A_n} . Si además de la familia $(A_n)_{n \in \mathbb{N}}$ se tuviera otra familia $(B_n)_{n \in \mathbb{N}}$ tal que sus distribuciones empíricas espectrales F^{B_n} convergen a una función de distribución F^B , ¿qué podríamos decir acerca de las *aproximaciones* correspondientes para $F^{A_n+B_n}$ o $F^{A_n B_n}$?

Un enfoque para responder a esta pregunta, el cual extiende lo realizado para las sumas normalizadas S_N^X y S_N^Y al caso de matrices aleatorias, es el siguiente.

1) Puesto que estamos interesados en sumas y productos, necesitamos crear un álgebra

²En general las matrices no tienen que ser autoadjuntas, sin embargo, aquí hacemos esa suposición por simplicidad en la exposición.

\mathcal{A} de «variables aleatorias» en la cual existan $a, b \in \mathcal{A}$ tales que la «distribución» de a está íntimamente relacionada con F^A y la «distribución» de b lo está con F^B .

- 2) Derivar reglas para el cálculo de la «distribución» de la suma y producto de variables aleatorias «independientes».
- 3) Encontrar en que casos las variables inducidas por las familias de matrices aleatorias resultan ser «independientes».

Las «» son para denotar que los objetos que buscamos son análogos al caso clásico, pero no necesariamente son los mismos.

Observemos que de crear un tal espacio y una forma de asignarle una «variable aleatoria» a cada distribución de probabilidad, estaremos equipando a cada familia de matrices aleatorias (cuyo espectro converge) una «variable aleatoria» en una cierta álgebra \mathcal{A} . Esto nos permite interpretar este procedimiento como dotar de una matriz *infinita* a cada familia de matrices. Después veremos que las C^* -álgebras son de particular utilidad al trabajar con estos nuevos «espacios de probabilidad», y por lo tanto estas matrices *infinitas* de hecho pueden ser pensadas como operadores.

Una forma de caracterizar a las variables aleatorias reales de soporte compacto es mediante sus momentos. Por lo tanto, $L^\infty(\Omega, \mathbb{P})$ es un espacio conmutativo el cual es un posible candidato a álgebra donde aplicar el procedimiento anterior. Específicamente, la «independencia» es la independencia usual y la «distribución» es el conjunto de momentos. Además, las reglas para calcular la «distribución» de la suma o producto de variables aleatorias independientes también sigue la técnica usual (basada en momentos). Por lo tanto, la «distribución» depende sólo del funcional lineal esperanza lo que nos permite trabajar únicamente con la pareja $(L^\infty(\Omega, \mathbb{P}), \mathbb{E})$. Es importante enfatizar que en este caso trabajamos con los momentos de la variable aleatoria y no su distribución de probabilidad.

Si bien es cierto que $(L^\infty(\Omega, \mathbb{P}), \mathbb{E})$ es un candidato para aplicar el procedimiento de tres pasos descrito anteriormente, es intuitivamente claro que no es un buen can-

didato ya que $L^\infty(\Omega, \mathbb{P})$ es un álgebra conmutativa mientras que las matrices son no conmutativas.

Para el caso de matrices aleatorias, observemos que para toda $m \in \mathbb{N}$ se tiene que

$$\mathbb{E} \left(\int x^m F^{A_n}(\mathrm{d}x) \right) = \mathbb{E} (\mathrm{tr}_n(A_n^m)). \quad (1.1)$$

En analogía a los momentos del párrafo anterior, para matrices aleatorias de dimensión finita podemos tomar el valor esperado de la traza normalizada para obtener la «distribución» de la distribución empírica espectral. Más aún, para encajar a una familia $(A_n)_{n \in \mathbb{N}}$ en el espacio de matrices *infinitas* tomaremos un álgebra no conmutativa \mathcal{A} y un elemento $a \in \mathcal{A}$ de tal manera que la «distribución» (momentos) de a coincidan con los límites de los momentos de F^{A_n} . Específicamente, si $\varphi : \mathcal{A} \rightarrow \mathbb{C}$ es el funcional lineal asociado a las «distribuciones» en \mathcal{A} , requerimos que

$$\varphi(a^m) = \lim_{N \rightarrow \infty} \mathbb{E} (\mathrm{tr}_N(A_N^M))$$

para toda $m \in \mathbb{N}$.

Las reglas para calcular los momentos de sumas o productos en el espacio de matrices *infinitas* no son las usuales, donde la principal diferencia radica en que el álgebra de «variables aleatorias» en cuestión es no conmutativa. Por esta misma razón la noción de «independencia» también difiere a la del caso clásico, tal y como se verá en los siguientes capítulos.

Por lo tanto, en lo sucesivo trabajaremos con «variables aleatorias independientes» en un espacio (\mathcal{A}, φ) , donde \mathcal{A} es un álgebra no conmutativas y $\varphi : \mathcal{A} \rightarrow \mathbb{C}$ es un funcional lineal. La existencia de un álgebra \mathcal{A} de las características requeridas no es clara a priori. Este es un hecho técnico que necesita demostración, análogo a demostrar la existencia de un espacio de probabilidad clásico en el que viven familias de variables aleatorias independientes. Este aspecto de la teoría no será tratado en esta tesis, remitimos al lector interesado al libro de Nica y Speicher [19].

Capítulo 2

Probabilidad libre escalar

En este capítulo se dan algunos elementos básicos de la probabilidad libre escalar, en particular, se da la definición de variables aleatorias (no conmutativas) libres y se muestra la relación entre momentos, cumulantes y ciertas transformadas de dichas variables aleatorias. Comenzamos dando las definiciones básicas de la probabilidad libre escalar, a saber, las de espacio de probabilidad no conmutativo y variable aleatoria no conmutativa. A manera de aplicación demostramos el Teorema del Límite Central Libre (TLCL), para lo cual es necesaria la noción de convergencia en distribución de variables aleatorias no conmutativas.

2.1. Definiciones básicas

Antes de estudiar a nuestros principales objetos de estudio, recordemos las siguientes definiciones.

Definición 2.1. Definimos un álgebra \mathcal{A} sobre un campo \mathbb{K} como un espacio vectorial sobre \mathbb{K} el cual tiene definida una operación $\cdot : \mathcal{A} \times \mathcal{A} \rightarrow \mathcal{A}$ tal que para cualesquiera $\alpha \in \mathbb{K}$ y $x, y, z \in \mathcal{A}$ se tiene que

- (1) $\alpha(xy) = (\alpha x)y = x(\alpha y)$,
- (2) $x(y + z) = xy + xz$,

$$(3) (x + y)z = xz + yz.$$

Si para todas $x, y \in \mathcal{A}$ se tiene que $xy = yx$, se dice que \mathcal{A} es un álgebra conmutativa.

Definición 2.2. Definimos una $*$ -álgebra \mathcal{A} como un álgebra la cual tiene una operación antilineal $*$: $\mathcal{A} \rightarrow \mathcal{A}$ ($a \mapsto a^*$) tal que $(a^*)^* = a$ y $(ab)^* = b^*a^*$ para todas $a, b \in \mathcal{A}$.

Definición 2.3. Decimos que una $*$ -álgebra \mathcal{A} es una C^* -álgebra si es un espacio de Banach respecto a una norma $\|\cdot\|: \mathcal{A} \rightarrow [0, \infty)$ la cual satisface que $\|x^*x\| = \|x^*\| \|x\|$ para toda $x \in \mathcal{A}$.

En este contexto definimos los espacios de probabilidad no conmutativos.

Definición 2.4. Un espacio de probabilidad no conmutativo es un par (\mathcal{A}, φ) donde \mathcal{A} es un álgebra unital compleja y $\varphi : \mathcal{A} \rightarrow \mathbb{C}$ es un funcional lineal con $\varphi(\mathbf{1}_{\mathcal{A}}) = 1$, donde $\mathbf{1}_{\mathcal{A}}$ es la identidad en \mathcal{A} . A los elementos de \mathcal{A} se les llama variables aleatorias no conmutativas.

Decimos que φ es tipo traza si $\varphi(ab) = \varphi(ba)$ para cualesquiera $a, b \in \mathcal{A}$. En este caso decimos que el espacio de probabilidad no conmutativo es tracial.

Comentario 2.1. El término *no conmutativo* es para enfatizar que el álgebra en principio no es conmutativa. Siempre que no cause confusión, en lo sucesivo omitiremos el término *no conmutativo* cuando hablemos de espacios de probabilidad o variables aleatorias no conmutativas. El adjetivo *escalar* en *probabilidad libre escalar* se debe a que φ es \mathbb{C} -valuada (o escalar valuada). En el Capítulo 4 estudiaremos una clase de espacios de probabilidad en los que el funcional lineal en cuestión no es \mathbb{C} -valuado sino \mathcal{B} -valuado, donde \mathcal{B} es una subálgebra unital de \mathcal{A} .

Definición 2.5. Se define un $*$ -espacio de probabilidad como un espacio de probabilidad (\mathcal{A}, φ) tal que \mathcal{A} es una $*$ -álgebra y φ es un funcional positivo, i.e. $\varphi(a^*a) \geq 0$ para toda $a \in \mathcal{A}$.

Decimos que φ es fiel si $\varphi(a^*a) = 0$ si y sólo si $a = \mathbf{0}_{\mathcal{A}}$, donde $\mathbf{0}_{\mathcal{A}}$ denota al elemento neutro en \mathcal{A} .

Ejemplo 2.1. Sea $\mathcal{A} = \text{Mat}_d(\mathbb{C})$ el álgebra de matrices $d \times d$ con coeficientes complejos y $\varphi = \text{tr}_d = \frac{1}{d} \text{tr}$. Es claro que (\mathcal{A}, φ) es un $*$ -espacio de probabilidad, con φ tracial y

fiel.

Ejemplo 2.2. Observemos que el Ejemplo 2.1 puede ser extendido considerando las matrices $d \times d$ cuyos coeficientes son variables aleatorias en $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$, el álgebra de variables aleatorias sobre $(\Omega, \mathcal{F}, \mathbb{P})$ que tienen todos sus momentos. Específicamente, si tomamos $\mathcal{A} = \text{Mat}_d(L^\infty(\Omega, \mathcal{F}, \mathbb{P}))$ y $\varphi = \mathbb{E} \circ \text{tr}_d$ entonces (\mathcal{A}, φ) forma un *-espacio de probabilidad con φ tracial. Además, se tiene que φ es fiel si \mathcal{A} se toma como $\text{Mat}_d(L^\infty(\Omega, \mathcal{F}, \mathbb{P}))$ módulo la relación de equivalencia definida por la igualdad casi segura. En particular, para $d = 1$ se obtiene que $(L^\infty(\Omega, \mathcal{F}, \mathbb{P}), \mathbb{E})$ es un *-espacio de probabilidad donde el álgebra es de hecho conmutativa.

Ejemplo 2.3. Sea \mathcal{A} el álgebra de operadores lineales acotados en un espacio de Hilbert \mathcal{H} . Tomemos $u \in \mathcal{H}$ fijo con $\|u\|^2 = 1$, y definamos $\varphi_u : \mathcal{A} \rightarrow \mathbb{C}$ al mapeo dado por $a \mapsto \langle au, u \rangle$. Es claro que para cada $u \in \mathcal{H}$ la dupla (\mathcal{A}, φ_u) es un *-espacio de probabilidad no necesariamente tracial y no necesariamente fiel. Esto nos muestra que dotar de diferentes funcionales lineales a una misma álgebra puede dar lugar a diferentes espacios de probabilidad.

Ejemplo 2.4. Sea (\mathcal{A}, φ) un *-espacio de probabilidad y $\text{Mat}_d(\mathcal{A})$ el conjunto de matrices $d \times d$ con entradas¹ en \mathcal{A} . Definimos las operaciones $\cdot : \mathbb{C} \times \mathcal{A} \rightarrow \mathcal{A}$, $+$: $\mathcal{A} \times \mathcal{A} \rightarrow \mathcal{A}$, \times : $\mathcal{A} \times \mathcal{A} \rightarrow \mathcal{A}$ y $*$: $\mathcal{A} \rightarrow \mathcal{A}$ mediante

$$\begin{aligned} (\alpha A)_{i,j} &= \alpha A_{i,j}, \\ (A + B)_{i,j} &= A_{i,j} + B_{i,j}, \\ (AB)_{i,j} &= \sum_{k=1}^d A_{i,k} B_{k,j}, \\ (A^*)_{i,j} &= A_{j,i}^*, \end{aligned}$$

para cualesquiera $\alpha \in \mathbb{C}$ y $A, B \in \text{Mat}_d(\mathcal{A})$. Es fácil demostrar que $(\text{Mat}_d(\mathcal{A}), \varphi \circ \text{tr}_d)$ es un *-espacio de probabilidad tracial, y que además $\varphi \circ \text{tr}_d$ es fiel si y sólo si φ es

¹En general, siempre que una matriz tenga variables aleatorias no conmutativas por coeficientes, denominaremos a estos últimos «entradas».

fiel. Este tipo de espacios serán retomados con más detalle en el Capítulo 4, cuando trabajemos con los llamados espacios de probabilidad valuados en operadores.

Para finalizar con los ejemplos de $*$ -espacios de probabilidad, mostraremos uno de los ejemplos clásicos que motivaron el desarrollo de la probabilidad libre.

Ejemplo 2.5. Sea G un grupo. El álgebra de grupo $\mathbb{C}G$ de G es un álgebra compleja la cual tiene una base indexada por los elementos de G , i.e.

$$\mathbb{C}G := \left\{ \sum_{g \in G} \alpha_g g : \alpha_g \in \mathbb{C}, |\{g \in G \mid \alpha_g \neq 0\}| < \infty \right\}.$$

Más aún, $\mathbb{C}G$ tiene estructura de $*$ -álgebra con las operaciones dadas por

$$\begin{aligned} \left(\sum_{g \in G} \alpha_g g \right)^* &= \sum_{g \in G} \overline{\alpha_g} g^{-1}, \\ \left(\sum_{g \in G} \alpha_g g \right) \cdot \left(\sum_{g \in G} \beta_g g \right) &= \sum_{g, h \in G} \alpha_g \beta_h (gh), \\ \left(\sum_{g \in G} \alpha_g g \right) + \left(\sum_{g \in G} \beta_g g \right) &= \sum_{g \in G} (\alpha_g + \beta_g) g. \end{aligned}$$

Si denotamos por $\varphi : \mathbb{C}G \rightarrow \mathbb{C}$ al funcional lineal dado por

$$\varphi \left(\sum_{g \in G} \alpha_g g \right) = \alpha_e$$

con e la identidad en G , entonces $(\mathbb{C}G, \varphi)$ es un $*$ -espacio de probabilidad con φ tracial y fiel.

El Teorema de Gelfand-Naimark afirma que toda C^* -álgebra es isométricamente $*$ -isomorfa a una subálgebra cerrada de los operadores lineales acotados de un cierto espacio de Hilbert [22]. En este sentido, cuando se trabaja con una C^* -álgebra en esencia se está trabajando con un álgebra de operadores. En este tenor se definen los siguientes tipos de variables aleatorias.

Definición 2.6. Sea (\mathcal{A}, φ) un $*$ -espacio de probabilidad y $a \in \mathcal{A}$, decimos que:

- a es autoadjunta si $a^* = a$.
- a es unitaria si $a^*a = aa^* = \mathbf{1}_{\mathcal{A}}$.
- a es normal si $a^*a = aa^*$.
- a es Haar unitaria si es unitaria, i.e. $aa^* = a^*a = \mathbf{1}_{\mathcal{A}}$, y además para $m \in \mathbb{Z}$ se tiene que

$$\varphi(a^m) = \delta_0(m).$$

Las siguientes son algunas propiedades básicas de los espacios de probabilidad.

Proposición 2.1. *Sea (\mathcal{A}, φ) un $*$ -espacio de probabilidad. Entonces las siguientes proposiciones se cumplen:*

- (1) Si $a \in \mathcal{A}$ es una variable aleatoria autoadjunta entonces $\varphi(a) \in \mathbb{R}$.
- (2) $\varphi(x^*) = \overline{\varphi(x)}$ para toda $x \in \mathcal{A}$.

Demostración. Puesto que $\mathbf{1}_{\mathcal{A}} = (\mathbf{1}_{\mathcal{A}}^*)^* = (\mathbf{1}_{\mathcal{A}}\mathbf{1}_{\mathcal{A}}^*)^* = \mathbf{1}_{\mathcal{A}}\mathbf{1}_{\mathcal{A}}^* = \mathbf{1}_{\mathcal{A}}$, tenemos que $u := \frac{a + \mathbf{1}_{\mathcal{A}}}{2}$ y $v := \frac{a - \mathbf{1}_{\mathcal{A}}}{2}$ satisfacen la ecuación $a = u^*u - v^*v$. De la positividad de φ concluimos que

$$\begin{aligned} \varphi(a) &= \varphi(u^*u - v^*v) \\ &= \varphi(u^*u) - \varphi(v^*v) \in \mathbb{R}. \end{aligned}$$

Para demostrar (2), escribamos $x = a + ib$ con $a = \frac{x + x^*}{2}$ y $b = i\frac{x^* - x}{2}$. Observemos a y b son autoadjuntas. Por el primer inciso se tiene que $\varphi(a), \varphi(b) \in \mathbb{R}$, y por lo tanto

$$\begin{aligned} \varphi(x^*) &= \varphi(a - ib) \\ &= \varphi(a) - i\varphi(b) \\ &= \overline{\varphi(a) + i\varphi(b)} \\ &= \overline{\varphi(x)}, \end{aligned}$$

tal y como se quería demostrar. □

Proposición 2.2. Sea (\mathcal{A}, φ) un $*$ -espacio de probabilidad y $a, b \in \mathcal{A}$. Entonces se tiene que

$$|\varphi(b^*a)|^2 \leq \varphi(a^*a)\varphi(b^*b).$$

Demostración. Sea $t \in \mathbb{R}$, la positividad de φ implica que

$$\begin{aligned} 0 &\leq \varphi((tb - a)^*(tb - a)) \\ &= \varphi((tb^* - a^*)(tb - a)) \\ &= t^2\varphi(b^*b) - t\varphi((b^*a)^*) - t\varphi(b^*a) + \varphi(a^*a) \\ &= t^2\varphi(b^*b) - 2t\Re\{\varphi(b^*a)\} + \varphi(a^*a). \end{aligned}$$

Puesto que la forma cuadrática anterior es no negativa definida tenemos que

$$\Re\{\varphi(b^*a)\}^2 \leq \varphi(a^*a)\varphi(b^*b).$$

Utilizando $tb - ia$ en lugar de $tb - a$ obtenemos la desigualdad análoga

$$\Im\{\varphi(b^*a)\}^2 \leq \varphi(a^*a)\varphi(b^*b).$$

Si $\varphi(b^*b) = 0$, las dos desigualdades anteriores implican que

$$\Re\{\varphi(b^*a)\} = 0 \text{ y } \Im\{\varphi(b^*a)\} = 0,$$

dando lugar a la igualdad $|\varphi(b^*a)|^2 = \varphi(a^*a)\varphi(b^*b) = 0$. Supongamos que $\varphi(b^*b) > 0$, entonces

$$\begin{aligned} 0 &\leq \varphi((\varphi(b^*b)a - \varphi(b^*a)b)^*(\varphi(b^*b)a - \varphi(b^*a)b)) \\ &= \varphi\left((\varphi(b^*b)a^* - \overline{\varphi(b^*a)b^*})(\varphi(b^*b)a - \varphi(b^*a)b)\right) \\ &= \varphi(b^*b)^2\varphi(a^*a) - \varphi(b^*b)\overline{\varphi(b^*a)}\varphi(b^*a) - \varphi(b^*b)\varphi(b^*a)\overline{\varphi(b^*a)} + |\varphi(b^*a)|^2\varphi(b^*b) \\ &= \varphi(b^*b)\left[\varphi(b^*b)\varphi(a^*a) - |\varphi(b^*a)|^2\right]. \end{aligned}$$

Puesto que $\varphi(b^*b) > 0$, la última desigualdad implica el resultado deseado. \square

Comentario 2.2. La desigualdad del teorema anterior es conocida como la desigualdad de Cauchy-Schwarz. Si (\mathcal{A}, φ) es un $*$ -espacio de probabilidad tal que existe $a \in \mathcal{A}$ con $a \neq \mathbf{0}_{\mathcal{A}}$ y $\varphi(a^*a) = 0$, entonces la proposición anterior implica que $\varphi(b^*a) = \varphi(ba) = 0$ para toda $b \in \mathcal{A}$. Observemos que estos casos degenerados no suceden en espacios de probabilidad donde φ es fiel.

A continuación damos la definición de *distribución de una variable aleatoria no conmutativa*. Dicha definición tiene dos versiones: la algebraica y la analítica. Como veremos más adelante, estas dos versiones son compatibles para variables aleatorias normales en C^* -espacios de probabilidad (véase la Definición 2.9).

Definición 2.7. Sea (\mathcal{A}, φ) un $*$ -espacio de probabilidad y a una variable aleatoria en \mathcal{A} . La $*$ -distribución algebraica de a es el funcional lineal $\mu : \mathbb{C}\langle X, X^* \rangle \rightarrow \mathbb{C}$ el cual satisface que para cualesquiera $k \in \mathbb{N}$ y $\epsilon(1), \dots, \epsilon(k) \in \{1, *\}$ se tiene que

$$\mu(X^{\epsilon(1)} \dots X^{\epsilon(k)}) = \varphi(a^{\epsilon(1)} \dots a^{\epsilon(k)}).$$

A las expresiones de la forma $\varphi(a^{\epsilon(1)} \dots a^{\epsilon(k)})$ para $k \in \mathbb{N}$ y $\epsilon(j) \in \{1, *\}$ se les denomina los $*$ -momentos de a .

Comentario 2.3. Puesto que

$$\mathbb{C}\langle X, X^* \rangle = \bigoplus_{\substack{k \in \mathbb{N} \\ \epsilon(1), \dots, \epsilon(k) \in \{1, *\}}} \langle X^{\epsilon(1)} \dots X^{\epsilon(k)} \rangle,$$

la $*$ -distribución algebraica está bien definida.

Definición 2.8. Sea (\mathcal{A}, φ) un $*$ -espacio de probabilidad, a una variable aleatoria normal en \mathcal{A} y μ una medida de probabilidad sobre \mathbb{C} con soporte compacto. Llamaremos a μ la $*$ -distribución analítica de a si para toda $j, k \in \mathbb{N}$ se tiene que

$$\int z^j \bar{z}^k d\mu(z) = \varphi(a^j (a^*)^k).$$

Comentario 2.4. Supongamos que μ y ν son $*$ -distribuciones analíticas de un elemento normal a en \mathcal{A} . Por definición μ y ν tienen soporte compacto, y por lo tanto existe $K > 0$ tal que

$$\text{supp}(\mu), \text{supp}(\nu) \subset \mathcal{B}_0(K).$$

Definamos $\Phi_\mu : C(\mathcal{B}_0(K), \mathbb{C}) \rightarrow \mathbb{C}$ mediante

$$\Phi_\mu(f) = \int f(z) d\mu(z).$$

Análogamente definamos Φ_ν . Dado que $|\Phi_\mu(f)| \leq \int K d\mu(z) = K$, es claro que Φ_μ y Φ_ν son funcionales lineales acotados. Puesto que μ y ν son $*$ -distribuciones analíticas de a , se tiene que $\Phi_\mu(p) = \Phi_\nu(p)$ para todo $p \in \mathbb{C}\langle z, \bar{z} \rangle$. Dado que $\mathcal{B}_0(K)$ es un subconjunto compacto de \mathbb{C} y los polinomios son una $*$ -álgebra que contiene a las constantes y separa puntos (ya que la identidad $z \mapsto z$ lo hace), el Teorema de Stone-Weierstrass [16] implica que $\mathbb{C}\langle z, \bar{z} \rangle$ es denso en $C(\mathcal{B}_0(K), \mathbb{C})$. Por lo tanto, Φ_μ está determinada por los valores que toma en $\mathbb{C}\langle z, \bar{z} \rangle$, y al coincidir estos últimos con los de Φ_ν se tiene que $\Phi_\mu = \Phi_\nu$. Observemos que $\Phi_\mu = \Phi_\nu$ implica que $\int f(z) d\mu(z) = \int f(z) d\nu(z)$ para toda $f \in C(\mathcal{B}_0(K), \mathbb{C})$, sin embargo, la última igualdad se puede extender trivialmente para toda $f \in C(\mathbb{C}, \mathbb{C})$ y por lo tanto $\mu = \nu$. Esto nos muestra que de existir una $*$ -distribución analítica esta es única, y por lo tanto la definición anterior es consistente.

Definición 2.9. Un C^* -espacio de probabilidad es un $*$ -espacio de probabilidad (\mathcal{A}, φ) tal que \mathcal{A} es una C^* -álgebra.

Definición 2.10. Sea \mathcal{A} una C^* -álgebra unital y a una variable aleatoria en \mathcal{A} , el espectro $\text{Sp}(a)$ de a está definido por

$$\text{Sp}(a) = \{z \in \mathbb{C} \mid z\mathbf{1}_{\mathcal{A}} - a \text{ no es invertible}\}.$$

Los siguientes tres teoremas, provenientes de la teoría de C^* -álgebras, constituyen la herramienta necesaria para mostrar la conexión entre ciertos espacios de probabilidad no conmutativos y ciertos espacios de probabilidad clásicos. Los detalles acerca de dicha

conexión se encuentran en el Comentario 2.5. Una exposición más detallada de estos tres teoremas puede ser encontrada en [19].

Teorema 2.1. *Sea \mathcal{A} una C^* -álgebra unital, entonces las siguientes proposiciones se cumplen:*

- *Si $a \in \mathcal{A}$ entonces el conjunto $\text{Sp}(a)$ es un subconjunto compacto no vacío de \mathbb{C} . Además, se tiene que $\text{Sp}(a) \subset \mathcal{B}_0(\|a\|)$.*
- *Si $a \in \mathcal{A}$ es una variable aleatoria normal entonces $\|a\| = \|a^*\|$.*
- *$a \in \mathcal{A}$ es una variable aleatoria autoadjunta si y sólo si $\text{Sp}(a) \subset \mathbb{R}$.*
- *$a \in \mathcal{A}$ es una variable aleatoria unitaria si y sólo si $\text{Sp}(a) \subset \mathcal{S}^1 \subset \mathbb{C}$, donde \mathcal{S}^1 representa a los números complejos de norma unitaria.*

Teorema 2.2. *Sea \mathcal{A} una C^* -álgebra unital y $a \in \mathcal{A}$ una variable aleatoria normal. Entonces existe un mapeo $\Phi : C(\text{Sp}(a), \mathbb{C}) \rightarrow \mathcal{A}$ el cual satisface las siguientes propiedades:*

- *Φ es un homomorfismo de $*$ -álgebras unitales.*
- *$\|\Phi(f)\| = \|f\|_\infty$ para toda $f \in C(\text{Sp}(a), \mathbb{C})$, donde $\|f\|_\infty = \sup_{z \in \text{Sp}(a)} |f(z)|$.*
- *$\Phi(z \mapsto z) = a$.*

Además, Φ también satisface las siguientes propiedades:

- *Φ es inyectivo.*
- *Φ es continuo.*
- *Φ es único.*

Para $f \in C(\text{Sp}(a), \mathbb{C})$ se define $f(a) := \Phi(f)$. A Φ se le conoce como el cálculo funcional para funciones continuas de a .

Teorema 2.3. *Sea (\mathcal{A}, φ) un C^* -espacio de probabilidad y $a \in \mathcal{A}$ una variable aleatoria normal. Entonces a tiene una $*$ -distribución analítica μ la cual satisface que:*

- (1) *Para toda $f \in C(\text{Sp}(a), \mathbb{C})$ se tiene que*

$$\int f(z) d\mu(z) = \varphi(f(a)).$$

(2) $\text{supp}(\mu) \subseteq \text{Sp}(a)$, con igualdad si φ es fiel.

Teorema 2.4. *Sea μ una medida de probabilidad en \mathbb{R} . Si μ tiene soporte compacto entonces está determinada por sus momentos, es decir, si ν es otra medida de probabilidad en \mathbb{R} tal que $\int x^m d\mu = \int x^m d\nu$ para toda $m \in \mathbb{N}$ entonces $\nu = \mu$.*

La demostración del teorema anterior es un ejercicio de rutina. Basta observar que existe $K > 0$ tal que para toda $m \in \mathbb{N}$ se tiene que

$$\left| \int_{\mathbb{R}} x^m d\mu(x) \right| \leq K^m,$$

y después escribir la transformada de Fourier como una serie de potencias utilizando los momentos de μ (que coinciden con los de ν).

Comentario 2.5. Sea (\mathcal{A}, φ) un C^* -espacio de probabilidad y $a \in \mathcal{A}$ una variable aleatoria autoadjunta. El Teorema 2.1 implica que $\text{Sp}(a)$ es un subconjunto compacto de \mathbb{R} , mientras que el Teorema 2.3 asegura la existencia de la distribución analítica, digamos μ , tal que $\text{supp}(\mu) \subset \text{Sp}(a)$. Por lo tanto, el teorema anterior nos muestra que la distribución analítica y la distribución algebraica contienen la misma información, i.e. una determina a la otra y viceversa. En resumen, en el contexto de variables aleatorias normales en una C^* -álgebra, es suficiente trabajar con las distribuciones algebraicas ya que estas contienen implícitamente toda la información acerca de las medidas de probabilidad correspondientes. En esta tesis, generalmente trabajaremos con variables aleatorias autoadjuntas en C^* -espacios de probabilidad.

2.2. Variables aleatorias libres

El objetivo de esta sección es presentar una noción de «independencia» entre variables aleatorias en un espacio de probabilidad no conmutativo. En particular, daremos la definición de variables aleatorias libres y calcularemos algunos ejemplos relacionados. Estos ejemplos nos ayudarán a mostrar las similitudes y diferencias entre las expresiones

para los momentos mixtos de variables aleatorias independientes en el sentido clásico y en el sentido libre.

Comentario 2.6. Sean X, Y variables aleatorias independientes en un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$. Entonces la siguiente igualdad se cumple para cualesquiera $j, k \in \mathbb{N}$

$$\mathbb{E}(X^j Y^k) = \mathbb{E}(X^j) \mathbb{E}(Y^k).$$

En particular, tenemos que los momentos mixtos, i.e. el valor esperado de productos que contienen a X y Y , se pueden calcular via los momentos individuales de X y Y . En esta sección mostraremos que si x, y son variables aleatorias libres en un espacio de probabilidad (\mathcal{A}, φ) entonces la distribución algebraica mixta está determinada por las distribuciones algebraicas individuales.

Definición 2.11. Sea I un conjunto arbitrario, (\mathcal{A}, φ) un $*$ -espacio de probabilidad y $(a_i)_{i \in I}$ variable aleatorias en \mathcal{A} . La $*$ -distribución algebraica de las variables aleatorias $(a_i)_{i \in I}$ es el funcional lineal $\mu : \mathbb{C}\langle \{X_i, X_i^* \mid i \in I\} \rangle \rightarrow \mathbb{C}$ el cual satisface que para cualesquiera $k \in \mathbb{N}$, $i(1), \dots, i(k) \in I$ y $\epsilon(1), \dots, \epsilon(k) \in \{1, *\}$ se tiene que

$$\mu \left(X_{i(1)}^{\epsilon(1)} \cdots X_{i(k)}^{\epsilon(k)} \right) = \varphi \left(a_{i(1)}^{\epsilon(1)} \cdots a_{i(k)}^{\epsilon(k)} \right).$$

A las expresiones de la forma $\varphi \left(a_{i(1)}^{\epsilon(1)} \cdots a_{i(k)}^{\epsilon(k)} \right)$ para $k \in \mathbb{N}$, $i(j) \in I$ y $\epsilon(j) \in \{1, *\}$ se les denomina los $*$ -momentos mixtos de $(a_i)_{i \in I}$.

Definición 2.12. Sea (\mathcal{A}, φ) un $*$ -espacio de probabilidad, I un conjunto arbitrario y \mathcal{A}_i una subálgebra unital de \mathcal{A} para cada $i \in I$. Decimos que las álgebras $\{\mathcal{A}_i \mid i \in I\}$ son libres si: para todo $n \in \mathbb{N}$ y variables aleatorias $a_j \in \text{Ker}(\varphi) \cap \mathcal{A}_{i(j)}$ ($i(j) \in I, j \in \{1, \dots, n\}$) tales que² $i(1) \neq i(2), i(2) \neq i(3), \dots, i(n-1) \neq i(n)$, se tiene que

$$\varphi(a_1 \dots a_n) = 0.$$

²Observemos que está permitido que $i(1) = i(3)$ y que $i(1) = i(n)$.

Decimos que las familias de variables aleatorias $\{a_1^{(i)}, \dots, a_{n_i}^{(i)}\}$ para $i \in I$ son libres si las álgebras $\langle \mathbf{C}\mathbf{1}_{\mathcal{A}}, a_1^{(i)}, \dots, a_{n_i}^{(i)} \rangle$ para $i \in I$ son libres.

Ejemplo 2.6. Sean $\{a\}, \{b\} \subset \mathcal{A}$ libres. Puesto que $\varphi(a - \varphi(a)\mathbf{1}) = 0$, de la definición de libertad se tiene que

$$\begin{aligned} 0 &= \varphi((a - \varphi(a)\mathbf{1})(b - \varphi(b)\mathbf{1})) \\ &= \varphi(ab - \varphi(b)a - \varphi(a)b + \varphi(a)\varphi(b)\mathbf{1}) \\ &= \varphi(ab) - \varphi(a)\varphi(b). \end{aligned}$$

Por lo tanto, para $\{a\}, \{b\}$ libres se tiene que $\varphi(ab) = \varphi(a)\varphi(b)$.

Comentario 2.7. Sean X, Y variables aleatorias independientes en un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$, entonces $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$. Por lo tanto, la expresión encontrada en el inciso anterior coincide con la fórmula del caso clásico.

Notación 2.1. En lo sucesivo, si z es un número complejo denotaremos a $z\mathbf{1}_{\mathcal{A}}$ únicamente por z . Notemos que este abuso de notación es consistente con la identificación natural $\iota: \mathbb{C} \rightarrow \mathcal{A}$ dada por $\iota(z) = z\mathbf{1}_{\mathcal{A}}$. Además, para $a, b \in \mathcal{A}$ utilizaremos la expresión “ $a, b \in \mathcal{A}$ libres” en lugar de “ $\{a\}, \{b\} \subset \mathcal{A}$ libres”.

Ejemplo 2.7. Sean $\{b\}$ y $\{a_1, a_2\}$ libres, entonces

$$\begin{aligned} 0 &= \varphi((a_1 - \varphi(a_1))(b - \varphi(b))(a_2 - \varphi(a_2))) \\ &= \varphi(a_1(b - \varphi(b))(a_2 - \varphi(a_2))) - \varphi(a_1)\varphi((b - \varphi(b))(a_2 - \varphi(a_2))). \end{aligned} \quad (2.1)$$

De la definición de libertad, el segundo término de la última expresión se anula. Por lo tanto, la ecuación (2.1) implica que

$$\begin{aligned} 0 &= \varphi(a_1(b - \varphi(b))(a_2 - \varphi(a_2))) \\ &= \varphi(a_1ba_2) - \varphi(a_2)\varphi(a_1b) - \varphi(b)\varphi(a_1a_2) + \varphi(b)\varphi(a_1)\varphi(a_2) \\ &= \varphi(a_1ba_2) - \varphi(b)\varphi(a_1a_2), \end{aligned}$$

o equivalentemente $\varphi(a_1ba_2) = \varphi(b)\varphi(a_1a_2)$.

Comentario 2.8. Observemos que la expresión encontrada en el ejemplo anterior coincide nuevamente con la del caso clásico, i.e. $\mathbb{E}(Y_1XY_2) = \mathbb{E}(X)\mathbb{E}(Y_1Y_2)$ siempre que X sea independiente de $\{Y_1, Y_2\}$.

Ejemplo 2.8. Sean a, b elementos en \mathcal{A} libres. Procediendo como en el ejemplo anterior se obtiene que $\varphi(abab) = \varphi(a^2)\varphi(b)^2 + \varphi(a)^2\varphi(b^2) - \varphi(a)^2\varphi(b)^2$. Esta fórmula ya no es igual a la del caso clásico: $\mathbb{E}(abab) = \mathbb{E}(a^2b^2) = \mathbb{E}(a^2)\mathbb{E}(b^2)$.

Comentario 2.9. Los ejemplos anteriores nos muestran que la libertad imita a la independencia en el sentido que los momentos mixtos de variables aleatorias libres pueden ser calculados a partir de los momentos marginales.

Al menos en el caso en el que φ es tracial, es natural que la primer expresión cuyo «valor esperado» difiera del caso clásico sea $abab$. Esto debido a que no hay manera de llevar $\varphi(abab)$ a $\varphi(a^2b^2)$, no así para aba donde se tiene que $\varphi(aba) = \varphi(a^2b)$.

2.3. Teorema del límite central libre

Definición 2.13. Sean (\mathcal{A}, φ) y $(\mathcal{A}_N, \varphi_N)$ para $N \in \mathbb{N}$ espacios de probabilidad. Consideremos variables aleatorias $a \in \mathcal{A}$ y $a_N \in \mathcal{A}_N$ para cada $N \in \mathbb{N}$. Decimos que a_N converge en distribución a a cuando $N \rightarrow \infty$ si

$$\lim_{N \rightarrow \infty} \varphi_N(a_N^m) = \varphi(a^m)$$

para toda $m \in \mathbb{N}$. Denotamos a esta convergencia por $a_N \xrightarrow{\text{distr}} a$.

Similarmente, decimos que a_N converge en *-distribución a a (denotado por $a_N \xrightarrow{*-\text{dist}} a$) si en lugar de la convergencia de momentos se tiene la convergencia de los *-momentos.

Comentario 2.10. Esta convergencia lo único que implica es la convergencia de momentos, sin embargo, si la distribución límite esta determinada por sus momentos entonces esta convergencia en distribución algebraica implica la convergencia de las dis-

tribuciones analíticas.

Para dar una aplicación a la teoría expuesta hasta el momento, a continuación bosquejamos la demostración del Teorema del Límite Central en el caso libre.

Definición 2.14. Sea (\mathcal{A}, φ) un espacio de probabilidad. Decimos que $s \in \mathcal{A}$ es un elemento semicircular estandar si sus momentos impares son cero y sus momentos pares están dados por los números de Catalan, i.e. para toda $m \in \mathbb{N}$ se tiene que

$$\varphi(s^m) = \begin{cases} 0 & \text{si } m \text{ es impar,} \\ \frac{1}{k+1} \binom{2k}{k} & \text{si } m = 2k \text{ con } k \in \mathbb{N}. \end{cases}$$

Decimos que $c \in \mathcal{A}$ es un elemento circular estandar si

$$c = \frac{s + it}{\sqrt{2}},$$

donde s y t son elementos semicirculares estandar libres.

Una forma de obtener los números de Catalan es mediante el número de emparejamientos que no se cruzan del conjunto $\{1, \dots, m\}$. Este resultado se utiliza al demostrar el Teorema del Límite Central Libre.

Teorema 2.5 (Teorema del Límite Central Libre). *Sea (\mathcal{A}, φ) un $*$ -espacio de probabilidad y $(a_n)_{n \in \mathbb{N}} \subset \mathcal{A}$ una sucesión de variables aleatorias autoadjuntas libres idénticamente distribuidas. Más aún, supongamos que todas las variables son centradas y con varianza unitaria, i.e. $\varphi(a_n) = 0$ y $\varphi(a_n^2) = 1$ para toda $n \in \mathbb{N}$. Entonces*

$$\frac{a_1 + \dots + a_N}{\sqrt{N}} \xrightarrow{\text{distr}} s,$$

donde s es una variable aleatoria semicircular estandar.

Definición 2.15. Sea $\pi = \{V_1, \dots, V_l\}$ una partición del conjunto $\{1, \dots, n\}$, i.e.

V_1, \dots, V_l son subconjuntos no vacíos de $\{1, \dots, n\}$ tales que

$$\{1, \dots, n\} = \bigcup_{k=1}^l V_k$$

y $V_i \cap V_j = \emptyset$ siempre que $i \neq j$. Decimos que π es una partición que no se cruza si siempre que $a_1, b_1 \in V_{i_1}$ y $a_2, b_2 \in V_{i_2}$ con $a_1 < a_2 < b_1 < b_2$ se tiene que $i_1 = i_2$.

Decimos que una partición $\pi = \{V_1, \dots, V_l\}$ es un emparejamiento si cada bloque V_k contiene exactamente dos elementos.

Bosquejo de la Demostración del Teorema 2.5. Para cada $N \in \mathbb{N}$ sea

$$S_N = N^{-1/2} \sum_{k=1}^N a_k,$$

entonces para $m \in \mathbb{N}$ se tiene que

$$\varphi(S_N^m) = N^{-m/2} \sum_{\substack{i(k)=1, \dots, N \\ k \in \{1, \dots, m\}}} \varphi(a_{i(1)} \cdots a_{i(m)}). \quad (2.2)$$

Si para algún término $a_{i(1)} \cdots a_{i(m)}$ de la suma anterior existe un índice único, i.e. existe $j \in \{1, \dots, m\}$ tal que $i(j) \neq i(k)$ para toda $k \neq j$, entonces por el Ejemplo 2.7 y el hecho que $a_{i(j)} \in \text{Ker}(\varphi)$ tenemos que dicho término se anula. Por lo tanto, en la suma anterior solo aparecen los términos cuyos índices están dos o mas veces.

Puesto que las variables $(a_n)_{n \in \mathbb{N}}$ tienen la misma distribución, para una partición $\pi = \{V_1, \dots, V_l\}$ de $\{1, \dots, m\}$ hay

$$N(N-1) \cdots (N-|\pi|+1) \leq N^{|\pi|}$$

términos en la suma de la ecuación (2.2) con valor $\varphi(a_\pi) := \varphi(a_{p_1} \cdots a_{p_m})$ donde $p_k = j$ si $k \in V_j$. De las observaciones anteriores, en particular de la última desigualdad, se

obtiene que

$$\lim_{N \rightarrow \infty} \varphi(S_N^m) = \sum_{\pi \in P} \varphi(a_\pi) \quad (2.3)$$

donde P son todos los emparejamientos del conjunto $\{1, \dots, m\}$. Para finalizar, del Ejemplo 2.8 concluimos que las únicas $\varphi(a_\pi)$ que no se anulan son aquellas que provienen de los emparejamientos que no se cruzan. Más aún, los términos que no se anulan satisfacen que $\varphi(a_\pi) = 1$. Por lo tanto, concluimos que

$$\lim_{N \rightarrow \infty} \varphi(S_N^m) = \begin{cases} 0 & \text{si } m \text{ es impar,} \\ \frac{1}{k+1} \binom{2k}{k} & \text{si } m = 2k \text{ con } k \in \mathbb{N}, \end{cases}$$

tal y como queríamos demostrar. \square

Comentario 2.11. Es posible dar una prueba alternativa para el Teorema del Límite Central clásico siguiendo las ideas de la demostración anterior. Específicamente, es posible imitar la demostración hasta la ecuación (2.3). En este punto, la conmutatividad del caso clásico implicaría que $\varphi(a_\pi) = 1$ para todo emparejamiento $\pi \in P$, y por lo tanto

$$\lim_{N \rightarrow \infty} \varphi(S_N^m) = \begin{cases} 0 & \text{si } m \text{ es impar,} \\ \frac{(2k)!}{2^k k!} & \text{si } m = 2k \text{ con } k \in \mathbb{N}. \end{cases}$$

Puesto que las cantidades anteriores son los momentos de la variable aleatoria gaussiana estandar, y esta está determinada por sus momentos, el resultado se sigue. Es importante señalar que para demostrar el TLCL utilizamos los emparejamientos que no se cruzan, mientras que en el caso del TLC clásico utilizamos todos los emparejamientos (tanto los que se cruzan como los que no).

Es posible extender el teorema anterior al caso multidimensional.

Definición 2.16. Sea I un conjunto arbitrario y $(c_{ij})_{i,j \in I} \subset \mathbb{C}$. Una familia $(s_i)_{i \in I}$ de variables aleatorias autoadjuntas en un $*$ -espacio de probabilidad es llamada una familia semicircular de covarianza $(c_{ij})_{i,j \in I}$ si para toda $n \in \mathbb{N}$ y para todas $i(1), \dots, i(n) \in I$

se tiene que

$$\varphi \left(\prod_{j=1}^n s_{i(j)} \right) = \sum_{\pi \in \text{NC}_2(n)} \prod_{(p,q) \in \pi} c_{i(p), i(q)},$$

donde $\text{NC}_2(n)$ denota a los emparejamientos que no se cruzan del conjunto $\{1, \dots, n\}$.

Definición 2.17. Sean (\mathcal{A}, φ) y $(\mathcal{A}_N, \varphi_N)$ para $N \in \mathbb{N}$ espacios de probabilidad. Para I un conjunto arbitrario, consideremos familias de variables aleatorias $(a^{(i)})_{i \in I} \subset \mathcal{A}$ y $(a_N^{(i)})_{i \in I} \subset \mathcal{A}_N$ para cada $N \in \mathbb{N}$. Decimos que $a_N^{(i)}$ converge en distribución a $a^{(i)}$ cuando $N \rightarrow \infty$ si los momentos mixtos de $(a_N^{(i)})_{i \in I}$ convergen a los correspondientes de $(a^{(i)})_{i \in I}$.

Teorema 2.6. Sea (\mathcal{A}, φ) un $*$ -espacio de probabilidad, I un conjunto arbitrario y $(a_N^{(i)})_{i \in I} \subset \mathcal{A}$ para cada $N \in \mathbb{N}$. Supongamos que las familias $(a_N^{(i)})_{i \in I}$ son familias libres de variables aleatorias autoadjuntas con distribuciones conjuntas idénticas. Si además las variables aleatorias de cada familia son centradas con covarianza $c_{ij} := \varphi \left(a_N^{(i)} a_N^{(j)} \right)$ entonces

$$\left(\frac{a_1^{(i)} + \dots + a_N^{(i)}}{\sqrt{N}} \right)_{i \in I} \xrightarrow{\text{dist}} (s_i)_{i \in I},$$

donde $(s_i)_{i \in I}$ es una familia semicircular con covarianza $(c_{ij})_{i,j \in I}$.

2.4. Cumulantes y transformadas

Análogamente al caso clásico, en probabilidad libre se tiene la noción de cumulantes. Dichos cumulantes son de particular interés al trabajar con las transformadas que a continuación se definen.

Notación 2.2. Denotemos por $\text{NC}(n)$ al conjunto de particiones que no se cruzan del conjunto $\{1, \dots, n\}$, y por $\text{NC} := \bigcup_{n \geq 1} \text{NC}(n)$.

Ejemplo 2.9. Supongamos que buscamos funcionales multilineales $\kappa_\pi : \mathcal{A}^n \rightarrow \mathbb{C}$ para

cada $\pi \in \text{NC}$ tales que

$$\varphi(a_1 \cdots a_n) = \sum_{\pi \in \text{NC}(n)} \kappa_\pi(a_1, \dots, a_n).$$

Por simplicidad denotaremos por κ_n a $\kappa_{\{1, \dots, n\}}$. De existir dichos funcionales, a los cuales llamaremos cumulantes, de la fórmula recursiva anterior se debe tener que

$$\varphi(a) = \kappa_1(a),$$

y por lo tanto $\kappa_1(a) = \varphi(a)$. Procediendo de manera similar

$$\begin{aligned} \varphi(a_1 a_2) &= \sum_{\pi \in \text{NC}(2)} \kappa_\pi(a_1, a_2) \\ &= \kappa_2(a_1, a_2) + \kappa_{\{1\}\{2\}}(a_1, a_2). \end{aligned}$$

Una propiedad que es deseable es que κ sea multiplicativo respecto a π , i.e. que

$$\kappa_\pi(a_1, \dots, a_n) = \prod_{\substack{V \in \pi \\ V = \{i_1, \dots, i_{|V|}\}}} \kappa_{|V|}(a_{i_1}, \dots, a_{i_{|V|}}).$$

En tal caso, se tiene que $\kappa_{\{1\}\{2\}}(a_1, a_2) = \kappa_1(a_1)\kappa_2(a_2)$ y por lo tanto $\kappa_2(a_1, a_2) = \varphi(a_1 a_2) - \varphi(a_1)\varphi(a_2)$. Procediendo de manera análoga obtenemos que

$$\begin{aligned} \kappa_3(a_1, a_2, a_3) &= \varphi(a_1 a_2 a_3) - \varphi(a_1 a_2)\varphi(a_3) - \varphi(a_1 a_3)\varphi(a_2) \\ &\quad - \varphi(a_2 a_3)\varphi(a_1) + 2\varphi(a_1)\varphi(a_2)\varphi(a_3). \end{aligned}$$

Comentario 2.12. En el ejemplo anterior pudimos construir hasta κ_3 bajo la condición de que los funcionales κ_π sean multiplicativos respecto a π . Sin embargo, no tenemos la certeza de poder continuar *ad infinitum*. La existencia de tal familia de funcionless se demuestra vía la teoría general de la inversión de Möbius. En el Apéndice A se exponen brevemente algunas de las nociones principales de dicha teoría, la cual nos permite

demostrar el siguiente teorema.

Teorema 2.7. *Sea (\mathcal{A}, φ) un espacio de probabilidad. Entonces existen funcionales multilineales $(\kappa_\pi)_{\pi \in \text{NC}}$ los cuales son multiplicativos en π y satisfacen la siguiente ecuación*

$$\varphi(a_1 \dots a_n) = \sum_{\pi \in \text{NC}(n)} \kappa_\pi(a_1, \dots, a_n). \quad (2.4)$$

para toda $a_1, \dots, a_n \in \mathcal{A}$.

Ejemplo 2.10. Si $\pi = \{\{1, 8\}, \{2, 6\}, \{3, 4, 5\}, \{7\}\}$ entonces

$$\kappa_\pi(a_1, \dots, a_8) = \kappa_2(a_1, a_8) \kappa_2(a_2, a_6) \kappa_3(a_3, a_4, a_5) \kappa_1(a_7).$$

Notación 2.3. *Denotamos por κ_n^a a $\kappa_n(a, \dots, a)$.*

Observación 2.1. De la multiplicatividad de la familia $(\kappa_\pi)_{\pi \in \text{NC}}$ es claro que dicha familia está determinada por $(\kappa_n)_{n \geq 1}$ y viceversa.

A partir de la fórmula de momentos-cumulantes y las fórmulas de inversión de Möbius (ecuaciones (2.4), (A.1) y (A.2) respectivamente) se pueden derivar algunos resultados acerca de los cumulantes. De particular importancia son los siguientes dos teoremas.

Teorema 2.8. *Sea (\mathcal{A}, φ) un espacio de probabilidad y $(\kappa_n)_{n \in \mathbb{N}}$ los cumulantes respectivos. Si $(\mathcal{A}_i)_{i \in I}$ son subálgebras unitales de \mathcal{A} con I un conjunto arbitrario, entonces las familias $(\mathcal{A}_i)_{i \in I}$ son libres si y sólo si para toda $n \geq 2$ y para cualesquiera $a_j \in \mathcal{A}_{i(j)}$ con $j = 1, \dots, n$ se tiene que $\kappa_n(a_1, \dots, a_n) = 0$ siempre que exista $k \in \{1, \dots, n\}$ tal que $i(k) \neq i(j)$ para algún $j \neq k$.*

Comentario 2.13. El teorema anterior caracteriza la libertad de variables aleatorias mediante la cancelación de los cumulantes mixtos. El siguiente corolario es una consecuencia directa de este teorema.

Corolario 2.1. Si $a, b \in \mathcal{A}$ son variables aleatorias libres entonces para toda $n \in \mathbb{N}$ se tiene que

$$\kappa_n^{a+b} = \kappa_n^a + \kappa_n^b. \quad (2.5)$$

Definición 2.18. Sea (\mathcal{A}, φ) un espacio de probabilidad y a una variable aleatoria con distribución analítica μ sobre \mathbb{R} . Se define la transformada de Cauchy $G_a : \mathbb{C} \setminus \mathbb{R} \rightarrow \mathbb{C}$ de a como

$$G_a(z) = \int_{\mathbb{R}} \frac{\mu(dt)}{z - t}.$$

Puesto que G_a depende completamente de μ es usual denotar a G_a por G_μ .

Teorema 2.9. Sean $(m_n^a)_{n \geq 0}$ y $(\kappa_n^a)_{n \geq 1}$ los momentos y cumulantes de una variable aleatoria $a \in \mathcal{A}$ respectivamente. Definamos las series de potencias formales

$$M_a(z) := \sum_{n \geq 0} m_n^a z^n,$$

$$C_a(z) := 1 + \sum_{n \geq 1} \kappa_n^a z^n.$$

Entonces

$$C_a(zM_a(z)) = M_a(z)$$

y además

$$G_a(z) = \frac{1}{z} M_a\left(\frac{1}{z}\right).$$

Comentario 2.14. Más precisamente, la igualdad entre la transformada de Cauchy G_a y la función generadora de momentos M_a se da en una vecindad de infinito para z , o bien una vecindad de cero para $1/z$. Esto no representa ningún problema al trabajar con series de potencias formales. Sin embargo, no es posible recuperar la distribución analítica de a a partir de dicha serie de potencias *per se*, ya que para ello es usual evaluar a la transformada de Cauchy fuera de tal vecindad de infinito.

Definición 2.19. Sea (\mathcal{A}, φ) un espacio de probabilidad y $a \in \mathcal{A}$ una variable aleatoria con distribución analítica μ . Se define su transformada R por

$$\mathcal{R}_a(z) = \sum_{n \geq 0} \kappa_{n+1}^a z^n.$$

Al igual que antes, es usual denotar a \mathcal{R}_a mediante \mathcal{R}_μ .

De la definición anterior y la ecuación (2.5) se tiene la siguiente proposición.

Proposición 2.3. *Sea (\mathcal{A}, φ) un espacio de probabilidad y $a, b \in \mathcal{A}$ dos variables aleatorias. Si a, b son libres entonces*

$$\mathcal{R}_{a+b} = \mathcal{R}_a + \mathcal{R}_b.$$

La proposición anterior afirma que la transformada R linealiza la suma de variables aleatorias libres. Esto muestra que los cumulantes libres son análogos a los cumulantes clásicos en el sentido que la transformada R, o función generadora de cumulantes, linealiza la suma de variables aleatorias libres o independientes según sea el caso. Para finalizar este capítulo presentamos la transformada S, el análogo a la transformada R para la multiplicación de variables aleatorias libres.

Definición 2.20. Sea (\mathcal{A}, φ) un espacio de probabilidad y a una variable aleatoria con distribución analítica μ . Definimos la función generadora de momentos $\Psi_a : \mathbb{C} \setminus \mathbb{R} \rightarrow \mathbb{C}$ de a por

$$\Psi_a(z) = \int_{\mathbb{R}} \frac{zx}{1 - zx} \mu(dx).$$

Al igual que antes, es usual denotar a Ψ_a por Ψ_μ .

Comentario 2.15. La función generadora de momentos recibe su nombre del hecho que para una medida μ de soporte compacto tiene lugar la igualdad

$$\Psi_\mu(z) = \sum_{n \geq 1} m_n(\mu) z^n$$

para z en una vecindad del cero. De hecho, en dicha vecindad del cero se tiene que Ψ_μ está bien definida y tiene una única inversa siempre que $m_1(\mu) \neq 0$.

Definición 2.21. Sea (\mathcal{A}, φ) un espacio de probabilidad. Para una variable aleatoria a con media no nula y distribución analítica μ se define su transformada S , denotada por \mathcal{S}_a , mediante

$$\mathcal{S}_a(z) = \Psi_a^{-1}(z) \frac{1+z}{z},$$

para z en una vecindad del cero.

Proposición 2.4. Sea (\mathcal{A}, φ) un espacio de probabilidad y $a, b \in \mathcal{A}$ variables aleatorias libres, entonces

$$\mathcal{S}_{ab}(z) = \mathcal{S}_a(z)\mathcal{S}_b(z)$$

para z en una vecindad del cero.

2.5. Notas adicionales

Este capítulo está basado principalmente en el libro de Nica y Speicher [19], salvo lo que concierne a la transformada S lo cual está basado en el libro de Arizmendi, Barndorff-Nielsen, Franz, Perez-Abreu, Thorbjensen y Vargas [1].

La demostración de que la cantidad de emparejamientos que no se cruzan del conjunto $\{1, \dots, 2n\}$ es el n -ésimo término de los números de Catalan puede ser encontrada, junto con muchos resultados más acerca de estos números, en el libro de Koshy [15].

Capítulo 3

Probabilidad libre y matrices aleatorias

En el capítulo anterior expusimos algunos elementos básicos de la probabilidad libre, una teoría donde las variables aleatorias no conmutativas son fundamentales. En particular, a partir del concepto de variables aleatorias libres se obtuvieron reglas para calcular la distribución de la suma y producto de variables aleatorias no conmutativas. En el presente capítulo se mostrará brevemente como las matrices aleatorias pueden ponerse en el contexto de la probabilidad libre. Debido al carácter general de este capítulo, omitimos las demostraciones de los teoremas. Las referencias pertinentes se encuentran al final del capítulo.

3.1. Matrices aleatorias

Como veremos en la siguiente sección, una herramienta bastante utilizada en la literatura es la transformada de Cauchy¹.

Notación 3.1. Denotamos por \mathbb{C}^+ al conjunto de números complejos con parte imagi-

¹Otra transformada ampliamente utilizada en la literatura es la transformada de Stieltjes, la cual es el negativo de la transformada de Cauchy.

naria positiva, i.e.

$$\mathbb{C}^+ := \{z \in \mathbb{C} \mid \Im(z) > 0\}.$$

Definición 3.1. Sea μ una medida de probabilidad en \mathbb{R} . Definimos su transformada de Cauchy $G_\mu : \mathbb{C}^+ \rightarrow \mathbb{C}^+ \cup \mathbb{R}$ mediante

$$z \mapsto \int_{\mathbb{R}} \frac{1}{z - x} \mu(dx).$$

Teorema 3.1. Sea μ una medida de probabilidad en \mathbb{R} . Para cualesquiera $a < b$ puntos de continuidad de μ tenemos que

$$\mu[a, b] = \lim_{\epsilon \downarrow 0} \frac{1}{\pi} \int_a^b \Im G_\mu(x + i\epsilon) dx.$$

Teorema 3.2. Sean $(\mu_n)_{n \geq 1}$ y μ medidas de probabilidad en \mathbb{R} . Entonces $\mu_n \xrightarrow{d} \mu$ si y sólo si $G_{\mu_n}(z) \rightarrow G_\mu(z)$ para toda $z \in \mathbb{C}^+$.

El siguiente teorema es particularmente útil para obtener la densidad de una función de distribución de la cual únicamente se conoce su transformada de Cauchy. Esta fórmula es utilizada en las Secciones 5.2 y 5.4 para obtener la densidad de la distribución del espectro asintótico de ciertos canales de comunicación inalámbrica.

Teorema 3.3. Sea μ una medida de probabilidad en \mathbb{R} y $x_0 \in \mathbb{R}$. Si el límite

$$\lim_{\substack{z \in \mathbb{C}^+ \\ z \rightarrow x_0}} \Im G_\mu(z)$$

existe, entonces la función de distribución F^μ de μ es diferenciable en x_0 y además

$$F^\mu(dx_0) = -\frac{1}{\pi} \lim_{\substack{z \in \mathbb{C}^+ \\ z \rightarrow x_0}} \Im G_\mu(z).$$

A continuación definimos los principales objetos de estudio en la teoría de matrices aleatorias: matrices aleatorias y distribución empírica espectral.

Definición 3.2. Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad, $d, e \in \mathbb{N}$ y $M = \text{Mat}_{d \times e}(\mathbb{C})$. Decimos que $\mathbf{X} : (\Omega, \mathcal{F}) \rightarrow (M, \mathfrak{B}(M))$ es una matriz aleatoria si es una función medible.

Definición 3.3. Sea \mathbf{X} una matriz aleatoria en $(\Omega, \mathcal{F}, \mathbb{P})$ con valores en $\text{Mat}_d(\mathbb{C})$. Definimos su medida empírica espectral $\mu^{\mathbf{X}}$ mediante

$$\mu^{\mathbf{X}} = \frac{1}{d} \sum_{k=1}^d \delta_{\lambda_k},$$

donde δ_x es la medida de Dirac en x y $\lambda_1, \dots, \lambda_d$ son los eigenvalores de \mathbf{X} . Denotaremos por $F^{\mathbf{X}}$ a la función de distribución asociada, la cual llamaremos función de distribución empírica espectral de \mathbf{X} .

En general los eigenvalores $\lambda_1, \dots, \lambda_d$ pueden ser números complejos. Sin embargo, usualmente trabajaremos con matrices Hermitianas por lo que los eigenvalores $\lambda_1 \leq \dots \leq \lambda_d$ son reales y más aún, son continuos respecto a las entradas de \mathbf{X} [25]. Puesto que la definición de matriz aleatoria implica la medibilidad de los coeficientes (aleatorios) de \mathbf{X} tenemos que $\lambda_1 \leq \dots \leq \lambda_n$ son variables aleatorias. En particular, $\mu^{\mathbf{X}}$ es un proceso puntual.

Ejemplo 3.1. Para cada $n \in \mathbb{N}$ sea \mathbf{X}_n una matriz de $n \times n$ tal que:

- 1) \mathbf{X}_n es Hermitiana.
- 2) Los coeficientes en la diagonal son variables aleatorias reales independientes idénticamente distribuidas con media cero y varianza unitaria con media cero y varianza unitaria.
- 3) Los coeficientes arriba de la diagonal son variables aleatorias complejas independientes idénticamente distribuidas con media cero y varianza unitaria.

A la familia $(\mathbf{X}_n)_{n \geq 1}$ se le conoce como ensamble de Wigner.

Ejemplo 3.2. Para cada $n \in \mathbb{N}$ sea \mathbf{X}_n una matriz $p \times n$ donde $p = p(n)$ es tal que $p/n \rightarrow y \in (0, \infty)$. Supongamos que los coeficientes de \mathbf{X}_n son variables aleatorias

complejas independientes idénticamente distribuidas con media cero y varianza unitaria. Definamos la matriz \mathbf{H}_n mediante

$$\mathbf{H}_n = \frac{1}{n} \mathbf{X}_n \mathbf{X}_n^*.$$

A la familia $(\mathbf{H}_n)_{n \geq 1}$ se le conoce como ensamble de Wishart.

Comentario 3.1. Si $\mathbf{x}_1, \dots, \mathbf{x}_n$ son n vectores observados procedentes de un vector aleatorio \mathbf{x} , entonces la matriz de covarianza muestral está dada por

$$\frac{1}{n} \sum_{k=1}^n \mathbf{x}_k \mathbf{x}_k^* = \frac{1}{n} \mathbf{X} \mathbf{X}^*,$$

donde $\mathbf{X} = (\mathbf{x}_1 \cdots \mathbf{x}_n)$. Por este motivo a las matrices de la forma $\mathbf{X} \mathbf{X}^*$ se les suele denominar matrices de covarianza empírica.

Ejemplo 3.3. Para cada $n \in \mathbb{N}$ sea \mathbf{U}_n una matriz aleatoria Haar unitaria de $n \times n$, i.e. una matriz aleatoria unitaria cuya distribución es la medida de Haar en las matrices unitarias. A la familia $(\mathbf{U}_n)_{n \geq 1}$ se le conoce como el ensamble de Haar unitario.

Generalmente se llama ensamble a una colección de matrices aleatorias cuya dimensión crece. Vale la pena mencionar que salvo los supuestos distribucionales no se especifica ninguna relación entre las matrices que componen el ensamble. Dos de los supuestos más usuales son que las matrices de un mismo ensamble sean independientes o bien que la matriz de dimensión $(n+1) \times (n+1)$ se obtenga agregando una fila y una columna a la matriz de dimensión $n \times n$.

Proposición 3.1. *Sea \mathbf{X} una matriz aleatoria de $d \times d$ con distribución empírica espectral $F^{\mathbf{X}}$, entonces para toda $m \in \mathbb{N}$ se tiene que*

$$\int_{\mathbb{C}} x^m F^{\mathbf{X}}(dx) = \text{tr}_d(\mathbf{X}^m).$$

Demostración. Puesto que $\lambda_i(X^m) = \lambda_i(X)^m$ tenemos que

$$\begin{aligned} \int_{\mathbb{C}} x^m F^{\mathbf{X}}(dx) &= \frac{1}{d} \sum_{k=1}^d \lambda_k(\mathbf{X})^m \\ &= \frac{1}{d} \sum_{k=1}^d \lambda_k(\mathbf{X}^m) \\ &= \text{tr}_d(\mathbf{X}^m), \end{aligned}$$

tal y como se quería demostrar. □

En particular se tiene la siguiente fórmula.

Corolario 3.1. *Sea \mathbf{X} una matriz aleatoria de $d \times d$ con distribución empírica espectral $F^{\mathbf{X}}$, entonces para toda $m \in \mathbb{N}$ se tiene que*

$$\mathbb{E} \left(\int_{\mathbb{C}} x^m F^{\mathbf{X}}(dx) \right) = \mathbb{E}(\text{tr}_d(\mathbf{X}^m)).$$

3.2. Convergencia de matrices aleatorias clásicas

En esta sección enunciamos algunos de los teoremas clásicos acerca de la convergencia débil de la distribución empírica espectral de ciertos ensambles.

Definición 3.4. Llamamos *ley del semicírculo* a la medida de probabilidad μ^s determinada por

$$\mu^s(dx) = \frac{\sqrt{4-x^2}}{2\pi} \mathbb{1}_{[-2,2]}(x) dx.$$

Proposición 3.2. *La transformada de Cauchy G_{μ^s} de la ley del semicírculo está dada por*

$$G_{\mu^s}(z) = \frac{z - \sqrt{z^2 - 4}}{2}$$

para $z \in \mathbb{C}^+$.

Teorema 3.4 (Teorema de Wigner). *Supóngase que $(\mathbf{X}_n)_{n \geq 1}$ es un ensamble de Wigner. Entonces, con probabilidad uno, la distribución empírica espectral de $\frac{1}{\sqrt{n}}\mathbf{X}_n$ converge débilmente a la ley del semicírculo.*

Comentario 3.2 (Método de los momentos). Una posible demostración para el teorema anterior es vía el método de los momentos. Este método se puede describir a grandes rasgos como sigue:

Paso 0. Denotemos por F^s a la función de distribución asociada a la ley del semicírculo. Puesto que μ^s tiene soporte compacto, está determinada por sus momentos. Dichos momentos están dados por

$$\int x^m F^s(dx) = \begin{cases} 0 & \text{si } m \text{ es impar,} \\ \frac{1}{k+1} \binom{2k}{k} & \text{si } m = 2k \text{ con } k \in \mathbb{N}. \end{cases}$$

Paso 1. Para toda $k \geq 0$ se tiene que

$$\mathbb{E} \left(\int x^k F^{n-1/2 X_n}(dx) \right) \xrightarrow{n \rightarrow \infty} \int x^k F^s(dx).$$

Paso 2. $\text{Var} \left(\int x^k F^{n-1/2 X_n}(dx) \right) \leq \frac{C}{n^2}$ para alguna $C \in \mathbb{R}$.

Paso 3. Extender la convergencia en media (Paso 1) a convergencia c.s. vía la cota para las varianzas (Paso 2) y el Lema de Borel-Cantelli.

Paso 4. De la convergencia casi segura del paso anterior y la caracterización de μ^s por sus momentos (Paso 0) concluir el resultado.

La demostración del Paso 1 utiliza algunos argumentos de teoría de gráficas. Específicamente, del Corolario 3.1 tenemos que

$$\begin{aligned} \mathbb{E} \left(\int_{\mathbb{R}} x^m F^{n-1/2 X_n}(dx) \right) &= \frac{1}{n^{m/2}} \mathbb{E}(\text{tr}_n(\mathbf{X}^m)) \\ &= \frac{1}{n^{m/2+1}} \sum_{\substack{i_k=1, \dots, n \\ k \in \{1, \dots, m\}}} \mathbb{E}(X_{i_1, i_2} X_{i_2, i_3} \cdots X_{i_m, i_1}). \end{aligned}$$

Para cada término $\mathbb{E}(X_{i_1, i_2} X_{i_2, i_3} \cdots X_{i_m, i_1})$ podemos construir una gráfica cuyos vértices sean $V = \{i_1, \dots, i_m\}$ y cuyas aristas sean $E = \{(i_1, i_2), \dots, (i_m, i_1)\}$. Haciendo un conteo adecuado de las gráficas formadas de esta manera, el cual sigue hasta cierto punto las ideas de la demostración del Teorema de Límite Central Libre, se obtiene la convergencia buscada. El Paso 2 sigue ideas combinatorias similares, mientras que el resto de los pasos siguen argumentos usuales de teoría de probabilidad.

Como corolario del teorema anterior se tiene que para un ensamble de Wigner $(\mathbf{X}_n)_{n \geq 1}$ existe un espacio de probabilidad (\mathcal{A}, φ) y un elemento semicircular $s \in \mathcal{A}$ tal que $\mathbf{X}_n \xrightarrow{\text{dist}} s$ casi seguramente.

Similarmente, existen resultados acerca de la convergencia de la distribución empírica espectral de matrices de Wishart, i.e. matrices de la forma $\mathbf{H} = \mathbf{X}\mathbf{X}^*$.

Definición 3.5. Sea $y \in (0, \infty)$, $a = (1 - \sqrt{y})^2$ y $b = (1 + \sqrt{y})^2$. Llamamos *ley de Marchenko-Pastur de parámetro y* a la medida de probabilidad ρ_y determinada por

$$\rho_y(dx) = \left(1 - \frac{1}{y}\right) \mathbb{1}_{(1, \infty)}(y) \delta_0(dx) + \frac{1}{2\pi xy} \sqrt{(b-x)(x-a)} \mathbb{1}_{[a, b]}(x) dx.$$

Proposición 3.3. La transformada de Cauchy G_{ρ_y} de la ley de Marchenko-Pastur de parámetro y está dada por

$$G_{\rho_y}(z) = \frac{z + y - 1 - \sqrt{(z - 1 - y)^2 - 4y}}{2yz}$$

para $z \in \mathbb{C}^+$.

Teorema 3.5 (Teorema de Marchenko-Pastur). Sea $(\mathbf{H}_n)_{n \geq 1}$ un ensamble de Wishart tal que \mathbf{H}_n es de dimensión $p \times n$ donde $p = p(n)$ con $p/n \rightarrow y \in (0, \infty)$. Entonces, con probabilidad uno, la distribución empírica espectral de \mathbf{H}_n converge débilmente a la distribución de Marchenko-Pastur de parámetro y .

Es posible demostrar el teorema anterior siguiendo el método de los momentos, al igual que con el Teorema de Wigner. En este caso los pasos 1 y 2 también siguen argumentos combinatorios, sin embargo las gráficas a estudiar resultan ser diferentes.

En particular, en el caso del Teorema de Wigner las gráficas de interés son arboles mientras que en el caso del Teorema de Marchenko-Pastur son gráficas bipartitas. Por lo tanto las pruebas son similares mas no idénticas.

Teorema 3.6. *Sea $(\mathbf{U}_n)_{n \geq 1}$ un ensamble Haar unitario. Entonces existe un espacio de probabilidad (\mathcal{A}, φ) tal que $\mathbf{U}_n \xrightarrow{\text{dist}} u$ casi seguramente, donde $u \in \mathcal{A}$ es una variable aleatoria Haar unitaria.*

3.3. Libertad asintótica

Supongamos que $(\mathbf{X}_n)_{n \geq 1}$ es un ensamble cuya distribución empírica espectral converge débilmente casi seguramente. Los teoremas de la sección anterior nos permiten asociar a tal ensamble una variable aleatoria x en un cierto espacio de probabilidad (\mathcal{A}, φ) . La distribución algebraica de x debe estar dada por los momentos de la distribución espectral límite, esto es

$$\begin{aligned} \varphi(x^m) &= \int x^m F(dx) \\ &= \lim_{n \rightarrow \infty} \mathbb{E}(\text{tr}_d(\mathbf{X}_n^m)) \end{aligned}$$

para toda $m \in \mathbb{N}$ y F tal que $F^{\mathbf{X}_n} \xrightarrow{d} F$ c.s.

Si tuvieramos un par de ensambles $(\mathbf{X}_n)_{n \geq 1}$ y $(\mathbf{Y}_n)_{n \geq 1}$, el procedimiento anterior² nos proporcionaría un espacio de probabilidad (\mathcal{A}, φ) en el cual existen variables aleatorias x y y las cuales están asociadas a los ensambles correspondientes. Para utilizar los resultados del capítulo anterior es necesario agregar un ingrediente más: la libertad entre x y y . Sin la libertad, al menos a primera vista, el procedimiento de cambiar ensambles de matrices aleatorias por elementos en un espacio de probabilidad no conmutativo es poco útil.

²Además del procedimiento anterior es necesario garantizar la existencia del producto libre de espacios de probabilidad. Para la demostración de este hecho véase [19].

Definición 3.6. Sean (\mathcal{A}, φ) y $(\mathcal{A}_n, \varphi_n)$ para $n \in \mathbb{N}$ espacios de probabilidad. Si S es un conjunto arbitrario y para cada $n \in \mathbb{N}$ $(a(s, n))_{s \in S}$ son variables aleatorias tales que $\{a(s, n) \mid s \in S\} \subset \mathcal{A}_n$ y $(a(s, n))_{s \in S} \xrightarrow{\text{dist.}} (a(s))_{s \in S}$ con $\{a(s) \mid s \in S\} \subset \mathcal{A}$, decimos que $(\{a(s, n)\})_{s \in S}$ son asintóticamente libres cuando $n \rightarrow \infty$ si $(\{a(s)\})_{s \in S}$ son libres.

En notación de la definición anterior, el comentario del párrafo anterior a la Definición 3.6 resalta la importancia de la libertad asintótica entre matrices aleatorias. Los siguientes teoremas muestran que algunos ensambles clásicos son asintóticamente libres.

Teorema 3.7. *Para cada $n \in \mathbb{N}$ sea $(\mathbf{U}(s, n))_{s \in S}$ una familia de matrices aleatorias unitarias estandar de $n \times n$ independientes, y sea $(\mathbf{D}(t, n))_{t \in T}$ una familia de matrices no aleatorias de $n \times n$ tales que $\sup_n \|\mathbf{D}(t, n)\| < \infty$ para cada $t \in T$ y $(\mathbf{D}(t, n), \mathbf{D}(t, n)^*)_{t \in T}$ tiene una distribución límite conjunta. Entonces las familias*

$$(\{\mathbf{U}(s, n), \mathbf{U}(s, n)^*\})_{s \in S}, \{\mathbf{D}(t, n), \mathbf{D}(t, n)^* \mid t \in T\}$$

son asintóticamente libres cuando $n \rightarrow \infty$ casi seguramente.

Definición 3.7. Decimos que una matriz aleatoria \mathbf{T} autoadjunta de $n \times n$ es unitariamente invariante si

$$\mathbf{T} \stackrel{d}{=} \mathbf{V}\mathbf{T}\mathbf{V}^*$$

para cualquier matriz unitaria \mathbf{V} de $n \times n$.

Teorema 3.8. *Para cada $n \in \mathbb{N}$ sea $(\mathbf{H}(s, n))_{s \in S}$ una familia de matrices aleatorias autoadjuntas unitariamente invariantes de $n \times n$ independientes, y sea $(\mathbf{D}(t, n))_{t \in T}$ como en el Teorema 3.7. Si cuando $n \rightarrow \infty$ se tiene que $\mathbf{H}(s, n)$ converge en distribución casi seguramente a una medida de soporte compacto para cada $s \in S$, entonces la familia*

$$((\{\mathbf{H}(s, n)\})_{s \in S}, \{\mathbf{D}(t, n), \mathbf{D}(t, n)^* \mid t \in T\})$$

es asintóticamente libre cuando $n \rightarrow \infty$ casi seguramente.

Supóngase que se tienen dos ensambles de matrices constantes $(\mathbf{A}_n)_{n \geq 1}$ y $(\mathbf{B}_n)_{n \geq 1}$ tales que $\mathbf{A}_n \xrightarrow{\text{dist}} a$ y $\mathbf{B}_n \xrightarrow{\text{dist}} b$, con a y b variables aleatorias en un cierto espacio (\mathcal{A}, φ) . En tal situación no es necesariamente cierto que $(\mathbf{A}_n, \mathbf{B}_n) \xrightarrow{\text{dist}} (a, b)$, de hecho, podría ser que $(\mathbf{A}_n, \mathbf{B}_n)$ no converja. El siguiente corolario³ nos proporciona un par de ensambles asintóticamente libres los cuales convergen en distribución a los mismos elementos. Este recurso será particularmente útil en el Capítulo 5, cuando necesitemos de la relación de libertad para calcular el espectro asintótico de ciertos canales multiantena.

Corolario 3.2. *Sean $(\mathbf{A}_n)_{n \geq 1}$ y $(\mathbf{B}_n)_{n \geq 1}$ dos ensambles de matrices Hermitianas constantes, los cuales satisfacen las mismas condiciones que el ensamble $(\mathbf{D}(n))_{n \geq 1}$ del Teorema 3.7. Si $(\mathbf{U}_n)_{n \geq 1}$ es un ensamble de matrices Haar unitarias, entonces las familias*

$$\{\mathbf{U}_n \mathbf{A}_n \mathbf{U}_n^*\}, \{\mathbf{B}_n\}$$

son asintóticamente libres cuando $n \rightarrow \infty$ casi seguramente.

Como consecuencia del teorema anterior tenemos la siguiente generalización del Teorema de Wigner (Teorema 3.4) la cual es conocida como Teorema de Voiculescu.

Teorema 3.9 (Teorema de Voiculescu). *Para cada $n \in \mathbb{N}$ sea $(\mathbf{X}(s, n))_{s \in S}$ una familia de matrices gaussianas autoadjuntas estandar independientes de $n \times n$, y sea $(\mathbf{D}(t, n))_{t \in T}$ como en el Teorema 3.7. Entonces existe un espacio de probabilidad (\mathcal{A}, φ) tal que*

$$((\{\mathbf{X}(s, n)\})_{s \in S}, (\mathbf{D}(t, n))_{t \in T}) \xrightarrow{*-\text{dist}} ((x(s))_{s \in S}, (d(t))_{t \in T}),$$

$((\{\mathbf{X}(s, n)\})_{s \in S}, \{\mathbf{D}(t, n) \mid t \in T\})$ es asintóticamente libre y $x(s)$ es una variable aleatoria semicircular para toda $s \in S$.

Observemos que la parte medular del teorema anterior radica en la libertad asintótica de las matrices aleatorias en cuestión. Por lo tanto, podemos construir un teorema

³Es una consecuencia directal del hecho que $\mathbf{U}_n \mathbf{A}_n \mathbf{U}_n^*$ es unitariamente invariante.

análogo al teorema anterior para matrices Haar unitarias, el cual se sigue de los Teoremas 3.6 y 3.7. Terminamos esta sección enunciando el teorema análogo al Teorema 3.7 en el caso de matrices de covarianza.

Teorema 3.10. *Para $s \in S$ y $n \in \mathbb{N}$ sea $p(s, n) \in \mathbb{N}$ tal que $p(s, n)/n \rightarrow \lambda_s \in (0, \infty)$ cuando $n \rightarrow \infty$. Sea $\mathbf{Z}(s, n)$ una matriz aleatoria de $p(s, n) \times n$ tal que $\Re Z_{i,j}(s, n)$ y $\Im Z_{i,j}(s, n)$ son independientes con distribución común $N(0, 1/2n)$. Sea $\mathbf{B}(s, n)$ una matrix autoadjunta de $p(s, n) \times p(s, n)$. Supongamos que $(\mathbf{Z}(s, n))_{s \in S}, (\mathbf{B}(s, n))_{s \in S}$ son independientes para cada $n \in \mathbb{N}$. Supongamos además que para cada $s \in S$ se tiene que $\mathbf{B}(s, n)$ converge en distribución casi seguramente a una medida de soporte compacto. Si $(\mathbf{D}(t, n))_{t \in T}$ es como en el Teorema 3.7 entonces*

$$((\{\mathbf{Z}(s, n)^* \mathbf{B}(s, n) \mathbf{Z}(s, n)\})_{s \in S}, \{\mathbf{D}(t, n), \mathbf{D}(t, n)^* \mid t \in T\})$$

es asintóticamente libre cuando $n \rightarrow \infty$ casi seguramente.

3.4. Notas adicionales

Salvo el Teorema 3.6, las pruebas de las proposiciones y teoremas de las secciones 3.1 y 3.2 pueden ser encontradas en libro de Bai y Silverstein [2]. En particular, las demostraciones detalladas de los teoremas 3.4 y 3.5 se encuentra en los capítulos 2 y 3 respectivamente. El Teorema 3.6, así como la sección anterior, se encuentra en el libro de Hiai y Petz [13].

El Teorema de Voiculescu (Teorema 3.9) se puede extender a matrices aleatorias no autoadjuntas, en cuyo caso las matrices gaussianas convergen a una familia de variables aleatorias circulares libres (véase [10]).

Capítulo 4

Probabilidad libre valuada en operadores

En este capítulo daremos una breve introducción a la probabilidad libre valuada en operadores, poniendo especial énfasis en los espacios de probabilidad utilizados para estudiar las matrices aleatorias a bloques y rectangulares. Estos espacios son la base de las aplicaciones discutidas en el siguiente capítulo.

4.1. Definiciones básicas

Comencemos poniendo en contexto algunos elementos que utilizaremos después.

Definición 4.1. Sea \mathcal{A} una C^* -álgebra unital. Definimos el conjunto de elementos positivos de \mathcal{A} como

$$\mathcal{A}^+ := \{p \in \mathcal{A} \mid p = p^*, \text{Sp}(p) \subset [0, \infty)\}.$$

Escribimos $p \geq 0$ para denotar que $p \in \mathcal{A}^+$.

Teorema 4.1. *Sea \mathcal{A} una C^* -álgebra y a una variable aleatoria normal. Si $f \in C(\text{Sp}(a), \mathbb{C})$ entonces*

$$\text{Sp}(f(a)) = f(\text{Sp}(a)).$$

Demostración. Primero demostremos lo siguiente: “ $f(a)$ es invertible si y sólo si $f(\lambda) \neq 0$ para toda $\lambda \in \text{Sp}(a)$ ”. Si $f(\lambda) \neq 0$ para toda $\lambda \in \text{Sp}(a)$ entonces $h = \frac{1}{f}$ es continua en $\text{Sp}(a)$. De las propiedades del cálculo funcional definido en el Capítulo 2 tenemos que

$$\begin{aligned}\mathbf{1}_{\mathcal{A}} &= \Phi(1) \\ &= \Phi(fh) \\ &= \Phi(f)\Phi(h),\end{aligned}$$

y por lo tanto $\Phi(f)$ es invertible. Para demostrar el recíproco tomemos $f(a)$ invertible pero supongamos que $f(\lambda_0) = 0$ para algún $\lambda_0 \in \text{Sp}(a)$. Tomemos $\alpha > \|\Phi(f)^{-1}\|$, entonces por continuidad existe $r \in \mathbb{R}_+$ tal que para toda $\lambda \in \mathcal{B}_{\lambda_0}(r)$ se tiene que $|f(\lambda)| < 1/\alpha$. Construyamos una función continua¹ $h : \text{Sp}(a) \rightarrow [0, \alpha]$ tal que $\text{supp}(h) \subset \overline{\mathcal{B}_{\lambda_0}(r/2)}$ y $h(\lambda_0) = \alpha$. De la construcción de h se tiene que $\text{supp}(fh) \subset \mathcal{B}_{\lambda_0}(r/2)$ y además $\|fh\|_{\infty} \leq 1$. Claramente $\alpha = \|h\|_{\infty}$, y por lo tanto

$$\begin{aligned}\alpha &= \|\Phi(h)\| \\ &= \|\Phi(f)^{-1}\Phi(f)\Phi(h)\| \\ &= \|\Phi(f)^{-1}\| \|\Phi(fh)\| \\ &< \alpha,\end{aligned}$$

lo cual es una contradicción al hecho que $f(\lambda_0) = 0$ para algún $\lambda_0 \in \text{Sp}(a)$. Con esto queda demostrada la afirmación.

Por último, observemos que $\lambda \in \text{Sp}(f(a))$ si y sólo si $\Phi(\lambda - f)$ es no invertible. Más aún, del párrafo anterior tenemos que $\Phi(\lambda - f)$ es no invertible (en $C(\text{Sp}(a), \mathbb{C})$) si y sólo si $\lambda - f(\lambda_0) = 0$ para algún $\lambda_0 \in \text{Sp}(a)$, lo cual implica trivialmente el resultado. \square

Comentario 4.1. En la demostración anterior usamos implícitamente que $\Phi(C(\text{Sp}(a), \mathbb{C}))$

¹Podemos construir una función de prueba (bump function) por ejemplo.

es una $*$ -álgebra conmutativa, i.e. si $f, g \in C(\text{Sp}(a), \mathbb{C})$ entonces $\Phi(f)\Phi(g) = \Phi(fg) = \Phi(g)\Phi(f)$.

Definición 4.2. Sean \mathcal{A} y \mathcal{B} dos C^* -álgebras unitales. Un mapeo lineal $f : \mathcal{A} \rightarrow \mathcal{B}$ se dice positivo si $f(a) \geq 0$ para toda $a \geq 0$.

Definimos el mapeo lineal $I_d \otimes f : \text{Mat}_d(\mathcal{A}) \rightarrow \text{Mat}_d(\mathcal{B})$ de tal forma que para $M \in \text{Mat}_d(\mathbb{C})$ y $a \in \mathcal{A}$ se tiene que $(I_d \otimes f)(M \otimes a) = M \otimes f(a)$. Decimos que $f : \mathcal{A} \rightarrow \mathcal{B}$ es completamente positivo si $I_d \otimes f$ es positivo para toda $d \in \mathbb{N}$.

Definición 4.3. Sea \mathcal{A} una C^* -álgebra unital y \mathcal{B} una C^* -subálgebra unital de \mathcal{A} . Un mapeo $E : \mathcal{A} \rightarrow \mathcal{B}$ es una esperanza condicional si

- a) E es un mapeo lineal completamente positivo,
- b) $E(b) = b$ para toda $b \in \mathcal{B}$,
- c) $E(b_1 a b_2) = b_1 E(a) b_2$ para toda $a \in \mathcal{A}$ y cualesquiera $b_1, b_2 \in \mathcal{B}$.

Un espacio de probabilidad valuado en operadores consiste de una terna $(E, \mathcal{A}, \mathcal{B})$ donde $\mathcal{B} \subset \mathcal{A}$ son C^* -álgebras unitales y $E : \mathcal{A} \rightarrow \mathcal{B}$ es una esperanza condicional.

Definición 4.4. Un W^* -espacio de probabilidad es un C^* -espacio de probabilidad en el que el álgebra correspondiente es una W^* -álgebra, i.e. una C^* -álgebra la cual es el espacio de Banach dual de un cierto espacio de Banach.

Comentario 4.2. Si (\mathcal{A}, φ) es un W^* -espacio de probabilidad tracial, entonces para cada W^* -subálgebra \mathcal{B} existe una única esperanza condicional $E : \mathcal{A} \rightarrow \mathcal{B}$ tal que $\varphi \circ E = \varphi$. Por lo tanto, en el contexto de los W^* -espacios de probabilidad, la esperanza condicional E posee la propiedad de anidamiento de la esperanza condicional probabilista (véase [24], [3] y referencias allí). En este sentido, el término esperanza condicional es un abuso de notación ya que la esperanza condicional de la definición anterior no es lo misma que la esperanza condicional probabilista.

La siguiente proposición ilustra una consecuencia de la positividad de E .

Proposición 4.1. *Sea $(\mathcal{A}, \mathcal{B}, E)$ es un espacio de probabilidad valuado en operadores, entonces las siguientes afirmaciones se cumplen:*

- 1) Si $p \in \mathcal{A}^+$ entonces $E(p) \in \mathcal{A}^+$.
- 2) Si $a \in \mathcal{A}$ es autoadjunto entonces $E(a) = E(a)^*$.
- 3) Si $x \in \mathcal{A}$ entonces $E(x^*) = E(x)^*$.

Demostración.

- 1) Puesto que E es completamente positivo, en particular es positivo. Por lo tanto $E(p) \in \mathcal{A}^+$ para todo $p \in \mathcal{A}^+$.
- 2) Definamos $u = \frac{a + \mathbf{1}_{\mathcal{A}}}{2}$ y $v = \frac{a - \mathbf{1}_{\mathcal{A}}}{2}$. Puesto que u es autoadjunto existe un cálculo funcional Φ_u asociado a u . Si $f(z) = \bar{z}z$ entonces del Teorema 4.1 se tiene que

$$\begin{aligned}
\text{Sp}(u^*u) &= \text{Sp}(f(u)) \\
&= f(\text{Sp}(u)) \\
&= \{|\lambda|^2 \mid \lambda \in \text{Sp}(u)\} \subset [0, \infty),
\end{aligned}$$

es decir, $u^*u \in \mathcal{A}^+$. Similarmente $v^*v \in \mathcal{A}^+$ y por el inciso anterior se tiene que

$$\begin{aligned}
E(a)^* &= E(u^*u - v^*v)^* \\
&= E(u^*u)^* - E(v^*v)^* \\
&= E(u^*u) - E(v^*v) \\
&= E(a).
\end{aligned}$$

- 3) La demostración de este inciso es directa del inciso anterior y la técnica de la demostración de la Proposición 2.1.

□

Comentario 4.3. Observemos que la positividad de la esperanza condicional es de utilidad al trabajar con la involución y el espectro. Sin embargo, generalmente traba-

jaremos en un nivel más algebraico por lo que omitiremos la propiedad a) al demostrar que un funcional lineal es esperanza condicional.

Definición 4.5. Sea $(E, \mathcal{A}, \mathcal{B})$ un espacio de probabilidad valuado en operadores. La distribución valuada en operadores de una variable aleatoria $a \in \mathcal{A}$ es el funcional lineal $\mu : \mathcal{B}\langle X, X^* \rangle \rightarrow \mathbb{C}$ el cual está determinado por

$$\mu(X^{\epsilon_1} b_1 X^{\epsilon_2} \dots X^{\epsilon_{n-1}} b_{n-1} X^{\epsilon_n}) = E(a^{\epsilon_1} b_1 a^{\epsilon_2} \dots a^{\epsilon_{n-1}} b_{n-1} a^{\epsilon_n})$$

para $n \in \mathbb{N}$, $\epsilon_1, \dots, \epsilon_n \in \{1, *\}$ y $b_1, \dots, b_{n-1} \in \mathcal{B}$.

A las expresiones de la forma $E(a^{\epsilon_1} b_1 a^{\epsilon_2} \dots a^{\epsilon_{n-1}} b_{n-1} a^{\epsilon_n})$ se les llama momentos valuados en operadores.

Comentario 4.4. Observemos que la definición anterior es prácticamente la misma que en el caso escalar, salvo que φ fue remplazado por E y $\mathbb{C}\mathbf{1}_{\mathcal{A}}$ por \mathcal{B} . De hecho, puesto que $\mathbb{C}\mathbf{1}_{\mathcal{A}}$ es un subgrupo del centro de \mathcal{A} , en la definición del caso escalar no es necesario escribir el análogo a los b_1, \dots, b_{n-1} de la definición anterior. La definición de distribución valuada en operadores de varias variables se establece mediante esta misma analogía, al igual que la siguiente definición.

Definición 4.6. Sean \mathcal{A}_i subálgebras unitales de \mathcal{A} para $i \in I$, con I un conjunto arbitrario. Las álgebras $(\mathcal{A}_i)_{i \in I}$ son libres respecto a E si

$$E(a_1 \cdots a_n) = 0$$

siempre que $a_j \in \text{Ker}(E) \cap \mathcal{A}_{i(j)}$ para $j \in \{1, \dots, n\}$ y además² $i(1) \neq i(2), i(2) \neq i(3), \dots, i(n-1) \neq i(n)$.

Decimos que las familias de variables aleatorias $\{a_1^{(i)}, \dots, a_{n_i}^{(i)}\}$ para $i \in I$ son libres respecto a E si las álgebras $\langle \mathcal{B}, a_1^{(i)}, \dots, a_{n_i}^{(i)} \rangle$ para $i \in I$ son libres respecto a E .

Comentario 4.5. Al igual que en el caso escalar, la distribución valuada en operadores de una familia libre respecto a E depende únicamente de las distribuciones valuadas

²Como antes, es permitido que $i(1) = i(3)$ y que $i(1) = i(n)$.

en operadores individuales. Por lo tanto, en este contexto tendremos herramientas muy similares a las del caso escalar, i.e. cumulantes, transformada de Cauchy, transformada R, transformada S, etc.

Ejemplo 4.1. Supongamos que $x, y \in \mathcal{A}$ son libres respecto a $E : \mathcal{A} \rightarrow \mathcal{B}$. Entonces³

$$\begin{aligned}
0 &= E((x - E(x))(y - E(y))) \\
&= E(xy - xE(y) - E(x)y + E(x)E(y)) \\
&= E(xy) - E(xE(y)) - E(E(x)y) + E(E(x)E(y))) \\
&= E(xy) - E(x)E(y),
\end{aligned}$$

donde la última igualdad se sigue de la propiedad c) en la Definición 4.3. Por lo tanto $E(xy) = E(x)E(y)$, similar al caso escalar.

Ejemplo 4.2. Sean x y $\{y_1, y_2\}$ libres respecto a E . Entonces

$$\begin{aligned}
0 &= E((y_1 - E(y_1))(x - E(x))(y_2 - E(y_2))) \\
&= E(y_1(x - E(x))(y_2 - E(y_2))) - E(E(y_1)(x - E(x))(y_2 - E(y_2))) \\
&= E(y_1(x - E(x))(y_2 - E(y_2))) - E(y_1)E((x - E(x))(y_2 - E(y_2))).
\end{aligned}$$

Puesto que x y y_2 son libres se tiene que $E(y_1)E(x - E(x))E(y_2 - E(y_2)) = 0$, y procediendo como en el ejemplo anterior se tiene que

$$\begin{aligned}
0 &= E(y_1(x - E(x))(y_2 - E(y_2))) \\
&= E(y_1(x - E(x))(y_2 - E(y_2))) \\
&= E(y_1xy_2) - E(y_1)E(x)E(y_2) - E(y_1E(x)y_2) + E(y_1)E(x)E(y_2) \\
&= E(y_1xy_2) - E(y_1E(x)y_2).
\end{aligned}$$

Por lo tanto $E(y_1xy_2) = E(y_1E(x)y_2)$.

³Puesto que $E(x) \in \mathcal{B}$, en este caso no hay abuso de notación al escribir $x - E(x)$, a diferencia del caso escalar donde $x - \varphi(x)$ significaba $x - \varphi(x)\mathbf{1}_{\mathcal{A}}$.

Comentario 4.6. Observemos que la fórmula del ejemplo anterior es idéntica a la del caso escalar, salvo que en este contexto el término $E(x)$ no conmuta con el resto de la expresión. Específicamente, si $\mathcal{B} = \mathbb{C}$ y $E = \varphi$ entonces $E(x) = \varphi(x) \in \mathbb{C}$ y por lo tanto

$$\begin{aligned} E(y_1xy_2) &= E(y_1E(x)y_2) \\ &= E(x)E(y_1y_2). \end{aligned}$$

En particular, observamos que en el caso general $\mathcal{B} \subset \mathcal{A}$ se tiene que los términos en las expresiones para los momentos mixtos de variables aleatorias libres mantienen el orden inicial de las variables en cuestión, e.g. $E(xy) = E(x)E(y)$ y $E(y_1xy_2) = E(y_1E(x)y_2)$.

Ejemplo 4.3. Siguiendo la metodología anterior, para $\{x_1, x_2\}, \{y_1, y_2\}$ libres se tiene que

$$E(x_1y_1x_2y_2) = E(x_1E(y_1)x_2)E(y_2) + E(x_1)E(y_1E(x_2)y_2) - E(x_1)E(y_1)E(x_2)E(y_2).$$

En analogía al caso escalar se definen a continuación los cumulantes valuados en operadores.

Definición 4.7. Sea \mathcal{B} un álgebra y M un espacio vectorial sobre \mathbb{C} . Decimos que M es un \mathcal{B} -bimódulo si el álgebra \mathcal{B} actúa en él por derecha e izquierda.

Ejemplo 4.4. De la definición de álgebra es claro que \mathcal{B} es un \mathcal{B} -bimódulo.

Definición 4.8. El producto tensorial $M \otimes_{\mathcal{B}} N$ de dos \mathcal{B} -bimódulos M y N se define como el producto tensorial $M \otimes N$ módulo la identificación $(m \cdot b) \otimes n \sim m \otimes (b \cdot n)$ para todo $(m, b, n) \in M \times \mathcal{B} \times N$.

Comentario 4.7. Obsérvese que $M \otimes_{\mathcal{B}} N$ se puede dotar de una estructura de \mathcal{B} -bimódulo definiendo $b_1 \cdot (m \otimes n) \cdot b_2 = (b_1 \cdot m) \otimes (n \cdot b_2)$ para todos $b_1, b_2 \in \mathcal{B}, m \in M$

y $n \in \mathbb{N}$. Esto nos permite definir el \mathcal{B} -bimódulo

$$\mathcal{A}^{\otimes_{\mathcal{B}} n} = \underbrace{\mathcal{A} \otimes_{\mathcal{B}} \cdots \otimes_{\mathcal{B}} \mathcal{A}}_n$$

para $n \in \mathbb{N}$.

Definición 4.9. Sean M y N dos \mathcal{B} -bimódulos. Se dice que $f : M \rightarrow N$ es un morfismo entre \mathcal{B} -bimódulos si

$$f(b_1 m b_2 + b'_1 m' b'_2) = b_1 f(m) b_2 + b'_1 f(m') b'_2$$

para todos $b_1, b'_1, b_2, b'_2 \in \mathcal{B}$ y $m, m' \in M$.

Definición 4.10. Sea $(E, \mathcal{A}, \mathcal{B})$ un espacio de probabilidad valuado en operadores. Definimos los cumulantes valuados en operadores como la familia de morfismos entre \mathcal{B} -bimódulos

$$(\kappa_{\pi}^{\mathcal{B}} : \mathcal{A}^{\otimes_{\mathcal{B}} n} \rightarrow \mathcal{B})_{\pi \in \text{NC}(n), n \in \mathbb{N}}$$

la cual satisface que:

- i) Es multiplicativa respecto al argumento π .
- ii) Satisface la relación

$$E(a_1 \dots a_n) = \sum_{\pi \in \text{NC}(n)} \kappa_{\pi}^{\mathcal{B}}(a_1 \otimes \dots \otimes a_n).$$

Al igual que en el caso escalar, a la ecuación anterior se le conoce como la fórmula momentos-cumulantes.

Comentario 4.8. Observemos que \mathcal{B} en $\kappa_{\pi}^{\mathcal{B}}$ denota que el cumulante en cuestión es \mathcal{B} -valuado. Cuando no cause confusión escribiremos únicamente κ_{π} en lugar de $\kappa_{\pi}^{\mathcal{B}}$.

Comentario 4.9. La definición anterior necesita justificación puesto que la existencia de tales morfismos no es clara a priori. Sin embargo, podemos imitar la teoría desarrollada en el caso escalar para el caso valuado en operadores. Para ello, es necesario extender

la teoría para trabajar con funciones \mathcal{B} -valuadas en lugar de funciones \mathbb{C} -valuadas⁴, y a partir de E crear una familia de morfismos entre \mathcal{B} -bimódulos la cual sea multiplicativa.

Puesto que \mathcal{B} en principio no es conmutativa, para que una familia sea multiplicativa requiere que se respete el orden de los bloques de π . Para aclarar este punto desarrollemos el siguiente ejemplo.

Ejemplo 4.5. Si $\pi = \{\{1, 8\}, \{2, 6\}, \{3, 4, 5\}, \{7\}\}$ entonces

$$\kappa_\pi(a_1 \otimes \dots \otimes a_8) = \kappa_2(a_1 \kappa_2(a_2 \kappa_3(a_3 \otimes a_4 \otimes a_5) \otimes a_6) \kappa_1(a_7) \otimes a_8).$$

Comentario 4.10. Observemos que las expresiones para los cumulantes valuados en operadores son muy similares a las del caso escalar. Las diferencias, al igual que con el mapeo E , provienen de la no conmutatividad de \mathcal{B} . Si $\mathcal{B} = \mathbb{C}$ y $E = \varphi$ entonces para π como en el ejemplo anterior se tiene que

$$\kappa_\pi(a_1, \dots, a_8) = \kappa_2(a_1 \otimes a_8) \kappa_2(a_2 \otimes a_6) \kappa_3(a_3 \otimes a_4 \otimes a_5) \kappa_1(a_7),$$

lo cual coincide con el caso escalar.

El siguiente teorema es la versión valuada en operadores del Teorema 2.8.

Teorema 4.2. Sean $(E, \mathcal{A}, \mathcal{B})$ un espacio de probabilidad valuado en operadores e I un conjunto arbitrario. Las familias de variables aleatorias $(a_1^{(i)}, \dots, a_{n_i}^{(i)})_{i \in I}$ son libres respecto a E si y sólo si los cumulantes mixtos se anulan.

4.2. Variables aleatorias semicirculares y el teorema del límite central

Definición 4.11. Sea $(E, \mathcal{A}, \mathcal{B})$ un espacio de probabilidad valuado en operadores y s una variable aleatoria. Se dice que s es un elemento semicircular valuado en operadores

⁴En particular, es necesario extender la formula de inversión de Möbius bajo la consideración de una función zeta dada por $\zeta(\sigma, \pi) = \mathbf{1}_{\mathcal{B}}$ para toda $\sigma \leq \pi$.

si para toda $n \geq 1$ se tiene que

$$E(sb_1 \cdots sb_{n-1}s) = \sum_{\pi \in \text{NC}_2(n)} \kappa_\pi(sb_1 \otimes \cdots \otimes sb_{n-1} \otimes s), \quad (4.1)$$

para $b_1, \dots, b_{n-1} \in \mathcal{B}$.

Ejemplo 4.6. Sea s es un elemento semicircular en $(E, \mathcal{A}, \mathcal{B})$. Puesto que $\text{NC}_2(2n+1) = \emptyset$ para toda $n \in \mathbb{N}$ entonces

$$E(s^{2n+1}) = \sum_{\pi \in \text{NC}_2(2n+1)} \kappa_\pi(s) = \mathbf{0}_{\mathcal{A}},$$

y en particular, de la fórmula momentos-cumulantes concluimos que $\kappa_1(s) = \mathbf{0}_{\mathcal{A}}$. Además, de la ecuación (4.1) se tiene que para $b \in \mathcal{B}$

$$E(sbs) = \kappa_2(sb \otimes s).$$

Definición 4.12. Sea x un elemento en $(E, \mathcal{A}, \mathcal{B})$. Definimos la función de covarianza de x , $\eta : \mathcal{B} \rightarrow \mathcal{B}$, mediante la aplicación $b \mapsto E(xbx)$.

Comentario 4.11. Si $\pi \in \text{NC}_2(n)$ entonces $\kappa_\pi(s \otimes \cdots \otimes s)$ se puede calcular utilizando únicamente la función de covarianza de s . Por ejemplo, para $\pi = \{\{1, 4\}, \{2, 3\}, \{5, 6\}\} \in \text{NC}_2(6)$ se tiene que

$$\begin{aligned} \kappa_\pi(s \otimes \cdots \otimes s) &= \kappa_2(s\kappa_2(s \otimes s) \otimes s)\kappa_2(s \otimes s) \\ &= \kappa_2(s\eta(\mathbf{1}_{\mathcal{A}}) \otimes s)\eta(\mathbf{1}_{\mathcal{A}}) \\ &= \eta(\eta(\mathbf{1}_{\mathcal{A}}))\eta(\mathbf{1}_{\mathcal{A}}). \end{aligned}$$

Más aún, la representación de $\pi \in \text{NC}_2(n)$ mediante paréntesis muestra claramente la

forma de $\kappa_\pi(s \otimes \cdots \otimes s)$, e.g.

$$\begin{aligned}\pi = (())() &\implies \kappa_\pi(s \otimes \cdots \otimes s) = \eta(\eta(\mathbf{1}_A))\eta(\mathbf{1}_A) \\ \pi = (((()))()) &\implies \kappa_\pi(s \otimes \cdots \otimes s) = \eta(\eta(\eta(\mathbf{1}_A))\eta(\mathbf{1}_A)).\end{aligned}$$

Lema 4.1. *Si s es una variable aleatoria semicircular en $(E, \mathcal{A}, \mathcal{B})$ entonces*

$$\kappa_{2n+1}(sb_1 \otimes \cdots \otimes sb_{2n} \otimes s) = 0$$

para $n \in \mathbb{N}$ y $b_1, \dots, b_{2n} \in \mathcal{B}$.

Demostración. Procedamos por inducción. El caso $n = 0$ se calculó en el Ejemplo 4.6. Supongamos el resultado para n . También del Ejemplo 4.6 sabemos que $E(s^{2(n+1)+1}) = \mathbf{0}_A$, y de la fórmula momentos-cumulantes tenemos que

$$\mathbf{0}_A = \sum_{\pi \in \text{NC}(2(n+1)+1)} \kappa_\pi(sb_1 \otimes \cdots \otimes sb_{2(n+1)} \otimes s),$$

o equivalentemente

$$\kappa_{2(n+1)+1}(sb_1 \otimes \cdots \otimes sb_{2(n+1)} \otimes s) = - \sum_{\pi < \mathbf{1}_{\text{NC}(2(n+1)+1)}} \kappa_\pi(sb_1 \otimes \cdots \otimes sb_{2(n+1)} \otimes s).$$

Demostraremos que cada término de la suma anterior se anula. Sea $\pi < \mathbf{1}_{\text{NC}(2(n+1)+1)}$, entonces π tiene al menos un bloque, digamos V , el cual tiene cardinalidad impar. Puesto que π tiene más de un bloque concluimos que $|V| \leq 2n + 1$. Observemos que V aparecera en κ_π de la siguiente manera

$$\kappa_{|V|}(sk_1 \otimes \cdots \otimes sk_{|V|-1} \otimes s),$$

donde $k_1, \dots, k_{|V|-1} \in \mathcal{B}$ son obtenidos mediante los bloques anidados en V . De la hipótesis de inducción tal término se anula. Puesto que $\pi < \mathbf{1}_{\text{NC}(2(n+1)+1)}$ es arbitrario concluimos el resultado deseado. \square

Proposición 4.2. Si s es una variable aleatoria semicircular en $(E, \mathcal{A}, \mathcal{B})$ entonces

$$\kappa_n(sb_1 \otimes \cdots \otimes sb_{n-1} \otimes s) = 0$$

para $n \geq 3$ y $b_1, \dots, b_{n-1} \in \mathcal{B}$.

Demostración. Puesto que el lema anterior demostró el caso impar, bastará demostrar el caso par. De la fórmula de momentos-cumulantes y la ecuación (4.1) se tiene que

$$\begin{aligned} \kappa_4(sb_1 \otimes \cdots \otimes s) &= E(sb_1 \cdots s) - \sum_{\pi \in \text{NC}_2(4)} \kappa_\pi(sb_1 \otimes \cdots \otimes s) - \sum_{\substack{\pi \notin \text{NC}_2(4) \\ \pi < \mathbf{1}_{\text{NC}(4)}}} \kappa_\pi(sb_1 \otimes \cdots \otimes s) \\ &= - \sum_{\substack{\pi \notin \text{NC}_2(4) \\ \pi < \mathbf{1}_{\text{NC}(4)}}} \kappa_\pi(sb_1 \otimes \cdots \otimes s). \end{aligned}$$

Por el Lema 4.1 es claro que todos los términos de la última suma son cero. Esto último es el paso base de una inducción sobre $2(n+1)$ con $n \in \mathbb{N}$. Procediendo como en el caso base, el paso inductivo se sigue trivialmente de la base de inducción. \square

Definición 4.13. Sean $(E, \mathcal{A}_n, \mathcal{B})$ para $n \in \mathbb{N}$ espacios de probabilidad valuados en operadores, I un conjunto arbitrario y $(a_i(n))_{i \in I}$ familias de elementos en \mathcal{A}_n para cada $n \in \mathbb{N}$. Si $(E, \mathcal{A}, \mathcal{B})$ es un espacio de probabilidad valuado en operadores el cual contiene a la familia de variables aleatorias $(a_i)_{i \in I}$, decimos que las familias $(a_i(n))_{i \in I}$ convergen en \mathcal{B} -distribución a $(a_i)_{i \in I}$ cuando $n \rightarrow \infty$ si las distribuciones valuadas en operadores convergen.

Definición 4.14. Sean I_s conjuntos arbitrarios para cada $s \in S$ con S un conjunto arbitrario. Además, sean $(E, \mathcal{A}_n, \mathcal{B})$ espacios de probabilidad valuados en operadores para cada $n \in \mathbb{N}$. Si $(a_i(n))_{i \in I_s} \subset \mathcal{A}_n$ son familias de variables aleatorias para cualesquier $s \in S$ y $n \in \mathbb{N}$, decimos que son asintóticamente libres respecto a E si convergen en \mathcal{B} -distribución a familias de variables aleatorias libres respecto a E .

A continuación enunciamos el Teorema del Límite Central en su versión valuada en operadores.

Teorema 4.3. Sea $(E, \mathcal{A}, \mathcal{B})$ un espacio de probabilidad valuado en operadores. Sea $(X_n)_{n \geq 1}$ una familia de variables aleatorias autoadjuntas tales que

I) X_1, X_2, \dots son libres respecto a E .

II) Para cualesquiera $n \geq 1$ y $b \in \mathcal{B}$ se tiene que $E(X_n) = 0$ y $E(X_n b X_n) = \eta(b)$.

III) Para toda $k \in \mathbb{N}$ y cualesquiera $b_1, \dots, b_k \in \mathcal{B}$ se tiene que

$$\sup_{n \geq 1} \|E(X_n b_1 \cdots X_n b_k)\| < \infty.$$

Entonces $\frac{1}{\sqrt{N}} \sum_{n=1}^N X_n$ converge en \mathcal{B} -distribución a un elemento semicircular con covarianza η .

Bosquejo de la demostración. La demostración sigue esencialmente los mismos argumentos que en el caso escalar, con la salvedad que ahora hay que considerar las \mathcal{B} -distribuciones. Específicamente, definamos a

$$s_N := \frac{1}{\sqrt{N}} \sum_{n=1}^N X_n.$$

Entonces para $m \in \mathbb{N}$ y $b_1, \dots, b_{m-1} \in \mathcal{B}$ se tiene que

$$E(s_N b_1 \cdots b_{m-1} s_N) = N^{-m/2} \sum_{\substack{i(j)=1, \dots, N \\ j \in \{1, \dots, m\}}} E(X_{i_1} b_1 \cdots b_{m-1} X_{i_m}).$$

De la condición III) y un análisis similar al del caso escalar, utilizando el Ejemplo 4.2 concluimos que únicamente las particiones que no se cruzan aportan a la suma anterior asintóticamente, i.e.

$$\lim_{N \rightarrow \infty} E(s_N b_1 \cdots b_{m-1} s_N) = \sum_{\pi \in \text{NC}_2(m)} E(X_{j_1(\pi)} b_1 \cdots b_{m-1} X_{j_m(\pi)})$$

donde $j_1(\pi), \dots, j_m(\pi)$ son índices compatibles con la partición π . Por último, basta

probar la igualdad

$$E \left(X_{j_1(\pi)} b_1 \cdots b_{m-1} X_{j_m(\pi)} \right) = \kappa_\pi (s b_1 \otimes \cdots \otimes s b_{m-1} \otimes s)$$

donde s es un elemento semicircular cualquiera. Esta igualdad se sigue de la condición II) y un procedimiento iterativo de cancelación de pares basado en el Ejemplo 4.3. \square

4.3. Transformadas y ecuaciones de subordinación

En esta sección se dan las versiones valuadas en operadores para las transformadas definidas en la Sección 2.4. Además, se enuncian las llamadas «ecuaciones de subordinación», las cuales utilizaremos en el siguiente capítulo.

Definición 4.15. Sea $(E, \mathcal{A}, \mathcal{B})$ un espacio de probabilidad valuado en operadores. Para $a \in \mathcal{A}$ se define su transformada de Cauchy valuada en operadores $G_a : \mathcal{B} \rightarrow \mathcal{B}$ por

$$G_a(b) = E \left((b - a)^{-1} \right),$$

siempre que $b - a$ sea invertible.

Comentario 4.12. Si tomamos E de tal forma que sea compatible con φ , i.e. $\varphi \circ E = \varphi$, entonces para $z \in \mathbb{C}$ se tiene que

$$g_a(z) = \varphi \left(G_a(z \mathbf{1}_{\mathcal{A}}) \right),$$

donde g_a es la transformada de Cauchy escalar de a . Puesto que la transformada de Cauchy escalar determina de manera única la distribución analítica μ_a de a , es claro que la transformada de Cauchy valuada en operadores también determina dicha distribución de manera única. En otras palabras, si $a, \hat{a} \in \mathcal{A}$ son tales que $G_a(b) = G_{\hat{a}}(b)$ para toda b con $b - a$ o $b - \hat{a}$ invertible, entonces $\mu_a = \mu_{\hat{a}}$.

La siguiente proposición ilustra una representación alternativa de la transformada

de Cauchy valuada en operadores. Al igual que en el caso escalar, las representaciones por medio de series de potencias permiten utilizar herramientas combinatorias, enriqueciendo a la contraparte analítica.

Proposición 4.3. *Sea $(E, \mathcal{A}, \mathcal{B})$ un espacio de probabilidad valuado en operadores. Si $a \in \mathcal{A}$ es una variable aleatoria y $b \in \mathcal{B}$ es invertible tal que $\|b^{-1}\| < \|a\|^{-1}$, entonces*

$$G_a(b) = \sum_{n \geq 0} E(b^{-1}(ab^{-1})^n).$$

Demostración. Dado que b es invertible tenemos que

$$\begin{aligned} (b - a)^{-1} &= ((\mathbf{1}_{\mathcal{A}} - ab^{-1})b)^{-1} \\ &= b^{-1}(\mathbf{1}_{\mathcal{A}} - ab^{-1})^{-1}, \end{aligned}$$

y puesto que $\|b\| > \|a\|$ entonces

$$\begin{aligned} (b - a)^{-1} &= b^{-1} \sum_{n \geq 0} (ab^{-1})^n \\ &= \sum_{n \geq 0} b^{-1}(ab^{-1})^n. \end{aligned}$$

Puesto que E es completamente positivo, en particular es continuo⁵. Por lo tanto

$$\begin{aligned} G_a(b) &= E\left(\sum_{n \geq 0} b^{-1}(ab^{-1})^n\right) \\ &= \sum_{n \geq 0} E(b^{-1}(ab^{-1})^n), \end{aligned}$$

tal y como queríamos demostrar. □

Definición 4.16. Sea $a \in \mathcal{A}$ una variable aleatoria en $(E, \mathcal{A}, \mathcal{B})$. Se define la transfor-

⁵Este hecho es un resultado conocido de la teoría de C^* -álgebras. Véase [21] y referencias allí.

mada \mathcal{R} valuada en operadores de a para $b \in \mathcal{B}$ mediante

$$\mathcal{R}_a(b) = \sum_{n \geq 0} \kappa_{n+1}(ab \otimes \cdots \otimes ab \otimes a).$$

Comentario 4.13. Se puede demostrar que la serie de la definición anterior converge en un vecindad de cero. De hecho, existe toda una teoría acerca de los aspectos analíticos relacionados. En este trabajo no abordaremos el tema, remitimos al lector interesado al artículo de Belinschi, Mai y Speicher [3].

Ejemplo 4.7. Si s es un elemento semicircular con covarianza η , de la Proposición 4.2 tenemos que

$$\begin{aligned} \mathcal{R}_s(b) &= \sum_{n \geq 0} \kappa_{n+1}(sb \otimes \cdots \otimes sb \otimes s) \\ &= \kappa_2(sb \otimes s) \\ &= \eta(b). \end{aligned}$$

Teorema 4.4. *Sea a una variable aleatoria en $(E, \mathcal{A}, \mathcal{B})$, entonces las siguientes afirmaciones se cumplen:*

a) *La transformada de Cauchy y la transformada R de a satisfacen que*

$$bG_a(b) = \mathbf{1} + \mathcal{R}_a(G_a(b)) \cdot G_a(b). \quad (4.2)$$

b) *Si x, y son libre sobre \mathcal{B} , entonces los cumulantes mixtos \mathcal{B} -valuadas en x y y se anulan, lo cual implica la aditividad de la transformada R*

$$\mathcal{R}_{x+y}(b) = \mathcal{R}_x(b) + \mathcal{R}_y(b).$$

c) *Si x, y son libres respecto a E entonces*

$$G_{x+y}(b) = G_x(b - R_y(G_{x+y}(b))). \quad (4.3)$$

Comentario 4.14. La afirmación a) de la proposición anterior nos permite encontrar la transformada \mathcal{R} a partir de la transformada de Cauchy, específicamente se tiene que

$$\mathcal{R}_a(b) = G_a^{-1}(b) - b^{-1}.$$

Por lo tanto, si se tienen las transformadas de Cauchy G_x y G_y de dos variables aleatorias x y y entonces se tienen las transformadas \mathcal{R}_x y \mathcal{R}_y . De b) y c) se tiene entonces que la transformada de Cauchy de la suma de dos variables aleatorias libres es la solución a la ecuación de punto fijo (4.3). Sin embargo, calcularemos la transformada de Cauchy de la suma de variables aleatorias libres respecto a E mediante la ecuación de subordinación aditiva del siguiente teorema.

Definición 4.17. Sea $(E, \mathcal{A}, \mathcal{B})$ un espacio de probabilidad valuado en operadores con \mathcal{A} y \mathcal{B} W^* -álgebras. Para una variable aleatoria $x \in \mathcal{A}$ definimos su parte real $\Re x$ y su parte imaginaria $\Im x$ por

$$\begin{aligned}\Re x &= \frac{x + x^*}{2}, \\ \Im x &= \frac{x - x^*}{2i}.\end{aligned}$$

Decimos que $x > 0$ si x es invertible y $x \geq 0$. Definimos el semiplano superior de \mathcal{B} , denotado por $\mathbb{H}^+(\mathcal{B})$, mediante

$$\mathbb{H}^+(\mathcal{B}) := \{b \in \mathcal{B} \mid \Im b > 0\}.$$

Para $x = x^*$ en \mathcal{A} definimos, para toda $b \in \mathbb{H}^+(\mathcal{B})$, su transformada r y su transformada h mediante

$$\begin{aligned}r_x(b) &= G_x(b)^{-1} - b, \\ h_x(b) &= b^{-1} - G_x(b^{-1})^{-1}.\end{aligned}$$

Teorema 4.5. Sean $x, y \in \mathcal{A}$ variables aleatorias autoadjuntas libres respecto a E . Entonces para toda $b \in \mathbb{H}^+(\mathcal{B})$ se tiene que

$$G_{x+y}b = G_x(\omega_1(b)),$$

con $\omega_1(b) = \lim_{n \rightarrow \infty} f_b^n(\omega)$ para cualquier $\omega \in \mathcal{H}^+(\mathcal{B})$ y

$$f_b(\omega) = r_y(r_x(\omega) + b) + b.$$

Para finalizar, a continuación damos los elementos necesarios para calcular la transformada de Cauchy del producto de variables aleatorias libres. Remitimos al lector interesado en los detalles de la teoría analítica involucrada al artículo de Belinschi, Speicher, Treilhard y Vargas [4] del 2013.

Teorema 4.6. Sean $x > 0$ y $y = y^*$ variables aleatorias en \mathcal{A} libres respecto a $E : \mathcal{A} \rightarrow \mathcal{B}$ con esperanzas condicionales invertibles. Para toda $b \in \mathcal{B}$ con $\Im(bx) > 0$ definamos la función $g_b(\omega) = bh_x(h_y(\omega)b)$ para toda $\omega \in \mathbb{H}^+(\mathcal{B})$. Entonces existe una función ω_2 tal que

$$\omega_2(b) = \lim_{n \rightarrow \infty} g_b^n(\omega)$$

para toda $\omega \in \mathbb{H}^+(\mathcal{B})$, y además

$$\begin{aligned} G_{xy}(z\mathbf{1}_{\mathcal{B}}) &= (z\mathbf{1}_{\mathcal{B}} - h_{xy}(z^{-1}\mathbf{1}_{\mathcal{B}}))^{-1}, \\ zh_{xy}(z\mathbf{1}_{\mathcal{B}}) &= \omega_2(z\mathbf{1}_{\mathcal{B}})h_y(\omega_2(z\mathbf{1}_{\mathcal{B}})). \end{aligned}$$

Comentario 4.15. De las fórmulas anteriores es claro que únicamente son necesarias las transformadas de Cauchy valuadas en operadores de variables aleatorias libres para calcular la transformada de Cauchy de su producto.

4.4. Espacios de probabilidad rectangulares

Esta sección está dedicada a la exposición de una clase muy particular de espacios valuados en operadores, los llamados espacios rectangulares. Estos espacios de probabilidad nos permitirán trabajar de manera sistemática con matrices no cuadradas.

Definición 4.18. Para $d \in \mathbb{N}$ fijo, decimos que (\mathcal{A}, φ) es un (p_1, \dots, p_d) -espacio de probabilidad si (\mathcal{A}, φ) es un $*$ -espacio de probabilidad tracial y p_1, \dots, p_d son proyecciones autoadjuntas no cero, ortogonales a pares y tales que $p_1 + \dots + p_d = \mathbf{1}_{\mathcal{A}}$.

Ejemplo 4.8. Sean $P_1, \dots, P_d \in \mathbb{N}$ fijos y $N := P_1 + \dots + P_d$. Denotemos por \mathcal{A} a las matrices aleatorias de $N \times N$ y por $\varphi := \mathbb{E} \otimes \text{tr}_N : \mathcal{A} \rightarrow \mathbb{C}$. Para cada $k \in \{1, \dots, d\}$, denotemos por p_k al elemento dado por

$$(p_k)_{a,b} = \begin{cases} 1 & \text{si } a = b \text{ y } a \in (P_1 + \dots + P_{k-1}, P_1 + \dots + P_k] \\ 0 & \text{en otro caso} \end{cases}$$

para toda $a, b \in \{1, \dots, N\}$. En este contexto (\mathcal{A}, φ) es un (p_1, \dots, p_d) -espacio de probabilidad.

Más aún, si A es una matriz aleatoria de $P_i \times P_k$ para algunos $i, k \in \{1, \dots, d\}$ entonces existe un encajamiento canónico \bar{A} de A en \mathcal{A} el cual está determinado por

$$\bar{A}_{a,b} = \begin{cases} A_{a-(P_1+\dots+P_{i-1}), b-(P_1+\dots+P_{k-1})} & \text{si } a \in (P_1 + \dots + P_{i-1}, P_1 + \dots + P_i] \\ & \text{y } b \in (P_1 + \dots + P_{k-1}, P_1 + \dots + P_k], \\ 0 & \text{en otro caso.} \end{cases}$$

para $a, b \in \{1, \dots, N\}$.

De esta manera podemos trabajar con matrices aleatorias rectangulares vía matrices aleatorias a bloques. Más específicamente, podemos pensar que cada $A \in \mathcal{A}$ es de la

forma

$$A = \begin{pmatrix} A^{(1,1)} & \cdots & A^{(1,d)} \\ \vdots & \ddots & \vdots \\ A^{(d,1)} & \cdots & A^{(d,d)} \end{pmatrix}$$

con $A^{(i,k)}$ una matrices aleatoria de dimensión $P_i \times P_k$.

Abusando de la notación, denotaremos por A tanto al encajamiento como a la matriz original.

La siguiente definición nos proporciona una representación, que como veremos más adelante, nos permite llevar la intuición de las matrices a bloques a los espacios rectangulares.

Definición 4.19. Sea (\mathcal{A}, φ) un (p_1, \dots, p_d) -espacio de probabilidad. Para $x \in \mathcal{A}$ denotamos por \tilde{x} a la representación de x en $\text{Mat}_d(\mathcal{A})$ dada por

$$\tilde{x} = \begin{pmatrix} \tilde{x}_{11} & \cdots & \tilde{x}_{1d} \\ \vdots & \ddots & \vdots \\ \tilde{x}_{d1} & \cdots & \tilde{x}_{dd} \end{pmatrix},$$

donde $\tilde{x}_{ij} = p_i x p_j$ para $i, j \in \{1, \dots, d\}$.

Observación 4.1. El mapeo $\tilde{\cdot} : \mathcal{A} \rightarrow \text{Mat}_d(\mathcal{A})$ de la definición anterior es inyectivo. Más aún, es biyectivo si nos restringimos al subconjunto de $\text{Mat}_d(\mathcal{A})$ que contiene a las matrices tales que el elemento en la posición i, j es de la forma $p_i a p_j$ para algún $a \in \mathcal{A}$.

La siguiente proposición es una consecuencia directa de la definición anterior.

Proposición 4.4. Sean $x, y \in \mathcal{A}$, entonces

$$\begin{aligned} x &= \sum_{i,j=1}^d \tilde{x}_{ij}, \\ \widetilde{xy} &= \tilde{x}\tilde{y}, \\ \widetilde{x+y} &= \tilde{x} + \tilde{y} \\ \widetilde{x^*} &= \tilde{x}^*. \end{aligned}$$

Observación 4.2. La proposición anterior y la Observación 4.1 implican que podemos trabajar con la representación o con el elemento del álgebra indistintamente. Abusando de la notación, en lo sucesivo utilizaremos a x para representar a x y \tilde{x} indistintamente.

Definición 4.20. Definamos $\mathcal{A}_{i,j} := p_i \mathcal{A} p_j$ para $i, j \in \{1, \dots, d\}$. Llamaremos elementos simples a los elementos en

$$\bigcup_{i,j=1}^d \mathcal{A}_{i,j} \setminus \{0_{\mathcal{A}}\}.$$

Además, para $i \in \{1, \dots, d\}$ definamos $\varphi_i : \mathcal{A}_{i,i} \rightarrow \mathbb{C}$ mediante

$$\varphi_i = \frac{1}{\rho_i} \varphi|_{\mathcal{A}_{i,i}},$$

donde $\rho_i = \varphi(p_i)$.

Corolario 4.1. Para cada $i \in \{1, \dots, d\}$ se tiene que $(\mathcal{A}_{i,i}, \varphi_i)$ es un $*$ -espacio de probabilidad tracial.

Proposición 4.5. Sea $a \in \mathcal{A}_{i,j}$ y $b \in \mathcal{A}_{j,i}$, entonces

$$\rho_i \varphi_i(ab) = \rho_j \varphi_j(ba). \tag{4.4}$$

Demostración. Un cálculo muestra que

$$\begin{aligned}\rho_i \varphi_i(ab) &= \rho_i \frac{1}{\rho_i} \varphi|_{\mathcal{A}_{ii}}(ab) \\ &= \varphi(ab) \\ &= \varphi(ba) \\ &= \rho_j \varphi_j(ba),\end{aligned}$$

tal y como queríamos demostrar. □

Comentario 4.16. La proposición anterior aunque sencilla, refleja una de las características principales de los espacios rectangulares, a saber, es posible pasar de un subespacio $\mathcal{A}_{i,i}$ a otro mediante una ecuación de *balance*⁶. Encontraremos ecuaciones similares para esperanzas condicionales y cumulantes.

Definición 4.21. Definamos $\mathcal{D} := \langle p_1, \dots, p_d \rangle$ y $E : \mathcal{A} \rightarrow \mathcal{D}$ por

$$E(x) = \text{diag}(\varphi_1(x_{11}), \dots, \varphi_d(x_{dd})).$$

Corolario 4.2. *El espacio $(E, \mathcal{A}, \mathcal{D})$ es un espacio de probabilidad valuado en operadores.*

Comentario 4.17. Es claro que \mathcal{D} es una $*$ -subálgebra de \mathcal{A} , ya que la biyección dada por

$$\sum_{i=1}^d \lambda_i p_i \mapsto \text{diag}(\lambda_1, \dots, \lambda_d)$$

⁶Si X es una matriz gaussiana de $N \times M$ entonces

$$\begin{aligned}\text{tr}_N(XX^*) &= \frac{1}{N} \text{tr}(XX^*) \\ &= \frac{M}{N} \text{tr}_M(X^*X).\end{aligned}$$

Observemos pues que los momentos de XX^* y X^*X son prácticamente los mismos salvo un coeficiente, este tipo de coeficientes se maneja sistemáticamente con este tipo de ecuaciones de balance.

identifica a \mathcal{D} con las matrices diagonales, las cuales son una $*$ -subálgebra de $\text{Mat}_d(\mathcal{A})$.

Demostración del Corolario 4.2. Para demostrar este corolario basta demostrar que⁷

- $E(\mathbf{1}_{\mathcal{A}}) = \mathbf{1}_{\mathcal{A}}$,
- $E(d_1 a d_2) = d_1 E(a) d_2$ para cualesquieras $a \in \mathcal{A}$ y $d_1, d_2 \in \mathcal{D}$.

La primer propiedad se cumple trivialmente de las definiciones de φ_i y x_{ii} . Puesto que la representación en $\text{Mat}_d(\mathcal{A})$ de cualquier elemento de \mathcal{D} es una matriz diagonal con entradas complejas, un cálculo directo demuestra la segunda propiedad. \square

De la linealidad de los cumulantes, un cálculo directo nos muestra el siguiente corolario.

Corolario 4.3. Sean $a, b \in \mathcal{A}$, entonces

$$\kappa_2(a \otimes b) = \sum_{i,k,j,l=1}^d \kappa_2(a_{ik} \otimes b_{jl}).$$

Por inducción podemos extender el corolario anterior para cualquier cantidad finita de variables aleatorias. De dicha extensión concluimos que basta trabajar con elementos simples.

Corolario 4.4. Sean $a \in \mathcal{A}_{i,k}$ y $b \in \mathcal{A}_{j,l}$. Si $k \neq j$ entonces $\kappa_2(a \otimes b) = 0$.

Demostración. Observemos que

$$\begin{aligned} \kappa_2(a \otimes b) &= \kappa_2(p_i a p_k \otimes p_j b p_l) \\ &= \kappa_2(p_i a p_k p_j \otimes b p_l). \end{aligned}$$

De la ortogonalidad de las proyecciones el resultado se sigue. \square

⁷Como se mencionó anteriormente, puesto que estamos trabajando desde un punto de vista algebraico omitimos la demostración de que E es completamente positivo. Para la demostración de este hecho véase [21].

Por lo tanto, al trabajar con κ_n nos restringiremos a trabajar con elementos en

$$\mathcal{A}_{i_0, i_1} \otimes_{\mathcal{D}} \mathcal{A}_{i_1, i_2} \otimes_{\mathcal{D}} \cdots \otimes_{\mathcal{D}} \mathcal{A}_{i_{n-1}, i_n}$$

para algunos $i_0, \dots, i_n \in \{1, \dots, d\}$.

Corolario 4.5. *Sea $n \in \mathbb{N}$, $\pi \in \text{NC}(n)$, $i_0, \dots, i_n \in \{1, \dots, d\}$ y $a \in \mathcal{A}_{i_0, i_1} \otimes_{\mathcal{D}} \cdots \otimes_{\mathcal{D}} \mathcal{A}_{i_{n-1}, i_n}$. Entonces*

$$\kappa_\pi(a) = \widehat{\kappa}_\pi p_{i_n}$$

con $\widehat{\kappa}_\pi \in \mathbb{C}$ dada por

$$\widehat{\kappa}_\pi = \begin{cases} \prod_{\substack{V \in \pi \\ V = \{k_1 < \dots < k_m\}}} \widehat{\kappa}_m(a_{k_1} \otimes \cdots \otimes a_{k_m}) & i_0 = i_n \\ 0 & i_0 \neq i_n. \end{cases} \quad (4.5)$$

Observación 4.3. Si $\pi = \mathbf{1}_{\text{NC}(n)}$ entonces la fórmula (4.5) únicamente afirma que $\kappa_n(a) \in \langle p_{i_n} \rangle$, y que $\widehat{\kappa}_n$ queda determinado por el coeficiente que acompaña a p_{i_n} en $\kappa_n(a)$.

Comentario 4.18. La esencia de los espacios rectangulares reside principalmente en las ecuaciones de balance y el corolario anterior. Las ecuaciones de balance nos proporcionan una especie de tracialidad, mientras que el corolario anterior nos muestra que los cumulantes rectangulares obedecen prácticamente las mismas reglas que los escalares.

Demostración del Corolario 4.5. La demostración se realizará por inducción en n . Si $n = 1$, entonces $a \in \mathcal{A}_{i_0, i_1}$ y

$$\begin{aligned} \kappa_1(a) &= E_1(a) \\ &= E(a) \\ &= \text{diag}(\varphi_1(a_{11}), \dots, \varphi_d(a_{dd})). \end{aligned}$$

Esto implica que si $i_0 \neq i_1$ entonces $\kappa_1(a) = 0p_{i_1} = \mathbf{0}_{\mathcal{A}}$, mientras que $\kappa_1(a) = \varphi_{i_1}(a_{i_1 i_1})p_{i_1}$ si $i_0 = i_1$. Supongamos que se cumple para $n \in \mathbb{N}$. Si $a = p_{i_0}a_1p_{i_1} \otimes \cdots \otimes p_{i_n}a_{n+1}p_{i_{n+1}}$ con $a_i \in \mathcal{A}$ para $i = 1, \dots, n+1$, entonces de la fórmula momentos-cumulantes se tiene que para $\sigma = \mathbf{1}_{\text{NC}(n+1)}$

$$\begin{aligned}\kappa_{n+1}(a) &= \sum_{\pi \in \text{NC}(n+1)} E_{\pi}(a)\mu(\pi, \mathbf{1}_{\text{NC}(n+1)}) \\ &= E_n(a) + \sum_{\pi < \mathbf{1}_{\text{NC}(n+1)}} E_{\pi}(a)\mu(\pi, \mathbf{1}_{\text{NC}(n+1)}).\end{aligned}$$

Observemos que

$$\begin{aligned}E_n(a) &= E_n(p_{i_0}a_1p_{i_1} \otimes \cdots \otimes p_{i_n}a_{n+1}p_{i_{n+1}}) \\ &= E(p_{i_0}a_1p_{i_1} \cdots p_{i_n}a_{n+1}p_{i_{n+1}}) \\ &= \kappa_1(p_{i_0}a_1p_{i_1} \cdots p_{i_n}a_{n+1}p_{i_{n+1}}),\end{aligned}$$

y por el paso base tenemos que $E_n(a) \in \langle p_{i_{n+1}} \rangle$. Para $\pi < \mathbf{1}_{n+1}$, entonces existe un bloque $V \in \pi$ con $|V| \leq n$ tal que $V = \{i_j, i_j + 1, \dots, i_j + |V| - 1\}$ para algún $j \in \{1, \dots, n\}$. De la estructura multiplicativa de E y la hipótesis de inducción se tiene que

$$\begin{aligned}E_{\pi}(a) &= E_{\pi \setminus V}(p_{i_0}a_1p_{i_1} \otimes \cdots \otimes E_{|V|}(p_{i_{j-1}}a_jp_{i_j} \otimes \cdots \otimes p_{i_{j+|V|-2}}a_{j+|V|-1}p_{i_{j+|V|-1}})) \quad (4.6) \\ &\quad p_{i_{j+|V|-1}}a_{j+|V|}p_{i_{j+|V|}} \otimes \cdots \otimes p_{i_n}a_{n+1}p_{i_{n+1}}) \in \langle p_{i_{n+1}} \rangle,\end{aligned}$$

donde $\pi \setminus V$ representa a la partición π cuando se le quita el bloque V . Observemos que aquí supusimos que V es un bloque *interno* de π , si no es así, entonces $E_{\pi}(a)$ se factoriza como el producto de esperanzas de orden menor, y la conclusión sigue siendo válida. Por lo tanto, $\kappa_{n+1}(a) \in \langle p_{i_{n+1}} \rangle$.

Si $\sigma < \mathbf{1}_{\text{NC}(n+1)}$, entonces la ecuación (4.6) sigue siendo válida para κ en lugar E y

σ en lugar de π . Por la hipótesis de inducción tenemos que

$$\begin{aligned} \kappa_\pi(a) = & \kappa_{\pi \setminus V'}(p_{i_0} a_1 p_{i_1} \otimes \cdots \otimes \kappa_{|V|}(p_{i_{j-1}} a_j p_{i_j} \otimes \cdots \otimes p_{i_{j+|V|-2}} a_{j+|V|-1} p_{i_{j+|V|-1}}) \\ & p_{i_{j+|V|-1}} a_{j+|V|} p_{i_{j+|V|}} \otimes \cdots \otimes p_{i_n} a_{i_{n+1}} p_{i_{n+1}}). \end{aligned}$$

Puesto que $\mathcal{A}_{i_0, i_1} \otimes_{\mathcal{D}} \mathcal{A}_{i_1, i_2} \otimes_{\mathcal{D}} \cdots \otimes_{\mathcal{D}} \mathcal{A}_{i_{n-1}, i_n}$ para algunos $i_0, \dots, i_n \in \{1, \dots, d\}$, entonces

$$\begin{aligned} \kappa_\pi(a) = & \widehat{\kappa}_{|V|}(p_{i_{j-1}} a_j p_{i_j} \otimes \cdots \otimes p_{i_{j+|V|-2}} a_{j+|V|-1} p_{i_{j+|V|-1}}) \\ & \kappa_{\pi \setminus V}(p_{i_0} a_1 p_{i_1} \otimes \cdots \otimes p_{i_{j+|V|-1}} a_{j+|V|} p_{i_{j+|V|}} \otimes \cdots \otimes p_{i_n} a_{i_{n+1}} p_{i_{n+1}}). \end{aligned}$$

Aplicando el procedimiento anterior de manera iterativa con los bloques de σ obtenemos la fórmula (4.5), tal y como queríamos demostrar. \square

Finalizaremos esta sección con un teorema que enunciamos sin demostración, sin embargo, la prueba sigue la técnica del corolario anterior, i.e. utilizar la relación momentos-cumulantes y analizar partición por partición. Los detalles técnicos completos pueden ser vistos en [5].

Teorema 4.7. *Sea $a = a_1 \otimes \cdots \otimes a_n$ con $a_j \in \mathcal{A}_{i_{j-1}, i_j}$ y $i_0 = i_n$. Entonces*

$$\rho_{i_0} \widehat{\kappa}_n(a_1 \otimes \cdots \otimes a_n) = \rho_{i_1} \widehat{\kappa}_n(a_2 \otimes \cdots \otimes a_1) \quad (4.7)$$

y

$$\rho_{i_0} \widehat{E}_n(a_1 \otimes \cdots \otimes a_n) = \rho_{i_1} \widehat{E}_n(a_2 \otimes \cdots \otimes a_1). \quad (4.8)$$

Corolario 4.6. *Sea $a = a_1 \otimes \cdots \otimes a_n$ con $a_j \in \mathcal{A}_{i_{j-1}, i_j}$ y $i_0 = i_n$. Entonces*

4.5. Matrices a bloques

En esta sección estudiamos una clase de espacios de probabilidad valuados en operadores los cuales nos permitirán trabajar con matrices a bloques en las que todos sus bloques son cuadrados.

Proposición 4.6. *Sea (\mathcal{C}, φ) un $*$ -espacio de probabilidad. Definamos las álgebras $\mathcal{A} := \text{Mat}_d(\mathcal{C})$ y $\mathcal{B} := \text{Mat}_d(\mathbb{C})$. Denotemos por $E : \mathcal{A} \rightarrow \mathcal{B}$ a la función $\mathbf{1} \otimes \varphi$, la cual actúa en \mathcal{A} mediante*

$$(a_{i,j})_{i,j=1}^d \mapsto (\varphi(a_{i,j}))_{i,j=1}^d.$$

Entonces $(E, \mathcal{A}, \mathcal{B})$ es un espacios de probabilidad valuado en operadores.

La demostración de esta proposición es un ejercicio de rutina. En este contexto hay varios problemas de interés, como el de encontrar a los elementos semicirculares de \mathcal{A} . Al final de este capítulo enunciaremos un teorema derivado del estudio de estos elementos semicirculares el cual nos permite calcular el espectro asintótico de matrices aleatorias gaussianas a bloques con una cierta correlación.

Ejemplo 4.9. Sea $s \in \mathcal{A}$ una variable aleatoria autoadjunta tal que $\{s_{i,j} \mid i, j \in \{1, \dots, d\}\}$ es una familia semicircular con función de covarianza σ , i.e. $\varphi(s_{i,k}s_{j,l}) = \sigma(i, k; j, l)$. Para $i, j \in \{1, \dots, d\}$ se tiene que

$$\begin{aligned} [E(s^m)]_{i,j} &= \varphi([s^m]_{i,j}) \\ &= \sum_{\substack{i(k)=1,\dots,d \\ k \in \{2,\dots,m\}}} \varphi(s_{i,i(2)} \cdots s_{i(m),j}) \\ &= \sum_{\substack{i(k)=1,\dots,d \\ k \in \{2,\dots,m\}}} \sum_{\pi \in \text{NC}_2(m)} \prod_{(p,q) \in \pi} \sigma(i(p), i(p+1); i(q), i(q+1)) \\ &= \sum_{\pi \in \text{NC}_2(m)} [c_\pi(s \otimes \cdots \otimes s)]_{i,j}, \end{aligned}$$

donde

$$[c_\pi(s \otimes \cdots \otimes s)]_{i,j} = \sum_{\substack{i(k)=1,\dots,d \\ k \in \{2,\dots,m\}}} \prod_{(p,q) \in \pi} \sigma(i(p), i(p+1); i(q), i(q+1)).$$

Si demostramos que la matriz $c_\pi(s \otimes \cdots \otimes s) = ([c_\pi]_{i,j})_{i,j=1}^d$ es igual al cumulante $\kappa_\pi(s \otimes$

$\cdots \otimes s$) para toda $\pi \in \text{NC}$, habremos demostrado que s es un elemento semicircular en $(E, \mathcal{A}, \mathcal{B})$.

En el ejemplo anterior $c_\pi(s \otimes \cdots \otimes s)$ es la notación utilizada para la matriz allí definida y no hace referencia, en principio, a una función. Sin embargo, como veremos en los siguientes ejemplos, es posible dotar a $(c_\pi(s \otimes \cdots \otimes s))_{\pi \in \text{NC}}$ de una estructura de tal manera que c_π defina una función de $\mathcal{A}^{\otimes \mathcal{B}^n}$ a \mathcal{B} . Omitiremos el argumento $(s \otimes \cdots \otimes s)$ cuando esto no cause confusión.

Notación 4.1. Denotemos por η a la función de covarianza de s , i.e.

$$\eta(b) := E(sbs)$$

para toda $b \in \mathcal{B}$.

Observación 4.4. En términos de σ la función de covarianza está dada por

$$\begin{aligned} [\eta(b)]_{i,j} &= \varphi([sbs]_{i,j}) \\ &= \varphi\left(\sum_{k,l=1}^d s_{i,k} b_{k,l} s_{l,j}\right) \\ &= \sum_{k,l=1}^d b_{k,l} \sigma(i, k; l, j). \end{aligned}$$

Corolario 4.7. El función de covarianza es lineal.

Demostración. Sean $a, b \in \text{Mat}_d(\mathbb{C})$ y $\alpha \in \mathbb{C}$, entonces

$$\begin{aligned} \eta(\alpha a + b) &= E(s(\alpha a + b)s) \\ &= \alpha E(sas) + E(sbs) \\ &= \alpha \eta(a) + \eta(b), \end{aligned}$$

es decir, η es lineal. □

El siguiente teorema lo enunciamos sin demostración. En su lugar damos dos ejemplos para ilustrar la metodología de la prueba, a saber, la cancelación sucesiva de los pares de una partición que no se cruza.

Teorema 4.8. *Podemos construir una familia de funciones $(c_\pi)_{\pi \in \text{NC}}$ multiplicativa⁸ respecto a π la cual extiende la notación dada anteriormente para $c_\pi(s \otimes \cdots \otimes s)$.*

Comentario 4.19. Puesto que $c_1(s) = E(s)$ y $\varphi(s_{i,j}) = 0$ para toda i, j , tenemos que $c_1(s) = \mathbf{0}_B$. Más aún, se tiene que

$$\begin{aligned} [c_2(s \otimes s)]_{i,j} &= \sum_{i(2)=1,\dots,d} \sigma(i, i(2); i(2), j) \\ &= \eta(\mathbf{1}_B). \end{aligned}$$

Por lo tanto, el teorema anterior se puede reinterpretar de la siguiente manera: « $c_\pi(s \otimes \cdots \otimes s)$ puede ser calculado recursivamente a partir de η y la estructura de $\pi \in \text{NC}_2$.»

Ejemplo 4.10. Sea $\pi = \{(1, 2), (3, 4)\} \in \text{NC}_2(4)$. Observemos que para $i, j \in \{1, \dots, d\}$ se tiene que

$$\begin{aligned} [c_\pi]_{i,j} &= \sum_{\substack{i(k)=1,\dots,d \\ k \in \{2,3,4\}}} \prod_{(p,q) \in \pi} \sigma(i(p), i(p+1); i(q), i(q+1)) \\ &= \sum_{i(2), i(3), i(4)=1,\dots,d} \sigma(i, i(2); i(2), i(3)) \sigma(i(3), i(4); i(4), j). \end{aligned}$$

Puesto que la expresión $\sigma(i, i(2); i(2), i(3)) \sigma(i(3), i(4); i(4), j)$ sólo depende de $i(4)$ a través del segundo factor, podemos reescribir a $[c_\pi]_{i,j}$ como

$$[c_\pi]_{i,j} = \sum_{i(2), i(3)=1,\dots,d} \sigma(i, i(2); i(2), i(3)) \sum_{i(4)=1,\dots,d} \sigma(i(3), i(4); i(4), j). \quad (4.9)$$

⁸Por supuesto hablamos de multiplicatividad en el sentido no conmutativo, tal y como se ha trabajado a través de este capítulo.

Del comentario anterior es claro que

$$[c_\pi]_{i,j} = \sum_{i(2),i(3)=1,\dots,d} \sigma(i, i(2); i(2), i(3))[\eta(\mathbf{1}_B)]_{i(3),j}.$$

Aplicando este mismo argumento para $i(2)$ en lugar de $i(4)$ se tiene que

$$\begin{aligned} [c_\pi]_{i,j} &= \sum_{i(3)=1,\dots,d} [\eta(\mathbf{1}_B)]_{i,i(3)}[\eta(\mathbf{1}_B)]_{i(3),j} \\ &= [\eta(\mathbf{1}_B)\eta(\mathbf{1}_B)]_{i,j}. \end{aligned}$$

Puesto que la última expresión se cumple para toda $i, j \in \{1, \dots, d\}$ se tiene que $c_\pi = \eta(\mathbf{1}_B)\eta(\mathbf{1}_B)$.

Este ejemplo nos permite concluir que a la partición⁹ $\pi = ()()$ le corresponde el valor de $c_\pi = \eta(\mathbf{1}_B)\eta(\mathbf{1}_B)$. En lo sucesivo omitiremos el índice \mathcal{B} en $\mathbf{1}_B$.

Ejemplo 4.11. Sea $\pi = \{(1, 2), (3, 6), (4, 5)\} \in \text{NC}_2(6)$. Aplicando la misma metodología del ejemplo anterior, para $i, j \in \{1, \dots, d\}$ se tiene que

$$\begin{aligned} [c_\pi]_{i,j} &= \sum_{\substack{i(k)=1,\dots,d \\ k \in \{2,\dots,6\}}} \sigma(i, i(2); i(2), i(3))\sigma(i(3), i(4); i(6), j)\sigma(i(4), i(5); i(5), i(6)) \\ &= \sum_{\substack{i(k)=1,\dots,d \\ k \in \{3,\dots,6\}}} [\eta(\mathbf{1})]_{i,i(3)}\sigma(i(3), i(4); i(6), j)\sigma(i(4), i(5); i(5), i(6)) \\ &= \sum_{\substack{i(k)=1,\dots,d \\ k \in \{3,\dots,6\} \setminus \{5\}}} [\eta(\mathbf{1})]_{i,i(3)}\sigma(i(3), i(4); i(6), j)[\eta(\mathbf{1})]_{i(4),i(6)} \\ &= \sum_{i(3)=1,\dots,d} [\eta(\mathbf{1})]_{i,i(3)} \sum_{i(4),i(6)=1,\dots,d} [\eta(\mathbf{1})]_{i(4),i(6)}\sigma(i(3), i(4); i(6), j) \\ &= \sum_{i(3)=1,\dots,d} [\eta(\mathbf{1})]_{i,i(3)}[\eta(\eta(\mathbf{1}))]_{i(3),j} \\ &= [\eta(\mathbf{1})\eta(\eta(\mathbf{1}))]_{i,j}. \end{aligned}$$

⁹Escribimos a π como $()()$ en lugar de $\{(1, 2), (3, 4)\}$ puesto que en la primer expresión es más fácil observar lo que queremos mostrar.

Por lo tanto se tiene que $c_\pi = \eta(\mathbf{1})\eta(\eta(\mathbf{1}))$.

De los ejemplos anteriores y su metodología podemos concluir que

$$\begin{aligned}\pi = ()() &\rightarrow c_\pi = \eta(\mathbf{1})\eta(\mathbf{1}) \\ \pi = ()(()) &\rightarrow c_\pi = \eta(\mathbf{1})\eta(\eta(\mathbf{1})) \\ \pi = ((())) &\rightarrow c_\pi = \eta(\eta(\eta(\mathbf{1}))) \\ \pi = ()()() &\rightarrow c_\pi = \eta(\mathbf{1})\eta(\mathbf{1})\eta(\mathbf{1}).\end{aligned}$$

Esto muestra que c_π puede ser calculado a través de η y la estructura de π , mientras que el siguiente ejemplo muestra la multiplicatividad de c_π respecto a π .

Ejemplo 4.12. Sea $\pi = ((())())$. De los ejemplos anteriores se puede concluir que

$$c_\pi = \eta(\eta(\mathbf{1})\eta(\mathbf{1}))\eta(\mathbf{1}).$$

Si escribimos a π como $\pi = \{(1, 6)\} \cup \pi_1 \cup \pi_2$ donde $\pi_1 = \{(2, 3), (4, 5)\}$ y $\pi_2 = \{(7, 8)\}$. Definamos $\widehat{\pi}_1 = ()() \in \text{NC}_2(4)$ y $\widehat{\pi}_2 = () \in \text{NC}_2(2)$ a las identificaciones naturales de π_1 y π_2 como elementos de $\text{NC}_2(4)$ y $\text{NC}_2(2)$ respectivamente. Puesto que $c_{\widehat{\pi}_1} = \eta(\mathbf{1})\eta(\mathbf{1})$ y $c_{\widehat{\pi}_2} = \eta(\mathbf{1})$ concluimos que

$$c_\pi = \eta(c_{\widehat{\pi}_1})c_{\widehat{\pi}_2}.$$

Esto muestra la estructura multiplicatividad de c_π respecto a π .

Comentario 4.20. De esto es claro que $(c_\pi)_{\pi \in \text{NC}_2}$ es de hecho $(\kappa_\pi)_{\pi \in \text{NC}_2}$ y por lo tanto s es un elemento semicircular en $(E, \mathcal{A}, \mathcal{B})$.

Por compleción, a continuación demostraremos la afirmación a) del Teorema 4.4 en el caso de la variable semicircular s .

Corolario 4.8. *Los momentos valuados en operadores $(E(s^m))_{m \geq 0}$ satisfacen la siguiente recurrencia*

$$E(s^m) = \sum_{k=0}^{m-2} \eta(E(s^k)) \cdot E(s^{m-2-k}).$$

Demostración. Aplicando las ideas del Ejemplo 4.12 tenemos que

$$\begin{aligned}
E(s^m) &= \sum_{\pi \in \text{NC}_2(m)} \kappa_\pi \\
&= \sum_{r=2}^m \sum_{\pi_1 \in \text{NC}_2(r-2)} \sum_{\pi_2 \in \text{NC}_2(m-r)} \eta(\kappa_{\pi_1}) \cdot \kappa_{\pi_2} \\
&= \sum_{r=2}^m \sum_{\pi_1 \in \text{NC}_2(r-2)} \eta(\kappa_{\pi_1}) \cdot E(s^{m-r}) \\
&= \sum_{r=2}^m \eta \left(\sum_{\pi_1 \in \text{NC}_2(r-2)} \kappa_{\pi_1} \right) E(s^{m-r}) \\
&= \sum_{r=2}^m \eta(E(s^{r-2})) E(s^{m-r}) \\
&= \sum_{r=0}^{m-2} \eta(E(s^r)) E(s^{m-2-r}),
\end{aligned}$$

tal y como se quería demostrar. □

Definición 4.22. Sea $a \in \mathcal{A}$ una variable aleatoria en $(E, \mathcal{A}, \mathcal{B})$. Se define la función generadora de momentos de a mediante la serie de potencias formal

$$M_a(b) = \sum_{n \geq 0} E((ab)^n).$$

Corolario 4.9. La función generadora de momentos de s satisface la ecuación

$$M(z\mathbf{1}) = \mathbf{1} + z^2 \eta(M(z\mathbf{1})) \cdot M(z\mathbf{1})$$

para toda $z \in \mathbb{C}^+$.

Demostración. De la definición de $M(z)$ y el Corolario 4.8 se tiene que

$$M(z) = \mathbf{1} + E(s)z + \sum_{m \geq 2} \sum_{k=0}^{m-2} \eta(E(s^k)) \cdot E(s^{m-2-k}) z^m$$

$$\begin{aligned}
M(z) &= \mathbf{1} + \sum_{n \geq 0} \sum_{k=0}^n \eta(E(s^k)) \cdot E(s^{n-k}) z^n z^2 \\
&= \mathbf{1} + z^2 \sum_{k \geq 0} \sum_{n \geq k} \eta(E(s^k)) \cdot E(s^{n-k}) z^{n-k} z^k \\
&= \mathbf{1} + z^2 \sum_{k \geq 0} \sum_{m \geq 0} (\eta(E(s^k)) z^k) \cdot (E(s^m) z^m) \\
&= \mathbf{1} + z^2 \left(\sum_{k \geq 0} \eta(E(s^k) z^k) \right) \cdot \left(\sum_{m \geq 0} E(s^m) z^m \right) \\
&= \mathbf{1} + z^2 \eta \left(\sum_{k \geq 0} E(s^k) z^k \right) \cdot M(z\mathbf{1}) \\
&= \mathbf{1} + z^2 \eta(M(z\mathbf{1})) \cdot M(z\mathbf{1}),
\end{aligned}$$

donde utilizamos fuertemente el hecho que la función de covarianza es lineal. \square

Similarmente al caso escalar se tiene que $G(z\mathbf{1}) = \frac{1}{z} M \left(\frac{1}{z} \mathbf{1} \right)$. Por lo tanto, el corolario anterior implica trivialmente la afirmación a) del Teorema 4.4 para el caso de la variable semicircular s .

Corolario 4.10. *La transformada de Cauchy valuada en operadores de s satisface la siguiente ecuación*

$$zG(z\mathbf{1}) = \mathbf{1} + \eta(G(z\mathbf{1})) \cdot G(z\mathbf{1})$$

para toda $z \in \mathbb{C}^+$.

El siguiente teorema, el cual puede ser consultado en [11] y referencias allí, nos permite calcular el espectro asintótico de ciertas matrices aleatorias a bloques.

Teorema 4.9. *Sea $d \in \mathbb{N}$ fija y σ una función de covarianza tal que $\sigma(i, j; k, l) = \overline{\sigma(k, l; i, j)}$ para toda $i, j, k, l = 1, \dots, d$. Suponga, para $N \in \mathbb{N}$, que $\{a_{r,p}^{(i,j)} \mid i, j = 1, \dots, d; r, p = 1, \dots, N\}$ son variables aleatorias gaussianas con $a_{r,p}^{(i,j)} = \overline{a_{p,r}^{(j,i)}}$, media cero y covarianza*

$$E \left(a_{r,p}^{(i,j)} a_{q,s}^{(k,l)} \right) = \frac{1}{dN} \delta_{rs} \delta_{pq} \sigma(i, j; k, l).$$

Considere las matrices a bloques $X_N = (A^{(i,j)})_{i,j=1}^d$, donde los bloques están dados por

$$A^{(i,j)} = (a_{r,p}^{(i,j)})_{r,p=1}^N.$$

Entonces casi seguramente, cuando $N \rightarrow \infty$, la matriz X_N tiene una distribución espectral asintótica la cual tiene una transformada de Cauchy determinada por

$$G(z) = \text{tr}_d(\mathcal{G}(z))$$

donde $\mathcal{G}(z)$ es una función analítica en el semiplano complejo superior con valores en $\text{Mat}_d(\mathbb{C})$ la cual está determinada únicamente por

$$\lim_{\substack{|z| \rightarrow \infty \\ \Im z > 0}} z\mathcal{G}(z) = I_d$$

y que satisface, para toda z en el semiplano superior complejo, la ecuación matricial

$$z\mathcal{G}(z) = I_d + \eta(\mathcal{G}(z))\mathcal{G}(z)$$

donde $\eta : \text{Mat}_d(\mathbb{C}) \rightarrow \text{Mat}_d(\mathbb{C})$ es la función de covarianza

$$[\eta(D)_{i,j=1}^d]_{i,j} = \frac{1}{d} \sum_{k,l=1}^d D_{k,l} \sigma(i,k;l,j).$$

Para finalizar, damos un teorema el cual nos permite construir un criterio bastante útil para probar la libertad respecto a E .

Teorema 4.10. *Sea (\mathcal{C}, φ) un espacio de probabilidad, $d \in \mathbb{N}$ fijo, $\mathcal{A} = \text{Mat}_d(\mathcal{C})$, $\mathcal{B} = \text{Mat}_d(\mathbb{C})$ y $E = \mathbf{1} \otimes \varphi$. Denotemos a los cumulantes escalares respecto a φ por κ y a los cumulantes valuados en operadores por $\kappa^{\mathcal{B}}$. Sea J un conjunto arbitrario y $A^{(r)} = (a_{i,j}^{(r)})_{i,j=1}^d \in \mathcal{A}$ para cada $r \in J$. Entonces los cumulantes valuados en operadores*

de $A^{(r)}$ están dados por

$$[\kappa_n^E (A^{(r_1)} \otimes \cdots \otimes A^{(r_n)})]_{i,j} = \sum_{i_2, \dots, i_n=1}^d \kappa_n \left(a_{i, i_2}^{(r_1)} \otimes \cdots \otimes a_{i_n, j}^{(r_n)} \right).$$

Corolario 4.11. *En notación del teorema anterior, si las entradas de dos variables aleatorias en \mathcal{A} son dos familias libres en (\mathcal{C}, φ) entonces las dos matrices son libres respecto a E .*

Demostración del Teorema 4.10. Sean $r_1, \dots, r_n \in \{1, 2\}$ como en el Teorema 4.10. Si $r(\alpha) \neq r(\beta)$ para algunos $\alpha, \beta \in \{1, \dots, n\}$, de la libertad entre las entradas de $A^{(1)}$ y $A^{(2)}$ se tiene que los cumulantes mixtos de dichas entradas se anulan, por tanto

$$\begin{aligned} [\kappa_n^E (A^{(r_1)} \otimes \cdots \otimes A^{(r_n)})]_{i,j} &= \sum_{i_2, \dots, i_n=1}^d \kappa_n \left(a_{i, i_2}^{(r_1)} \otimes \cdots \otimes a_{i_n, j}^{(r_n)} \right) \\ &= \sum_{i_2, \dots, i_n=1}^d 0 \\ &= 0 \end{aligned}$$

para todo $i, j \in \{1, \dots, d\}$. De esto concluimos que $\kappa_n^E (A^{(r_1)} \otimes \cdots \otimes A^{(r_n)}) = \mathbf{0}_{\mathcal{B}}$ y por lo tanto $A^{(1)}$ y $A^{(2)}$ son libres sobre \mathcal{B} . \square

Comentario 4.21. El corolario anterior es de interés en términos de modelación ya que nos garantiza que la libertad asintótica entre los bloques de dos matrices a bloques es suficiente para tener la libertad sobre $\mathcal{B} = \text{Mat}_d(\mathbb{C})$ de las variables aleatorias en $\mathcal{A} = \text{Mat}_d(\mathcal{C})$ correspondientes a dichas matrices a bloques. Sin este hecho sería necesario probar o suponer la libertad respecto a $E : \mathcal{A} \rightarrow \mathcal{B}$ de las matrices en \mathcal{A} , lo cual aparenta ser difícil de justificar intuitivamente¹⁰.

Comentario 4.22. Sea $\mathcal{C} = L^\infty(\Omega, \mathcal{F})$ el conjunto de todas las variables aleatorias $X : \Omega \rightarrow \mathbb{C}$ tales que su soporte es acotado, y denotemos por φ al valor esperado

¹⁰Puesto que estamos hablando de modelación, en general, los supuestos se deben justificar desde el punto de vista de la aplicación.

asociado. Es claro que (\mathcal{C}, φ) es una $*$ -álgebra (con $f^* = \overline{f}$). El corolario anterior nos garantiza que es suficiente con tener libertad entre las entradas de dos matrices en $\mathcal{A} := \text{Mat}_d(\mathcal{C})$ para que estas sean libres sobre $\mathcal{B} := \text{Mat}_d(\mathbb{C})$. Es un ejercicio de rutina demostrar que si dos variables aleatorias son libres y conmutan entonces una debe ser un múltiplo escalar de $\mathbf{1}_{\mathcal{C}}$, y puesto que \mathcal{C} es un álgebra conmutativa la condición del corolario anterior implicaría que una de las matrices debe ser constante. Esto nos muestra una complicación al tratar de llevar la probabilidad libre valuada en operadores al caso de matrices de dimensión finita.

4.6. Notas adicionales

Para una exposición detallada de la teoría general de probabilidad libre valuada en operadores sugerimos el artículo de Helton, Far y Speicher [12] del 2007 y referencias allí. La sección 4.4 se basa fuertemente en las primeras dos secciones del artículo [5] de Benaych-Georges del 2005.

Capítulo 5

Análisis del espectro asintótico de canales multiantena

En este capítulo mostramos algunas aplicaciones de la teoría de probabilidad libre al análisis de los espectros asintóticos de ciertos canales multiantena. En particular, en la Sección 5.1 mostramos una aplicación de los espacios rectangulares a los canales multiacceso la cual se debe a Speicher, Vargas y Mai [24]. En la Sección 5.2 utilizamos los espacios rectangulares y las ecuaciones de subordinación para el análisis de ciertos enlaces en redes multisalto. Este problema fue resuelto por Fawaz, Zarifi, Debbah y Gesbert en 2011 [7] utilizando una gran variedad de herramientas, incluida la probabilidad libre. Nuestro principal aporte en este problema es que todos los cálculos se mantienen dentro del contexto de la probabilidad libre. En la Sección 5.3 mostramos, a manera de preambulo, una aplicación de la probabilidad libre valuada en operadores al cálculo de la capacidad en un sistema de comunicaciones en el que se considera la interferencia intersímbolo. Esta aplicación se debe a Far, Oraby, Bryc y Speicher [11]. Por último, en la Sección 5.4 adaptamos el modelo a bloques de la Sección 5.3 en el contexto de canales monousuario correlacionados. Tras una breve discusión acerca de porque dicha adaptación es poco realista, proponemos un nuevo modelo en el cual cambiamos la distribución (estadística) de los bloques. Este modelo se basa en introducir correlación en

los ángulos de arribo de las señales recibidas. Además, mostramos que los casos límite respecto al ángulo de arribo corresponden a una variación del modelo de correlación de Kronecker y al modelo canónico de los sistemas MIMO. Concluimos el capítulo dando un breve resumen del modelo propuesto, mostrando brevemente una metodología útil en ciertos casos particulares y proponiendo algunas líneas de investigación futuras.

5.1. Canal multiacceso

Un canal multiacceso es un canal de comunicación en el cual varios usuarios transmiten a un usuario común al mismo tiempo. Por lo tanto, en el contexto de dispositivos multiantena un canal multiacceso es aquel en el que hay $K \in \mathbb{N}$ transmisores y un único receptor. Supongamos que para cada $k \in \{1, \dots, K\}$ se tiene que el k -ésimo transmisor cuenta con N_k antenas transmisoras mientras que el receptor recibe con N antenas. Además, supongamos que el canal entre el k -ésimo transmisor y el receptor está dado por $R_k X_k T_k$, donde R_k^2 y T_k^2 son matrices cuadradas de correlación (no aleatorias) y X_k es una matriz rectangular gaussiana, esto es, cada canal transmisor-receptor experimenta correlación de Kronecker¹.

En esta situación es de interés determinar la capacidad del sistema, lo cual se puede obtener a partir del espectro de la matriz

$$G = \sum_{k=1}^K R_k X_k T_k^2 X_k^* R_k.$$

Este problema fue resuelto por Couillet, Debbah y Silverstein en [6]. Su método de resolución utiliza aproximaciones de ciertas transformadas de Stieltjes y de Shannon. A la técnica utilizada por ellos se le denomina equivalentes deterministas. Speicher, Vargas y Mai introdujeron una técnica llamada equivalentes deterministas libres para resolver alternativamente este problema [24]. Esta técnica a grandes rasgos consiste

¹La notación usual es tomar R_k y T_k a las matrices de correlación, por lo que el canal está dado por $R_k^{1/2} X_k T_k^{1/2}$. Aquí hacemos el cambio de notación por claridad.

en reemplazar las matrices en cuestión por operadores (en espacios rectangulares) que tengan las distribuciones de los elementos en cuestión y que además satisfagan ciertas relaciones de libertad. A continuación discutimos brevemente las ideas clave de este enfoque.

Una modificación realizada en [24] al problema original es que se analizó el espectro de la siguiente matriz aleatoria

$$G = \sum_{k=1}^K R_k U_k T_k^2 U_k^* R_k,$$

donde R_k y T_k son como antes pero las matrices U_k son matrices Haar unitarias. Al final comentaremos como es posible resolver el problema original utilizando este enfoque de equivalentes deterministas libres.

En el Ejemplo 4.8 del capítulo anterior encajamos matrices aleatorias rectangulares en espacios de matrices aleatorias de dimensión más grande, de tal forma que el espacio resultante tenía la estructura de un espacio de probabilidad rectangular. Puesto que las matrices X_k son rectangulares, trabajaremos con espacios rectangulares. Una de las aportaciones principales de [24] es el siguiente teorema, el cual extiende el Teorema 3.2 al caso rectangular.

Teorema 5.1. *Sea $K \in \mathbb{N}$ fijo y para cada $N \in \mathbb{N}$ sea $(\mathcal{A}_N, \tau_N := \mathbb{E} \circ \text{tr}_N)$ un espacio de matrices aleatorias de $N \times N$ con estructura de $(P_1^{(N)}, \dots, P_K^{(N)})$ -espacio de probabilidad con elementos simples $U_i^{(N)} \in (\mathcal{A}_N)_{i,i}$ tales que*

- $\lim_{N \rightarrow \infty} \frac{1}{N} \text{tr}(P_i^{(N)}) \rightarrow c_i > 0$.
- $U_1^{(N)}, \dots, U_K^{(N)}$ son independientes con distribución Haar unitaria en el espacio $((\mathcal{A}_N)_{i,i}, (\tau_N)_i)$.

Sea $U_N = U_1^{(N)} + \dots + U_K^{(N)}$ y $\mathcal{D}_1^{(N)}, \mathcal{D}_2^{(N)}$ colecciones de matrices simples deterministas con \mathcal{D} -distribución asintótica donde $\mathcal{D} = \langle p_1, \dots, p_d \rangle \subset \mathcal{A}$. Entonces $\mathcal{D}_1^{(N)}$ y $U_N \mathcal{D}_2^{(N)} U_N^$ son asintóticamente libres (respecto a $E : \mathcal{A} \rightarrow \mathcal{D}$).*

El teorema anterior nos permite aproximar nuestro problema original por el siguien-

te. Sea r_k un operador cuya distribución coincide con la distribución límite de R_k y t_k el respectivo para T_k . Por el teorema anterior podemos tomar a $(r_k)_{k \leq K}$ y t_1, \dots, t_k libres sobre \mathcal{D} . Por lo tanto, necesitamos calcular el espectro del operador

$$g := \sum_{k=1}^K r_k t_k r_k. \quad (5.1)$$

De las relaciones de libertad supuestas es posible demostrar que los elementos $r_k t_k r_k$ con $k \leq K$ son libres sobre $\mathcal{C} = \langle \mathcal{D}, r_1 r_1^*, \dots, r_K r_K^* \rangle$, donde $\mathcal{D} = \langle p_0, \dots, p_K \rangle$ y $r_k \in \mathcal{A}_{0,k}$ para toda $k \leq K$. Por lo tanto, basta obtener la transformada \mathcal{R} con valores en \mathcal{C} de los elementos $r_k t_k r_k$ para obtener la transformada \mathcal{R} con valores en \mathcal{C} de g , esto es

$$\mathcal{R}_g^{\mathcal{C}}(b) = \sum_{k=1}^K \mathcal{R}_{r_k t_k r_k}^{\mathcal{C}}(b). \quad (5.2)$$

A partir de la ecuación anterior es posible llegar a un sistema de ecuaciones para las funciones escalares

$$f_j(b) := \frac{N}{N_j} \tau_0 \left[\left(\sum_{k=1}^K r_k r_k^* \mathcal{R}_{t_k}^{(k)}(f_k(b)) - b \right)^{-1} r_j r_j^* \right],$$

donde τ_0 es la restricción normalizada de τ a $\mathcal{C}_0 := p_0 \mathcal{C} p_0$. Con estas funciones es posible calcular la transformada de Cauchy G_g con valores en \mathcal{C}_0 , la cual está dada por

$$G_g^{\mathcal{C}_0}(b) = \left(\sum_{k=1}^K r_k r_k^* \mathcal{R}_{t_k}^{(k)}(f_k(b)) - b \right)^{-1}.$$

Por último, basta remplazar a r_k de la expresión anterior por R_k , respectivamente para t_k y T_k , para obtener una expresión que dependa de las matrices dadas. El nombre equivalentes deterministas libres proviene del hecho que nuestras matrices deterministas fueron remplazadas por operadores libres.

Alternativamente al desarrollo de Speicher, Vargas y Mai, es importante observar que si se tienen las transformadas de Cauchy para las $r_k t_k r_k$ entonces la ecuación (4.2)

nos proporciona las transformadas R correspondientes², específicamente, se tiene la relación

$$\mathcal{R}(b) = G^{-1}(b) - b^{-1}.$$

De esta forma obtenemos las transformadas R de los elementos $r_k t_k r_k$ para cada $k \leq K$ y por ende la transformada R de g . Por último, tenemos que resolver la ecuación de punto fijo dada por (4.2) para obtener la transformada de Cauchy de g . En esta situación el único problema restante es calcular la transformada de Cauchy de $r_k t_k r_k$ a partir de G_{r_k} y G_{t_k} , sin embargo, este tipo de problema será tratado con detalle en la siguiente sección.

Observemos que es posible resolver el problema con matrices aleatorias gaussianas en lugar de matrices unitarias si extendemos el Teorema 5.1 para el caso de matrices gaussianas, esto es, la extensión del Teorema 3.10 al caso de espacios rectangulares. En tal situación estaremos interesados en el espectro de operadores de la forma $r_k h_k t_k h_k^* r_k$. Como ya se mencionó, en la siguiente sección mostraremos como calcular la transformada de Cauchy del producto de operadores libres en un espacio rectangular.

5.2. Redes retransmisoras con multisalto

Para mejorar la calidad de los enlaces algunas redes implementan relevos cuya función es recibir una señal, precodificarla y retransmitirla. En particular, las redes conocidas como retransmisoras multisalto implementan relevos sucesivos desde el transmisor hasta el receptor.

En el esquema de precodificación lineal, en el cual la señal recibida (vector) es multiplicada por una matriz para después ser retransmitida, la matriz equivalente del canal es una matriz compuesta del producto de otras matrices. Específicamente, si hay N relevos entre transmisor y receptor entonces la matriz equivalente del canal tiene la

²La expresión para la transformada R en términos de la transformada de Cauchy requiere la inversa de esta última, lo cual conlleva ciertos problemas numéricos.

forma

$$G_N = H_N P_{N-1} H_{N-1} \cdots H_1 P_0,$$

donde las P_i son las matrices cuadradas de precodificación de los relevos (la matriz P_0 corresponde a la matriz de precodificación del transmisor) y las H_i son las matrices rectangulares asociadas a los canales entre relevos. Es importante notar que las P_i son constantes mientras que las H_i son aleatorias.

En este contexto, la información mutua asintótica por antena receptora está dada por

$$I_\infty = \lim_{N \rightarrow \infty} \frac{1}{k_0} \log \det (I_{k_N} + \eta G_N G_N^*)$$

donde η es una constante asociada al ruido del sistema, k_0 es el número de antenas transmisoras y k_N el número de antenas receptoras. De la ecuación anterior vemos que la información mutua asintótica depende únicamente de la distribución empírica espectral asintótica de $G_N G_N^*$, por lo que nos concentraremos en calcular esta última. Por simplicidad, calcularemos el espectro de $G_N G_N^*$ cuando G_N tiene dos elementos los cuales son matrices aleatorias gaussianas rectangulares, i.e. cuando $N = 2$ y P_k es la identidad para $k = 0, 1$. Al final de esta sección discutiremos como extender la solución de este caso particular al caso general.

Sean, para cada $N \in \mathbb{N}$, X_N y Y_N matrices aleatorias gaussianas complejas independientes de dimensión $a(n) \times b(n)$ y $b(n) \times a(n)$ respectivamente. Sea $c(n) = a(n) + b(n)$ y supongamos que $c(n) \rightarrow \infty$ cuando $n \rightarrow \infty$. Supongamos además que $a(n)/c(n) \rightarrow \rho_a > 0$ y $b(n)/c(n) \rightarrow \rho_b > 0$ cuando $n \rightarrow \infty$.

Procediendo como en la sección anterior, por el Teorema 3.5 es claro que debemos encajar a $X_N/\sqrt{a(N)}$ en un espacio de probabilidad de tal forma que si x es dicho encajamiento entonces x^*x es una variable aleatoria Marchenko-Pastur de parámetro $\beta = \frac{\rho_b}{\rho_a}$. Similarmente podemos encajar a $Y_N/\sqrt{a(N)}$ a una variable aleatoria y de tal

forma que yy^* sea una variable aleatoria Marchenko-Pastur del mismo parámetro. Más aún, por el Teorema 3.10 las variables aleatorias x^*x y yy^* son libres.

Puesto que estamos en un contexto no cuadrado, es claro que el encajamiento de x y y debe ser realizado en un espacio de probabilidad rectangular de constantes ρ_a y ρ_b . Específicamente, sea (\mathcal{A}, φ) un (p_a, p_b) -espacio de probabilidad con esperanza condicional asociada E , y tal que $\varphi(p_a) = \rho_a$ y $\varphi(p_b) = \rho_b$. Tomemos x como una variable aleatoria tal que $x \in \mathcal{A}_{a,b}$ y satisface que x^*x es una variable Marchenko-Pastur de parámetro ρ_b/ρ_a en el espacio restringido $(\mathcal{A}_{b,b}, \varphi_b)$. Denotaremos esto diciendo que x^*x es una φ_b MP(ρ_b/ρ_a). Similarmente encajaremos a $y \in \mathcal{A}$ de tal forma que $y \in \mathcal{A}_{b,a}$ y yy^* es una φ_b MP(ρ_b/ρ_a).

Puesto que la normalización usual es dividir a la matriz gaussiana por la raíz cuadrada de sus columnas, tenemos que observar entonces que estamos interesados en

$$\begin{aligned} G_N G_N^* &= \frac{X_N Y_N Y_N^* X_N^*}{a(N)b(N)} \\ &= \frac{a(N)}{b(N)} \frac{X_N Y_N Y_N^* X_N^*}{a(N)a(N)} \\ &\xrightarrow{\text{dist}} \frac{\rho_a}{\rho_b} x y y^* x^*. \end{aligned}$$

Definamos pues al operador g dado por

$$g = \frac{\rho_a}{\rho_b} x y y^* x^*.$$

Obsevemos que la transformada de Cauchy valuada en operadores de g está dada por

$$\begin{aligned} G_g(rp_a + sp_b) &= G_{(\rho_a/\rho_b)xyy^*x^*}(rp_a + sp_b) \\ &= \sum_{n \geq 0} E \left((r^{-1}p_a + s^{-1}p_b) \left(\frac{\rho_a}{\rho_b} x y y^* x^* (r^{-1}p_a + s^{-1}p_b) \right)^n \right) \\ &= \frac{\rho_b}{\rho_a} \sum_{n \geq 0} \left(\frac{\rho_a}{\rho_b r} \right)^{n+1} E((xyy^*x^*)^n). \end{aligned}$$

Utilizando la fórmula de balance para E (ecuación 4.8) tenemos que

$$\begin{aligned}
G_g(rp_a + sp_b) &= \frac{\rho_b}{\rho_a} \sum_{n \geq 0} \left(\frac{\rho_a}{\rho_b r} \right)^{n+1} \widehat{E}((xyy^*x^*)^n) p_a \\
&= \frac{\rho_b^2}{\rho_a^2} \sum_{n \geq 0} \left(\frac{\rho_a}{\rho_b r} \right)^{n+1} \widehat{E}((yy^*x^*x)^n) p_a \\
&= \frac{\rho_b^2}{\rho_a^2} p_a \left(\sum_{n \geq 0} \left(\frac{\rho_a}{\rho_b r} \right)^{n+1} \widehat{E}((yy^*x^*x)^n) \right) \\
&= \widehat{G}_{yy^*x^*x} \left(\frac{r\rho_b}{\rho_a} p_b \right) \frac{\rho_b^2}{\rho_a^2} p_a,
\end{aligned}$$

o de manera más compacta

$$G_g(r\mathbf{1}_{\mathcal{A}}) = \widehat{G}_{yy^*x^*x} \left(\frac{r\rho_b}{\rho_a} \mathbf{1}_{\mathcal{A}} \right) \frac{\rho_b^2}{\rho_a^2} p_a. \quad (5.3)$$

Por lo tanto, para obtener G_g es necesario calcular G_{x^*x} y G_{yy^*} . Para ello, observemos que

$$\begin{aligned}
G_{x^*x}(rp_a + sp_b) &= \sum_{n \geq 0} E((r^{-1}p_a + s^{-1}p_b)(x^*x(r^{-1}p_a + s^{-1}p_b))^n) \\
&= \sum_{n \geq 0} \frac{1}{s^{n+1}} \text{diag}(0, \varphi_b((x^*x)^n)) \\
&= G_{\rho_b/\rho_a}(s)p_b,
\end{aligned}$$

donde $G_y(s)$ denota a la transformada de Cauchy escalar de una v.a. Marchenko-Pastur de parámetro y evaluada en $s \in \mathbb{C}^+$. Observemos que $G_{x^*x}(rp_a + sp_b)$ no es invertible, sin embargo, si es invertible en el álgebra $\langle p_b \rangle$. En lo sucesivo trabajaremos con la inversa en el álgebra restringida, ya que siempre es posible trabajar con el elemento $x^*x + \epsilon p_a$ con ϵ arbitrariamente pequeño y mediante ciertos argumentos de continuidad demostrar que en términos de trabajo numérico es suficiente con trabajar con las inversas restringidas³.

³En este caso es particularmente importante que el álgebra \mathcal{D} es el conjunto de matrices diagonales complejas, lo cual implica que la inversa es la suma de las inversas restringidas.

Por la digresión del párrafo anterior, trabajaremos con la transformada h de x^*x dada por

$$\begin{aligned} h_{x^*x}(rp_a + sp_b) &= r^{-1}p_a + s^{-1}p_b - G_{x^*x}(r^{-1}p_a + s^{-1}p_b)^{-1} \\ &= r^{-1}p_a + (s^{-1} - G_{\rho_b/\rho_a}(s^{-1})^{-1})p_b. \end{aligned}$$

Puesto que yy^* también es una φ_b MP(ρ_b/ρ_a) tenemos que

$$\begin{aligned} G_{yy^*}(rp_a + sp_b) &= G_{\rho_b/\rho_a}(s)p_b, \\ h_{yy^*}(rp_a + sp_b) &= r^{-1}p_a + (s^{-1} - G_{\rho_b/\rho_a}(s^{-1})^{-1})p_b. \end{aligned}$$

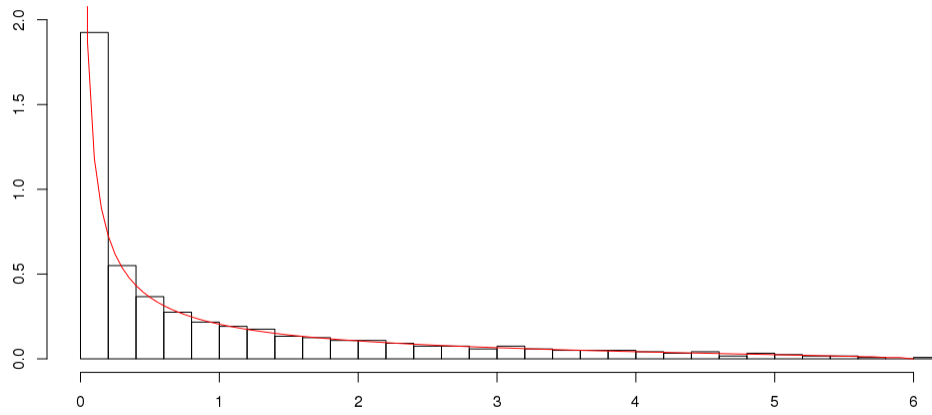
Además, recordemos que se tiene que $\omega_2(d) = \lim_{n \rightarrow \infty} g_d^n(\omega)$ para cualquier $\omega \in \mathbb{H}^+(\mathcal{D})$, donde $g_d(\omega) = dh_{yy^*}(h_{x^*x}(\omega)d)$ y

$$G_{yy^*x^*x}(z\mathbf{1}_{\mathcal{A}}) = (z\mathbf{1}_{\mathcal{A}} - z\omega_2(z^{-1}\mathbf{1}_{\mathcal{A}})h_{x^*x}(\omega_2(z^{-1}\mathbf{1}_{\mathcal{A}})))^{-1}.$$

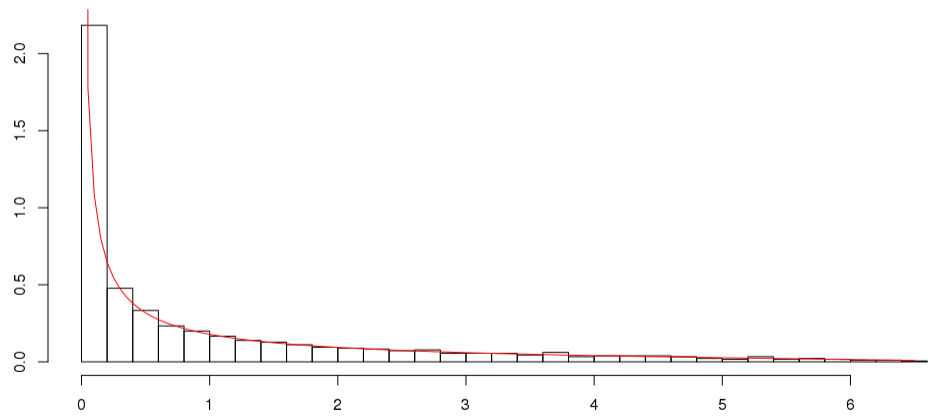
Despejando la última ecuación en la ecuación (5.3) obtenemos la transformada de Cauchy de g , tal y como estabamos buscando.

Para mostrar la concordancia entre las densidades estimadas a partir de las fórmulas recién deducidas y el espectro de las matrices aleatorias correspondientes, a continuación se muestra el histograma de los eigenvalores de $G_N G_N^*$ junto a la densidad asociada a la distribución de g para algunos valores de $a(N)$ y $b(N)$. En negro se muestra el histograma de los eigenvalores y en rojo la densidad estimada. Para mejorar la visualización, en el caso $a(N) > b(N)$ descartamos los eigenvalores nulos de la gráfica.

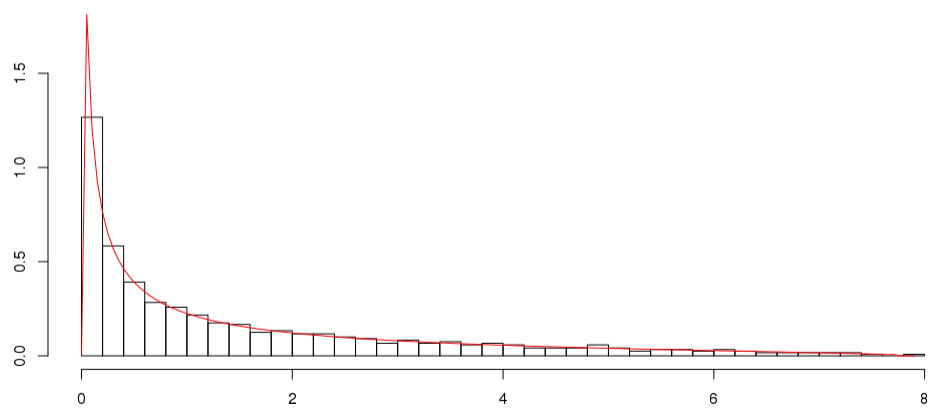
Histograma de los eigenvalores para $a=600$ y $b=900$ contra densidad teórica



Histograma de los eigenvalores para $a=900$ y $b=900$ contra densidad teórica



Histograma de los eigenvalores para $a=900$ y $b=600$ contra densidad teórica



5.3. Sistema monousuario con interferencia entre símbolos

Para hacer una modelación más realista del fenómeno de la transmisión de señales, es usual tomar en cuenta las versiones de una señal que llegan retrasadas en el tiempo. Esencialmente, la salida y_n al tiempo $n \in \mathbb{Z}$ depende de las señales transmitidas hasta ese tiempo. Específicamente, se suele suponer que la salida tienen la forma

$$y_n = \sum_{k \geq 0} a_k x_{n-k},$$

donde a_k es el coeficiente que acompaña a la señal transmitida k periodos antes. Físicamente es sensato suponer que $a_n \rightarrow 0$ cuando $n \rightarrow \infty$, por lo que es usual tomar $a_k = 0$ para $k \geq L$, donde $L \in \mathbb{N}$ es una constante.

En el contexto multiantena las a_k son matrices y las x_k son vectores. En tal situación la matriz equivalente del canal tiene la forma

$$H = \begin{pmatrix} A_1 & A_2 & \cdots & A_L & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & A_1 & A_2 & \cdots & A_L & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & A_1 & A_2 & \cdots & A_L & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \cdots & \mathbf{0} & A_1 & A_2 & \cdots & A_L \end{pmatrix},$$

donde las A_1, \dots, A_L son matrices de $n_R \times n_T$, y en cada fila hay $K - 1$ matrices cero con $K \in \mathbb{N}$ fija. En el caso en el que las matrices A_1, \dots, A_L son matrices gaussianas cuadradas independientes, la intuición nos dice que H se comporta asintóticamente de manera similar a un elemento $h \in \text{Mat}_{K \times L + K - 1}(\mathcal{A})$, para \mathcal{A} un álgebra de operadores apropiada, el cual tiene la misma estructura que H salvo que los elementos A_1, \dots, A_L son reemplazados por una familia circular libre.

Observemos que en esta situación estamos en el contexto de la Sección 4.5, y por lo

tanto podemos aplicar el tipo de cálculos allí realizados. Este enfoque fue tomado por Far, Oraby, Bryc y Speicher en [11] de 2008. Hasta donde sabemos, esta fue una de las primeras aplicaciones de la probabilidad libre valuada en operadores en comunicación inalámbrica. Remitimos al lector interesado en los detalles de esta aplicación a dicho artículo.

5.4. Sistema monousuario con correlación angular

En esta sección adaptaremos el modelo de la sección anterior para modelar la correlación en sistemas multiantena monousuario. Además, propondremos una extensión basada en la dispersión del ángulo de arribo. Esta extensión nos ayudará a describir de una forma más realista el comportamiento asintótico del espectro de un canal multiantena, y por ende el de su capacidad. Finalizaremos este capítulo planteando posibles líneas de investigación futuras acerca de este modelo que, hasta donde sabemos, no ha sido explorado en la literatura.

5.4.1. Modelos de correlación a bloques

Supóngase que un canal de comunicación inalámbrica se comporta como una matriz aleatoria

$$H = (h_{i,j})_{\substack{i=1,\dots,n_R \\ j=1,\dots,n_T}}$$

la cual posee una función de correlación $\sigma(i, k; j, l) := \mathbb{E}(h_{i,k} \overline{h_{j,l}})$ para $i, j \in \{1, \dots, n_R\}$ y $k, l \in \{1, \dots, n_T\}$. Una posibilidad para aproximar al espectro asintótico de este canal es remplazar los coeficientes $h_{i,j}$ por matrices aleatorias gaussianas $A^{(i,j)}$ de tamaño $N \times N$, i.e. construir una nueva matriz (a bloques)

$$H_N = \begin{pmatrix} A^{(1,1)} & \dots & A^{(1,n_T)} \\ \vdots & \ddots & \vdots \\ A^{(n_R,1)} & \dots & A^{(n_R,n_T)} \end{pmatrix},$$

y utilizar el espectro asintótico del ensamble $(H_n)_{n \geq 1}$ como la aproximación buscada. Puesto que estamos aproximando al canal representado por la matriz H , pediremos que las submatrices $\{A^{(i,j)}\}$ se correlacionen de acuerdo a la correlación de H . Específicamente, pediremos que la correlación de la matriz H_N esté dada por

$$\mathbb{E} \left((H_N)_{r,p}^{(i,k)} (H_N)_{s,q}^{(j,l)} \right) = \frac{1}{(n_R + n_T)N} \delta_{r,s} \delta_{p,q} \sigma(i, k; j, l).$$

Es importante señalar que únicamente los coeficientes correspondientes de dos bloques pueden estar correlacionadas. Más aún, la gaussianidad implica que las matrices $A^{(i,j)}$ tienen coeficientes independientes. Como señalaremos en la siguiente subsección, dicha independencia es relativamente artificial y poco realista.

Puesto que las matrices $(A^{(i,j)})_{i \leq n_R, j \leq n_T}$ convergen en distribución a una familia circular, es natural pensar en operar con los bloques de H_N en lugar de hacerlo con la matriz directamente, tal y como se hizo en la sección anterior. En este contexto es posible utilizar la teoría de probabilidad libre valuada en operadores, específicamente, tomando $\mathcal{A} = \text{Mat}_{n_R \times n_T}(\mathcal{C})$, $\mathcal{B} = \text{Mat}_{n_R \times n_T}(\mathbb{C})$ y $E = \mathbf{1} \otimes \varphi$, donde (\mathcal{C}, φ) es un espacio de probabilidad adecuado.

Es importante mencionar que la teoría desarrollada en el capítulo anterior está hecha para elementos semicirculares y en el presente contexto estamos trabajando con elementos circulares. Una matriz de elementos circulares C se puede escribir como una suma $A + iB$ con A y B dos matrices de elementos semicirculares tales que los elementos de una son libres de los de la otra. El Corolario 4.11 implica que las matrices A y B son libres sobre $\text{Mat}_{n_R \times n_T}(\mathbb{C})$, y de la linealidad de la transformada R valuada en operadores concluimos que la extensión al caso circular se sigue del caso semicircular⁴.

⁴Una variable aleatoria circular x es de la forma $x = 2^{-1/2}(s + it)$ con s, t variables aleatorias semicirculares libres. Puesto que $(it)^* = -it$, el caso circular no se puede obtener *per se* del semicircular. Sin embargo, en las siguientes subsecciones utilizamos algunos encajamientos que hacen posible pasar del caso semicircular al circular.

5.4.2. Extensión a elementos no circulares y correlación angular

En notación de la subsección anterior, trabajemos en la modelación de un bloque $A^{(i,j)}$ fijo. El hecho de tomar los coeficientes del bloque $A^{(i,j)}$ independientes es bastante artificial, ya que al aumentar el número de antenas en un mismo dispositivo éstas terminan por correlacionarse. Por lo tanto, tenemos que extender el modelo anterior al caso cuando las matrices $A^{(i,j)}$ no convergen a una familia circular. Por simplicidad en la notación denotemos a $A^{(i,j)}$ únicamente por A .

Si cada antena del sistema original, tanto receptora como transmisora, se reemplaza por N antenas colocadas en una vecindad de esta, entonces tiene lugar el remplazo de $h_{i,j}$ por $A^{(i,j)}$. Por lo tanto, cada bloque modela el comportamiento de alguna región espacial de los dispositivos que intervienen en el proceso de comunicación.

Si n_R y n_T son suficientemente grandes entonces la separación entre las antenas reales debería ser relativamente pequeña. Por lo tanto, los coeficientes que modelan a los canales entre las antenas virtuales⁵ correspondientes a la matriz A deberían de ser prácticamente iguales. En particular, podemos pensar que todos los coeficientes tienen la misma amplitud, la cual la tomaremos por simplicidad como uno.

Con base en lo anterior, podemos tomar a la matriz A de tal forma que para cada $i, j \in \{1, \dots, N\}$ se tenga que

$$A_{i,j} = \exp(\gamma i \theta_{i,j})$$

con $\theta_{i,j}$ una variable aleatoria real y $\gamma > 0$ un parámetro físico del sistema el cual modela la dispersión del ángulo de arribo.

Algunos de los factores que más impactan en la correlación de un arreglo de antenas están relacionados con parámetros físicos del arreglo o bien con dispersores locales [20]. Puesto que estos factores son diferentes para cada extremo del enlace de comunicación, proponemos modelar la correlación entre los ángulos $(\theta_{i,j})_{i,j=1}^N$ mediante el modelo de

⁵Virtuales puesto que las antenas pertenecientes al bloque A son antenas hipótéticas alrededor de un par de antenas reales.

correlación de Kronecker. Observese que este modelo no implica que la correlación de la matriz A sea de Kronecker. Específicamente, proponemos que la matriz $\theta = (\theta_{i,j})_{i,j=1}^N$ sea de la forma

$$\theta = RXT$$

donde R^2 y T^2 son matrices de correlación y X es una matriz gaussiana con entradas centradas y de varianza unitaria. Es importante mencionar que las matrices R^2 y T^2 modelan la dinámica de la correlación en el extremo receptor y transmisor respectivamente. Para mostrar algunas de las propiedades de este modelo, a continuación estudiaremos al espectro asintótico de la matriz A en los casos límites $\gamma \rightarrow \infty$ y $\gamma \rightarrow 0$.

Observemos que para N fija se tiene que

$$\begin{aligned} \mathbb{E} (A_{i,k} \overline{A_{j,l}}) &= \mathbb{E} (\exp(\gamma i(\theta_{i,k} - \theta_{j,l}))) \\ &= \mathbb{E} \left(\exp \left(\gamma i \left(\sum_{a,b=1}^N R_{i,a} X_{a,b} T_{b,k} - \sum_{a,b=1}^N R_{j,a} X_{a,b} T_{b,l} \right) \right) \right) \\ &= \mathbb{E} \left(\exp \left(\sum_{a,b=1}^N \gamma i X_{a,b} (R_{i,a} T_{b,k} - R_{j,a} T_{b,l}) \right) \right) \\ &= \prod_{a,b=1}^N \mathbb{E} (\exp(\gamma i X_{a,b} (R_{i,a} T_{b,k} - R_{j,a} T_{b,l}))) \\ &= \prod_{a,b=1}^N \exp \left(-\frac{1}{2} \gamma^2 (R_{i,a} T_{b,k} - R_{j,a} T_{b,l})^2 \right) \\ &\rightarrow 0 \end{aligned}$$

cuando $\gamma \rightarrow \infty$. Mutatis mutandis, del cálculo anterior se tiene que $\mathbb{E} (A_{i,k}^n \overline{A_{j,l}^m}) \rightarrow 0$ cuando $\gamma \rightarrow \infty$ para toda $n, m \geq 1$. Esto implica en particular que las variables aleatorias $(A_{i,j})_{i,j=1}^N$ se decorrelacionan y su media va a cero cuando $\gamma \rightarrow \infty$. Por lo tanto, para N suficientemente grande y $\gamma = \gamma(N)$ suficientemente grande se tiene que el espectro de A se comportará similarmente al de una variable aleatoria circular. Esto coincide con la intuición desde el punto de vista de modelación. Específicamente, el caso $\gamma \rightarrow \infty$ implica que la dispersión del ángulo de arribo es grande y por lo tanto se

está en presencia de un ambiente ideal en el cual hay una gran cantidad de dispersores locales.

Similarmente, desde el punto de vista de modelación el caso $\gamma \rightarrow 0$ es el peor. En este caso prácticamente todas las antenas correspondientes a la matriz A experimentan el mismo canal y por lo tanto el sistema MIMO no provee ninguna ganancia respecto al sistema SISO⁶. Observemos que $A_{i,j} = 1 + i\gamma\theta_{i,j} + o(\gamma)$, por lo que

$$\begin{aligned} A &= \mathbf{1} + i\gamma\theta + o(\gamma) \\ &= \mathbf{1} + i\gamma RXT + o(\gamma) \end{aligned}$$

donde $\mathbf{1}$ es la matrix de $N \times N$ que contiene únicamente unos. Por lo tanto tiene lugar la aproximación

$$A \stackrel{d}{\approx} \mathbf{1} + i\gamma RXT.$$

Por simplicidad en la notación, denotemos por A a la matriz $\mathbf{1} + i\gamma RXT$.

A continuación haremos una breve digresión acerca del comportamiento del espectro de A , y de porque este se comporta «similar» al espectro de γRXT . Puesto que los eigenvalores de AA^* son las normas de los eigenvalores de A y AA^* es hermitiana, trabajaremos con esta última por simplicidad.

Recordemos que $\lambda_1(AA^*)$ denota al eigenvalor más grande de AA^* , por lo cual

$$\begin{aligned} \lambda_1(AA^*) &\leq \|AA^*\|_{op} \\ &= \|N\mathbf{1} + i\gamma(RXT\mathbf{1} - \mathbf{1}TX^*R) + \gamma^2 RXT^2 X^*R\|_{op} \\ &\leq \|N\mathbf{1}\|_{op} + \|i\gamma(RXT\mathbf{1} - \mathbf{1}TX^*R)\|_{op} + \|\gamma^2 RXT^2 X^*R\|_{op} \\ &\leq N^2 + 2N|\gamma| \|RXT\|_{op} + \gamma^2 \|RXT^2 X^*R\|_{op}. \end{aligned}$$

Bajo algunos supuestos, el eigenvalor dominante de una canal con correlación de Kronecker converge a un número real cuando $N \rightarrow \infty$, y en particular se tiene que

⁶Del inglés *Single Input Single Output*. En este contexto se refiere a la situación en la que se transmite y se recibe con una sola antena.

$\|RXT^2X^*R\|_{op}$ converge cuando $N \rightarrow \infty$ (véase [8] y referencias allí). Por lo tanto, el eigenvalor dominante de la matriz AA^* es de orden $O(N^2)$. De esto es claro que este eigenvalor es despreciable al calcular la eficiencia espectral ergódica asintótica, la cual está dada por el límite

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \log(\lambda_k(AA^*)).$$

Similarmente podemos ignorar los primeros tres eigenvalores⁷ al calcular dicha eficiencia espectral ergódica asintótica.

Obsevemos que las matrices $N\mathbf{1}$, $RXT\mathbf{1}$ y $\mathbf{1}TX^*R$ tienen rango unitario, y por lo tanto

$$\text{Rango}(N\mathbf{1} + i\gamma(RXT\mathbf{1} - \mathbf{1}TX^*R)) \leq 3.$$

De esto se tiene que a lo mas tres eigenvalores de $N\mathbf{1} + i\gamma(RXT\mathbf{1} - \mathbf{1}TX^*R)$ son no nulos, por lo que $\lambda_k(N\mathbf{1} + i\gamma(RXT\mathbf{1} - \mathbf{1}TX^*R)) = 0$ para $k \geq 4$. De la desigualdad de estabilidad para los eigenvalores⁸, para $k \geq 2$ se tiene que

$$\begin{aligned} |\lambda_k(AA^*)| &= |\lambda_k(AA^*) - \lambda_k(N\mathbf{1} + i\gamma(RXT\mathbf{1} - \mathbf{1}TX^*R))| \\ &\leq \|\gamma^2 RXT^2X^*R\|_{op} \\ &= \gamma^2 \|RXT^2X^*R\|_{op}. \end{aligned}$$

Puesto que γ es pequeño concluimos que la mayoría de los eigenvalores de AA^* son similares a los de $\gamma^2 RXT^2X^*R$. Por lo tanto, en lo sucesivo utilizaremos la siguiente aproximación⁹ $A = \gamma RXT$. Lo anterior implica que en el peor caso, i.e. cuando $\gamma \rightarrow 0$, la correlación de la matriz A es de Kronecker. Adicionalmente observemos que bajo esta aproximación, la serie de Mercator implica que la eficiencia espectral ergódica asintótica

⁷Es claro que podemos ignorar cualquier cantidad finita de eigenvalores una vez que el primero es despreciable, sin embargo con tres será suficiente.

⁸Véase el libro de Terence Tao [25].

⁹Puesto que X es una matriz gaussiana compleja, se tiene que X y iX tienen la misma distribución. Por lo tanto γRXT y $i\gamma RXT$ son idénticamente distribuidas.

decae linealmente con γ a medida que $\gamma \rightarrow 0$.

Puesto que la distribución empírica espectral de $U_0AU_0^*$ es la misma que la de A para toda matriz unitaria U_0 , la distribución empírica espectral de UAU^* es la misma que la de A con U una matriz aleatoria Haar unitaria independiente de X . De la invarianza biunitaria de X concluimos que la distribución empírica espectral URU^*XUTU^* es la misma que la de A para U como antes.

Basados en la discusión anterior proponemos el siguiente modelo para aproximar al espectro asintótico de la matriz H . Reemplazaremos a H por una variable aleatoria $h \in \text{Mat}_{n_R \times n_T}(\mathcal{A})$ donde (\mathcal{A}, φ) es un espacio de probabilidad adecuado. Además, cada operador $h_{i,j}$ será de la forma $h_{i,j} = r_{i,j}x_{i,j}t_{i,j}$, donde $r_{i,j}$ y $t_{i,j}$ son los operadores asociados a la correlación de cada bloque y $x_{i,j}$ una variable aleatoria circular. En sistemas donde se presente una alta homogeneidad es sensato que las $r_{i,j}$ tengan la misma distribución, y similarmente para las $t_{i,j}$. De la digresión del párrafo anterior y el Teorema¹⁰ 3.9 podemos suponer que las familias

$$(\{r_{i,j}, t_{i,j}\})_{\substack{i=1,\dots,n_R \\ j=1,\dots,n_T}}, \{h_{i,j} \mid i = 1, \dots, n_R; j = 1, \dots, n_T\}$$

son libres. Supondremos además que las variables aleatorias $(x_{i,j})_{i,j}$ es una familia circular cuya correlación está dada por la de H , i.e. $\varphi(x_{i,k}x_{j,l}^*) = \mathbb{E}(h_{i,k}\overline{h_{j,l}})$. Obsérvese que las $h_{i,j}$ del lado derecho de la última ecuación no son las mismas que los componentes del operador h , sin embargo, incurrimos en este abuso de notación por simplicidad. También es importante observar que los parámetros de dispersión angular correspondientes a los bloques del operador h están implícitos en la correlación de la familia $(x_{i,j})_{i,j}$.

5.4.3. Un caso de estudio: familia circular libre

Supongamos que los coeficientes $h_{i,j}$ de la matriz H no están correlacionados y que se conoce la dinámica de correlación al aumentar el número de antenas. En este

¹⁰Por supuesto nos referimos a la versión circular del Teorema de Voiculescu.

caso es sensato modelar a la familia de variables aleatorias no conmutativas $\{h_{i,j} \mid i \in \{1, \dots, n_r\}, j \in \{1, \dots, n_T\}\}$ mediante una familia circular libre. En esta situación la probabilidad libre valuada en operadores nos permiten calcular la distribución de h como sigue.

Consideremos el caso particular $n_R = n_T = 2$. Sea (\mathcal{A}, φ) un espacio de probabilidad. Supongamos que $h \in \text{Mat}_2(\mathcal{A})$ está dada por

$$h = \begin{pmatrix} r_1 h_1 t_1 & r_2 h_2 t_2 \\ r_3 h_3 t_3 & r_4 h_4 t_4 \end{pmatrix}$$

con $r_k, t_k, h_k \in \mathcal{A}$ para $k = 1, \dots, 4$. En esta situación tenemos que obtener la transformada de Cauchy escalar de h^*h la cual está dada por

$$G_{h^*h}(z) = (\text{tr}_2 \otimes \varphi) ((zI_2 - h^*h)^{-1}).$$

Sea $E := (I_2 \otimes \varphi) : \text{Mat}_2(\mathcal{A}) \rightarrow \text{Mat}_2(\mathbb{C})$, entonces podemos escribir la expresión anterior como

$$G_{h^*h}(z) = \text{tr}_2(G_{h^*h}(zI_2))$$

donde $G_{h^*h}(zI_2) = E((zI_2 - h^*h)^{-1})$. Para calcular $G_{h^*h}(zI_2)$ definamos el operador auxiliar $\widehat{h} \in \text{Mat}_4(\mathcal{A})$ el cual está dado por

$$\widehat{h} = \begin{pmatrix} \mathbf{0} & h^* \\ h & \mathbf{0} \end{pmatrix}.$$

Sea $\widehat{E} := (I_2 \otimes E) = (I_4 \otimes \varphi)$, puesto que $\{h_{i,j}\}$ es una familia circular tenemos entonces que $\widehat{E}(\widehat{h}^{2n+1}) = \mathbf{0}$ para toda $n \in \mathbb{N}$. Puesto que

$$\widehat{h}^{2n} = \begin{pmatrix} (h^*h)^n & \mathbf{0} \\ \mathbf{0} & (hh^*)^n \end{pmatrix}$$

para toda $n \in \mathbb{N}$, de las observaciones anteriores se tiene que

$$\begin{aligned}
G_{\widehat{h}}(zI_4) &= \sum_{n \geq 0} \frac{1}{z^{n+1}} \widehat{E}(\widehat{h}^n) \\
&= \sum_{n \geq 0} \frac{1}{z^{2n+1}} \widehat{E}(\widehat{h}^{2n}) \\
&= z \sum_{n \geq 0} \frac{1}{(z^2)^{n+1}} \widehat{E}((\widehat{h}^2)^n) \\
&= zG_{\widehat{h}^2}(z^2I_4).
\end{aligned}$$

La fórmula anterior coincide con una fórmula análoga bastante conocida del caso escalar.

Además observemos que

$$\begin{aligned}
G_{\widehat{h}^2}(zI_4) &= \sum_{n \geq 0} \frac{1}{z^{n+1}} \widehat{E}(\widehat{h}^{2n}) \\
&= \sum_{n \geq 0} \frac{1}{z^{n+1}} \begin{pmatrix} E((h^*h)^n) & \mathbf{0} \\ \mathbf{0} & E((hh^*)^n) \end{pmatrix} \\
&= \begin{pmatrix} G_{h^*h}(zI_2) & \mathbf{0} \\ \mathbf{0} & G_{hh^*}(zI_2) \end{pmatrix}.
\end{aligned}$$

Por lo tanto, $G_{h^*h}(z)$ se puede calcular vía $G_{h^*h}(zI_2)$, y para obtener esta última basta calcular la función $G_{\widehat{h}^2}(zI_4)$ la cual se puede obtener a su vez de $G_{\widehat{h}}(zI_4)$. Para obtener esta última haremos uso de la teoría de subordinación aditiva libre. Sean

$$J := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad E_1 := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad E_2 := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Escribamos $\widehat{h} = \widehat{h}_1 + \widehat{h}_2 + \widehat{h}_3 + \widehat{h}_4$ donde

$$\begin{aligned}
\widehat{h}_1 &= (J \otimes E_1)\widehat{h}(J \otimes E_1), & \widehat{h}_2 &= (J \otimes E_1)\widehat{h}(J \otimes E_2), \\
\widehat{h}_3 &= (J \otimes E_2)\widehat{h}(J \otimes E_1), & \widehat{h}_4 &= (J \otimes E_2)\widehat{h}(J \otimes E_2).
\end{aligned}$$

Para ser más específicos, se tiene que

$$\widehat{h}_1 = \begin{pmatrix} 0 & 0 & r_1 h_1^* t_1 & 0 \\ 0 & 0 & 0 & 0 \\ r_1 h_1 t_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Procediendo de manera similar al cálculo de $G_{\widehat{h}}$, y el hecho que (\mathcal{A}, φ) es tracial, se obtiene que

$$\begin{aligned} G_{\widehat{h}_1}(zI_4) &= zG_{t_1^2 h_1^* r_1^2 h_1}(z^2) \text{diag}(1, 0, 1, 0), \\ G_{\widehat{h}_2}(zI_4) &= zG_{t_2^2 h_2^* r_2^2 h_2}(z^2) \text{diag}(0, 1, 1, 0), \\ G_{\widehat{h}_3}(zI_4) &= zG_{t_3^2 h_3^* r_3^2 h_3}(z^2) \text{diag}(1, 0, 0, 1), \\ G_{\widehat{h}_4}(zI_4) &= zG_{t_4^2 h_4^* r_4^2 h_4}(z^2) \text{diag}(0, 1, 0, 1). \end{aligned}$$

Observemos que \widehat{h}_1 es libre de \widehat{h}_2 , que $\widehat{h}_1 + \widehat{h}_2$ es libre de \widehat{h}_3 y $\widehat{h}_1 + \widehat{h}_2 + \widehat{h}_3$ es libre de \widehat{h}_4 . Por lo tanto, podemos obtener la transformada de Cauchy de \widehat{h} mediante la obtención sucesiva de las transformadas de Cauchy de \widehat{h}_1 , $\widehat{h}_1 + \widehat{h}_2$, $\widehat{h}_1 + \widehat{h}_2 + \widehat{h}_3$ y \widehat{h} vía las ecuaciones de subordinación aditiva dadas en el capítulo anterior.

Observemos que las transformadas de Cauchy $G_{t_k^2 h_k r_k^2 h_k^2}$ para $k = 1, \dots, 4$ se pueden obtener vía las ecuaciones de subordinación multiplicativa del capítulo anterior (para el caso escalar basta tomar $\mathbf{1}_{\mathcal{A}} = 1$). Además, si suponemos que $(r_k, t_k)_{k=1, \dots, 4}$ son idénticamente distribuidos entonces las transformadas de Cauchy $G_{t_k^2 h_k r_k^2 h_k^2}$ para $k = 1, \dots, 4$ son iguales.

5.4.4. Conclusiones y trabajo futuro

El problema principal que queda por resolver es encontrar la distribución de la variable aleatoria h dadas las distribuciones de las variables aleatorias de correlación y la correlación de la familia circular (no libre) $(x_{i,j})_{i,j}$. La libertad de la familia circular im-

plicó que los encajamientos \widehat{h}_k tienen transformadas de Cauchy valuadas en operadores con valores en las matrices diagonales, lo cual deja de ser cierto en el caso general.

En el artículo de Far, Oraby, Bryc y Speicher [11] se dió un método para calcular la transformada de Cauchy valuada en operadores de una matriz compuesta por una familia semicircular con correlación arbitraria¹¹. Esto nos permite trabajar con una función arbitraria de correlación $\sigma(i, k; j, l) = \mathbb{E}(h_{i,k}\overline{h_{j,l}})$.

Sin embargo, es necesario encontrar modelos para las variables aleatorias de correlación $r_{i,j}$ y $t_{i,j}$ tales que

- Tengan transformadas de Cauchy o transformadas R^{12} valuadas en operadores con valores en¹³ $\text{Mat}_n(\mathbb{C})$ las cuales sean explícitas o por lo menos calculables con una computadora.
- Que sean lo suficientemente versátiles, de tal manera que permitan aproximar una gran gama de distribuciones de correlación.

Este tipo de modelos han sido explorados con anterioridad en el contexto del cálculo de la capacidad ergódica asintótica de sistemas MIMO con correlación de Kronecker. Específicamente, en [18] Mestre, Fonollosa y Pagès-Zamora modelan las matrices de correlación mediante matrices Toeplitz, las cuales resultan ser una familia versátil y manejable. Sin embargo, dicho trabajo se llevó por la vía de la probabilidad libre escalar, por lo que es necesario llevar este tipo de modelos al nivel valuado en operadores.

5.5. Notas adicionales

Al final de la Sección 5.1 se mencionó que es necesario resolver la ecuación de punto fijo dada por (4.2). En [12] Helton, Far y Speicher resuelven dicho problema cuando la transformada R es lineal y preserva cierta positividad.

El modelo propuesto en la Sección 5.4 supone que todos los coeficientes de la matriz

¹¹Utilizando el tipo de encajamientos mostrados en esta sección, el método se puede extender para una familia circular.

¹²Recordemos que la transformada de Cauchy puede ser obtenida a partir de la transformada R vía la ecuación (4.2).

¹³No solamente matrices diagonales.

A tienen la misma amplitud y por lo tanto la modelación del canal se hace respecto a los ángulos de arribo. Esta idea es clásica en el contexto de la modelación de canales inalámbricos, véase por ejemplo el artículo de Shiu, Foschini, Gans y Kahn (2000) [23] y referencias allí.

Apéndice A

Teoría general de la inversión de Möbius

En el presente apéndice se dan brevemente algunos elementos de la teoría general de la inversión de Möbius. En particular, se da la fórmula de inversión de Möbius y se discute someramente como esta implica la existencia de los cumulantes libres.

Definición A.1. Sea P un conjunto parcialmente ordenado finito, se define $P^{(2)}$ como

$$P^{(2)} := \{(\pi, \sigma) \mid \pi, \sigma \in P; \pi \leq \sigma\},$$

donde \leq es la relación de orden parcial definida en P . Sean $F, G : P^{(2)} \rightarrow \mathbb{C}$, denotamos la convolución de ellos por

$$(F * G)(\pi, \sigma) = \sum_{\substack{\rho \in P \\ \pi \leq \rho \leq \sigma}} F(\pi, \rho)G(\rho, \sigma).$$

Además, si $f : P \rightarrow \mathbb{C}$ entonces denotamos la convolución de f y G por

$$(f * G)(\sigma) = \sum_{\substack{\rho \in P \\ \rho \leq \sigma}} f(\rho)G(\rho, \sigma).$$

Definición A.2. Sea P un conjunto parcialmente ordenado finito. La función zeta $\zeta : P^{(2)} \rightarrow \mathbb{C}$ de P se define por $\zeta(\pi, \sigma) = 1$ para toda $(\pi, \sigma) \in P^{(2)}$. La inversa de ζ bajo convolución¹ es llamada la función de Möbius de P , denotada por μ .

Comentario A.1. La última línea de la definición anterior necesita justificación, ya que no es claro que la función ζ tenga una inversa bajo convolución. Una forma de ver que tal inversa existe es: (1) asociar matrices diagonales superiores a las funciones que van de $P^{(2)}$ a \mathbb{C} , y (2) observar que la matriz asociada a ζ tiene únicamente unos en la diagonal, y por tanto invertible².

Proposición A.1 (Fórmula de Inversión de Möbius). *Sea P un conjunto parcialmente ordenado finito y μ su función de Möbius. Sean $f, g : P \rightarrow \mathbb{C}$, entonces*

$$f(\pi) = \sum_{\substack{\sigma \in P \\ \sigma \leq \pi}} g(\sigma) \quad (\text{A.1})$$

es equivalente a

$$g(\pi) = \sum_{\substack{\sigma \in P \\ \sigma \leq \pi}} f(\sigma)\mu(\sigma, \pi). \quad (\text{A.2})$$

Demostración. Observemos que la ecuación (A.1) se puede escribir como $f = g * \zeta$. Esto es equivalente a $f * \mu = g$, es decir, equivalente a la ecuación (A.2). \square

Comentario A.2. La proposición anterior aún no nos garantiza la existencia de los cumulantes, ya que estos últimos son una familia de funcionales multiplicativos $(\kappa_\pi \mid n \in \mathbb{N}; \pi \in \text{NC}(n))$. Sin embargo, si nos permite construir una familia $(\kappa_\pi)_{\pi \in \text{NC}(n), n \in \mathbb{N}}$ tal que

$$\kappa_\pi(a_1, \dots, a_n) = \sum_{\substack{\sigma \in \text{NC}(n) \\ \sigma \leq \pi}} \varphi_\sigma(a_1, \dots, a_n)\mu(\sigma, \pi),$$

donde $(\varphi_\pi \mid \pi \in \text{NC}(n); n \in \mathbb{N})$ es una familia tal que

$$\varphi_n(a_1, \dots, a_n) = \varphi(a_1 \cdots a_n)$$

¹ f y g son inversas bajo convolución si $(f * g)(\pi, \sigma) = \delta_\pi(\sigma)$.

²Recordemos que la inversa de una matriz triangular superior es triangular superior también.

para toda $n \in \mathbb{N}$ y es multiplicativa respecto a π , esto es

$$\varphi_\pi(a_1, \dots, a_n) = \prod_{\substack{V \in \pi \\ V = \{i_1, \dots, i_{|V|}\}}} \varphi_{|V|}(a_{i_1}, \dots, a_{i_{|V|}}).$$

Por lo tanto, solo basta ver la multiplicatividad de $\kappa = \varphi * \mu$. La demostración de este hecho se basa en demostrar la «multiplicatividad» de μ respecto a sus argumentos y sustituir en la Fórmula de Momentos-Cumulantes dada al principio de este comentario.

Comentario A.3. De manera similar a la construcción de la familia multiplicativa $(\kappa_n)_{n \geq 1}$ a partir de $(\varphi_n)_{n \geq 1}$, en el contexto no conmutativo de los espacios valuados en operadores es posible construir una familia multiplicativa $(\kappa_n)_{n \geq 1}$ de morfismos entre \mathcal{B} -bimódulos a partir $(E_n)_{n \geq 1}$, donde esta última se obtiene de extender multiplicativamente a E .

Terminamos este apéndice con el siguiente ejemplo, el cual tiene como objetivo mostrar con más detalles lo expuesto en el comentario anterior.

Ejemplo A.1. De la fórmula de inversión de Möbius se obtiene la existencia de una familia $(\kappa_\pi)_{\pi \in \text{NC}(n), n \in \mathbb{N}}$ de morfismos entre \mathcal{B} -bimódulos tales que

$$\kappa_\pi(a_1 \otimes \dots \otimes a_n) = \sum_{\sigma \leq \pi} E_\sigma(a_1 \otimes \dots \otimes a_n) \mu(\sigma, \pi).$$

En particular, de la fórmula anterior se obtiene que

$$\begin{aligned} \kappa_1(a) &= E(a), \\ \kappa_2(a \otimes b) &= E(ab) - E(a)E(b). \end{aligned}$$

Demostremos pues que para $\pi = \{\{1, 4\}, \{2, 3\}\}$ se tiene que

$$\kappa_\pi(a \otimes b \otimes c \otimes d) = \kappa_2(a\kappa_2(b \otimes c) \otimes d).$$

De la fórmula entre momentos-cumulantes se tiene que

$$\kappa_\pi(a \otimes b \otimes c \otimes d) = \sum_{\sigma \leq \pi} E_\sigma(a \otimes b \otimes c \otimes d) \mu(\sigma, \pi).$$

Es posible demostrar que $\sigma \leq \pi$ es equivalente a que $\sigma_1 \leq \pi_1$ y $\sigma_2 \leq \pi_2$, donde $\pi_1 = \{1, 4\}$, $\pi_2 = \{2, 3\}$ y $\sigma = \sigma_1 \cup \sigma_2$. Por lo tanto, la ecuación anterior toma la forma

$$\kappa_\pi(a \otimes b \otimes c \otimes d) = \sum_{\sigma_1 \leq \pi_1} \sum_{\sigma_2 \leq \pi_2} E_{\sigma_1 \cup \sigma_2}(a \otimes b \otimes c \otimes d) \mu(\sigma_1 \cup \sigma_2, \pi_1 \cup \pi_2).$$

Es posible demostrar que μ es multiplicativa respecto a sus argumentos, específicamente, se tiene que $\mu(\sigma_1 \cup \sigma_2, \pi_1 \cup \pi_2) = \mu(\sigma_1, \pi_1) \mu(\sigma_2, \pi_2)$. De la multiplicatividad de la familia (E_π) se tiene entonces que

$$\begin{aligned} \kappa_\pi(a \otimes b \otimes c \otimes d) &= \sum_{\sigma_1 \leq \pi_1} \sum_{\sigma_2 \leq \pi_2} E_{\sigma_1}(a E_{\sigma_2}(b \otimes c) \otimes d) \mu(\sigma_1, \pi_1) \mu(\sigma_2, \pi_2) \\ &= \sum_{\sigma_1 \leq \pi_1} E_{\sigma_1} \left(a \left(\sum_{\sigma_2 \leq \pi_2} E_{\sigma_2}(b \otimes c) \mu(\sigma_2, \pi_2) \right) \otimes d \right) \mu(\sigma_1, \pi_1). \end{aligned}$$

Aplicando nuevamente la fórmula entre momentos y cumulantes se tiene que

$$\begin{aligned} \kappa_\pi(a \otimes b \otimes c \otimes d) &= \sum_{\sigma_1 \leq \pi_1} E_{\sigma_1}(a \kappa_{\pi_2}(b \otimes c) \otimes d) \mu(\sigma_1, \pi_1) \\ &= \kappa_{\pi_1}(a \kappa_{\pi_2}(b \otimes c) \otimes d) \\ &= \kappa_2(a \kappa_2(b \otimes c) \otimes d), \end{aligned}$$

tal y como se quería mostrar. Observemos cómo la metodología de este ejemplo puede ser extendida fácilmente a cualquier π , lo que probaría lo afirmado en el comentario anterior.

En el ejemplo anterior se incurrió en algunos abusos de notación con el fin de mantener la notación lo más limpia posible. Por ejemplo, cuando $\sigma_1 = \{\{1\}, \{4\}\}$ la notación κ_{σ_1} realmente denota a $\kappa_{\widehat{\sigma_1}}$ donde $\widehat{\sigma_1} = \{\{1\}, \{2\}\}$.

Remitimos a los lectores interesados en los detalles relacionados con esta teoría al libro de Nica y Speicher [19].

Bibliografía

- [1] O. Arizmendi, O. Barndorff-Nielsen, U. Franz, V. Perez-Abreu, S. Thorbjensen y C. Vargas: *Independence and Infinite Divisibility*, en preparación.
- [2] Z. Bai y J. Silverstein: *Spectral Analysis of Large Dimensional Random Matrices*, Springer, 2010.
- [3] S. Belinschi, T. Mai y R. Speicher: Analytic subordination theory of operator-valued free additive convolution and the solution of a general random matrix problem, sometido a revisión (2013).
- [4] S. Belinschi, R. Speicher, J. Treilhard y C. Vargas: Operator-valued free multiplicative convolution: analytic subordination theory and applications to random matrix theory, sometido a revisión (2013).
- [5] F. Benaych-Georges: Rectangular random matrices, related free entropy and free Fisher's information, *Journal of Operator Theory* **62** (2009), 371-419.
- [6] R. Couillet, M. Debbah y J. Silverstein: A deterministic equivalent for the analysis of correlated MIMO multiple access channels, *IEEE Transactions on Information Theory* **57** (2011), 3493-3514.

- [7] M. Debbah, N. Fawaz, D. Gesbert y K. Zarifi: Asymptotic capacity and optimal precoding in MIMO multi-hop relay networks, *IEEE Transactions on Information Theory* **7** (2011), 2050-2069.
- [8] M. Diaz: *Análisis de la Eficiencia Espectral Ergódica Asintótica de Sistemas MIMO con Correlación de Kronecker*, Tesis de Licenciatura, 2011.
- [9] K. Dykema: On the S-Transform over a Banach algebra, *Journal of Functional Analysis* **231** (2006), 90-110.
- [10] A. Edelman: The probability that a random real Gaussian matrix has k real eigenvalues, related distributions and the circular law, *J. Multivariate Anal.* **60** (1997), 701-712.
- [11] R. Far, T. Oraby, W. Bryc y R. Speicher: On slow-fading MIMO systems with nonseparable correlation, *IEEE Trans. on Information Theory* **54** (2008), 544-553.
- [12] J. Helton, R. Far y R. Speicher: Operator-valued semicircular elements: solving a quadratic matrix equation with positivity constraints, *International Mathematics Research Notices* **2007** (2007), ID:rnm086.
- [13] F. Hiai y D. Petz: *The Semicircle Law, Free Random Variables and Entropy*, American Mathematical Society, 2000.
- [14] V. Kargin: *Limit Theorems in Free Probability Theory*, ProQuest, 2008.
- [15] T. Koshy: *Catalan Numbers with Applications*, Oxford University Press, 2008.
- [16] S. Lang: *Real and Functional Analysis*, Springer, 1993.
- [17] A. Lozano, A. Tulino y S. Verdú: Impact of antenna correlation on the capacity of multiantenna channels, *IEEE Transactions on Information Theory* **51** (2005), 2491-2509.

- [18] X. Mestre, J. Fonollosa y A. Pagès-Zamora: Capacity of MIMO channels: asymptotic evaluation under correlated fading, *IEEE Journal on Selected Areas in Communications* **21** (2003), 829-838.
- [19] A. Nica y R. Speicher: *Lectures on the Combinatorics of Free Probability*, Cambridge University Press, 2006.
- [20] M. Özdemir, E. Arvas y H. Arslan: Dynamics of spatial correlation and implications on MIMO systems, *IEEE Radio Communications Junio* (2004), S14-S19.
- [21] V. Paulsen: *Completely bounded maps and operator algebras*, Cambridge University Press, 2002.
- [22] W. Rudin: *Functional Analysis*, McGraw-Hill, 1973.
- [23] D. Shiu, G. Foschini, M. Gans y J. Kahn: Fading correlation and its effect on the capacity of multielement antenna systems, *IEEE Transaction on Communications* **48** (2000), 502-513.
- [24] R. Speicher, C. Vargas y T. Mai: Free deterministic equivalents, rectangular random matrix models and operator-valued free probability theory, *Random Matrices: Theory and Applications* **1** (2012), 1-26.
- [25] T. Tao: *Topics in Random Matrix Theory*, American Mathematical Society, 2012.
- [26] E. Telatar: Capacity of multi-antenna Gaussian channels, *European Transactions on Telecommunications* **10** (1999), 585-595.

