



**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

FACULTAD DE CIENCIAS

CUMULANTES EN PROBABILIDAD LIBRE FINITA

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

MATEMÁTICO

P R E S E N T A:

DANIEL PERALES ANAYA



**DIRECTOR DE TESIS:
DR. OCTAVIO ARIZMENDI ECHEGARAY
2016**

CIUDAD UNIVERSITARIA, CDMX



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

Índice general

1. Probabilidad No Conmutativa	7
1.1. *-espacios de probabilidad	7
1.2. Distribución	9
1.3. Independencia	13
1.4. Convoluciones aditivas	15
1.4.1. Convolución aditiva clásica	16
1.4.2. La transformada de Cauchy	17
1.4.3. Otras transformadas	18
1.5. Teorema del Límite Central	19
2. La Función De Möbius En Retículas De Particiones	23
2.1. Conjuntos parcialmente ordenados y retículas	23
2.2. Álgebras de incidencia	26
2.3. Inversión de Möbius	28
2.4. Particiones	31
2.5. La función de Möbius en retículas de particiones	34
2.5.1. La función de Möbius en $\mathcal{I}(n)$	37
2.5.2. La función de Möbius en $\mathcal{P}(n)$	38
2.5.3. La función de Möbius en $\mathcal{NC}(n)$	40
2.6. Familias multiplicativas de funciones	44
3. Cumulantes	49
3.1. Cumulantes generalizados	49
3.2. Definición vía particiones	51
3.3. Relación entre ambas definiciones de cumulantes	53
3.3.1. Cumulantes de productos como argumentos	55
3.3.2. Reformulación de independencia en términos de cu- mulantes	58
3.4. Cumulantes monótonos	65
3.4.1. Particiones monótonas	69
3.5. Relación de cumulantes con transformadas	71
3.5.1. Fórmulas que relacionan recursiones	71
3.5.2. Relación con las transformadas	76

3.6. Demostración del Teorema del Límite Central	80
4. Probabilidad Finita	83
4.1. Preliminares	83
4.1.1. Transformada de Laplace	83
4.1.2. Transformada de Legendre	85
4.1.3. Normas L^p	86
4.1.4. Determinante de Fuglede-Kadison	87
4.2. Convolución de polinomios	88
4.2.1. Convolución aditiva simétrica	88
4.2.2. Transformada U	90
4.3. Transformada \mathcal{R} m -finita	93
4.3.1. Definición y propiedades básicas	94
4.3.2. Relación con probabilidad libre	95
4.3.3. Relación con la convolución de polinomios	96
4.4. Cumulantes finitos	100
4.4.1. Definición y notación	100
4.4.2. Los cumulantes finitos son lineales	102
4.4.3. Fórmulas entre momentos y cumulantes finitos	105
4.4.4. Análisis de las fórmulas obtenidas	108

Introducción

Los polinomios están íntimamente relacionados con la teoría de matrices aleatorias. Por ejemplo, la distribución espectral de matrices aleatorias coincide con la distribución asintótica de las raíces de varios polinomios ortogonales, y muchos de los resultados que involucran distribuciones espectrales se han demostrado utilizando esta relación. Por otro lado, la conexión entre probabilidad libre [NS06] y matrices aleatorias es bien conocida en el régimen asintótico. Una pregunta natural es si se tiene también un análogo de la independencia libre en el nivel finito.

Recientemente Marcus, Spielman y Srivastava [MSS15a] proporcionaron otro enfoque para una teoría de la probabilidad libre finita, basado en la noción de convoluciones polinómicas. La conexión con probabilidad libre es que a medida que $d \rightarrow \infty$ esta convolución de polinomios aproxima a la convolución libre habitual. La motivación original para definir estas convoluciones polinómicas no proviene de probabilidad libre, sino a partir del estudio de familias de polinomios entrelazados en el contexto de la prueba de la existencia de gráficas de Ramanujan [MSS15b], [MSS15d]. Estas investigaciones llevaron también a la solución de uno de los grandes problemas abiertos en matemática, el problema Kadison-Singer [MSS15c].

La probabilidad finita trata de encontrar un vínculo directo entre la probabilidad libre y la teoría de polinomios, no es la primera vez que se trata de relacionar estas dos áreas, sin embargo es la primera en la que se trabaja directamente con los polinomios y no con sus propiedades asintóticas. La principal ventaja de la probabilidad finita, es que proporciona un puente para relacionar dos áreas aparentemente lejanas, y mediante esta relación se ven beneficiados ambas partes, la probabilidad libre ofrece una visión más amplia sobre polinomios y puede ayudar a generar una mejor intuición sobre cómo se comportan, mientras que las convoluciones de polinomios proporcionan una herramienta para trabajar conceptos desde un punto de vista más amigable y concreto como son los polinomios, para después poder extenderlo a temas de probabilidad libre.

El objetivo principal de esta tesis es estudiar el concepto de probabilidad finita libre presentado por Adam Marcus [Mar15] y entender su relación con la probabilidad libre. Hasta el momento, la relación que hay entre estas dos áreas es sólo a nivel analítico. La principal aportación de este trabajo será

abordar esta relación desde un enfoque combinatorio. Para ello, será necesario un estudio exhaustivo de las herramientas combinatorias previamente utilizadas en la probabilidad no conmutativa.

En el Capítulo 1 describiremos los objetos principales del marco teórico que se estudiará en esta tesis. En las primeras secciones definiremos lo que es un espacio de probabilidad y una distribución. Posteriormente, hablaremos de las cuatro principales nociones de independencia y definiremos una convolución aditiva para cada una de ellas. Finalmente, introduciremos las principales herramientas analíticas que se han desarrollado para entender el comportamiento de las medidas de probabilidad bajo estas convoluciones y presentaremos el Teorema del Límite Central para cada noción de independencia. El atractivo principal de este capítulo respecto a trabajos previos, es que abordaremos de la manera más general posible las cuatro nociones principales de independencia para tener una mejor idea de las similitudes y diferencias entre ellas. Esto nos dará una idea más clara de lo que deberíamos esperar de la probabilidad libre finita.

En el Capítulo 2 nos familiarizaremos con todos los conceptos y las herramientas combinatorias básicas que serán necesarios en el siguiente capítulo. Lo más importante será introducir las retículas de particiones y calcular la función de Möbius para estos casos específicos.

En el Capítulo 3 hablaremos de los cumulantes y su relación con los momentos, una herramienta combinatoria que ha resultado bastante útil para trabajar con los distintos tipos de independencia. De nuevo, una diferencia de este trabajo respecto a otros, será que abordaremos los distintos cumulantes desde una perspectiva común, lo cual nos permitirá ver las similitudes y diferencias entre las definiciones y las herramientas que se tienen para demostrar sus principales características. Además, esto nos ayudará a tener una idea más clara de cómo debemos trabajar con los cumulantes en probabilidad libre finita. Por último, utilizaremos los cumulantes para demostrar de manera sencilla dos teoremas importantes en probabilidad y que presentaremos desde un enfoque analítico en el Capítulo 1.

En el Capítulo 4 daremos una breve introducción al concepto de probabilidad libre finita presentado por Adam Marcus [Mar15] y presentaremos los principales resultados de este artículo en el ámbito de la convolución aditiva libre y su relación con la convolución de polinomios. Posteriormente daremos la aportación principal de este trabajo, que será utilizar lo visto en los tres primeros capítulos para definir una noción de cumulantes en probabilidad finita. Veremos que estos cumulantes linealizan la convolución aditiva simétrica y demostraremos algunos de los teoremas del artículo de Marcus [Mar15] desde un enfoque combinatorio. Además, encontraremos una fórmula que relaciona los momentos con los cumulantes finitos la cual permite entender mejor la relación entre cumulantes finitos y cumulantes libres.

Capítulo 1

Probabilidad No Conmutativa

En este capítulo describiremos los objetos principales del marco teórico que se estudiará en esta tesis. En las primeras secciones definiremos lo que es un espacio de probabilidad y una distribución, estos conceptos son fundamentales para el desarrollo de este trabajo. Posteriormente, hablaremos de las cuatro principales nociones de independencia y definiremos una convolución aditiva para cada una de ellas. Finalmente, introduciremos las principales herramientas analíticas que se han desarrollado para entender el comportamiento de las medidas de probabilidad bajo estas convoluciones y presentaremos el Teorema del Límite Central. En general, lo que presentaremos en este capítulo sólo es una recopilación de lo básico que se ha hecho con cada una de las nociones de independencia. Sin embargo, el atractivo principal de este trabajo respecto a otros será abordar de la manera más general posible las cuatro nociones principales de independencia para tener una mejor idea de las similitudes y diferencias entre ellas. Además, nos dará una idea más clara de lo que deberíamos esperar de la probabilidad finita.

1.1. *-espacios de probabilidad

En esta sección introduciremos los espacios dónde estaremos trabajando y los conceptos básicos. Para las definiciones de esta sección nos basaremos en la monografía de Nica y Speicher [NS06].

Definición 1.1.1. Un *espacio de probabilidad no conmutativo*, (\mathcal{A}, ϕ) , consiste de un álgebra unitaria \mathcal{A} sobre \mathbb{C} y un funcional lineal unitario

$$\phi : \mathcal{A} \rightarrow \mathbb{C}, \quad \phi(1_{\mathcal{A}}) = 1.$$

A los elementos de \mathcal{A} se les conoce como *variables aleatorias no conmutativas* en (\mathcal{A}, ϕ) . Por simplicidad, sólo les llamaremos *variables aleatorias* $a \in \mathcal{A}$.

Diremos que el espacio de probabilidad no conmutativo (\mathcal{A}, ϕ) es *tracial* si cumple que

$$\phi(ab) = \phi(ba), \quad \forall a, b \in \mathcal{A}.$$

La definición anterior sólo pide que \mathcal{A} sea un álgebra unitaria, sin embargo en muchas ocasiones será útil trabajar con una **-álgebra*. Es decir, un álgebra \mathcal{A} equipada con una operación antilineal $*$: $\mathcal{A} \rightarrow \mathcal{A}$ con la propiedad de que $(a^*)^* = a$ y $(ab)^* = b^*a^*$ para todo $a, b \in \mathcal{A}$. A la operación $*$ se le conoce como *involución*. Esta involución nos permite hablar de elementos positivos en un álgebra. Más específicamente, decimos que un elemento $b \in \mathcal{A}$ es *positivo*, si se cumple que $b = a^*a$, para alguna variable $a \in \mathcal{A}$. A veces es deseable que ϕ mande elementos positivos a \mathbb{R}^+ , esto para poder hablar de distribuciones asociadas a elementos del álgebra.

Definición 1.1.2. Sea \mathcal{A} una *-álgebra y (\mathcal{A}, ϕ) un espacio de probabilidad no conmutativo. Si se cumple que

$$\phi(a^*a) \geq 0, \quad \forall a \in \mathcal{A},$$

entonces decimos que el funcional ϕ es *positivo* y que (\mathcal{A}, ϕ) es un **-espacio de probabilidad*.

En el contexto de *-álgebras, debemos resaltar algunas variables aleatorias no conmutativas que son importantes debido a las propiedades que satisfacen.

Definición 1.1.3. Sea \mathcal{A} una *-álgebra.

- Diremos que $a \in \mathcal{A}$ es *autoadjunto* si cumple que $a = a^*$.
- Diremos que $u \in \mathcal{A}$ es *unitario* si cumple que $u^*u = uu^* = 1$.
- Diremos que $a \in \mathcal{A}$ es *normal* si cumple que $a^*a = a^*a$.

Estaremos más interesados en los *-espacios de probabilidad ya que tienen mayor estructura y en ellos existen muchos ejemplos interesantes. A continuación mencionaremos algunas propiedades elementales de estos espacios.

Proposición 1.1.4. Dado (\mathcal{A}, ϕ) un *-espacio de probabilidad. Sabemos que el funcional lineal ϕ es autoadjunto, es decir, cumple que

$$\phi(a^*) = \overline{\phi(a)}, \quad \forall a \in \mathcal{A}.$$

Otra consecuencia de la positividad de ϕ es que tenemos una desigualdad del tipo Cauchy-Schwarz para *-álgebras.

Teorema 1.1.5. Sea (\mathcal{A}, ϕ) un $*$ -espacio de probabilidad. Entonces se cumple la desigualdad

$$|\phi(b^*a)|^2 \leq \phi(a^*a)\phi(b^*b), \quad \forall a, b \in \mathcal{A}.$$

Esta desigualdad se puede probar de la manera estándar en que se prueban las desigualdades del tipo Cauchy-Schwarz.

Observación 1.1.6. Si un elemento $a \in \mathcal{A}$ cumple que $\phi(a^*a) = 0$, entonces la desigualdad de Cauchy-Schwarz implica que $\phi(ba) = 0$ para todo $b \in \mathcal{A}$, esto quiere decir que a , de cierta forma, es un elemento degenerado para el funcional ϕ . En muchos casos nos gustaría evitar que existan estos elementos degenerados y eso motiva la siguiente definición.

Definición 1.1.7. Sea (\mathcal{A}, ϕ) un $*$ -espacio de probabilidad. Diremos que ϕ es *fiel* si para $a \in \mathcal{A}$ se tiene que:

$$\phi(a^*a) = 0 \quad \Rightarrow \quad a = 0.$$

Observación 1.1.8. En esta tesis sólo trabajaremos con $*$ -espacios de probabilidad, así que para facilitar la lectura omitiremos el símbolo $*$ en lo que resta de este trabajo. Es decir, a menos que se indique lo contrario, al hablar de un **espacio de probabilidad**, asumiremos que tiene una involución $*$ y que el funcional ϕ es positivo.

1.2. Distribución

Hay una fuerte relación entre los momentos de una variable aleatoria y su distribución. Sin embargo, al trabajar con espacios de probabilidad en general, no existe una estructura analítica amigable que se pueda relacionar con la distribución y los momentos. En consecuencia, tendremos que definir la distribución de una manera puramente algebraica.

Notación 1.2.1. Sea a un elemento en un espacio de probabilidad (\mathcal{A}, ϕ) . La *subálgebra de \mathcal{A} generada por a* es

$$\mathcal{A}_0 = \text{span}\{a^{\epsilon(1)} \cdots a^{\epsilon(k)} \mid k \geq 0, \epsilon(1), \dots, \epsilon(k) \in \{1, *\}\}.$$

Es decir, es la expansión lineal de las “palabras” que podemos hacer usando las “letras” a y a^* . Los valores de ϕ para esas palabras son conocidos como momentos.

Definición 1.2.2. Sea a una variable aleatoria en un espacio de probabilidad (\mathcal{A}, ϕ) . Una expresión de la forma

$$\phi(a^{\epsilon(1)} \cdots a^{\epsilon(k)}) \quad \text{con } k \geq 0 \text{ y } \epsilon(1), \dots, \epsilon(k) \in \{1, *\}$$

es un *momento de a* .

Notación 1.2.3. Denotaremos como $\mathbb{C}\langle X, X^* \rangle$ al álgebra unitaria generada libremente por dos indeterminadas no conmutativas X y X^* . De manera concreta, los monomios de la forma

$$X^{\epsilon(1)} \dots X^{\epsilon(k)}, \quad \text{donde } k \geq 0 \text{ y } \epsilon(1), \dots, \epsilon(k) \in \{1, *\},$$

son una base lineal de $\mathbb{C}\langle X, X^* \rangle$, y la multiplicación de dos monomios se hace por concatenación. Además, $\mathbb{C}\langle X, X^* \rangle$ tiene una involución natural, $*$, determinada al pedir que la involución aplicada a X nos de X^* .

Definición 1.2.4 (Distribución algebraica). Sea a una variable aleatoria en un espacio de probabilidad (\mathcal{A}, ϕ) . La *distribución de a* se define como el funcional lineal $\mu : \mathbb{C}\langle x, x^* \rangle \rightarrow \mathbb{C}$ determinado por:

$$\mu(x^{\epsilon(1)} \dots x^{\epsilon(k)}) = \phi(a^{\epsilon(1)} \dots a^{\epsilon(k)}),$$

para todo $k \geq 0$ y todo $\epsilon(1), \dots, \epsilon(k) \in \{1, *\}$.

La ventaja de definir la distribución de esta manera es que sin importar el espacio de probabilidad, la distribución siempre está definida en el mismo espacio $\mathbb{C}\langle X, X^* \rangle$. Esto hace posible comparar las distribuciones de elementos que viven en espacios distintos.

A pesar de que en general no tenemos una estructura analítica, observemos que en el caso particular de los elementos normales $a \in \mathcal{A}$ (donde tenemos la igualdad $aa^* = a^*a$), la $*$ -álgebra unitaria generada por a se reduce a

$$\mathcal{A}_0 = \text{span}\{a^k (a^*)^l \mid k, l \geq 0\}$$

En este caso es posible definir una distribución analítica que se parece más a la noción clásica de distribución.

Definición 1.2.5 (Distribución analítica). Sea (\mathcal{A}, ϕ) un espacio de probabilidad y sea a un elemento normal de \mathcal{A} . Si existe una medida de probabilidad, μ , con soporte compacto en \mathbb{C} tal que:

$$\int z^k \bar{z}^l d\mu(z) = \phi(a^k (a^*)^l) \quad \forall k, l \in \mathbb{N}. \quad (1.1)$$

Entonces μ es única y diremos que es la *distribución analítica de a* .

Observación 1.2.6. 1. La unicidad es una consecuencia del teorema de Stone-Weierstrass.

2. No necesariamente cualquier elemento normal en un espacio de probabilidad tiene una distribución analítica. Afortunadamente, lo anterior sí es válido para varios ejemplos importantes.

3. En el caso aún más particular en el que estemos trabajando con un elemento autoadjunto, $a \in \mathcal{A}$, se puede ver que el soporte de la distribución de a está contenido en los reales, y podemos transformar la ecuación (1.1) en:

$$\int t^n = \phi(a^n), \quad \forall n \in \mathbb{N}.$$

Además de tener distribuciones para cada elemento, también es importante poder ver cómo se relacionan los elementos entre sí, por ello también tenemos el concepto de distribución conjunta. Lo necesario para esta definición es una generalización bastante directa de lo que hicimos para una variable, así que no daremos tantos detalles.

Definición 1.2.7. Sean a_1, \dots, a_s variables aleatorias en un espacio de probabilidad (\mathcal{A}, ϕ) . Una expresión de la forma

$$\phi(a_{r_1}^{\epsilon(1)} \dots a_{r_k}^{\epsilon(k)}), \quad \text{con } k \geq 1, 1 \leq r_1, \dots, r_k \leq s \text{ y } \epsilon(1), \dots, \epsilon(k) \in \{1, *\},$$

es un *momento mixto* de a_1, \dots, a_s .

Notación 1.2.8. Denotaremos por $\mathbb{C}\langle X_1, X_1^*, \dots, X_s, X_s^* \rangle$ al álgebra unitaria generada libremente por las $2s$ indeterminadas no conmutativas, $X_1, X_1^*, \dots, X_s, X_s^*$. De nuevo, los monomios de la forma

$$X_{r_1}^{\epsilon(1)} \dots X_{r_k}^{\epsilon(k)}, \quad \text{donde } k \geq 0, 1 \leq r_1, \dots, r_k \leq s \text{ y } \epsilon(1), \dots, \epsilon(k) \in \{1, *\},$$

son una base lineal de $\mathbb{C}\langle X_1, X_1^*, \dots, X_s, X_s^* \rangle$, y la multiplicación se hace por concatenación. Además, $\mathbb{C}\langle X_1, X_1^*, \dots, X_s, X_s^* \rangle$ tiene una involución natural, $*$, determinada al pedir que la involución aplicada a X_r nos de X_r^* , para $1 \leq r \leq s$.

Definición 1.2.9 (Distribución conjunta). La distribución conjunta de una familia de elementos $a_1, \dots, a_n \in \mathcal{A}$ en un espacio de probabilidad no conmutativo (\mathcal{A}, ϕ) , es el funcional lineal $\mu_{a_1, \dots, a_n} : \mathbb{C}\langle X_1, X_1^*, \dots, X_n, X_n^* \rangle \rightarrow \mathbb{C}$ tal que $\mu_{a_1, \dots, a_n}(1) = 1$ y

$$\mu_{a_1, \dots, a_n}(X_{i_1}^{\epsilon(1)} \dots X_{i_k}^{\epsilon(k)}) = \phi(a_{i_1}^{\epsilon(1)} \dots a_{i_k}^{\epsilon(k)}),$$

para $k \geq 1, 1 \leq i_1, \dots, i_k \leq n$ y $\epsilon(1), \dots, \epsilon(k) \in \{1, *\}$.

Para que nos quede un poco más claro cómo funciona la distribución analítica daremos un ejemplo con un tipo especial de elemento llamado Haar unitario.

Definición 1.2.10. Sea (\mathcal{A}, ϕ) un espacio de probabilidad. Un elemento $u \in \mathcal{A}$ es *Haar unitario* si $uu^* = u^*u = 1$ (es unitario) y cumple que

$$\phi(u^k) = 0 \quad \forall k \in \mathbb{Z} \setminus \{0\}. \quad (1.2)$$

Observación 1.2.11. El nombre “Haar unitario” viene del hecho de que si u es un Haar unitario en un espacio de probabilidad, entonces la medida de Lebesgue normalizada (también llamada medida de Haar) en un círculo sirve como distribución de u . En efecto, para todo $k, l \in \mathbb{N} \cup \{0\}$ tenemos que

$$\phi(u^k (u^*)^l) = \phi(u^{k-l}) = \begin{cases} 0 & \text{si } k \neq l \\ 1 & \text{si } k = l \end{cases}$$

es igual a la integral

$$\int_{\mathbb{T}} z^k \bar{z}^l dz = \int_0^{2\pi} \frac{e^{i(k-l)t}}{2\pi} dt$$

donde $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$ y dz es la medida de Haar normalizada en \mathbb{T} .

Ejemplo 1.2.12. Sea (\mathcal{A}, ϕ) un espacio de probabilidad, y sea $u \in \mathcal{A}$ un Haar unitario. Entonces el elemento autoadjunto $u + u^* \in \mathcal{A}$ si tiene distribución.

Demostración. Si calculamos los momentos de $u + u^*$ obtenemos que:

$$\phi((u + u^*)^k) = \int_{\mathbb{T}} (z + \bar{z})^k dz = \frac{1}{\pi} \int_0^\pi (2 \cos t)^k dt.$$

Y si hacemos el cambio de variable $r = 2 \cos t$, concluimos que

$$\phi((u + u^*)^k) = \int_{-2}^2 r^k \frac{dr}{\pi \sqrt{4 - r^2}} = \int_{\mathbb{R}} t^k \mu(t) dt \quad k \geq 0,$$

donde la distribución $\mu(t)$ es la *densidad del arcoseno* en $[-2, 2]$:

$$\mu(t) = \begin{cases} \frac{1}{\pi \sqrt{4 - r^2}} & \text{si } |t| < 2 \\ 0 & \text{si } |t| \geq 2. \end{cases}$$

Además, a partir de la ecuación (1.2) y el hecho de que $u^* = u^{-1}$ podemos derivar la siguiente fórmula para los momentos:

$$\phi((u + u^*)^k) = \sum_{j=0}^k \binom{k}{j} \phi(u^{2j-k}) = \begin{cases} 0 & \text{si } k \text{ es impar} \\ \binom{k}{k/2} & \text{si } k \text{ es par.} \end{cases}$$

□

La distribución que acabamos de obtener es conocida como la *ley del arcoseno* y será importante para resultados futuros. Otras distribuciones específicas que también serán importantes después son las siguientes:

Definición 1.2.13. Sea (\mathcal{A}, ϕ) un espacio de probabilidad y $x \in \mathcal{A}$ un elemento autoadjunto con varianza σ^2 .

1. Diremos que x es una variable *Gaussiana* si tiene distribución normal, es decir, si su distribución es igual a $\frac{e^{-t^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}}$.
2. Diremos que x es una variable *semicircular* centrada de radio 2σ si su distribución es igual a $\frac{2}{\pi(2\sigma)^2}\sqrt{(2\sigma)^2 - t^2}$ en el intervalo $[-2\sigma, 2\sigma]$.
3. Diremos que x es una variable *Bernoulli* simétrica si su distribución es la medida de probabilidad $\frac{1}{2}(\delta_{-\sigma} + \delta_{\sigma})$.

Observación 1.2.14. 1. Los momentos de una variable Gaussiana, z , son:

$$\phi(z^n) = \begin{cases} \sigma^n(n-1)(n-3)\cdots 3 \cdot 1 & \text{si } n \text{ es par} \\ 0 & \text{si } n \text{ es impar.} \end{cases}$$

2. Los momentos de una variable semicircular, s , son:

$$\phi(s^n) = \begin{cases} \sigma^n C_{n/2} & \text{si } n \text{ es par} \\ 0 & \text{si } n \text{ es impar,} \end{cases}$$

dónde $C_k = \frac{1}{k+1} \binom{2k}{k}$ es el k -ésimo número de catalán.

3. Los momentos de una variable Bernoulli simétrica, b , son:

$$\phi(b^n) = \begin{cases} \sigma^n & \text{si } n \text{ es par} \\ 0 & \text{si } n \text{ es impar.} \end{cases}$$

4. Los momentos de una variable, x , cuya distribución es la ley del arco-seno son:

$$\phi(x^n) = \begin{cases} \left(\frac{\sigma}{2}\right)^n \binom{n}{n/2} & \text{si } n \text{ es par} \\ 0 & \text{si } n \text{ es impar.} \end{cases}$$

1.3. Independencia

Después de definir los conceptos de distribución de un elemento y la distribución conjunta de varios elementos, estamos listos para introducir uno de los conceptos fundamentales cuando trabajamos con espacios de probabilidad, el concepto de independencia. Observemos que si conocemos la distribución conjunta de a_1, \dots, a_n (dónde a_1, \dots, a_n son variables aleatorias de un espacio de probabilidad (\mathcal{A}, ϕ)) entonces podemos obtener la distribución de cada una de las a_i , con $1 \leq i \leq n$. Basta con fijarnos exclusivamente en los momentos que contienen a_i o a_i^* . Sin embargo, si quisiéramos recobrar la distribución conjunta de a_1, \dots, a_n a partir de la distribución de cada elemento, en general no será posible, ya que los momentos que contienen distintas variables aleatorias no necesariamente están relacionados con las distribuciones de las variables aleatorias involucradas.

Si analizamos la noción de independencia en la probabilidad clásica, en el fondo, se puede entender como una estructura universal que justamente nos proporciona una regla para calcular los momentos mixtos a partir de los momentos de cada variable. Aunque trabajar sólo con variables independientes nos limita un poco, simplifica mucho el manejo de distribuciones conjuntas. Además, las variables independientes encajan bastante bien con muchos procesos naturales en los que se aplica la probabilidad clásica. Por supuesto, la independencia clásica no es la única que se puede aplicar a procesos importantes, y es por eso que a lo largo de la historia se han definido y estudiado otros tipos de independencia, es decir, otras reglas para obtener estos momentos mixtos. A continuación abordaremos las nociones más importantes de independencia de un manera general, con el mismo enfoque que se da en [HS11b], esto con el objetivo de tener un terreno común para compararlas y tener una visión más clara de las similitudes y diferencias entre las nociones.

Definición 1.3.1 (Independencia). Sea (\mathcal{A}, ϕ) un espacio de probabilidad y $\{\mathcal{A}_\lambda : \lambda \in \Lambda\}$ una familia de *-subálgebras de \mathcal{A} . Sea $\phi(a_1 a_2 \cdots a_n)$ un momento mixto tal que $a_i \in \mathcal{A}_{\lambda_i}$, $a_i \notin C1$ y tal que elementos consecutivos son de distintas subálgebras ($\lambda_1 \neq \lambda_2 \neq \cdots \neq \lambda_{k-1} \neq \lambda_k$). Entonces decimos que:

1. (Independencia clásica o tensorial). $\{\mathcal{A}_\lambda\}$ es *independiente en el sentido clásico* (o tensorial) si se cumple que

$$\phi(a_1 a_2 \cdots a_n) = \phi(a_1) \phi(a_2 \cdots a_n)$$

cuando $\lambda_1 \neq \lambda_r$ para todo $2 \leq r \leq n$. O si tenemos que

$$\phi(a_1 a_2 \cdots a_n) = \phi(a_2 \cdots a_{r-1} (a_1 a_r) a_{r+1} \cdots a_n),$$

cuando r es el menor número tal que $\lambda_1 = \lambda_r$.

2. (Independencia libre [Voi85]). $\{\mathcal{A}_\lambda\}$ es *libre* si se cumple que

$$\phi(a_1 a_2 \cdots a_n) = 0$$

cuando $\phi(a_1) = \phi(a_2) = \dots = \phi(a_n) = 0$.

3. (Independencia booleana [SW97]). $\{\mathcal{A}_\lambda\}$ es *independiente en el sentido booleano* si se cumple que

$$\phi(a_1 a_2 \cdots a_n) = \phi(a_1) \phi(a_2 \cdots a_n).$$

4. (Independencia monótona [Mur00]). Supongamos que el índice Λ tiene un orden lineal $<$. $\{\mathcal{A}_\lambda\}$ es *monótonamente independiente* si se cumple que

$$\phi(a_1 \cdots a_i \cdots a_n) = \phi(a_i) \phi(a_1 \cdots a_{i-1} a_{i+1} \cdots a_n)$$

cuando i cumple que $\lambda_{i-1} < \lambda_i > \lambda_{i+1}$ (en los casos $i = 1$ o $i = n$ la desigualdad correspondiente no se considera).

Observación 1.3.2. 1. Si aplicamos sucesivamente la regla para la independencia clásica, es sencillo llegar a una regla equivalente para la independencia clásica. $\{\mathcal{A}_\lambda\}$ es independiente en el sentido clásico si se cumple que

$$\phi(a_1 a_2 \cdots a_n) = \prod_{\lambda \in \Lambda} \phi \left(\overrightarrow{\prod_{i: a_i \in \mathcal{A}_\lambda} a_i} \right)$$

donde $\overrightarrow{\prod_{i \in V} a_i}$ es el producto de los a_i , con $i \in V$, en el mismo orden en el que aparecen en $a_1 a_2 \cdots a_n$.

2. También existe una quinta noción de independencia conocida como independencia antimonótona. La regla para calcular los momentos, es básicamente la misma que para la monótona, pero se debe cumplir $\lambda_{i-1} > \lambda_i < \lambda_{i+1}$ en lugar de $\lambda_{i-1} < \lambda_i > \lambda_{i+1}$. Esta noción es completamente análoga a la de independencia monótona y por lo tanto no la abordaremos en este trabajo.

Podemos extender cualquiera de las cuatro nociones de independencia a subconjuntos arbitrarios de \mathcal{A} , como nos muestra la siguiente definición.

- Definición 1.3.3.** 1. Diremos que $\{X_\lambda \subseteq \mathcal{A} : \lambda \in \Lambda\}$ son independientes (en el sentido clásico, libre, booleano o monótono) si $\{\mathcal{A}_\lambda : \lambda \in \Lambda\}$ son independientes (en el sentido clásico, libre, booleano o monótono), donde $\mathcal{A}_\lambda = \text{alg}(1, X_\lambda)$ es la $*$ -álgebra unitaria generada por X_λ .
2. Diremos que una familia de elementos $\{a_\lambda \in \mathcal{A} : \lambda \in \Lambda\}$ es independiente si la familia de subconjuntos $\{\{a_\lambda\} \subseteq \mathcal{A} : \lambda \in \Lambda\}$ es independiente.

1.4. Convoluciones aditivas

Ya que tenemos una noción de independencia, cualquiera de las cuatro que queramos elegir, la primera duda que nos podría surgir es cómo se comporta la distribución de la suma de dos variables aleatorias independientes. Para poder hablar de ello de una manera más concisa primero definiremos la suma de dos medidas de probabilidad.

Definición 1.4.1. Sean μ y ν medidas de probabilidad en \mathbb{R} con soporte compacto. Sean x y y variables aleatorias autoadjuntas en algún espacio de probabilidad C^* tales que x tiene distribución μ y y tiene distribución ν , y tales que x y y son independientes (en el sentido clásico, libre, booleano o monótono). Entonces la *convolución aditiva* clásica, libre, booleana o monótona de μ y ν será la distribución de la suma $x + y$ y la denotaremos como $\mu \star \nu$, $\mu \boxplus \nu$, $\mu \boxdot \nu$ y $\mu \triangleright \nu$, respectivamente.

- Observación 1.4.2.** 1. Es importante mencionar que dadas μ y ν , y para cualquier tipo de independencia, siempre podemos encontrar variables x y y como se piden en la definición. Además, la suma de la distribución sólo depende de la distribución μ de x y de la distribución ν de y , es decir, no importa cuáles realizaciones x y y nos tomemos. Por lo tanto, la convolución aditiva está bien definida para cualquier independencia.
2. Como $x + y$ es autoadjunto y acotado, entonces su distribución también es una medida de probabilidad con soporte compacto en \mathbb{R} . Por lo tanto, las convoluciones serán operaciones binarias sobre el espacio de medidas de probabilidad con soporte compacto en \mathbb{R} . Estas convoluciones se pueden extender a medidas con soporte no compacto. Como esta tesis estudia principalmente los aspectos combinatorios, nos restringiremos a medidas con soporte compacto.

En cada caso se han desarrollado herramientas analíticas para poder encontrar de manera sistemática la convolución aditiva de dos medidas de probabilidad. La idea es codificar la información de los momentos de μ y de ν en una serie formal de potencias, luego transformar estas series en otras que se comporten de manera lineal con respecto a la convolución, de esta forma obtener la suma es un paso directo. Por último, debemos utilizar la inversa de la transformación anterior para regresar a la serie de potencias que codifica los momentos de la suma y recobrar la medida a partir de esta serie.

1.4.1. Convolución aditiva clásica

El caso de la convolución clásica es el más conocido y ha sido ampliamente estudiado, para su análisis debemos codificar la información de los momentos utilizando la transformada de Fourier.

Definición 1.4.3. Sea μ una medida de probabilidad en \mathbb{R} . Definimos la *transformada de Fourier* de μ como

$$\mathcal{F}_\mu(z) = \int_{\mathbb{R}} e^{-izt} d\mu(t).$$

Observación 1.4.4. En el caso en que μ tiene soporte compacto podemos tomar la expansión en serie de potencias de la función exponencial

$$e^{-izt} = \sum_{n=0}^{\infty} \frac{(-izt)^n}{n!}$$

y luego integrar término a término, con ello obtenemos que

$$\mathcal{F}_\mu(z) = \sum_{n=0}^{\infty} \frac{(-iz)^n}{n!} m_n,$$

donde $m_n := \int_{\mathbb{R}} t^n d\mu(t)$ es el n -ésimo momento de μ , para $n \geq 0$.

El siguiente teorema explica la relación entre la transformada de Fourier y la convolución aditiva clásica.

Teorema 1.4.5. Si μ y ν son medidas de probabilidad en \mathbb{R} con soporte compacto e independientes en el sentido clásico, tenemos que

$$\log \mathcal{F}_{\mu \star \nu}(z) = \log \mathcal{F}_{\mu}(z) + \log \mathcal{F}_{\nu}(z).$$

Demostración. Tomemos x y y variables aleatorias autoadjuntas e independientes en el sentido clásico, en algún espacio de probabilidad C^* tales que x tiene distribución μ y y tiene distribución ν . De la observación anterior y la definición de independencia clásica podemos obtener que:

$$\begin{aligned} \mathcal{F}_{\mu \star \nu}(z) &= \sum_{n=0}^{\infty} \frac{(-iz)^n}{n!} \phi((x+y)^n) \\ &= \sum_{n=0}^{\infty} \frac{(-iz)^n}{n!} \sum_{j=0}^n \binom{n}{j} \phi(x^j) \phi(y^{n-j}) \\ &= \sum_{n=0}^{\infty} \sum_{i=0}^n \frac{(-iz)^j}{j!} \phi(x^j) \frac{(-iz)^{n-j}}{(n-j)!} \phi(y^{n-j}) \\ &= \left(\sum_{n=0}^{\infty} \frac{(-iz)^n}{n!} \phi(x^n) \right) \left(\sum_{n=0}^{\infty} \frac{(-iz)^n}{n!} \phi(y^n) \right) \\ &= \mathcal{F}_{\mu}(z) \mathcal{F}_{\nu}(z). \end{aligned}$$

Por lo tanto, si aplicamos logaritmo de ambos lados, obtenemos el resultado deseado. \square

1.4.2. La transformada de Cauchy

Para los otros tipos de independencia la transformada de Cauchy resulta ser más útil para codificar la información de los momentos.

Definición 1.4.6. Sea μ una medida de probabilidad en \mathbb{R} . La *transformada de Cauchy* de μ es la función G_{μ} definida en el semiplano superior $\mathbb{C}^+ = \{s + it \mid s, t \in \mathbb{R}, t > 0\}$ por la fórmula:

$$G_{\mu}(z) = \int_{\mathbb{R}} \frac{d\mu(t)}{z - t}, \quad z \in \mathbb{C}^+.$$

Observación 1.4.7. 1. Es sencillo verificar que G_{μ} es analítica en \mathbb{C}^+ y que toma valores en $\mathbb{C}^- = \{s + it \mid s, t \in \mathbb{R}, t < 0\}$.

2. Supongamos que μ tiene soporte compacto, y denotemos $r := \sup\{|t| \mid t \in \text{sop}(\mu)\}$. Para $|z| > r$ podemos expandir:

$$\frac{1}{z-t} = \sum_{n=0}^{\infty} \frac{t^n}{z^{n+1}}, \quad \forall t \in \text{sop}(\mu).$$

La convergencia de esta serie es uniforme para $t \in \text{sop}(\mu)$ y por lo tanto podemos integrar la serie término a término y así obtenemos la siguiente expansión en serie de potencias:

$$G_\mu(z) = \sum_{n=0}^{\infty} \frac{m_n}{z^{n+1}}, \quad |z| > r,$$

donde $m_n := \int_{\mathbb{R}} t^n d\mu(t)$ es el n -ésimo momento de μ , para $n \geq 0$.

3. Observemos que si tomamos la expansión anterior alrededor del punto al infinito, una consecuencia obvia es que:

$$\lim_{z \in \mathbb{C}^+, |z| \rightarrow \infty} zG_\mu(z) = 1.$$

Observación 1.4.8. Podemos recuperar la medida de probabilidad μ a partir de su transformada de Cauchy, G_μ , utilizando la fórmula de inversión de Stieltjes. Si denotamos

$$h_\epsilon(t) := -\frac{1}{\pi} \text{Im}[G_\mu(t + i\epsilon)], \quad \forall \epsilon > 0, \quad \forall t \in \mathbb{R}$$

(donde Im es la operación de tomar la parte imaginaria de un número complejo), entonces la fórmula de inversión de Stieltjes dice que

$$d\mu(t) = \lim_{\epsilon \rightarrow 0} h_\epsilon(t) dt.$$

El límite lo estamos considerando en la topología débil sobre el espacio de medidas de probabilidad en \mathbb{R} , y por lo tanto, se refiere al hecho de que

$$\int_{\mathbb{R}} f(t) d\mu(t) = \lim_{\epsilon \rightarrow \infty} \int_{\mathbb{R}} f(t) h_\epsilon(t) dt,$$

para toda función continua y acotada $f : \mathbb{R} \rightarrow \mathbb{C}$.

1.4.3. Otras transformadas

Definición 1.4.9. Sea μ una medida de probabilidad en \mathbb{R} .

1. La *transformada \mathcal{K}* de μ es la función inversa (con respecto a la composición) de $G_\mu(z)$:

$$\mathcal{K}_\mu(z) := G_\mu^{-1}(z).$$

2. La *transformada* \mathcal{R} de μ es la serie formal de potencias tal que:

$$\mathcal{R}_\mu(z) := \mathcal{K}_\mu(z) - \frac{1}{z} = G_\mu^{-1}(z) - \frac{1}{z}.$$

3. La *transformada* F de μ es la recíproca de la transformada de Cauchy:

$$F_\mu(z) := \frac{1}{G_\mu(z)}.$$

4. La *transformada* \mathcal{B} de μ es la serie formal de potencias tal que:

$$\mathcal{B}_\mu(z) := \frac{1}{z} - F_\mu(1/z) = \frac{1}{z} - \frac{1}{G_\mu(\frac{1}{z})}.$$

Teorema 1.4.10. Sean μ y ν medidas de probabilidad en \mathbb{R} con soporte compacto.

1. Si μ y ν son libres, tenemos que

$$\mathcal{R}_{\mu \boxplus \nu}(z) = \mathcal{R}_\mu(z) + \mathcal{R}_\nu(z).$$

2. Si μ y ν son independientes en el sentido booleano, tenemos que

$$\mathcal{B}_{\mu \uplus \nu}(z) = \mathcal{B}_\mu(z) + \mathcal{B}_\nu(z).$$

3. Si μ y ν son monótonamente independientes, tenemos que

$$F_{\mu \triangleright \nu}(z) = F_\mu(F_\nu(z)).$$

A diferencia del caso clásico, para dar una demostración analítica de estos teoremas es necesario introducir varias herramientas analíticas que quedan fuera de los objetivos de esta tesis, por ello, nos limitaremos a dar una demostración combinatoria en el Capítulo 3.

1.5. Teorema del Límite Central

Una vez que hemos entendido cómo funciona la suma de medidas de probabilidad, una pregunta bastante natural es cómo se comporta la suma cuando tenemos muchas variables con la misma distribución. Para el caso clásico, es bastante conocido que tenemos el Teorema del Límite Central (de ahora en adelante abreviaremos TLC) que justamente nos habla del comportamiento de

$$\frac{a_1 + \dots + a_N}{\sqrt{N}}$$

cuando N tiende a infinito y donde $a_1, a_2 \dots$ son variables aleatorias independientes e idénticamente distribuidas. En esta sección recordaremos este teorema y hablaremos sobre los teoremas análogos que se obtienen con los distintos tipos de independencia.

Primero, necesitamos definir el concepto de convergencia. El TLC clásico normalmente aparece en términos de convergencia débil, sin embargo para poder hablar del TLC en probabilidad no conmutativa es mejor definir convergencia de otra forma. A continuación mencionaremos ambas definiciones de convergencia y veremos porque en el caso que estamos trabajando, no importa cuál de las dos elijamos.

Definición 1.5.1. Sean (\mathcal{A}_N, ϕ_N) (con $N \in \mathbb{N}$) y (\mathcal{A}, ϕ) espacios de probabilidad no conmutativos y consideremos variables aleatorias $a \in \mathcal{A}$ y $a_N \in \mathcal{A}_N$ para cada $N \in \mathbb{N}$.

1. Si se cumple que

$$\lim_{N \rightarrow \infty} \phi_N(a_N^n) = \phi(a^n) \quad \forall n \in \mathbb{N},$$

entonces decimos que a_N converge en distribución a a cuando $N \rightarrow \infty$ y lo denotamos como

$$a_N \xrightarrow{\text{distr}} a.$$

2. Si μ_N y μ son las distribuciones en el sentido analítico de a_N y a , respectivamente, entonces diremos que a_N converge débilmente en distribución a a si μ_N converge débilmente a μ , es decir:

$$\lim_{N \rightarrow \infty} \int f(t) d\mu_N(t) = \int f(t) d\mu(t) \quad \forall f \text{ continua y acotada.}$$

Observación 1.5.2. Utilizando el teorema de Stone-Weierstrass, la convergencia de todos los momentos es suficiente para asegurar la convergencia de todas las funciones continuas f en el soporte compacto de μ . Por lo tanto, en este caso, ambas nociones de convergencia coinciden.

La observación anterior sólo aplica en situaciones en el que el elemento límite a tiene una distribución con soporte compacto (que es lo que se pide en la Definición 1.2.5). Debido a que la densidad normal no tiene soporte compacto, el teorema clásico no cae en el caso anterior. Sin embargo, la distribución normal tiene la propiedad de que está determinada por sus momentos, y eso nos permite llegar a la misma conclusión.

Definición 1.5.3. Sea μ una medida de probabilidad en \mathbb{R} con momentos

$$m_n := \int_{\mathbb{R}} t^n d\mu(t) \quad \forall n \in \mathbb{N}.$$

Decimos que μ está *determinada por sus momentos*, si μ es la única medida de probabilidad en \mathbb{R} con estos momentos, es decir, si para cualquier medida de probabilidad ν en \mathbb{R} tenemos que

$$\int_{\mathbb{R}} t^n d\nu(t) = m_n \quad \forall n \in \mathbb{N} \quad \iff \quad \nu = \mu.$$

Observación 1.5.4. 1. La importancia de las medidas de probabilidad determinadas por sus momentos radica en los siguientes dos hechos bastante conocidos de la probabilidad clásica:

- La distribución normal está determinada por sus momentos.
- Sean μ y μ_N (con $N \in \mathbb{N}$) medidas de probabilidad en \mathbb{R} tales que μ está determinada por sus momentos y μ_N tiene momentos de todos los órdenes. Si tenemos que

$$\lim_{N \rightarrow \infty} \int_{\mathbb{R}} t^n d\mu_N(t) = \int_{\mathbb{R}} t^n d\mu(t) \quad \forall n \in \mathbb{N},$$

entonces μ_N converge débilmente a μ .

Estos dos hechos implican que para demostrar la convergencia débil de variables aleatorias clásicas a una distribución normal, es suficiente con mostrar la convergencia de todos los momentos. Por lo tanto, para demostrar el TLC clásico, solo hará falta demostrar la convergencia de todos los momentos, que es lo mismo que trabajar sólo con la nueva noción de convergencia en distribución.

2. Con la definición anterior, tiene sentido ampliar un poco nuestra definición de distribución analítica y permitir medidas de probabilidad que están determinadas por sus momentos aún cuando no tengan soporte compacto. Esto nos da una mayor flexibilidad para conectar nuestras consideraciones combinatorias con las consideraciones analíticas clásicas.

Con la introducción anterior, ya estamos listos para presentar el TLC para cada una de las independencias desde una plataforma común.

Teorema 1.5.5 (Teorema del Límite Central). Sea (\mathcal{A}, ϕ) un espacio de probabilidad y sea $a_1, a_2, \dots \in \mathcal{A}$ una sucesión de variables aleatorias auto-adjuntas idénticamente distribuidas. Además, supongamos que las variables son centradas ($\phi(a_n) = 0$ para toda $n \in \mathbb{N}$) y denotemos por $\sigma^2 := \phi(a_n^2)$ a la varianza común de las variables. Entonces tenemos que:

1. (Clásico) Si las variables aleatorias a_1, a_2, \dots son independientes en el sentido clásico entonces

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

donde z es una variable Gaussiana de varianza σ^2 .

2. (Libre) Si las variables aleatorias a_1, a_2, \dots son libres entonces

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

donde s es una variable semicircular de varianza σ^2 .

3. (Booleano) Si las variables aleatorias a_1, a_2, \dots son independientes en el sentido booleano entonces

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

donde b es una variable Bernoulli de varianza σ^2 .

4. (Monótono) Si las variables aleatorias a_1, a_2, \dots son independientes en el sentido clásico entonces

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

donde x es una variable aleatoria con distribución ley del arco seno de varianza σ^2 .

Esta demostración también la dejaremos para el Capítulo 3, ya que contamos con las herramientas suficientes para dar una demostración combinatoria clara y concisa.

Capítulo 2

La Función De Möbius En Retículas De Particiones

El objetivo de este capítulo es familiarizarnos con todos los conceptos y las herramientas combinatorias básicas que serán necesarios para el siguiente capítulo. Una de las herramientas más importantes es la fórmula de inversión de Möbius, para hablar de ella, primero necesitamos introducir el concepto de álgebra de incidencia, que a su vez necesita de otros conceptos más básicos. Posteriormente hablaremos de particiones, que será un concepto fundamental para lo que resta de este trabajo. Luego, nos concentraremos en las retículas de particiones y calcularemos la función de Möbius para estos casos específicos. Por último, hablaremos de familias multiplicativas de funciones, que nos ayuda a formalizar una propiedad importante que cumple la función de Möbius.

2.1. Conjuntos parcialmente ordenados y retículas

Las nociones de conjunto parcialmente ordenado y de retícula son unas de las más básicas en matemáticas, así que nos limitaremos a dar las definiciones de los conceptos que utilizaremos y las observaciones básicas para que el lector se familiarice con la notación que utilizaremos en lo que resta de este trabajo.

Definición 2.1.1. Un *conjunto parcialmente ordenado* (copo), (P, \leq) , es un conjunto, P , junto con una relación binaria, \leq (escribiremos \leq_P si puede haber confusión), que satisface los siguientes tres axiomas:

1. Reflexividad: $t \leq t$ para todo $t \in P$.
2. Antisimetría: si $s \leq t$ y $t \leq s$, entonces $s = t$.
3. Transitividad: si $s \leq t$ y $t \leq u$, entonces $s \leq u$.

Para simplificar, nos referiremos a (P, \leq) sólo como P , (teniendo en cuenta que por lo general le asignaremos un solo orden, \leq).

Utilizaremos la notación $s < t$ para referirnos a $s \leq t$ y $s \neq t$. Diremos que dos elementos s y t de P son comparables si $s \leq t$ ó $t \leq s$. En caso contrario, diremos que s y t son incomparables.

Definición 2.1.2. 1. Dos copos P y Q son *isomorfos*, y escribimos $P \sim Q$, si existe una biyección $\phi : P \rightarrow Q$ que preserve el orden y cuya inversa también preserve el orden. Es decir,

$$s \leq_P t \iff \phi(s) \leq_Q \phi(t).$$

2. El producto directo (o cartesiano) de los ordenes parciales sobre los conjuntos P_1, \dots, P_n es el orden parcial sobre el producto cartesiano de los conjuntos, $P_1 \times \dots \times P_n$, definido como

$$(s_1, \dots, s_n) \leq_{P_1 \times \dots \times P_n} (t_1, \dots, t_n) \iff s_i \leq_{P_i} t_i \quad \forall i = 1, \dots, n.$$

Definición 2.1.3. Sea (P, \leq) un copo.

1. Diremos que (Q, \leq_Q) es un *subconjunto parcialmente ordenado* o *subcopo* de (P, \leq_P) si $Q \subset P$ y \leq_Q es tal que para todo $s, t \in Q$ tenemos que

$$s \leq_Q t \iff s \leq_P t.$$

2. Un tipo especial de subcopo de P es el *intervalo cerrado* que se define como $[s, t] = \{u \in P : s \leq u \leq t\}$ siempre que $s \leq t$. (El conjunto vacío no es cerrado). Denotaremos por $P^{(2)}$ al conjunto de intervalos cerrados de P , es decir,

$$P^{(2)} := \{[x, y] \mid x, y \in P, x \leq y\}.$$

3. Si todos los intervalos de P son finitos, entonces decimos que P es un copo *localmente finito*.

4. Decimos que P tiene un *mínimo* si existe un elemento $0_P \in P$ tal que $0_P \leq t$ para todo $t \in P$. Análogamente, P tiene un *máximo* si existe un elemento $1_P \in P$ tal que $t \leq 1_P$ para todo $t \in P$.

Definición 2.1.4. Sea (P, \leq) un copo.

1. Una *cadena* (o conjunto totalmente ordenado, coto) es un copo en el que cualesquiera dos elementos son comparables. Diremos que un subconjunto C de un copo P es una cadena, si C es una cadena cuando la vemos como subcopo de P .

2. La *longitud*, $l(C)$, de una cadena finita se define como $l(C) := \#C - 1$. La *longitud o rango* de un copo finito es

$$l(P) := \text{máx}\{l(C) : C \text{ es una cadena de } P\}.$$

Por simplicidad, a la longitud de un intervalo $[s, t]$ la denotaremos como $l(s, t)$.

Para este trabajo solo nos importará un tipo muy específico de copo, conocido como retícula.

Definición 2.1.5. Sea P un copo finito. Y sean $s, t \in P$.

1. Si el conjunto $U = \{r \in P : r \geq s \text{ y } r \geq t\}$ es no vacío y tiene un mínimo r_0 , entonces decimos que r_0 es el *supremo* de s y t , y lo denotamos como $s \vee t$.
2. Si el conjunto $L = \{r \in P : r \leq s \text{ y } r \leq t\}$ es no vacío y tiene un mínimo r_0 , entonces decimos que r_0 es el *ínfimo* de s y t , y lo denotamos como $s \wedge t$.
3. Decimos que el copo P es una *retícula* si cualesquiera dos elementos de $s, t \in P$ tienen un ínfimo y un supremo.

Observación 2.1.6. 1. Sea P una retícula finita. Inductivamente podemos ver que cualquier familia finita de elementos $t_1, \dots, t_n \in P$ tiene un supremo (menor cota superior común), $t_1 \vee \dots \vee t_n$, y un ínfimo (mayor cota inferior común), $t_1 \wedge \dots \wedge t_n$. En particular, si tomamos todos los elementos de P , entonces tenemos que existe un elemento mínimo 0_P y un elemento máximo 1_P .

2. Sea P un copo finito. Es útil observar que si P tiene un elemento máximo y cualesquiera dos elementos de P tienen ínfimo, entonces P es una retícula. Análogamente, si P tiene un elemento mínimo y cualesquiera dos elementos de P tienen supremo, entonces P es una retícula.
3. Si P_1, \dots, P_n son retículas, entonces $P = P_1 \times \dots \times P_n$ también es una retícula. Además, el supremo y el ínfimo en P se pueden obtener encontrando componente por componente el supremo y el ínfimo:

$$(t_1, \dots, t_n) \vee (s_1, \dots, s_n) = (t_1 \vee s_1, \dots, t_n \vee s_n),$$

$$(t_1, \dots, t_n) \wedge (s_1, \dots, s_n) = (t_1 \wedge s_1, \dots, t_n \wedge s_n).$$

2.2. Álgebras de incidencia

En esta sección introduciremos las álgebras de incidencia, nos basaremos en las definiciones y propiedades básicas que se pueden encontrar en [Sta12].

Definición 2.2.1. Sea P un copo localmente finito y K un campo. El *álgebra de incidencia*, $I(P, K)$, de P sobre K , es la K -álgebra de todas las funciones

$$F : P^{(2)} \rightarrow K$$

(con la estructura usual de espacio vectorial sobre K), donde definimos la multiplicación o *convolución* como

$$(F * G)(s, u) = \sum_{s \leq t \leq u} F(s, t)G(t, u),$$

(escribiremos $F(x, y)$ en lugar de $F([x, y])$ para simplificar).

Observación 2.2.2. 1. Como P es localmente finito entonces la suma anterior es finita y por lo tanto la convolución $F * G$ está bien definida.

2. Si P es finito, podemos etiquetar a los elementos de P con t_1, t_2, \dots, t_p de tal forma que $t_i < t_j \Rightarrow i < j$. Entonces, tenemos que $I(P, K)$ es isomorfo al álgebra de las matrices triangulares superiores sobre K , $M = (m_{ij})_{i,j=1}^p$, que cumplen que $m_{ij} = 0$ si t_i y t_j son incomparables. Para el isomorfismo hay que asignarle a cada función $F \in I(P, K)$ la matriz $M_F = (m_{ij})_{i,j=1}^p$ tal que

$$m_{ij} = \begin{cases} F(t_i, t_j) & \text{si } t_i \leq t_j \\ 0 & \text{en cualquier otro caso.} \end{cases}$$

3. La convolución que acabamos de definir corresponde al producto de las matrices triangulares superiores correspondientes, es decir,

$$M_{F * G} = M_F M_G.$$

Por lo tanto, podemos deducir que la convolución será asociativa, tendrá elemento identidad (que será la identidad por ambos lados y es única), los inversos por un lado serán únicos (y también serán inversos por el otro lado), y la convolución será distributiva respecto a las operación de tomar combinaciones lineales de funciones en $P^{(2)}$.

Definición 2.2.3. A la identidad de $I(P, K)$ la denotamos como δ y está definida de la siguiente manera:

$$\delta(s, t) = \begin{cases} 1 & \text{si } s = t \\ 0 & \text{si } s \neq t. \end{cases}$$

Es claro que δ es la identidad en $I(P, K)$ debido a que M_δ es la matriz identidad.

Proposición 2.2.4. Sea $f \in I(P, K)$, entonces tenemos que f es invertible respecto a la convolución si y sólo si $f(t, t) \neq 0$ para todo $t \in P$.

Demostración. La proposición $f * g = \delta$ implica que

$$f(s, s)g(s, s) = 1 \quad \forall s \in P. \quad (2.1)$$

Esto nos dice que si f es invertible, entonces $f(t, t) \neq 0$. Por otra parte, de la ecuación

$$g(s, u) = -\frac{1}{f(s, s)} \sum_{s < t \leq u} f(s, t)g(t, u) \quad \forall [s, u] \in P \text{ con } s < u, \quad (2.2)$$

podemos observar que si $s < u$ entonces $g(s, u)$ depende sólo del intervalo $[s, u]$. Y por lo tanto, si $f(s, s)$ es distinto de cero, entonces podemos definir $g(s, u)$ recursivamente sobre la longitud de $[s, u]$. Es decir, empezamos definiendo los valores para los intervalos en los que $s = u$, luego en los que u cubre a s , luego en los que $l(s, u) = 2$ y así sucesivamente. Como la proposición $f * g = \delta$ es equivalente a las ecuaciones (2.1) y (2.2), podemos concluir que la inversa, g , de f existe si y sólo si $f(s, s) \neq 0$ para todo $s \in P$. \square

A continuación presentamos las dos funciones más importantes de un álgebra de incidencia, $I(P, K)$. La función zeta y su inversa la función de Möbius.

Definición 2.2.5. 1. La *función zeta* de P denotada como $\zeta \in I(P, K)$ está definida como

$$\zeta(s, u) = 1, \quad \forall (s, u) \in P^{(2)}.$$

2. La inversa de ζ respecto a la convolución es conocida como la *función de Möbius* de P y se denota por μ .

Observación 2.2.6. También podemos definir μ recursivamente. Esto es porque la relación $\mu * \zeta = \delta$ es equivalente a

$$\begin{aligned} \mu(s, s) &= 1 \quad \forall s \in P, \\ \mu(s, u) &= - \sum_{s \leq t < u} \mu(s, t) \quad \forall s < u \text{ en } P. \end{aligned} \quad (2.3)$$

La convolución con la que hemos estado trabajando es para funciones cuyo dominio es $P^{(2)}$, es decir, funciones que guardan información dada por los intervalos de P . Sin embargo, también será muy útil trabajar con funciones que sólo guardan información de P , es decir, funciones de la forma $f : P \rightarrow K$.

A continuación presentaré otro tipo de convolución para cuando tenemos una función con dominio en P y una con dominio en $P^{(2)}$ y cuya convolución nos dará otra función con dominio en P .

Definición 2.2.7. Para una función $f : P \rightarrow K$ y una función $G : P^{(2)} \rightarrow K$ se define la convolución $f * G : P \rightarrow K$ dada por la fórmula:

$$(f * G)(s) := \sum_{t \leq s} f(t)G(t, s).$$

Utilizando la misma lista del segundo inciso en la Observación 2.2.2 podemos asignarle a las funciones $f : P \rightarrow K$ un vector de tamaño n y de esta forma el segundo tipo de convolución que definimos se puede ver también como un producto de matrices con un vector. Sin embargo, para el caso en el que P tiene un elemento mínimo, 0_P , resultará más sencillo ver a las funciones con dominio en P como un subconjunto especial de las funciones con dominio en $P^{(2)}$.

Definición 2.2.8. Sea P un copo con elemento mínimo 0_P .

1. Diremos que una función $F : P^{(2)} \rightarrow K$ es unidimensional si cumple que $F(u, r) = 0$ para toda $u \neq 0_P$.
2. A cada función $f : P \rightarrow K$ le podemos asignar una función unidimensional $\bar{f} : P^{(2)} \rightarrow K$ tal que

$$\bar{f}(u, r) := \begin{cases} f(r) & \text{si } u = 0_P \\ 0 & \text{si } u \neq 0_P. \end{cases}$$

Observación 2.2.9. De la definición que acabamos de dar es sencillo observar que la operación \bar{f} que acabamos de definir es una biyección entre las funciones de P en K y las funciones unidimensionales. Además con esta operación queda mucho más clara la relación entre los dos tipos de convolución que definimos, ya que tenemos que:

$$f * G = h \Leftrightarrow \bar{f} * G = \bar{h},$$

(observemos que la convolución del lado izquierdo es la segunda que definimos, mientras que la del lado derecho es la primera).

2.3. Inversión de Möbius

La teoría de las álgebras de incidencia ha sido ampliamente desarrollada de manera general, sin embargo, en este caso sólo será necesario trabajar con álgebras bastante específicas, en las cuales P es una retícula finita. De ahora en adelante nos concentraremos en el caso específico en el que el copo P es una retícula finita. Esto nos dice que en particular podemos usar la operación \bar{f} que acabo de definir, (para el caso general en el que P es finito pero no tiene mínimo, también se puede definir una operación, un poco más compleja, para ver que ambas convoluciones están relacionadas).

Proposición 2.3.1 (Fórmula de inversión de Möbius). Sea P un copo finito con elemento mínimo. Y sean $f, g : P \rightarrow K$, dónde K es un campo. Entonces

$$g(t) = \sum_{s \leq t} f(s) \quad \forall t \in P$$

si y sólo si

$$f(t) = \sum_{s \leq t} g(s)\mu(s, t) \quad \forall t \in P.$$

Demostración. La primera fórmula es equivalente a $g = \zeta * f$ mientras que la segunda es equivalente a $f = \mu * g$. Por lo tanto, el resultado se sigue directamente del hecho de que μ es la función inversa de ζ :

$$g = \zeta * f \Leftrightarrow \bar{g} = \zeta * \bar{f} \Leftrightarrow \mu * \bar{g} = \bar{f} \Leftrightarrow f = \mu * g.$$

(Nota: la proposición anterior también se vale cuando quitamos la hipótesis del elemento mínimo) \square

Proposición 2.3.2 (Teorema del producto). Sean P y Q dos copo localmente finitos, y sea $P \times Q$ el producto directo. Si $(s, t) \leq (s', t')$ en $P \times Q$ entonces

$$\mu_{P \times Q}((s, t), (s', t')) = \mu_P(s, s')\mu_Q(t, t').$$

Demostración. Sea $(s, t) \leq (s', t')$. Tenemos que

$$\begin{aligned} \sum_{(s,t) \leq (u,v) \leq (s',t')} \mu_P(s, u)\mu_Q(t, v) &= \left(\sum_{s \leq u \leq s'} \mu_P(s, u) \right) \left(\sum_{t \leq v \leq t'} \mu_Q(t, v) \right) \\ &= \delta_{ss'}\delta_{tt'} = \delta_{(s,t), (s',t')}. \end{aligned}$$

Para terminar sólo debemos comparar la ecuación anterior con la ecuación (2.3), que sabemos que determina a μ . \square

Proposición 2.3.3. 1. Sean P y Q copos finitos y supongamos que $\Phi : P \rightarrow Q$ es un isomorfismo de copos. Entonces $\mu_Q(\Phi(s), \Phi(t)) = \mu_P(s, t)$ para todo $s, t \in P$ tal que $s \leq t$ (donde μ_P y μ_Q son las funciones de Möbius de las retículas P y Q , respectivamente).

2. Sean P_1, P_2, \dots, P_k copos finitos y considera el producto directo $P = P_1 \times P_2 \times \dots \times P_k$. Entonces para $s_1 \leq t_1$ en $P_1, \dots, s_k \leq t_k$ en P_k tenemos que

$$\mu_P((s_1, \dots, s_k), (t_1, \dots, t_k)) = \mu_{P_1}(s_1, t_1) \cdots \mu_{P_k}(s_k, t_k).$$

Demostración. 1. Este resultado se debe a que la función de Möbius sólo depende de la estructura del copo y se puede ver como consecuencia directa de la definición recursiva de μ que mencionamos en la Observación 2.2.6.

2. Para este inciso sólo hace falta aplicar varias veces el Teorema del Producto (Teorema 2.3.2). □

La siguiente proposición muestra que, en el contexto de las retículas, se pueden escribir versiones parciales de la fórmula de inversión de Möbius.

Proposición 2.3.4. Sea P una retícula finita y sea μ la función de Möbius de P . Consideremos dos funciones $f, g : P \rightarrow K$ que están relacionadas de la siguiente manera:

$$g(t) = \sum_{s \leq t} f(s) \quad \forall t \in P.$$

Entonces, para todo $u, t \in P$ con $u \leq t$, tenemos la relación:

$$\sum_{u \leq r \leq t} g(r)\mu(r, t) = \sum_{s \vee u = t} f(s).$$

Demostración. Tenemos que

$$\sum_{u \leq r \leq t} g(r)\mu(r, t) = \sum_{u \leq r \leq t} \sum_{s \leq r} f(s)\mu(r, t) = \sum_{s \leq t} \sum_{s \vee u \leq r \leq t} \mu(r, t)f(s).$$

Ahora, consideremos $s \in P$ con $s \leq t$. Como también tenemos que $u \leq t$, entonces $s \vee u \leq t$. Hay dos posibles casos:

- El primer caso es $s \vee u = t$, en donde la suma correspondiente sobre r se reduce sólo al término $r = s \vee u = t$ y contribuye con $\mu(r, t)f(s) = \mu(t, t)f(s) = f(s)$.
- El segundo caso es $s \vee u < t$, en donde la suma correspondiente es igual a cero debido a la fórmula recursiva (2.3) para la función de Möbius:

$$\sum_{s \vee u \leq r \leq t} \mu(r, t) = 0 \quad \text{si } s \vee u < t.$$

Entonces la suma que nos quedaba en los cálculos anteriores es igual a

$$\sum_{\substack{s \vee u = t \\ s \leq t}} f(s).$$

Como tenemos que $s \vee u = t$ implica que $s \leq t$, concluimos que

$$\sum_{u \leq r \leq t} g(r)\mu(r, t) = \sum_{s \vee u = t} f(s).$$

□

Corolario 2.3.5. Sea P una retícula finita y sea μ la función de Möbius de P . Entonces, para cada $u \neq 0_P$ tenemos que

$$\sum_{s \vee u = 1_P} \mu(0_P, s) = 0.$$

Demostración. Considera la función $f : P \rightarrow K$ definida por

$$f(s) := \mu(0_P, s).$$

Fijémonos en la función $g := f\zeta$ (obtenida a partir de f haciendo sumas parciales), para cada $t \in P$ tenemos que

$$\begin{aligned} g(r) = \sum_{s \leq r} f(s) &= \sum_{0_P \leq s \leq r} \mu(0_P, s) \zeta(s, r) \\ &= \mu\zeta(0_P, r) = \begin{cases} 1 & \text{si } r = 0_P \\ 0 & \text{en otro caso.} \end{cases} \end{aligned}$$

Si utilizamos la proposición anterior con las funciones f y g , elegimos la u que queremos y tomamos $t = 1_P$. Entonces obtenemos que

$$\sum_{u \leq r \leq 1_P} g(r) \mu(r, 1_P) = \sum_{s \vee u = 1_P} f(s).$$

El lado izquierdo de esta ecuación es igual a 0 ya que tenemos que $g(r) = 0$ para todos los elementos s involucrados en la suma. Por lo tanto el lado derecho también es 0. \square

2.4. Particiones

En esta sección introduciremos todos los conceptos necesarios para poder definir formalmente las retículas con las que trabajaremos.

Definición 2.4.1. Sea S un conjunto finito totalmente ordenado.

1. Decimos que $\pi = \{V_1, \dots, V_r\}$ es una *partición* del conjunto S si y sólo si los V_i ($1 \leq i \leq r$) son subconjuntos no vacíos de S , son disjuntos por parejas y cumplen que $V_1 \cup \dots \cup V_r = S$. Decimos que V_1, \dots, V_r son los *bloques* de π . Denotamos por $|\pi|$ al número de bloques de π . Al conjunto de todas las particiones de S lo denotamos como $\mathcal{P}(S)$.
2. Cualquier partición define una relación de equivalencia en S y viceversa. Dados dos elementos $p, q \in S$ y una partición $\pi \in \mathcal{P}(s)$ escribiremos $p \sim_\pi q$ si y sólo si p y q pertenecen al mismo bloque de π .

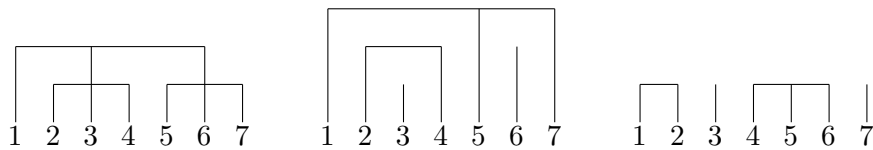
3. Diremos que una partición π del conjunto S se *cruza* si existen $p_1 < q_1 < p_2 < q_2$ en S tales que $p_1 \sim_\pi p_2 \not\sim_\pi q_1 \sim_\pi q_2$. En caso contrario diremos que π es una *partición que no se cruza*. Al conjunto de las particiones que no se cruzan de S lo denotamos por $\mathcal{NC}(S)$.
4. Diremos que un bloque V de una partición $\pi \in \mathcal{P}(n)$ es un *intervalo*, si es de la forma $V = \{a, a+1, a+2, \dots, b-1, b\}$ donde $1 \leq a \leq b \leq n$. Denotamos por $IB(n)$ al conjunto de todos los intervalos de $[n]$.
5. Una *partición de intervalos* es una partición π tal que todos sus bloques son intervalos. Al conjunto de las particiones de intervalos de S lo denotamos por $\mathcal{I}(S)$.

Observación 2.4.2. Es claro que $\mathcal{I}(S)$, $\mathcal{NC}(S)$ y $\mathcal{P}(S)$ sólo dependen del número de elementos en S . Por lo tanto, en el caso especial en el que $S = [n] = \{1, \dots, n\}$, los denotaremos como $\mathcal{I}(n)$, $\mathcal{NC}(n)$ y $\mathcal{P}(n)$, respectivamente. En el futuro, si $|S| = n$ utilizaremos la identificación natural $\mathcal{P}(S) \cong \mathcal{P}(n)$ (respect. $\mathcal{NC}(S) \cong \mathcal{NC}(n)$, $\mathcal{I}(S) \cong \mathcal{I}(n)$) que a cada elemento $i = 1, \dots, n$ le asigna el elemento $s \in S$ que está en la i -ésima posición (de menor a mayor) en el orden de S .

Definición 2.4.3. Sean $\pi, \sigma \in \mathcal{P}(n)$. Escribimos $\pi \leq \sigma$ si cada bloque de π está completamente contenido en alguno de los bloques de σ (es decir, si podemos obtener a π como un refinamiento de la estructura de bloques de σ). El orden parcial que obtenemos de esta manera en $\mathcal{P}(n)$ es conocido como el *orden de refinamiento inverso*. Como $\mathcal{I}(n)$ y $\mathcal{NC}(n)$ son subconjuntos de $\mathcal{P}(n)$, también podemos definir un orden parcial en ellos si nos restringimos sólo a sus elementos.

Observación 2.4.4. $\mathcal{I}(n)$, $\mathcal{NC}(n)$ y $\mathcal{P}(n)$ tienen un elemento máximo y uno mínimo respecto al orden de refinamiento inverso y resulta ser el mismo para los tres. El máximo es la partición que consiste de un sólo bloque y la denotamos como 1_n . El mínimo es la partición que consiste de n bloques, cada bloque con un elemento, y la denotamos como 0_n .

Figura 2.1: Ejemplos de una partición, una partición que no se cruza y una partición de intervalos.



Proposición 2.4.5. El orden parcial de refinamiento inverso induce una estructura de retícula en $\mathcal{I}(n)$, $\mathcal{NC}(n)$ y $\mathcal{P}(n)$.

Demostración. Sea $L = \mathcal{I}$, \mathcal{NC} o \mathcal{P} . Dadas $\pi, \sigma \in L(n)$, tales que $\pi = \{V_1, \dots, V_r\}$ y $\sigma = \{W_1, \dots, W_s\}$, es inmediato observar que

$$\pi \wedge \sigma = \{V_i \cap W_j \mid 1 \leq i \leq r, 1 \leq j \leq s, V_i \cap W_j \neq \emptyset\}.$$

Además, como $L(n)$ tiene un elemento máximo, por la Observación 2.1.6 podemos concluir que $L(n)$ es una retícula. \square

Definición 2.4.6. 1. Si S (totalmente ordenado) es la unión de dos conjuntos disjuntos S_1 y S_2 , entonces, para $\pi_1 \in \mathcal{P}(S_1)$ y $\pi_2 \in \mathcal{P}(S_2)$, denotaremos por $\pi_1 \cup \pi_2$ a la partición de S cuyos bloques son los bloques de π_1 y los bloques de π_2 .

2. Sea S un conjunto totalmente ordenado. Sea W un subconjunto no vacío de S en el que consideramos el orden inducido por S . Para $\pi \in \mathcal{P}(S)$ denotaremos por $\pi|_W$ la restricción de π a W . Es decir:

$$\pi|_W = \{V \cap W \mid V \text{ es un bloque de } \pi\}.$$

Observemos que en el caso particular en el que W es la unión de algunos de los bloques de π , la igualdad anterior se reduce a

$$\pi|_W = \{V \in \pi \mid V \subset W\}.$$

Observación 2.4.7. 1. Es importante observar que si π_1 y π_2 son particiones que no se cruzan, $\pi_1 \cup \pi_2$ no necesariamente es una partición que no se cruza. Por ejemplo, en la Figura 2.2 se puede apreciar que

$$\pi_1 = \{(1, 3), (4, 6)\} \in \mathcal{NC}(\{1, 3, 4, 6\})$$

y

$$\pi_2 = \{(2, 9, 10), (5, 8), (7)\} \in \mathcal{NC}(\{2, 5, 7, 8, 9, 10\}),$$

mientras que

$$\pi_1 \cup \pi_2 = \{(1, 3), (4, 6), (2, 9, 10), (5, 8), (7)\} \notin \mathcal{NC}(10).$$

2. En el caso de la restricción, es claro que si $\pi \in \mathcal{NC}(S)$ entonces $\pi|_W \in \mathcal{NC}(W)$. En la Figura 2.3 podemos ver que la restricción de $\{(1), (5), (7), (4, 6), (2, 9, 10), (3, 8), \} \in \mathcal{NC}(10)$ respecto a $\{2, 4, 6, 7, 8, 10\}$ resulta ser $\{(2, 10), (4, 6), (7), (8)\}$.

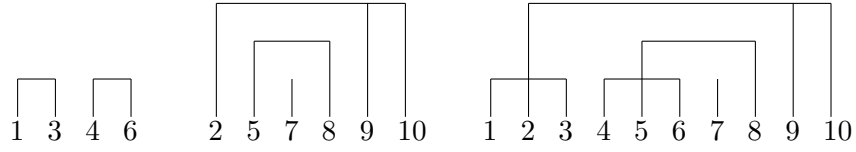


Figura 2.2: Ejemplo de la unión de dos particiones.

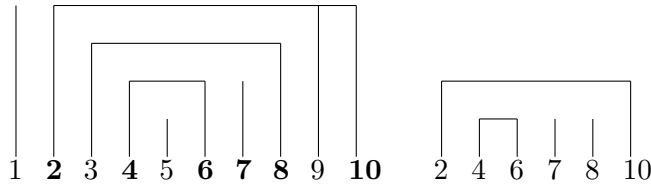


Figura 2.3: Ejemplo de restricción.

2.5. La función de Möbius en retículas de particiones

El objetivo de esta sección es calcular los valores de la función de Möbius para los tres tipos de retículas de particiones que conocemos, $\mathcal{I}(n)$, $\mathcal{P}(n)$ y $\mathcal{NC}(n)$. En toda esta sección escribiremos $L(n)$, $n \geq 1$, para referirnos indistintamente a cualquiera de las retículas $\mathcal{I}(n)$, $\mathcal{P}(n)$ o $\mathcal{NC}(n)$, a menos que se indique lo contrario. Hay que recordar que para cada uno de los tres casos tenemos una retícula distinta para cada $n \geq 1$, es decir, $\mathcal{P}(1)$ es un retícula, $\mathcal{P}(2)$ es otra, $\mathcal{P}(3)$ es otra, etc. A pesar de que tenemos muchas retículas distintas, esto no representará ningún problema ya que, de hecho, es más conveniente analizar la función de Möbius de las retículas para todos los valores de n al mismo tiempo. En los tres casos seguiremos la misma línea:

- Primero, veremos a cada intervalo como producto de retículas completas $L(n)$, es decir,

$$[\pi, \sigma] \cong L(1)^{k_1} \times L(2)^{k_2} \times \dots \times L(n)^{k_n}.$$

- Después, calcularemos los valores de la función de Möbius para las retículas completas, es decir, encontraremos $\mu_{L(n)}(0_n, 1_n)$, para $n \geq 1$.
- Por último, utilizaremos la Proposición 2.3.3 para calcular el valor de la función de Möbius en cualquier intervalo $[\sigma, \pi] \in L(n)$ a partir de la factorización del primer inciso y los valores encontrados en el segundo inciso.

Cabe mencionar que debido a la estructura de las retículas $\mathcal{NC}(n)$, necesitaremos introducir herramientas más complejas para esta retícula, además, la factorización en intervalos será más complicada y por lo tanto no daremos una fórmula explícita para calcular los valores de la función de Möbius, sino que daremos un algoritmo sencillo para obtenerlos de manera general.

Antes de trabajar de manera particular con cada tipo de retícula, veamos una factorización en intervalos que se cumple para los tres tipos de retículas y que además motiva una definición de la siguiente sección.

Lema 2.5.1 (Factorización en intervalos). Sea $[\pi, \sigma] \in L(n)^{(2)}$ y supongamos que $\sigma = \{V_1, \dots, V_r\}$ y que

$$\pi = \{W_{1,1}, \dots, W_{1,j_1}, W_{2,1}, \dots, W_{2,j_2}, \dots, W_{r,1}, \dots, W_{r,j_r}\}$$

donde $W_{i,1} \cup \dots \cup W_{i,j_i} = V_i$, para $1 \leq i \leq r$. Entonces tenemos el siguiente isomorfismo de retículas:

$$[\pi, \sigma] \cong \prod_{i=1}^r [\tau_i, 1_{|V_i|}]$$

donde para cada $1 \leq i \leq r$ tenemos que τ_i es la imagen de $\tau|_{V_i}$ bajo la biyección que preserva el orden entre V_i y $\{1, \dots, |V_i|\}$.

Demostración. Si tomamos un elemento $\pi \leq \rho \leq \sigma$, por definición necesitamos que los bloques de ρ estén contenidos en los bloques de σ , por lo tanto podemos partir a $\rho = \rho_1 \cup \dots \cup \rho_r$ de forma que $\rho_i \in L(V_i)$ para $1 \leq i \leq r$, además es claro que $\{W_{i,1}, \dots, W_{i,j_i}\} \leq \rho_i$. Es decir, tendremos que $\pi|_{V_i} \leq \rho|_{V_i} \leq \sigma|_{V_i}$, y para cada $1 \leq i \leq r$ se deben cumplir estas condiciones de manera independiente, por lo tanto tendremos que

$$[\pi, \sigma] \cong \prod_{i=1}^r [\pi|_{V_i}, \sigma|_{V_i}].$$

La igualdad deseada se sigue de ver que para cada $1 \leq i \leq r$, la biyección que preserva el orden entre V_i y $\{1, \dots, |V_i|\}$ manda a $[\pi|_{V_i}, \sigma|_{V_i}]$ en $[\tau_i, 1_{|V_i|}]$. \square

Definición 2.5.2. 1. Dado un segmento $[\pi, \sigma] \in \mathcal{P}(n)^{(2)}$. La *clase del segmento* $[\pi, \sigma]$ es una sucesión (k_1, \dots, k_n) , donde k_i es el número de bloques en π que son la unión de exactamente i bloques de σ .

2. Para una partición $\pi \in \mathcal{P}(n)$, la *clase de la partición* π se define como la clase del segmento $[0_n, \pi]$.

Observación 2.5.3. 1. De la definición anterior es claro que

$$|\sigma| = \sum_{i=1}^n ik_i \quad \text{y} \quad |\pi| = \sum_{i=1}^n k_i.$$

2. Observemos que si $\pi \in \mathcal{P}(n)$ es una partición de clase (k_1, \dots, k_n) , entonces eso quiere decir que π tiene exactamente k_i bloques de tamaño i para $i = 1, 2, \dots, n$.
3. Si tenemos que $\sigma = \{V_1, \dots, V_r\}$ y que

$$\pi = \{W_{1,1}, \dots, W_{1,j_1}, W_{2,1}, \dots, W_{2,j_2}, \dots, W_{r,1}, \dots, W_{r,j_r}\}$$

donde $W_{i,1} \cup \dots \cup W_{i,j_i} = V_i$, para $1 \leq i \leq r$. Entonces $[\pi, \sigma]$ es un intervalo de clase (k_1, \dots, k_n) donde k_i cuenta la cantidad de j 's que son iguales a i , es decir:

$$k_i = |\{t \in \{1, \dots, r\} : j_t = i\}|.$$

Antes de continuar veamos un lema que tiene que ver con la cantidad de particiones de cierta clase y que nos será útil en el último capítulo.

Lema 2.5.4. Dado un entero positivo n y enteros $r_1 \geq 0, \dots, r_n \geq 0$ tales que $r_1 + 2r_2 + \dots + nr_n = n$ entonces:

1. La cantidad de particiones con puros intervalos ($\pi \in I(n)$) de clase (r_1, \dots, r_n) es:

$$\binom{r_1 + r_2 + \dots + r_n}{r_1, r_2, \dots, r_n} = \frac{(r_1 + r_2 + \dots + r_n)!}{r_1! r_2! \dots r_n!}$$

2. La cantidad de particiones que no se cruzan ($\pi \in NC(n)$) de clase (r_1, \dots, r_n) es:

$$\frac{n!}{r_1! r_2! \dots r_n! ((n+1) - (r_1 + r_2 + \dots + r_n))!}$$

3. La cantidad de particiones ($\pi \in P(n)$) de clase (r_1, \dots, r_n) es:

$$\frac{n!}{r_1! r_2! \dots r_n! (1!)^{r_1} (2!)^{r_2} \dots (n!)^{r_n}}$$

Demostración. 1. Como son intervalos, la partición quedará definida una vez que sepamos el tamaño del primer bloque, el tamaño del segundo bloque y así con cada uno de los $r_1 + r_2 + \dots + r_n$ bloques, (esto es lo mismo que ver en qué posición están los bloques de tamaño 1, en qué posición están los bloques de tamaño 2 y así sucesivamente). Es claro que esto se puede hacer de $\binom{r_1 + r_2 + \dots + r_n}{r_1, r_2, \dots, r_n}$ formas.

2. Una demostración de este resultado se puede consultar en [Ari08]. La idea importante de la demostración es observar que si resolvemos el problema para cuando los bloques son distinguibles, el resultado

sólo depende del número de bloques $m := r_1 + \dots + r_n$. Para llegar a eso, debemos hacer una biyección entre las particiones de la forma (a_1, a_2, \dots, a_m) y las particiones de la forma $(a_1 + 1, a_2 - 1, a_3, \dots, a_m)$, donde a_i indica el tamaño del bloque A_i . Como podemos elegir el orden de los bloques como queramos, si aplicamos el proceso anterior varias veces, obtenemos que hay una biyección entre las particiones de la forma (a_1, a_2, \dots, a_m) y las particiones de la forma $(n - m + 1, 1, 1, \dots, 1)$. La cantidad de particiones de esta forma es fácil de calcular:

$$\frac{n!}{(n + 1 - m)!} = \frac{n!}{((n + 1) - (r_1 + r_2 + \dots + r_n))!}.$$

Por último, si los bloques no son distinguibles, solo hace falta dividir entre $r_1!r_2! \dots r_n!$ y de esta forma obtenemos la fórmula buscada. En [NS06] se puede consultar otra demostración de este hecho que utiliza caminos de Lukasiewicz.

3. $\binom{n}{r_1, 2r_2, 3r_3, \dots, nr_n}$ son las formas de definir los r_1 elementos que estarán en bloques de tamaño 1, los $2r_2$ elementos que estarán en bloques de tamaño 2, \dots , los nr_n elementos que estarán en bloques de tamaño n . Además, $\frac{(ir_i)!}{(i!)^{r_i}(r_i)!}$ son las maneras en las que puedes formar r_i bloques de i elementos cada uno, cuando dispones de ir_i elementos. Si multiplicamos todo obtenemos la fórmula buscada. □

2.5.1. La función de Möbius en $\mathcal{I}(n)$

Lema 2.5.5. Sea $[\pi, \sigma] \in \mathcal{I}(n)^{(2)}$ un intervalo de clase (k_1, \dots, k_n) . Tenemos el siguiente isomorfismo de retículas:

$$[\pi, \sigma] \cong \mathcal{I}(1)^{k_1} \times \dots \times \mathcal{I}(n)^{k_n}.$$

Demostración. Supongamos que $\sigma = \{V_1, \dots, V_r\}$ y que

$$\pi = \{W_{1,1}, \dots, W_{1,j_1}, W_{2,1}, \dots, W_{2,j_2}, \dots, W_{r,1}, \dots, W_{r,j_r}\}$$

donde $W_{i,1} \cup \dots \cup W_{i,j_i} = V_i$, para $1 \leq i \leq r$. Por el Lema 2.5.1 tenemos que

$$[\pi, \sigma] \cong \prod_{i=1}^r [\tau_i, 1_{|V_i|}],$$

donde τ_i es una partición de $\{1, \dots, |V_i|\}$ con j_i bloques (las imágenes de $W_{i,1}, \dots, W_{i,j_i}$). Además, podemos observar que τ_i es una partición de intervalos y si tomamos $\tau_i \leq \rho \leq 1_{|V_i|}$ tendremos que los bloques de τ_i deben estar contenidos en los bloques de ρ y a final de cuentas podemos tratar a

cada bloque (intervalo) de τ_i como un solo elemento. Formalmente, si identificamos a la imagen del bloque $W_{i,j}$ con el elemento j tendremos que

$$[\tau_i, 1_{|V_i|}] \cong [\{\{1\}, \dots, \{j_i\}\}, \{\{1, \dots, j_i\}\}] \cong [0_{j_i}, 1_{j_i}] \cong \mathcal{I}(j_i).$$

Por lo tanto, concluimos que

$$[\pi, \sigma] \cong \prod_{i=1}^r [\tau_i, 1_{|V_i|}] \cong \prod_{i=1}^r \mathcal{I}(j_i) \cong \mathcal{I}(1)^{k_1} \times \dots \times \mathcal{I}(n)^{k_n},$$

donde la última igualdad se debe al tercer inciso de la Observación 2.5.3. \square

Lema 2.5.6. Para cada $n \geq 1$, se tiene que $\mu_{\mathcal{I}(n)}(0_n, 1_n) = (-1)^{n-1}$.

Demostración. Para cada $n \geq 1$, denotaremos por $b_n := \mu_{\mathcal{I}(n)}(0_n, 1_n)$. Procederemos por inducción sobre n . Se puede comprobar fácilmente que $b_1 = 1$, $b_2 = -1$ y $b_3 = 1$. Ahora tomemos un $n \geq 4$ fijo y veamos que es cierta la fórmula suponiendo que es cierta para $1 \leq k \leq n-1$.

Utilizaremos el Corolario 2.3.5 con $P = \mathcal{I}(n)$ y $u = \{\{1\}, \{2\}, \dots, \{n-2\}, \{n-1, n\}\}$. Observemos que si $\pi \in \mathcal{P}(n)$ es una partición que cumple que $\pi \vee u = 1_n$ entonces sólo hay dos opciones, $\pi = 1_n$ ó $\pi = \{\{1, \dots, n-1\}, \{n\}\}$ (para unir al 1 con el n , π debe tener al bloque $\{1, \dots, n-1\}$). Por lo tanto, si utilizamos el Corolario 2.3.5 y la hipótesis de inducción concluimos que

$$\begin{aligned} b_n = \mu_{\mathcal{I}(n)}(0_n, 1_n) &= -\mu_{\mathcal{I}(n)}(0_n, \{\{1, \dots, n-1\}, \{n\}\}) \\ &= -a_1 a_{n-1} \\ &= (-1)^{n-1}. \end{aligned}$$

Con esto terminamos la inducción y obtenemos la fórmula deseada. \square

Corolario 2.5.7. Sea $[\pi, \sigma] \in \mathcal{I}^{(2)}(n)$ un intervalo de clase (k_1, \dots, k_n) . Entonces tenemos que

$$\mu_{\mathcal{I}(n)}(\pi, \sigma) = (-1)^{k_2+k_4+k_6+\dots}.$$

Demostración. Consecuencia directa de los dos lemas anteriores y la Proposición 2.3.3. \square

2.5.2. La función de Möbius en $\mathcal{P}(n)$

Lema 2.5.8. Sea $[\pi, \sigma] \in \mathcal{P}(n)^{(2)}$ un intervalo de clase (k_1, \dots, k_n) . Tenemos el siguiente isomorfismo de retículas:

$$[\pi, \sigma] \cong \mathcal{P}(1)^{k_1} \times \dots \times \mathcal{P}(n)^{k_n}.$$

Demostración. Esta demostración es idéntica a la que hicimos para \mathcal{I} excepto en un pequeño paso. Así que sólo explicaré con detalle ese paso. Al igual que en la demostración del Lema 2.5.5, tenemos que

$$[\pi, \sigma] \cong \prod_{i=1}^r [\tau_i, 1_{|V_i|}]$$

donde τ_i es una partición de $\{1, \dots, |V_i|\}$ con j_i bloques (las imágenes de $W_{i,1}, \dots, W_{i,j_i}$). Ahora, si tomamos $\tau_i \leq \rho \leq 1_{|V_i|}$, tendremos que los bloques de τ_i deben estar contenidos en los bloques de ρ y a final de cuentas podemos tratar a cada bloque de τ_i como un solo elemento. Formalmente, si identificamos a la imagen del bloque $W_{i,j}$ con el elemento j tendremos que

$$[\tau_i, 1_{|V_i|}] \cong [\{\{1\}, \dots, \{j_i\}\}, \{\{1, \dots, j_i\}\}] \cong [0_{j_i}, 1_{j_i}] \cong \mathcal{P}(j_i).$$

A partir de aquí, concluimos de la misma forma que en la demostración del Lema 2.5.5. \square

Este isomorfismo de intervalos es el paso diferente en ambas demostraciones, ya que la estructura del intervalo $[\tau_i, 1_{|V_i|}]$ depende de si estamos tomando todas las particiones o sólo las que no se cruzan o sólo las particiones en intervalos. Tal vez en estas dos demostraciones el cambio es imperceptible ya que este isomorfismo coincide para el caso de \mathcal{I} y \mathcal{P} , sin embargo, para \mathcal{NC} será distinto.

Lema 2.5.9. Para cada $n \geq 1$, se tiene que $\mu_{\mathcal{P}(n)}(0_n, 1_n) = (-1)^{n-1}(n-1)!$.

Demostración. La idea de la demostración es similar a la que hicimos para \mathcal{I} . Para cada $n \geq 1$, denotaremos por $a_n := \mu_{\mathcal{P}(n)}(0_n, 1_n)$. Procederemos por inducción sobre n . Se puede comprobar fácilmente que $a_1 = 1$, $a_2 = -1$ y $a_3 = 2$. Ahora tomemos un $n \geq 4$ fijo y veamos que es cierta la fórmula suponiendo que es cierta para $1 \leq k \leq n-1$.

Utilizaremos el Corolario 2.3.5 con $P = \mathcal{P}(n)$ y $u = \{\{1\}, \{2\}, \dots, \{n-2\}, \{n-1, n\}\}$. Observemos que si $\pi \in \mathcal{P}(n)$ es una partición que cumple que $\pi \vee u = 1_n$ entonces hay dos opciones, $\pi = 1_n$ ó $\pi = \{U, V\}$ donde $n-1 \in U$ y $n \in V$ (en cualquier otro caso, π tendría un bloque, W , que no contiene ni a $n-1$ ni a n y entonces W también sería bloque de $\pi \vee u$). Por lo tanto tenemos que

$$\begin{aligned} \{\pi \in \mathcal{P}(n) \mid \pi \vee u = 1_n\} &= \{1_n\} \cup \{\pi = \{U, V\} \in \mathcal{P}(n) \mid n-1 \in U \text{ y } n \in V\} \\ &= \{1_n\} \cup \{\pi_A \in \mathcal{P}(n) \mid A \subset \{1, \dots, n-2\}\} \end{aligned}$$

donde $\pi_A = \{U, V\}$ con $U = A \cup \{n-1\}$ y $V = \{1, \dots, n\} \setminus U$.

Además, sabemos que $[0_n, \{U, V\}] \cong L(|U|) \times L(|V|)$ y por lo tanto tenemos que

$$\mu_{\mathcal{P}(n)}(0_n, \pi_A) = a_{1+|A|} a_{n-1-|A|}.$$

Entonces, por el Corolario 2.3.5 y utilizando la hipótesis de inducción concluimos que

$$\begin{aligned}
a_n = \mu_{\mathcal{P}(n)}(0_n, 1_n) &= - \sum_{A \subset \{1, \dots, n-2\}} \mu_{\mathcal{P}(n)}(0_n, \pi_A) \\
&= - \sum_{A \subset \{1, \dots, n-2\}} a_{1+|A|} a_{n-1-|A|} \\
&= - \sum_{i=0}^{n-2} \sum_{\substack{A \subset \{1, \dots, n-2\} \\ |A|=i}} a_{i+1} a_{n-1-i} \\
&= - \sum_{i=0}^{n-2} \binom{n-2}{i} (-1)^i (i)! (-1)^{n-i-2} (n-i-2)! \\
&= (-1)^{n-1} \sum_{i=0}^{n-2} (n-2)! \\
&= (-1)^{n-1} (n-1)!.
\end{aligned}$$

Con esto terminamos la inducción y obtenemos la fórmula buscada. \square

Corolario 2.5.10. Sea $[\pi, \sigma] \in \mathcal{P}^{(2)}(n)$ un intervalo de clase (k_1, \dots, k_n) . Entonces tenemos que

$$\mu_{\mathcal{P}(n)}(\pi, \sigma) = \prod_{i=1}^n ((-1)^{i-1} (i-1)!)^{k_i}.$$

Demostración. Consecuencia directa de los dos lemas anteriores y la Proposición 2.3.3. \square

2.5.3. La función de Möbius en $\mathcal{NC}(n)$

Las retículas de las particiones que no se cruzan son un poco más complicadas y truculentas que las dos anteriores, así que primero hablaremos de algunas propiedades básicas de las particiones que no se cruzan y de las retículas $\mathcal{NC}(n)$ para familiarizarnos un poco con el concepto y posteriormente encontraremos los valores de la función de Möbius en estas retículas. Para empezar, será útil tener en cuenta la siguiente descripción recursiva de las particiones que no se cruzan:

Observación 2.5.11. Una partición π de $\{1, \dots, n\}$ no se cruza si y sólo si al menos un bloque $V \in \pi$ es un intervalo y $\pi \setminus V$ es una partición que no se cruza. Es decir, para algunos $a, b \in \mathbb{N}$ tales que $1 \leq a \leq b \leq n$ tenemos que $V = \{a, a+1, \dots, b-1, b\}$ y $\pi \setminus V \in \mathcal{NC}(\{1, \dots, a-1, b+1, \dots, n\}) \cong \mathcal{NC}(n - (b-a+1))$.

Proposición 2.5.12. El número de elementos en $\mathcal{NC}(n)$ es igual al n -ésimo número de Catalán, C_n .

Demostración. Para $n \geq 1$ denotamos por $d_n := |\mathcal{NC}(n)|$ y definimos $d_0 = 1$. Para $n \geq 1$ y $1 \leq i \leq n$ denotamos como $\mathcal{NC}^{(i)}(n)$ al conjunto de las particiones que no se cruzan $\pi \in \mathcal{NC}(n)$ tales que i es el elemento más grande del bloque que contiene al elemento 1. Por la condición de que no se cruza, una partición $\pi \in \mathcal{NC}^{(i)}(n)$ se puede descomponer como $\pi = \pi_1 \cup \pi_2$, donde $\pi_1 \in \mathcal{NC}^{(i)}(i)$ y $\pi_2 \in \mathcal{NC}(\{i+1, \dots, n\})$. Entonces tenemos que

$$\mathcal{NC}^{(i)}(n) \cong \mathcal{NC}^{(i)}(i) \times \mathcal{NC}(n-i).$$

Sin embargo, si restringimos π_1 a $\{1, \dots, i-1\}$ podemos ver que $\mathcal{NC}^{(i)}(i)$ está en biyección con $\mathcal{NC}(i-1)$. Se sigue que

$$\mathcal{NC}^{(i)}(n) \cong \mathcal{NC}(i-1) \times \mathcal{NC}(n-i).$$

Como $\mathcal{NC}(n) = \bigcup_{i=1}^n \mathcal{NC}^{(i)}(n)$ y esta es una unión disjunta, entonces tenemos que

$$d_n = \sum_{i=1}^n |\mathcal{NC}^{(i)}(n)| = \sum_{i=1}^n |\mathcal{NC}(i-1)| |\mathcal{NC}(n-i)| = \sum_{i=1}^n d_{i-1} d_{n-i}.$$

Como los números de Catalán cumplen esta misma recursión y $C_0 = 1 = d_0$ entonces podemos concluir que $d_n = C_n$ como queríamos. \square

Definición 2.5.13. La función complemento de Kreweras $K : \mathcal{NC}(n) \rightarrow \mathcal{NC}(n)$ se define de la siguiente manera. Consideramos los números adicionales $\bar{1}, \dots, \bar{n}$ y los entrelazamos con $1, \dots, n$ de manera que queden en el siguiente orden:

$$1 < \bar{1} < 2 < \bar{2} < \dots < n < \bar{n}.$$

Sea π una partición que no se cruza de $\{1, \dots, n\}$. Entonces su complemento de Kreweras, $K(\pi) \in \mathcal{NC}(\bar{1}, \dots, \bar{n}) \cong \mathcal{NC}(n)$ se define como la mayor partición $\sigma \in \mathcal{NC}(\bar{1}, \dots, \bar{n})$ que tiene la propiedad de que

$$\pi \cup \sigma \in \mathcal{NC}(1, \bar{1}, 2, \bar{2}, \dots, n, \bar{n}).$$

Observación 2.5.14. Es sencillo observar que si $\pi \leq \sigma$, entonces tendremos que $K(\sigma) \leq K(\pi)$. También tenemos que $K(0_n) = 1_n$ y que $K(1_n) = 0_n$. Además, si ponemos atención, $K^2(\pi)$ es muy parecida a π excepto por una permutación cíclica. Si iteramos esta permutación, deducimos que K^{2n} es el operador identidad en $\mathcal{NC}(n)$ y por lo tanto K es una biyección (también podemos describir a la inversa explícitamente al copiar la definición de K pero pidiendo que $\bar{1} < 1 < \bar{2} < 2 < \dots < \bar{n} < n$ en lugar de $1 < \bar{1} < 2 < \bar{2} < \dots < n < \bar{n}$). Gracias a las observaciones anteriores, podemos decir que el operador K es un *anti-isomorfismo de retículas*.

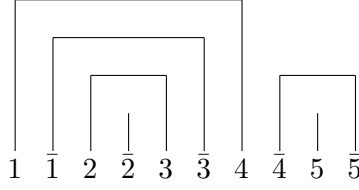


Figura 2.4: Ejemplo de complemento de Kreweras.

Con esta introducción a las retículas $\mathcal{NC}(n)$ es suficiente para seguir el mismo camino que seguimos con los otros dos tipos de retículas. Lo primero que queremos es obtener una factorización de los intervalos de $\mathcal{NC}(n)$ como producto de otros $\mathcal{NC}(k)$.

Teorema 2.5.15. Para cualquier $[\pi, \sigma] \in \mathcal{NC}(n)^{(2)}$ existe una sucesión canónica (r_1, \dots, r_n) de enteros no negativos, tal que tenemos el isomorfismo de retículas

$$[\pi, \sigma] \cong \mathcal{NC}(1)^{r_1} \times \mathcal{NC}(2)^{r_2} \times \dots \times \mathcal{NC}(n)^{r_n}.$$

Es importante recalcar que en este caso, (r_1, \dots, r_n) no necesariamente es la clase del segmento $[\pi, \sigma]$.

Demostración. Al igual que en las demostraciones para \mathcal{I} y \mathcal{P} . Supongamos que $\sigma = \{V_1, \dots, V_r\}$ y que

$$\pi = \{W_{1,1}, \dots, W_{1,j_1}, W_{2,1}, \dots, W_{2,j_2}, \dots, W_{r,1}, \dots, W_{r,j_r}\}$$

donde $W_{i,1} \cup \dots \cup W_{i,j_i} = V_i$, para $1 \leq i \leq r$. Por el Lema 2.5.1 tenemos que

$$[\pi, \sigma] \cong \prod_{i=1}^r [\tau_i, 1_{|V_i|}]$$

donde τ_i es una partición de $\{1, \dots, |V_i|\}$. Aquí es dónde todo se complica debido a que $[\tau_i, 1_{|V_i|}] \not\cong \mathcal{NC}(j_i)$ como sucedía en los dos casos anteriores. Sin embargo, aún podemos factorizar a los intervalos de la forma $[\tau, 1_k]$. Para ello utilizamos la función complemento de Kreweras en $\mathcal{NC}(k)$ y observamos que $[\tau, 1_k]$ es anti-isomorfo a $[K(1_k), K(\tau)] = [0_k, K(\tau)]$. Pero para el segundo intervalo, de nuevo sabemos que

$$[0_k, K(\tau)] \cong \prod_{W \in K(\tau)} [0_k|_W, K(\tau)|_W].$$

Como cada $[0_k|_W, K(\tau)|_W]$ es simplemente $\mathcal{NC}(W) \cong \mathcal{NC}(|W|)$, entonces tenemos que $[\tau, 1_k]$ es anti-isomorfo al producto $\prod_{W \in K(\tau)} \mathcal{NC}(|W|)$. Finalmente, este producto es anti-isomorfo consigo mismo, y así obtenemos la factorización deseada de $[\tau, 1_k]$. □

Definición 2.5.16. La descomposición que acabamos de observar en la demostración anterior es conocida como la *factorización canónica* de $[\pi, \sigma]$.

Lema 2.5.17. Para cada $n \geq 1$, se tiene que $\mu_{\mathcal{NC}(n)}(0_n, 1_n) = (-1)^{n-1}C_n$.

Demostración. La idea de la demostración es similar a la que hicimos para \mathcal{I} y \mathcal{P} . Denotemos $s_n := \mu_{\mathcal{NC}(n)}(0_n, 1_n)$ y veamos que las s_n cumplen la misma recursión que los números de Catalán salvo por un signo. Se puede comprobar fácilmente que $s_1 = 1$, $s_2 = -1$ y $s_3 = 2$. Ahora tomemos un $n \geq 4$ fijo. Utilizaremos el Corolario 2.3.5 tomando $P = \mathcal{NC}(n)$ y

$$u = \{\{1\}, \{2\}, \dots, \{n-2\}, \{n-1, n\}\}.$$

Observemos que podemos describir de manera explícita al conjunto $\{\pi \in \mathcal{NC}(n) \mid \pi \vee u = 1_n\}$. Este es $\{1_n, \pi_1, \pi_2, \dots, \pi_{n-1}\}$, donde definimos $\pi_1 := \{\{1, \dots, n-1\}, \{n\}\}$ y para cada $2 \leq i \leq n-1$ definimos $\pi_i := \{\{i, \dots, n-1\}, \{1, \dots, i-1, n\}\}$.

Lo anterior se debe a que si $\pi \in \mathcal{NC}(n)$ es una partición que cumple que $\pi \vee u = 1_n$ entonces hay dos opciones, $\pi = 1_n$ o $\pi = \{U, V\}$ donde $n-1 \in U$ y $n \in V$ (en cualquier otro caso, π tendría un bloque, W , que no contiene ni a $n-1$ ni a n y entonces W también sería bloque de $\pi \vee u$). Sea i el menor elemento de U , entonces forzosamente $U = \{i, \dots, n-1\}$ (o de lo contrario se cruzaría con V) y por lo tanto $\pi = \pi_i$.

Por el Corolario 2.3.5 tenemos que

$$\mu_n(0_n, 1_n) + \sum_{i=1}^{n-1} \mu_n(0_n, \pi_i) = 0.$$

Para cada $1 \leq i \leq n-1$ es inmediato que la factorización canónica del intervalo $[0_n, \pi_i]$ es

$$[0_n, \pi_i] \cong \mathcal{NC}(i) \times \mathcal{NC}(n-i),$$

y por lo tanto tenemos que $\mu_n(0_n, \pi_i) = s_i s_{n-i}$. Y sustituyendo obtenemos que

$$s_n = - \sum_{k=1}^{n-1} s_k s_{n-k}.$$

Esto se vale para toda $n \geq 4$ pero es sencillo ver que para los valores s_1 , s_2 y s_3 también se cumple. Por lo tanto tenemos una fórmula recursiva que salvo el signo menos, es la misma que cumplen los números de Catalán. Como además sabemos que $s_1 = 1$, entonces $s_n = (-1)^{n-1}C_{n-1}$. \square

Corolario 2.5.18. Sea $[\pi, \sigma] \in \mathcal{NC}(n)^{(2)}$ un intervalo cuya factorización canónica es

$$[\pi, \sigma] \cong \mathcal{NC}(1)^{k_1} \times \mathcal{NC}(2)^{k_2} \times \dots \times \mathcal{NC}(n)^{k_n}.$$

Entonces tenemos que

$$\mu_n(\pi, \sigma) = s_1^{k_1} s_2^{k_2} \cdots s_n^{k_n},$$

donde $s_n := \mu_{\mathcal{NC}(n)}(0_n, 1_n)$.

Demostración. Se sigue directo de la Proposición 2.3.3 y los dos lemas anteriores. \square

2.6. Familias multiplicativas de funciones

Nos hace falta un pequeño resultado para poder trabajar sin problemas con el enfoque combinatorio de los cumulantes. Lo que nos falta tiene que ver con el concepto de familias multiplicativas de funciones y el hecho de que estas familias sean cerradas bajo la convolución que definimos en 2.2.1. En la literatura, la definición de familia multiplicativa varía dependiendo del tipo de retículas con las que se trabaja [DRS72], [NS06]. Esto se debe principalmente a que la factorización en intervalos es distinta. Para el propósito de este trabajo, no es necesaria una definición tan específica en cada caso y será más ilustrativo trabajarlas de forma general tomando como base la factorización del Lema 2.5.1 que se cumple para los tres tipos de retículas.

Definición 2.6.1. Diremos que $F_n : L(n)^{(2)} \rightarrow \mathbb{C}$ es una familia multiplicativa de funciones en $L^{(2)}$ si para cada $[\tau, \sigma] \in L(n)^{(2)}$ se cumple que

$$F_n(\tau, \sigma) = \prod_{i=1}^r F_{|V_i|}(\tau_i, 1_{|V_i|})$$

donde $\sigma = \{V_1, \dots, V_r\}$ y para cada $1 \leq i \leq r$ tenemos que τ_i es la imagen de $\tau|_{V_i}$ bajo la biyección que preserva el orden entre V_i y $\{1, \dots, |V_i|\}$.

Observación 2.6.2. Por la Proposición 2.3.3 y el Lema 2.5.1 es claro que las familias de funciones de Möbius $\mu_{L(n)} : L(n)^{(2)} \rightarrow \mathbb{C}$, para $n = 1, 2, \dots$, forman una familia multiplicativa en $L^{(2)}$, para $L = \mathcal{I}, \mathcal{P}$ o \mathcal{NC} .

Definición 2.6.3. Diremos que $f_n : L(n) \rightarrow \mathbb{C}$ es una familia multiplicativa de funciones en L si $\bar{f}_n : L(n)^{(2)} \rightarrow \mathbb{C}$ es una familia multiplicativa de funciones en $L^{(2)}$.

Observación 2.6.4. Si tomamos una n fija y un intervalo $[\tau, \sigma] \in L(n)^{(2)}$ con $\tau \neq 0$, por definición, sabemos que $\bar{f}(\tau, \sigma) = 0$, además, $\tau \neq 0$ implica que $\tau_i \neq 0$ para algún i , y por lo tanto $\prod_{i=1}^r \bar{f}_{|V_i|}(\tau_i, 1_{|V_i|}) = 0$. Entonces para cualquier $f_n : L(n) \rightarrow \mathbb{C}$ tendremos que la condición

$$\bar{f}_n(\tau, \sigma) = \prod_{i=1}^r \bar{f}_{|V_i|}(\tau_i, 1_{|V_i|})$$

siempre se cumple para $\tau \neq 0$. Esto quiere decir que el hecho de que una familia de funciones $f_n : L(n) \rightarrow \mathbb{C}$ sea multiplicativa radica sólo en que cumpla la condición

$$\bar{f}_n(0, \sigma) = \prod_{i=1}^r \bar{f}_{|V_i|}(0, 1_{|V_i|})$$

para cada $[0, \sigma] \in L(n)^{(2)}$. Por como definimos la función \bar{f} , lo anterior se traduce en la condición

$$f_n(\sigma) = \prod_{i=1}^r f_{|V_i|}(1_{|V_i|}).$$

Entonces, la multiplicatividad de una familia $(f_n)_{n \geq 1}$ en L significa que tenemos una factorización de las f_n de acuerdo con la estructura de los bloques de la partición a la cual le estamos aplicando la f_n . Escribiré esto formalmente como un lema.

Lema 2.6.5. Sea $f_n : L(n) \rightarrow \mathbb{C}$ una familia de funciones. Denotemos por $\alpha_n := f_n(1_n)$, para $n \geq 1$. Entonces, $f_n : L(n) \rightarrow \mathbb{C}$ es una familia multiplicativa de funciones en $L(n)$ si y sólo si para toda partición $\pi \in L(n)$ de clase (r_1, \dots, r_n) se cumple que

$$f_n(\pi) := \alpha_1^{r_1} \alpha_2^{r_2} \cdots \alpha_n^{r_n}.$$

Notación 2.6.6. Sea $(\alpha_n)_{n \geq 1}$ una sucesión de números complejos, y sea $(f_n)_{n \geq 1}$ la familia multiplicativa de funciones en L tal que $\alpha_n = f_n(1_n)$ para $n \geq 1$, (por el lema anterior, la familia de funciones queda completamente definida al conocer los valores de las α_n). Utilizaremos la notación

$$\alpha_\pi := f_n(\pi) \quad \text{para } \pi \in L(n)$$

y nos referiremos a la familia de números $(\alpha_\pi)_{\pi \in L(n)}$ como la extensión multiplicativa de $(\alpha_n)_{n \geq 1}$.

Proposición 2.6.7. Sean $(H_n)_{n \geq 1}$ y $(F_n)_{n \geq 1}$ dos familias multiplicativas de funciones en $L^{(2)}$. Entonces la familia $(H_n * F_n)_{n \geq 1}$ también es multiplicativa en $L^{(2)}$.

Demostración. Para cada $n \geq 1$ denotaremos $G_n := H_n * F_n$, y utilizaremos la notación

$$\beta_\tau := G_n(\tau, 1_n) = \sum_{\substack{\pi \in L(n) \\ \pi \geq \tau}} H_n(\tau, \pi) F_n(\pi, 1_n).$$

Fijemos una $n \geq 1$ y una partición $\sigma = \{V_1, \dots, V_r\} \in L(n)$. Consideramos el isomorfismo de retículas canónico

$$\Phi : [\tau, \sigma] \rightarrow L(|V_1|) \times \cdots \times L(|V_r|)$$

$$\pi \mapsto (\pi_1, \dots, \pi_r)$$

donde π_k es la imagen de $\pi|_{V_k}$ bajo la biyección que preserva el orden entre V_k y $\{1, \dots, |V_k|\}$, para $1 \leq k \leq r$. Queremos verificar que

$$G_n(\tau, \sigma) = \prod_{i=1}^r \beta_{\tau_i}.$$

Tomemos un $\pi \in [\tau, \sigma]$ y supongamos que

$$\pi = \{W_{1,1}, \dots, W_{1,j_1}, W_{2,1}, \dots, W_{2,j_2}, \dots, W_{r,1}, \dots, W_{r,j_r}\}$$

donde $W_{i,1} \cup \dots \cup W_{i,j_i} = V_i$, para $1 \leq i \leq r$.

Por la multiplicatividad de H_n tenemos que

$$H_n(\tau, \pi) = \prod_{i=1}^r \prod_{j=1}^{j_i} H_{|W_{i,j}|}(\tau_{i,j}, 1_{|W_{i,j}|}) = \prod_{i=1}^r H_{|V_i|}(\tau_i, \pi_i)$$

donde $\tau_{i,j}$ es la imagen de $\tau|_{W_{i,j}}$ bajo la biyección que preserva el orden entre $W_{i,j}$ y $\{1, \dots, |W_{i,j}|\}$. Por la multiplicatividad de F_n tenemos que

$$F_n(\pi, \sigma) = \prod_{i=1}^r F_{|V_i|}(\pi_i, 1_{|V_i|})$$

donde π_k es la imagen de $\pi|_{V_k}$ bajo la biyección que preserva el orden entre V_k y $\{1, \dots, |V_k|\}$.

Por lo tanto, utilizando las fórmulas anteriores, el cambio de variable dado por la biyección Φ y factorizando concluimos que:

$$\begin{aligned} G_n(\tau, \sigma) &= \sum_{\pi \in [\tau, \sigma]} H_n(\tau, \pi) F_n(\pi, \sigma) \\ &= \sum_{\substack{(\pi_1, \dots, \pi_r) \in L(|V_1|) \times \dots \times L(|V_r|) \\ \pi_1 \geq \tau_1, \dots, \pi_r \geq \tau_r}} \left(\prod_{i=1}^r H_{|V_i|}(\tau_i, \pi_i) \right) \left(\prod_{i=1}^r F_{|V_i|}(\pi_i, 1_{|V_i|}) \right) \\ &= \sum_{\substack{(\pi_1, \dots, \pi_r) \in L(|V_1|) \times \dots \times L(|V_r|) \\ \pi_1 \geq \tau_1, \dots, \pi_r \geq \tau_r}} \prod_{i=1}^r H_{|V_i|}(\tau_i, \pi_i) F_{|V_i|}(\pi_i, 1_{|V_i|}) \\ &= \prod_{i=1}^r \left(\sum_{\substack{\pi \in L(|V_i|) \\ \pi \geq \tau_i}} H_{|V_i|}(\tau_i, \pi) F_{|V_i|}(\pi, 1_{|V_i|}) \right) \\ &= \prod_{i=1}^r \beta_{\tau_i}, \end{aligned}$$

que es lo que queríamos demostrar. \square

Corolario 2.6.8. Sea $(f_n)_{n \geq 1}$ una familia multiplicativa de funciones en L y sea $(F_n)_{n \geq 1}$ una familia multiplicativa de funciones en $L^{(2)}$. Entonces la familia $(f_n * F_n)_{n \geq 1}$ también es multiplicativa en L .

Capítulo 3

Cumulantes

En este capítulo hablaremos de los cumulantes, que son la principal herramienta combinatoria que tenemos para trabajar con los distintos tipos de independencia. Empezaremos por dar una idea general de lo que se espera de un cumulante. Después utilizaremos toda la maquinaria del capítulo anterior para definir los cumulantes libres, booleanos y clásicos de una manera simple y veremos que con esa definición, efectivamente se cumplen las propiedades deseadas. Posteriormente utilizaremos algunas ideas un poco distintas para definir los cumulantes monótonos y veremos que también cumplen las propiedades deseadas. Por último, utilizaremos los cumulantes para demostrar los teoremas que quedaron pendientes en el Capítulo 1.

3.1. Cumulantes generalizados

En [Leh04] Lehner menciona condiciones deseables que deberían tener los cumulantes, en [HS11b] Hasebe y Saigo generalizan un poco esta idea para incluir a los cumulantes monótonos. Así, en esta tesis trabajaremos con la definición más general de cumulantes, y sólo mencionaremos la propiedad más fuerte que cumplen los otros tres tipos de cumulantes.

Notación 3.1.1. Sea (\mathcal{A}, ϕ) un espacio de probabilidad y sean $a, a_1, \dots, a_N \in \mathcal{A}$ variables aleatorias independientes e idénticamente distribuidas en el sentido de momentos (es decir, $m_n^a = m_n^{a_i}$ para $1 \leq i \leq N$ y para toda n), entonces escribiremos:

$$N \cdot a := a_1 + \dots + a_N.$$

Si $N = 0$ pensaremos que $0 \cdot a = 0$.

Definición 3.1.2 (Cumulantes generalizados). Dada una noción de independencia en un espacio de probabilidad no-conmutativo (\mathcal{A}, ϕ) decimos que una sucesión de aplicaciones $\kappa_n : \mathcal{A} \rightarrow \mathbb{C}$, que manda $a \mapsto \kappa_n^a$, para

$n = 1, 2, 3, \dots$ se llama sucesión de cumulantes (con respecto a la independencia) si se cumplen las siguientes tres propiedades:

Propiedad 1. Para cualquier n , existe un polinomio Q_n de $n - 1$ variables tal que

$$m_n^a = \kappa_n^a + Q_n(\kappa_1^a, \dots, \kappa_{n-1}^a).$$

Propiedad 2. Homogeneidad: para cualquier $\lambda > 0$ y cualquier n , se tiene $\kappa_n^{\lambda a} = \lambda^n \kappa_n^a$.

Propiedad 3. Extensividad: $\kappa_n^{N \cdot a} = N \kappa_n^a$.

Observación 3.1.3. 1. La propiedad 1 nos dice que la sucesión de cumulantes contiene la misma información que la sucesión de momentos. En resumen nos dice que κ_n^a es un polinomio en los primeros n momentos de a con término mayor m_n^a .

2. La propiedad 2 nos dice que los cumulantes se comportan bien respecto a la multiplicación por una escalar. Y por lo tanto, el polinomio es homogéneo si el grado de κ_i es igual a i .

Observación 3.1.4. En el caso de la independencia clásica, libre o booleana, la propiedad 3 se puede cambiar por una aún más fuerte:

Propiedad 3'. Aditividad con respecto a la independencia: si a y b son variables aleatorias independientes, entonces $\kappa_n^{a+b} = \kappa_n^a + \kappa_n^b$.

Teorema 3.1.5. Los cumulantes generalizados son únicos y el n -ésimo cumulante está dado por el coeficiente lineal en $m_n^{N \cdot a}$.

Demostración. Por las propiedades 1 y 3 obtenemos que

$$\begin{aligned} m_n^{N \cdot a} &= \kappa_n^{N \cdot a} + Q_n(\kappa_1^{N \cdot a}, \dots, \kappa_{n-1}^{N \cdot a}) \\ &= N \kappa_n^a + Q_n(N \kappa_1^a, \dots, N \kappa_{n-1}^a). \end{aligned}$$

Por lo tanto $m_n^{N \cdot a}$ es un polinomio en N y m_k^a con $1 \leq k \leq n$. Por la propiedad 2, el polinomio Q_n no contiene términos lineales ni constantes para ninguna n . Por lo tanto, el coeficiente del término lineal N es precisamente κ_n^a . Ahora supongamos que hay otros cumulantes κ'_n . Por la propiedad 1 existen polinomios Q'_n tales que

$$m_n^{N \cdot a} = N \kappa'_n + Q'_n(N \kappa'_1, \dots, N \kappa'_{n-1}).$$

Al igualarlo con la primer ecuación, obtenemos una igualdad de polinomios en N y de manera inductiva podemos concluir que $\kappa_n^a = \kappa'_n$. \square

3.2. Definición vía particiones

Para definir los cumulantes utilizando la función de Möbius, primero convertiremos nuestro funcional ϕ en una *sucesión* de funcionales multilineales y luego extenderemos esta sucesión a una *familia* de funcionales multilineales. Para esto debemos introducir un poco de notación.

Notación 3.2.1. Sea \mathcal{A} un álgebra unitaria y sea $\phi : \mathcal{A} \rightarrow \mathbb{C}$ un funcional lineal unitario. A partir de este funcional, definimos una sucesión de funcionales multilineales en \mathcal{A} , $(\phi_n)_{n \in \mathbb{N}}$, de la siguiente manera:

$$\begin{aligned} \phi_n : \mathcal{A}^n &\rightarrow \mathbb{C} \\ (a_1, \dots, a_n) &\mapsto \phi(a_1 \cdots a_n). \end{aligned}$$

Luego, extendemos esta sucesión a una familia de funcionales multilineales ϕ_π ($n \geq 1$, $\pi \in \mathcal{P}(n)$),

$$\begin{aligned} \phi_\pi : \mathcal{A}^n &\rightarrow \mathbb{C} \\ (a_1, \dots, a_n) &\mapsto \phi_\pi[a_1, \dots, a_n], \end{aligned}$$

utilizando la siguiente fórmula. Si $\pi = \{V_1, \dots, V_r\} \in \mathcal{P}(n)$, entonces:

$$\phi_\pi[a_1, \dots, a_n] := \phi(V_1)[a_1, \dots, a_n] \cdots \phi(V_r)[a_1, \dots, a_n],$$

donde utilizamos la notación

$$\phi(V)[a_1, \dots, a_n] := \phi_s(a_{i_1}, \dots, a_{i_s}), \quad \text{para } V = \{i_1, \dots, i_s\}, \text{ con } i_1 < \cdots < i_s.$$

Utilizaremos paréntesis para los ϕ_n y corchetes para los ϕ_π . Los funcionales ϕ_π en realidad son una extensión de los funcionales ϕ_n ya que $\phi_n = \phi_{1_n}$, es decir,

$$\phi_n(a_1, \dots, a_n) = \phi_{1_n}[a_1, \dots, a_n]$$

para toda $n \geq 1$ y cualesquiera a_1, \dots, a_n .

Ahora sí, ya teniendo esta definición podemos definir los cumulantes multivariados booleanos, clásicos y libres. Hay que resaltar que esta es una definición más general de cumulante que la hecha en la Definición 3.1.2 ya que también nos permite hablar de cumulantes con distintas variables.

Definición 3.2.2. Sea (\mathcal{A}, ϕ) un espacio de probabilidad no conmutativo.

1. Los *cumulantes booleanos multivariados* $(b_\pi)_{\pi \in \mathcal{I}}$ son, para cada $n \in \mathbb{N}$ y $\pi \in \mathcal{I}(n)$, el funcional multilineal $b_\pi : \mathcal{A}^n \rightarrow \mathbb{C}$ que manda a (a_1, \dots, a_n) en

$$b_\pi[a_1, \dots, a_n] = \sum_{\substack{\sigma \in \mathcal{I}(n) \\ \sigma \leq \pi}} \phi_\sigma[a_1 \cdots a_n] \mu_{\mathcal{I}(n)}(\sigma, \pi).$$

Para cada $n \geq 1$ definimos $b_n := b_{1_n}$.

2. Los *cumulantes clásicos multivariados* $(c_\pi)_{\pi \in \mathcal{P}}$ son, para cada $n \in \mathbb{N}$ y $\pi \in \mathcal{P}(n)$, el funcional multilinear $c_\pi : \mathcal{A}^n \rightarrow \mathbb{C}$ que manda a (a_1, \dots, a_n) en

$$c_\pi[a_1, \dots, a_n] = \sum_{\substack{\sigma \in \mathcal{P}(n) \\ \sigma \leq \pi}} \phi_\sigma[a_1 \cdots a_n] \mu_{\mathcal{P}(n)}(\sigma, \pi).$$

Para cada $n \geq 1$ definimos $c_n := c_{1_n}$.

3. Los *cumulantes libres multivariados* $(r_\pi)_{\pi \in \mathcal{NC}}$ son, para cada $n \in \mathbb{N}$ y $\pi \in \mathcal{NC}(n)$, el funcional multilinear $r_\pi : \mathcal{A}^n \rightarrow \mathbb{C}$ que manda a (a_1, \dots, a_n) en

$$r_\pi[a_1, \dots, a_n] = \sum_{\substack{\sigma \in \mathcal{NC}(n) \\ \sigma \leq \pi}} \phi_\sigma[a_1 \cdots a_n] \mu_{\mathcal{NC}(n)}(\sigma, \pi).$$

Para cada $n \geq 1$ definimos $r_n := r_{1_n}$.

Notación 3.2.3. 1. Sean $(a_\lambda)_{\lambda \in \Lambda}$ variables aleatorias en un espacio de probabilidad (\mathcal{A}, ϕ) y sean $(\kappa_n)_{n \in \mathbb{N}}$ los correspondientes cumulantes multivariados (clásicos, libres o booleanos). Diremos que los *cumulantes de* $(a_\lambda)_{\lambda \in \Lambda}$ son todas las expresiones de la forma

$$\kappa_n(a_{\lambda_1}^{\epsilon(1)} \cdots a_{\lambda_n}^{\epsilon(n)}),$$

para $n \in \mathbb{N}$, $\lambda_1, \dots, \lambda_n \in \Lambda$ y $\epsilon(1), \dots, \epsilon(k) \in \{1, *\}$.

2. Si sólo tenemos una variable aleatoria, a , escribimos

$$\kappa_n^a := \kappa_n(a, \dots, a)$$

y decimos que $(\kappa_n^a)_{n \geq 1}$ son los cumulantes (booleanos, clásicos o libres) de a .

3. Por lo general es bastante clara la diferencia cuando estamos trabajando con una sola variable o con muchas, así que a partir de ahora sólo escribiremos *cumulantes* para referirnos a los cumulantes multivariados.

La maquinaria que desarrollamos sobre funciones multiplicativas e inversión de Möbius, es válida para cuando trabajamos con los cumulantes, $(\kappa_n)_{n \geq 1}$, de una variable aleatoria a . Sin embargo, también podemos observar que toda esta maquinaria continúa siendo válida en el contexto más general de familias multiplicativas de funcionales. Así que, a partir de los resultados que obtuvimos en el capítulo anterior, podemos obtener los siguientes resultados básicos sobre cumulantes multivariados (que en particular se valen para cuando sólo tenemos una variable).

Proposición 3.2.4. Dada $L = \mathcal{I}, \mathcal{P}$ o \mathcal{NC} . Tomamos $(\kappa_\pi)_{\pi \in L}$ la familia de cumulantes booleanos, clásicos o libres, respectivamente. Entonces tenemos que:

1. La familia de cumulantes $(\kappa_\pi)_{\pi \in L}$ es una familia multiplicativa de funcionales, es decir,

$$\kappa_\pi[a_1, \dots, a_n] = \prod_{V \in \pi} \kappa(V)[a_1, \dots, a_n].$$

2. En particular, toda la información sobre los cumulantes está contenida en la sucesión de cumulantes $(\kappa_n)_{n \in \mathbb{N}}$ donde $\kappa_n = \kappa_{1_n}$. De hecho, las definiciones en 3.2.2 se pueden reducir a que los cumulantes son la familia multiplicativa de funcionales tal que para todo $n \in \mathbb{N}$ y todo $a_1, \dots, a_n \in \mathcal{A}$ tenemos que

$$\kappa_n(a_1, \dots, a_n) = \sum_{\sigma \in L(n)} \phi_\sigma[a_1 \cdots a_n] \mu_{L(n)}(\sigma, 1_n).$$

3. Las definiciones de cumulantes en 3.2.2 son equivalentes a decir que los cumulantes $(\kappa_\pi)_{\pi \in L}$ son la familia multiplicativa de funcionales que cumple que para toda $n \in \mathbb{N}$ y cualesquiera $a_1, \dots, a_n \in \mathcal{A}$ tenemos que:

$$\phi(a_1 \cdots a_n) = \sum_{\pi \in L(n)} \kappa_\pi[a_1, \dots, a_n].$$

Demostración. Por la Proposición 2.6.8 sabemos que la multiplicatividad de ϕ y la multiplicatividad de κ son equivalentes. Y la inversión de Möbius nos da la equivalencia entre las fórmulas en el inciso 2 y el inciso 3. \square

Las ecuaciones de arriba se conocen como las *fórmulas entre momentos y cumulantes*.

Observación 3.2.5. Todo lo que vimos en la Proposición 3.2.4 sigue siendo válido cuando hablamos específicamente de los cumulantes de una sucesión de variables aleatorias $(a_\lambda)_{\lambda \in \Lambda}$ (ver Notación 3.2.3). Por ejemplo, podemos observar que la familia de cumulantes de $(a_\lambda)_{\lambda \in \Lambda}$ contiene la misma información que la familia de momentos mixtos de $(a_\lambda)_{\lambda \in \Lambda}$.

3.3. Relación entre ambas definiciones de cumulantes

En las secciones pasadas, definimos a los cumulantes booleanos, clásicos y libres de dos maneras distintas, en esta sección veremos que ambas definiciones coinciden. Para ello, veremos que los cumulantes $(\kappa_n^a)_{n \in \mathbb{N}}$ de la Notación 3.2.3 efectivamente cumplen con las tres propiedades dadas en la Definición 3.1.2. Las primeras dos propiedades son prácticamente directas.

Lema 3.3.1. Dada $L = \mathcal{I}, \mathcal{P}$ ó \mathcal{NC} . Sea a una variable en un espacio de probabilidad con sucesión de momentos $(m_n^a)_{n \in \mathbb{N}}$ y sea $(\kappa_n^a)_{n \in \mathbb{N}}$ su sucesión de cumulantes booleanos, clásicos o libres, respectivamente. Entonces tenemos que:

1. Para cualquier n , existe un polinomio Q_n de $n - 1$ variables tal que

$$m_n^a = \kappa_n^a + Q_n(\kappa_1^a, \dots, \kappa_{n-1}^a).$$

2. Para cualquier $\lambda > 0$ y cualquier n , se tiene $\kappa_n^{\lambda a} = \lambda^n \kappa_n^a$.

Demostración. 1. Por las fórmulas entre momentos y cumulantes tenemos que

$$m_n^a = \phi(a^n) = \sum_{\pi \in L(n)} \kappa_\pi^a = \kappa_n^a + \sum_{\substack{\pi \in L(n) \\ \pi \neq 1_n}} \kappa_\pi^a.$$

Y es claro que la segunda suma se puede ver como un polinomio Q_n de $n - 1$ variables, que toma cómo valores los primeros $n - 1$ cumulantes.

2. Tomamos cualquier $\lambda > 0$ y cualquier n , de nuevo por las fórmulas entre momentos y cumulantes tenemos que

$$\begin{aligned} \kappa_n^{\lambda a} &= \sum_{\sigma \in L(n)} \phi_\sigma(\lambda a, \dots, \lambda a) \mu_{L(n)}(\sigma, 1_n) \\ &= \lambda^n \sum_{\sigma \in L(n)} \phi_\sigma(a, \dots, a) \mu_{L(n)}(\sigma, 1_n) \\ &= \lambda^n \kappa_n^a. \end{aligned}$$

Donde la segunda igualdad se debe a la multilinealidad del funcional: $\phi_\sigma^{\lambda a, \dots, \lambda a} = \lambda^n \phi_\sigma(a, \dots, a)$.

□

En lugar de demostrar la propiedad 3, veremos que cumplen con una propiedad más fuerte:

Propiedad 3'. Aditividad con respecto a la independencia: si a y b son variables aleatorias independientes, entonces $\kappa_n^{a+b} = \kappa_n^a + \kappa_n^b$.

Para ver esta propiedad, primero demostraremos un teorema importante de cumulantes multivariados, que nos dice que la independencia es equivalente a que los cumulantes mixtos se anulen. Ya que tengamos este teorema, la propiedad 3' será consecuencia directa. La demostración de este teorema se simplifica una vez que conocemos el teorema de cumulantes de productos como argumentos.

3.3.1. Cumulantes de productos como argumentos

El comportamiento de los cumulantes con respecto a la estructura lineal de \mathcal{A} es bastante claro debido a que los cumulantes son funcionales multilineales. Sin embargo, también existe una relación entre los cumulantes y la estructura multiplicativa del álgebra. Una de las más importantes es conocida como la fórmula para productos como argumentos.

La formulación y demostración del teorema están basados en la que se da para el caso libre en el libro de Nica y Speicher [NS06]. Sin embargo le daremos un enfoque un poco más general para poder realizar la demostración en paralelo del caso booleano, libre y clásico.

Notación 3.3.2. Dados dos números naturales fijos $m < n$, y una partición $\omega = \{V_1, \dots, V_m\}$ en $\mathcal{P}(n)$. Definimos la inclusión de $\mathcal{P}(m)$ en $\mathcal{P}(n)$ dada por ω ,

$$\begin{aligned} \hat{\cdot} : \mathcal{P}(m) &\rightarrow \mathcal{P}(n) \\ \tau &\mapsto \hat{\tau}, \end{aligned}$$

de tal forma que $\hat{\tau} = f^{-1} \circ \tau$, donde $f : \{1, \dots, n\} \rightarrow \{1, \dots, m\}$ se define como $f(l) = k$ para todo $l \in V_k$.

También podemos ver a $\hat{\tau}$ como la partición que obtenemos a partir de τ , reemplazando cada $j \in \{1, \dots, m\}$ por $V_j \subset \{1, \dots, n\}$. Es decir, dados $a, b \in \{1, \dots, n\}$ con $a \in V_k$ y $b \in V_l$, tenemos que $a \sim_{\hat{\tau}} b$ si y sólo si $k \sim_{\tau} l$.

Observación 3.3.3. 1. Para el caso particular en el que $\omega \in \mathcal{I}(n)$, es sencillo observar que $\widehat{\mathcal{I}(m)} = \mathcal{I}(n)$ y que $\widehat{\mathcal{NC}(m)} = \mathcal{NC}(n)$. Es decir, la inclusión es cerrada en \mathcal{I} ó \mathcal{NC} y entonces podemos definir de manera análoga la inclusión para \mathcal{I} ó \mathcal{NC} en lugar de \mathcal{P} si nos restringimos a $\omega \in \mathcal{I}(n)$.

2. A partir de la definición anterior es importante resaltar que la función $\tau \mapsto \hat{\tau}$ es inyectiva y tenemos que

$$\begin{aligned} \hat{1}_m &= 1_n \\ \hat{0}_m &= \omega \in \mathcal{P}(m) \end{aligned}$$

Además, para $L = \mathcal{I}$, \mathcal{P} ó \mathcal{NC} , esta función preserva el orden parcial ($\sigma \leq \pi$ implica que $\hat{\sigma} \leq \hat{\pi}$) y la imagen de $L(m)$ es

$$\widehat{L(m)} = [\hat{0}_m, \hat{1}_m] = [\omega, 1_n].$$

En resumen, la función $\tau \mapsto \hat{\tau}$ es un isomorfismo de retículas entre $L(m)$ y $[\omega, 1_n] \subset L(n)$.

3. Como el valor $\mu(\sigma, \pi)$ de la función de Möbius depende exclusivamente del intervalo $[\sigma, \pi]$, el isomorfismo de retículas entre $[\sigma, \pi] \subset L(m)$ y $[\hat{\sigma}, \hat{\pi}] \subset L(n)$ implica que para todo $\sigma, \pi \in L(m)$ tenemos que $\mu(\sigma, \pi) = \mu(\hat{\sigma}, \hat{\pi})$.

Teorema 3.3.4 (Cumulantes clásicos de productos como argumentos). Sea (\mathcal{A}, ϕ) un espacio de probabilidad no conmutativo y sean $(\kappa_n)_{n \in \mathbb{N}}$ los correspondientes cumulantes clásicos. Dados dos números naturales fijos $m < n$ y una partición $\omega = \{V_1, \dots, V_m\}$ en $\mathcal{P}(n)$, consideremos la inclusión $\tau \mapsto \hat{\tau}$, definida anteriormente. Sean $a_1, \dots, a_n \in \mathcal{A}$ variables aleatorias, para $1 \leq k \leq m$, denotaremos por $A_k := a_{j_1} \cdots a_{j_s}$, suponiendo que $V_k = \{j_1, \dots, j_s\}$ con $j_1 < \cdots < j_s$.

1. Para cualquier partición $\tau \in \mathcal{P}(m)$ tenemos la siguiente ecuación:

$$\kappa_\tau[A_1, A_2, \dots, A_m] = \sum_{\substack{\pi \in \mathcal{P}(n) \\ \pi \vee \omega = \hat{\tau}}} \kappa_\pi[a_1, \dots, a_n]$$

2. En particular, para $\tau = 1_m$ tenemos que:

$$\kappa_m(A_1, A_2, \dots, A_m) = \sum_{\substack{\pi \in \mathcal{P}(n) \\ \pi \vee \omega = 1_n}} \kappa_\pi[a_1, \dots, a_n]$$

Demostración. Por las propiedades básicas de la Observación 3.3.3, obtenemos que:

$$\begin{aligned} \kappa_\tau[A_1, \dots, A_m] &= \sum_{\substack{\pi \in \mathcal{P}(m) \\ \pi \leq \tau}} \phi_\pi[A_1, \dots, A_m] \mu_{\mathcal{P}(m)}(\pi, \tau) \\ &= \sum_{\substack{\pi \in \mathcal{P}(m) \\ \pi \leq \tau}} \phi_{\hat{\pi}}[a_1, \dots, a_n] \mu_{\mathcal{P}(n)}(\hat{\pi}, \hat{\tau}) \\ &= \sum_{\substack{\sigma \in \mathcal{P}(n) \\ \hat{0}_m \leq \sigma \leq \hat{\tau}}} \phi_\sigma[a_1, \dots, a_n] \mu_{\mathcal{P}(n)}(\sigma, \hat{\tau}) \\ &= \sum_{\substack{\pi \in \mathcal{P}(n) \\ \pi \vee \hat{0}_m = \hat{\tau}}} \kappa_\pi[a_1, \dots, a_n] \end{aligned}$$

Donde la última igualdad se obtiene directamente de la Proposición 2.3.4. \square

Ahora enunciaremos el teorema que se vale también para las retículas de particiones que no se cruzan y de intervalos.

Teorema 3.3.5 (Productos como argumentos). Sea (\mathcal{A}, ϕ) un espacio de probabilidad no conmutativo y sean $(\kappa_n)_{n \in \mathbb{N}}$ los correspondientes cumulantes booleanos o libres. Dados dos números naturales fijos $m < n$ y una partición $\omega = \{V_1, \dots, V_m\}$ en $\mathcal{I}(n)$, consideremos la inclusión $\tau \mapsto \hat{\tau}$ restringida a $L = \mathcal{I}$ ó \mathcal{NC} . Sean $a_1, \dots, a_n \in \mathcal{A}$ variables aleatorias, para $1 \leq k \leq m$, denotaremos por $A_k := a_{j_1} \cdots a_{j_s}$, suponiendo que $V_k = \{j_1, \dots, j_s\}$ con $j_1 < \cdots < j_s$.

1. Para cualquier partición $\tau \in L(m)$ tenemos la siguiente ecuación:

$$\kappa_\tau[A_1, A_2, \dots, A_m] = \sum_{\substack{\pi \in L(n) \\ \pi \vee \omega = \hat{\tau}}} \kappa_\pi[a_1, \dots, a_n]$$

2. En particular, para $\tau = 1_m$ tenemos que:

$$\kappa_m(A_1, A_2, \dots, A_m) = \sum_{\substack{\pi \in L(n) \\ \pi \vee \omega = 1_n}} \kappa_\pi[a_1, \dots, a_n]$$

Demostración. La demostración es exactamente igual a la del teorema anterior, sólo debemos cambiar \mathcal{P} por \mathcal{I} ó \mathcal{NC} . Observemos que como ω es una partición de intervalos entonces podemos escribirla de la siguiente forma, $\omega = \{(1, \dots, i_1), (i_1 + 1 \dots i_2), \dots, (i_{m-1} + 1 \dots i_m)\}$ y entonces tendríamos que $A_j := a_{i_{j-1}+1} \dots a_{i_j}$ (para $1 \leq j \leq m$). \square

Ahora, veamos un pequeño lema que nos ayudará en la demostración del teorema importante de esta sección.

Lema 3.3.6. Sea (\mathcal{A}, ϕ) un espacio de probabilidad no conmutativo, tomemos $L = \mathcal{P}$, \mathcal{I} ó \mathcal{NC} y sean $(\kappa_n)_{n \in \mathbb{N}}$ los correspondientes cumulantes clásicos, booleanos o libres, respectivamente. Consideremos una $n \geq 2$ y $a_1, \dots, a_n \in \mathcal{A}$. Si $a_n = 1$, entonces tendremos que $\kappa_n(a_1, \dots, a_n) = 0$.

Demostración. Procederemos por inducción sobre n . Para $n = 2$, la proposición es cierta ya que en los tres casos tenemos que

$$\kappa_2(a, 1) = \phi(a1) - \phi(a)\phi(1) = \phi(a) - \phi(a) = 0.$$

Ahora supongamos que ya demostramos la proposición para $k \leq n$ y veamos que es cierta para n . Tenemos que:

$$\begin{aligned} \phi(a_1 \cdots a_{n-1}) &= \phi(a_1 \cdots a_{n-1} 1) \\ &= \sum_{\pi \in L(n)} \kappa_\pi[a_1, \dots, a_{n-1}, 1] \\ &= \kappa_n(a_1, \dots, a_{n-1}, 1) + \sum_{\substack{\pi \in L(n) \\ \pi \neq 1_n}} \kappa_\pi[a_1, \dots, a_{n-1}, 1] \\ &= \kappa_n(a_1, \dots, a_{n-1}, 1) + \sum_{\sigma \in L(n-1)} \kappa_\sigma[a_1, \dots, a_{n-1}] \\ &= \kappa_n(a_1, \dots, a_{n-1}, 1) + \phi(a_1 \cdots a_{n-1}). \end{aligned}$$

Y por lo tanto $\kappa_n(a_1, \dots, a_{n-1}, 1) = 0$. La penúltima igualdad se debe a que, por hipótesis de inducción, una partición $\pi \neq 1_n$ contribuye a la suma

anterior sólo si (n) es un bloque de un elemento de π (en caso contrario, el bloque que contiene a n vale cero por tener un 1 al final y el producto se anula), entonces tendríamos que $\pi = \sigma \cup (n)$ con $\sigma \in L(n-1)$ y por lo tanto

$$\kappa_\pi[a_1, \dots, a_{n-1}, 1] = \kappa_\sigma[a_1, \dots, a_{n-1}] \kappa_1(1) = \kappa_\sigma[a_1, \dots, a_{n-1}].$$

Si hacemos lo anterior para cada $\pi \neq 1_n$ podemos concluir la penúltima igualdad. \square

De hecho, en el caso clásico y libre, el lema anterior se vale de manera más general, para cuando cualquier a_i es igual a 1. La demostración se sigue exactamente igual que la del lema anterior, sólo que ahora el 1 aparece en cualquier posición del cumulante.

Lema 3.3.7. Sea (\mathcal{A}, ϕ) un espacio de probabilidad no conmutativo, tomemos $L = \mathcal{P}$ ó \mathcal{NC} y sean $(\kappa_n)_{n \in \mathbb{N}}$ los correspondientes cumulantes clásicos o libres, respectivamente. Consideremos una $n \geq 2$ y $a_1, \dots, a_n \in \mathcal{A}$. Si existe al menos una i , $1 \leq i \leq n$, tal que $a_i = 1$, entonces tendremos que $\kappa_n(a_1, \dots, a_n) = 0$.

Nota: el teorema anterior no se vale para $L = \mathcal{I}$ debido a su estructura. Específicamente, en el paso en que separamos al bloque con el 1, debido a que los bloques son intervalos, estamos partiendo al conjunto en dos partes independientes, y entonces no es cierto que $\pi = \sigma \cup (n)$ con $\sigma \in \mathcal{I}(n-1)$.

3.3.2. Reformulación de independencia en términos de cumulantes

La razón por la que los cumulantes son tan importantes en el enfoque combinatorio es debido a que nos proporcionan una forma más sencilla para verificar independencia, como nos muestra el siguiente teorema.

Teorema 3.3.8. Sea (\mathcal{A}, ϕ) un espacio de probabilidad no conmutativo y sean $(\kappa_n)_{n \in \mathbb{N}}$ los correspondientes cumulantes booleanos, clásicos o libres (dependiendo de la independencia que tomemos). Consideremos $(\mathcal{A}_\lambda)_{\lambda \in \Lambda}$ subálgebras unitarias de \mathcal{A} . Entonces las siguientes dos proposiciones son equivalentes.

1. $(\mathcal{A}_\lambda)_{\lambda \in \Lambda}$ son independientes (en el sentido booleano, clásico o libre).
2. Para todo $n \geq 2$ y para todo $a_j \in \mathcal{A}_{\lambda_j}$ ($j = 1, \dots, n$) con $\lambda_1, \dots, \lambda_n \in \Lambda$ tenemos que si existen $1 \leq l, k \leq n$ con $\lambda_l \neq \lambda_k$ entonces se cumple que $\kappa_n(a_1, \dots, a_n) = 0$.

Demostración. $2 \Rightarrow 1$. Recordemos que para ver que las variables son independientes (en cualquiera de los tres casos), debemos considerar los $\phi(a_1 \cdots a_n)$

tales que $a_j \in \mathcal{A}_{\lambda_j}$ ($j = 1, \dots, n$) con $\lambda_1 \neq \lambda_2 \neq \dots \neq \lambda_n$. Por la fórmula entre momentos y cumulantes, tenemos que:

$$\phi(a_1 \cdots a_n) = \sum_{\pi \in L(n)} \kappa_\pi[a_1, \dots, a_n],$$

donde $L = \mathcal{P}$, \mathcal{I} ó \mathcal{NC} dependiendo de la independencia que estemos tomando. Observemos que si alguna partición π contiene un bloque, V , con dos variables de distintas subálgebras, tendremos que $\kappa(V)[a_1, \dots, a_n] = 0$ (por la hipótesis de que los cumulantes mixtos se anulan). Esto nos dice que $\kappa_\pi[a_1, \dots, a_n] = \prod_{V \in \pi} \kappa(V)[a_1, \dots, a_n] = 0$, es decir, la partición π no aporta nada a la suma. Cómo tenemos que $\lambda_1 \neq \lambda_2 \neq \dots \neq \lambda_n$ en particular sabemos que si un bloque contiene dos elementos consecutivos, entonces la partición no aporta nada a la suma. Después de estas observaciones generales es mejor separar en casos:

1. Caso booleano. Tenemos que

$$\phi(a_1 \cdots a_n) = \sum_{\pi \in \mathcal{I}(n)} \kappa_\pi[a_1, \dots, a_n] = \kappa_{0_n}[a_1, \dots, a_n],$$

ya que cualquier bloque con más de un elemento es un intervalo y por lo tanto apareará dos elementos consecutivos. Para concluir sólo hace falta observar que

$$\kappa_{0_n}[a_1, \dots, a_n] = \kappa_1(a_1)\kappa_1(a_2) \cdots \kappa_1(a_n) = \phi(a_1)\phi(a_2) \cdots \phi(a_n).$$

Por lo tanto $\phi(a_1 \cdots a_n) = \phi(a_1)\phi(a_2) \cdots \phi(a_n)$ y eso nos dice que las $(\mathcal{A}_\lambda)_{\lambda \in \Lambda}$ son independientes en el sentido booleano.

2. Caso clásico. Tenemos dos subcasos:

- Si $\lambda_1 \neq \lambda_r$ para todo $2 \leq r \leq n$ entonces tenemos que

$$\phi(a_1 \cdots a_n) = \sum_{\pi \in \mathcal{P}(n)} \kappa_\pi[a_1, \dots, a_n] = \sum_{\substack{\pi \in \mathcal{P}(n) \\ \pi \leq \sigma}} \kappa_\pi[a_1, \dots, a_n],$$

donde $\sigma = \{(1), (2, 3, \dots, n)\}$, (las demás particiones se anulan por que contienen dos variables de distintas subálgebras). Además es sencillo observar que

$$\{\pi \in \mathcal{P}(n) \mid \pi \leq \sigma\} \cong \mathcal{P}(1) \times \mathcal{P}(n-1)$$

y por lo tanto tenemos que

$$\sum_{\substack{\pi \in \mathcal{P}(n) \\ \pi \leq \sigma}} \kappa_\pi[a_1, \dots, a_n] = \kappa_1(a_1) \left(\sum_{\pi \in \mathcal{P}(n-1)} \kappa_\pi[a_2, \dots, a_n] \right) = \phi(a_1)\phi(a_2 \cdots a_n).$$

- Si r es el menor número tal que $\lambda_1 = \lambda_r$, entonces tenemos que

$$\phi(a_1 \cdots a_n) = \sum_{\pi \in \mathcal{P}(n)} \kappa_\pi[a_1, \dots, a_n] = \sum_{\substack{\pi \in \mathcal{P}(n) \\ \pi \leq \sigma}} \kappa_\pi[a_1, \dots, a_n],$$

dónde $\sigma = \{\{1, r\}, \{2, \dots, r-1, r+1, \dots, n\}\}$, (las demás particiones se anulan por que contienen dos variables de distintas subálgebras). Ahora, tomamos $\pi \leq \sigma$ y sean $\pi', \sigma' \in \mathcal{P}(n-1)$ las particiones que se obtienen de quitarle a π y σ el elemento 1, respectivamente. Podemos observar que

$$\begin{aligned} \kappa_\pi[a_1, \dots, a_n] &= \prod_{V \in \pi} \kappa(V)[a_1, \dots, a_n] \\ &= \prod_{V \in \pi'} \kappa(V)[a_2, \dots, a_{r-1}, a_1 a_r, a_{r+1}, \dots, a_n] \\ &= \kappa_{\pi'}[a_2, \dots, a_{r-1}, a_1 a_r, a_{r+1}, \dots, a_n]. \end{aligned}$$

Por lo tanto tenemos que

$$\begin{aligned} \phi(a_1 \cdots a_n) &= \sum_{\substack{\pi \in \mathcal{P}(n) \\ \pi \leq \sigma}} \kappa_\pi[a_1, \dots, a_n] \\ &= \sum_{\substack{\pi \in \mathcal{P}(n-1) \\ \pi \leq \sigma'}} \kappa_\pi[a_2, \dots, a_{r-1}, a_1 a_r, a_{r+1}, \dots, a_n] \\ &= \phi(a_2 \cdots a_{r-1} (a_1 a_r) a_{r+1} \cdots a_n). \end{aligned}$$

En ambos casos $\phi(a_1 \cdots a_n)$ se factoriza como queremos y concluimos que las $(\mathcal{A}_\lambda)_{\lambda \in \Lambda}$ son independientes en el sentido clásico.

3. Caso libre. Queremos ver que para cuando todas las variables aleatorias son centradas ($\phi(a_j) = 0$ para todo $j = 1, \dots, n$), se tiene que $\phi(a_1 \cdots a_n) = 0$. Por la Observación 2.5.11 sabemos que cualquier partición que no se cruza, π , tiene al menos un bloque que es un intervalo. Si el intervalo es de tamaño 1 entonces $\kappa_\pi[a_1, \dots, a_n] = 0$ por la condición de ser centradas. Si el intervalo es de tamaño mayor a 1, entonces apareja dos elementos consecutivos y de nuevo $\kappa_\pi[a_1, \dots, a_n] = 0$. Concluimos que

$$\phi(a_1 \cdots a_n) = \sum_{\pi \in \mathcal{NC}(n)} \kappa_\pi[a_1, \dots, a_n] = 0,$$

y eso nos dice que las $(\mathcal{A}_\lambda)_{\lambda \in \Lambda}$ son independientes en el sentido libre.

1 \Rightarrow 2. Para esta implicación será más sencillo hacer cada caso por separado, sin embargo, en los tres casos seguiremos la misma línea. Primero

centraremos las variables en los cumulantes sin cambiar su valor, luego veremos que ciertos cumulantes mixtos valen cero y por último utilizamos inducción para ver que cualquier cumulante mixto vale cero.

1. Caso booleano. Por la multilinealidad de los cumulantes y el Lema 3.3.6 podemos centrar la última variable sin cambiar el valor del cumulante, es decir:

$$\begin{aligned}\kappa_n(a_1, \dots, a_{n-1}, a_n - \phi(a_n)1) &= \kappa_n(a_1, \dots, a_{n-1}, a_n) - \kappa_n(a_1, \dots, a_{n-1}, \phi(a_n)1) \\ &= \kappa_n(a_1, \dots, a_n).\end{aligned}$$

Y ahora sabemos que $\phi(a_n - \phi(a_n)1) = \phi(a_n) - \phi(a_n) = 0$. Por lo tanto, sólo tenemos que demostrar la proposición para el caso $\phi(a_n) = 0$.

Ahora veamos que $\kappa_n(a_1, \dots, a_n) = 0$ si las variables son alternantes, es decir, si tenemos que $\lambda_1 \neq \lambda_2 \neq \dots \neq \lambda_n$. Por la fórmula entre momentos y cumulantes sabemos que

$$\kappa_n(a_1, \dots, a_n) = \sum_{\sigma \in \mathcal{I}(n)} \phi_\sigma[a_1 \cdots a_n] \mu_{\mathcal{I}(n)}(\sigma, 1_n).$$

Para cualquier $\sigma \in \mathcal{I}(n)$ tenemos que sus bloques son intervalos y como las variables son alternantes, también lo serán en cada bloque. Como estamos suponiendo independencia booleana, eso nos dice que cada bloque de σ se puede factorizar y por lo tanto tendremos que $\phi_\sigma[a_1 \cdots a_n] = \phi(a_1) \cdots \phi(a_n)$ para todo $\sigma \in \mathcal{I}(n)$. Como podemos suponer que $\phi(a_n) = 0$, concluimos que todas las particiones se anulan y por lo tanto $\kappa_n(a_1, \dots, a_n) = 0$.

Para concluir, utilicemos inducción sobre la longitud de los cumulantes. Base: para $n = 2$ y a_1, a_2 libres, es claro que $\phi(a_1 a_2) = \phi(a_1)\phi(a_2)$ y por lo tanto

$$\kappa_2(a_1, a_2) = \phi(a_1 a_2) - \phi(a_1)\phi(a_2) = 0.$$

Ahora tomemos $n \geq 3$ y supongamos que ya demostramos el inciso 2 para todo κ_r con $r < n$. Queremos demostrar que $\kappa_n(a_1, \dots, a_n) = 0$. Ya lo hicimos para cuando $\lambda_1 \neq \lambda_2 \neq \dots \neq \lambda_n$, así que el único problema sería si tenemos variables consecutivas que pertenezcan a la misma subálgebra. Para lidiar con ello, vamos a pegar estas variables consecutivas usando el teorema de productos como argumentos. Escribimos $a_1 \cdots a_n = A_1 \cdots A_m$ de manera que las A consecutivas provengan de diferentes subálgebras. Por el Teorema 3.3.5, tenemos que

$$\begin{aligned}\kappa_m(A_1, \dots, A_m) &= \sum_{\substack{\pi \in \mathcal{I}(n) \\ \pi \vee \omega = 1_n}} \kappa_\pi[a_1, \dots, a_n] \\ &= \kappa_n(a_1, \dots, a_n) + \sum_{\substack{\pi \in \mathcal{I}(n), \pi \neq 1_n \\ \pi \vee \omega = 1_n}} \kappa_\pi[a_1, \dots, a_n]\end{aligned}$$

donde $\omega \in \mathcal{I}(n)$ está codificando la información de qué elementos a_j consecutivos pertenecen a la misma subálgebra. Como las A son alternantes, ya sabemos que $\kappa_m(A_1, \dots, A_m) = 0$. Además, para $\pi \neq 1_n$, el término $\kappa_\pi[a_1, \dots, a_n]$ es producto de cumulantes con longitud menor a n . Entonces, por hipótesis de inducción sabemos que $\kappa_\pi[a_1, \dots, a_n]$ es distinto de cero sólo si cada bloque de π agrupa elementos de la misma subálgebra. Por cómo lo definimos, sabemos que ω también agrupa elementos de la misma subálgebra y esto implica que los bloques de $\pi \vee \omega$ también agrupan elementos de la misma subálgebra. Por lo tanto, $\pi \vee \omega \neq 1_n$ ya que estamos suponiendo que el cumulante es mixto, es decir, tiene al menos dos elementos de subálgebras distintas. Entonces, tenemos que $\kappa_\pi[a_1, \dots, a_n] = 0$, para todo $\pi \in \mathcal{I}(n)$ tal que $\pi \neq 1_n$ y $\pi \vee \omega = 1_n$, lo cuál quiere decir que todos los sumandos de arriba se anulan. Con esto terminamos la inducción y concluimos que $\kappa_n(a_1, \dots, a_n) = 0$ para cualquier cumulante mixto.

2. Caso libre. A diferencia del caso booleano, ahora tenemos el Lema 3.3.7 que es más general y nos permite centrar todas las variables sin cambiar el valor del cumulante, es decir:

$$\kappa_n(a_1 - \phi(a_1)1, a_2 - \phi(a_2)1, \dots, a_n - \phi(a_n)1) = \kappa_n(a_1, \dots, a_n).$$

Por lo tanto, podemos suponer que todas las variables son centradas. Ahora veamos que $\kappa_n(a_1, \dots, a_n) = 0$ si las variables son alternantes. Por la fórmula entre momentos y cumulantes tenemos que

$$\kappa_n(a_1, \dots, a_n) = \sum_{\sigma \in \mathcal{NC}(n)} \phi_\sigma[a_1 \cdots a_n] \mu_{\mathcal{NC}(n)}(\sigma, 1_n).$$

Recordemos que cualquier $\sigma \in \mathcal{NC}(n)$ tiene al menos un bloque, V , que es un intervalo (Observación 2.5.11), y como las variables son alternantes, entonces también lo serán en V . Además, como las variables son centradas, la regla de independencia libre nos dice que $\phi(V)[a_1 \cdots a_n] = 0$, y por lo tanto $\phi_\sigma[a_1 \cdots a_n] = 0$ para todo σ . Esto nos dice que la suma anterior es cero.

Para concluir, sólo hace falta hacer exactamente la misma inducción que utilizamos en el caso booleano.

3. Caso clásico. Al igual que el caso libre, podemos utilizar el Lema 3.3.7 para ver que

$$\kappa_n(a_1 - \phi(a_1)1, a_2 - \phi(a_2)1, \dots, a_n - \phi(a_n)1) = \kappa_n(a_1, \dots, a_n).$$

Por lo tanto, podemos suponer que todas las variables son centradas. Ahora veamos que $\kappa_n(a_1, \dots, a_n) = 0$ si todas las variables provienen de distintas subálgebras, (en los dos casos anteriores sólo pedíamos

que las variables consecutivas provinieran de distintas subálgebras, la razón por la que ahora necesitamos que todas provengan de distintas subálgebras es que al trabajar con todas las particiones, no necesariamente tendremos un bloque que sea un intervalo, y eso era lo que nos aseguraba que podíamos encontrar un bloque alternante). Por la fórmula entre momentos y cumulantes tenemos que

$$\kappa_n(a_1, \dots, a_n) = \sum_{\sigma \in \mathcal{P}(n)} \phi_\sigma[a_1 \cdots a_n] \mu_{\mathcal{P}(n)}(\sigma, 1_n).$$

Como estamos suponiendo independencia clásica y que todas las variables provienen de distintas subálgebras, entonces para cualquier $\sigma \in \mathcal{P}(n)$ tendremos que todos sus bloques se pueden factorizar (utilizando la primer regla de la independencia). Factorizando varias veces cada bloque, obtenemos que $\phi_\sigma[a_1 \cdots a_n] = \phi(a_1) \cdots \phi(a_n)$ para toda $\sigma \in \mathcal{P}(n)$. Como las variables son centradas, concluimos que todos las particiones se anulan y por lo tanto $\kappa_n(a_1, \dots, a_n) = 0$.

Para concluir, utilizaremos una inducción análoga a la de los dos casos anteriores, sólo que ahora necesitamos el teorema más general de productos como argumentos. El caso base es el mismo. Para el paso inductivo tomemos $n \geq 3$ y supongamos que ya demostramos el inciso 2 para todo κ_r con $r < n$. Queremos demostrar que $\kappa_n(a_1, \dots, a_n) = 0$. Ya lo hicimos para cuando $\lambda_1, \lambda_2, \dots, \lambda_n$ son todos distintos, así que el único problema sería si tenemos más de una variable de la misma subálgebra. Tomamos la partición $\omega = \{V_1, \dots, V_n\}$ en $\mathcal{P}(n)$, que agrupa a los elementos de la misma subálgebra, es decir, tal que $k \sim_\omega l$ si y sólo si $\lambda_k = \lambda_l$. Por el Teorema 3.3.4, tenemos que

$$\kappa_m(A_1, \dots, A_m) = \kappa_n(a_1, \dots, a_n) + \sum_{\substack{\pi \in \mathcal{P}(n), \pi \neq 1_n \\ \pi \vee \omega = 1_n}} \kappa_\pi[a_1, \dots, a_n]$$

Como las A_j provienen de distintas subálgebras, entonces ya sabemos que $\kappa_m(A_1, \dots, A_m) = 0$. Y por el mismo razonamiento que usamos en el caso booleano, podemos llegar a que $\kappa_\pi[a_1, \dots, a_n] = 0$, para todo $\pi \in \mathcal{P}(n)$ tal que $\pi \neq 1_n$ y $\pi \vee \omega = 1_n$, es decir, que todos los sumandos de arriba se anulan. Con esto terminamos la inducción y concluimos que $\kappa_n(a_1, \dots, a_n) = 0$ para cualquier cumulante mixto.

Entonces la segunda implicación se vale para los tres casos y con esto terminamos la demostración del teorema. \square

Hasta ahora, hemos relacionado a la independencia de subálgebras (Definición 1.3.1) con los cumulantes multivariados (Definición 3.2.2). Sin embargo, nos gustaría ver si podemos decir algo más para cuando tenemos

independencia de variables aleatorias (Definición 1.3.3). Resulta que podemos dar una equivalencia aún más fuerte usando los cumulantes de variables aleatorias (Notación 3.2.4), cómo lo muestra el siguiente teorema.

Teorema 3.3.9. Sea (\mathcal{A}, ϕ) un espacio de probabilidad no conmutativo y sean $(\kappa_n)_{n \in \mathbb{N}}$ los correspondientes cumulantes booleanos, clásicos o libres (dependiendo de la independencia que tomemos). Consideremos $(a_\lambda)_{\lambda \in \Lambda}$ variables aleatorias de \mathcal{A} . Entonces las siguientes dos proposiciones son equivalentes.

1. $(a_\lambda)_{\lambda \in \Lambda}$ son independientes (en el sentido booleano, clásico o libre).
2. Para todo $n \geq 2$ y para toda $\lambda_1, \dots, \lambda_n \in \Lambda$ tenemos que si existen $1 \leq l, k \leq n$ con $\lambda_l \neq \lambda_k$ entonces se cumple que $\kappa_n(a_{\lambda_1}, \dots, a_{\lambda_n}) = 0$.

Demostración. Por definición, las $(a_\lambda)_{\lambda \in \Lambda}$ son independientes si las $*$ -subálgebras generadas por las variables a_λ son independientes. Por el Teorema 3.3.8, lo anterior es equivalente a que los cumulantes mixtos se anulen. Así que lo que queremos demostrar se reduce a ver que las siguientes dos proposiciones son equivalentes:

1. Los cumulantes mixtos formados con cualquier elemento de las $*$ -subálgebras generadas por las variables a_λ se anulan.
2. Los cumulantes mixtos formados sólo con las variables a_λ se anulan.

Claramente 1 implica 2 de forma directa. Lo interesante es precisamente ver que el inciso 2 es suficiente. Sea \mathcal{A}_λ la $*$ -álgebra unitaria generadas por el elemento a_λ y consideremos elementos $b_j \in \mathcal{A}_{\lambda_j}$. Podemos ver a b_j como un polinomio en a_{λ_j} , entonces, por la multilinealidad de los cumulantes podemos expresar a cualquier $\kappa_m(b_1, \dots, b_m)$ como suma de cumulantes de la forma $\alpha \kappa_m(b'_1, \dots, b'_m)$ en dónde cada b'_j es una potencia de a_{λ_j} y α es una escalar. Así que sólo hace falta ver que $\kappa_m(b_1, \dots, b_m) = 0$ dónde cada b_j es una potencia de a_{λ_j} . Esto es bastante similar a un paso en la demostración anterior y podemos proceder de la misma manera. Primero escribimos a $b_1 \cdots b_m$ como $a_{\lambda_1} \cdots a_{\lambda_n}$ y utilizamos el teorema productos como argumentos (3.3.5 y 3.3.4, respectivamente) para ver que

$$\kappa_m(b_1, \dots, b_m) = \sum_{\substack{\pi \in L(n) \\ \pi \vee \omega = 1_n}} \kappa_\pi[a_{\lambda_1}, \dots, a_{\lambda_n}]$$

dónde $\omega \in \mathcal{I}(n)$ está codificando la información de qué elementos a_{λ_j} consecutivos son iguales (pertenecen a la misma subálgebra). Sabemos que $\kappa_\pi[a_{\lambda_1}, \dots, a_{\lambda_n}]$ es distinto de cero sólo si cada bloque de π agrupa elementos iguales. Y por cómo lo definimos, sabemos que ω también agrupa elementos iguales y esto implica que los bloques de $\pi \vee \omega$ también agrupan

elementos iguales. Por lo tanto $\pi \vee \omega \neq 1_n$ ya que estamos suponiendo que el cumulante $\kappa_m(b_1, \dots, b_m)$ es mixto, es decir, tiene al menos dos elementos de subálgebras distintas. Y concluimos que todos los sumandos de arriba se anulan. \square

Ya que tenemos este teorema, la propiedad 3' se sigue de forma directa.

Proposición 3.3.10. Sean a y b variables aleatorias independientes (en el sentido booleano, clásico o libre) en algún espacio de probabilidad y sean $(\kappa_n^a)_{n \geq 1}$ y $(\kappa_n^b)_{n \geq 1}$ los cumulantes (booleanos, clásicos o libres) de a y b respectivamente. Entonces tenemos que

$$\kappa_n^{a+b} = \kappa_n^a + \kappa_n^b \quad \forall n \geq 1.$$

Demostración. Tenemos que

$$\begin{aligned} \kappa_n^{a+b} &= \kappa_n(a+b, \dots, a+b) \\ &= \kappa_n(a, \dots, a) + \kappa_n(a, \dots, a) \\ &= \kappa_n^a + \kappa_n^b, \end{aligned}$$

ya que los cumulantes que tienen como argumento una a y una b al mismo tiempo se anulan por el Teorema 3.3.9. \square

3.4. Cumulantes monótonos

Si observamos, todo lo que desarrollamos en el Capítulo 2 fue para definir a los cumulantes de una forma que cumplieran directamente con la propiedad 1 y 2, y gran parte de la sección anterior (que también está basada en el capítulo 2) fue corroborar que esa definición de cumulantes efectivamente cumplía la propiedad 3. A los cumulantes monótonos también podemos relacionarlos con los momentos utilizando retículas de particiones, pero estas retículas son un poco distintas. El problema con estas retículas es que no cumplen con ser multiplicativas, que es algo bastante importante en los otros tres casos. El que no sean multiplicativas tiene que ver con que los cumulantes monótonos no cumplen con la propiedad 3'. Así, que este caso lo abordaremos de manera distinta, seguiremos el artículo [HS11b]. La idea es definir los cumulantes de manera que cumplan directamente con la propiedad 2 y 3, y posteriormente veremos que también cumplen con la propiedad 1. Antes de dar la definición de cumulantes monótonos debemos observar unas cosas.

Proposición 3.4.1. Si a y b son variables aleatorias monótonamente independientes, entonces se cumple que:

$$\begin{aligned} m_n^{a+b} &= \sum_{k=0}^n \sum_{\substack{j_0+\dots+j_k=n-k \\ j_0 \geq 0, \dots, j_k \geq 0}} m_k^a m_{j_0}^b \cdots m_{j_k}^b \\ &= m_n^a + m_n^b + \sum_{k=1}^{n-1} \sum_{\substack{j_0+\dots+j_k=n-k \\ j_0 \geq 0, \dots, j_k \geq 0}} m_k^a m_{j_0}^b \cdots m_{j_k}^b. \end{aligned} \quad (3.1)$$

Demostración. Tomamos la expansión

$$(a+b)^n = a^n + b^n + \sum_{k=1}^{n-1} \sum_{\substack{j_0+\dots+j_k=n-k \\ j_0 \geq 0, \dots, j_k \geq 0}} b^{j_0} a b^{j_1} a \cdots a b^{j_k}$$

y aplicamos ϕ de ambos lados. Por la definición de independencia monótona, llegamos a la ecuación buscada (por la regla que nos da la independencia monótona, primero debemos factorizar los b^{j_i} y al final nos quedarán todas las a juntas). \square

Podemos observar que la suma de la segunda igualdad en la proposición anterior se puede expresar en términos de los primeros $n-1$ momentos de a y de los primeros $n-1$ momentos de b , lo cual nos lleva al siguiente resultado.

Corolario 3.4.2. Para cualquier $n \geq 1$ existe un polinomio P_n en $2n-2$ variables tal que se cumple

$$m_n^{a+b} = m_n^a + m_n^b + P_n(m_1^a, \dots, m_{n-1}^a, m_1^b, \dots, m_{n-1}^b)$$

si a y b son monótonamente independientes.

Si aplicamos el resultado anterior varias veces para variables aleatorias idénticamente distribuidas podemos obtener el siguiente resultado.

Proposición 3.4.3. $m_n^{N \cdot a}$ es un polinomio en N (sin término constante) para cualquier $n \geq 0$.

Demostración. Procedemos por inducción sobre n . Para $n=1$ es trivial por la linealidad de ϕ . Supongamos que la proposición es cierta para $n < l$. Por el corolario anterior, sabemos que

$$m_l^{N \cdot a} - m_l^{(N-1) \cdot a} = m_l^a + P_l(m_1^a, \dots, m_{l-1}^a, m_1^{(N-1) \cdot a}, \dots, m_{l-1}^{(N-1) \cdot a}). \quad (3.2)$$

Por hipótesis de inducción, P_l es un polinomio en N . Entonces $m_l^{N \cdot a} - m_l^{(N-1) \cdot a}$ es un polinomio en N , si sumamos los polinomios $m_l^{M \cdot a} - m_l^{(M-1) \cdot a}$ para $M = 1, \dots, N$ y analizamos cada coeficiente, es sencillo concluir que $m_l^{N \cdot a}$ es un polinomio en N . Además, es claro que no tiene término constante porque $m_l^{0 \cdot a} = 0$. \square

Gracias a la proposición anterior, ahora podemos pensar $m_n^{N \cdot a}$ como un polinomio sobre todos los reales y no sólo en los enteros, es decir, para cada variable aleatoria a podemos definir el polinomio $M(t) = m_n^{t \cdot a}$ reemplazando N con $t \in \mathbb{R}$. Entonces ahora tenemos un polinomio con respecto a t que cumple que $M_n(1) = m_n^a$. Además, de la ecuación (3.1) obtenemos que

$$M_n(t+s) = M_n(t) + M_n(s) + \sum_{k=1}^{n-1} \sum_{\substack{j_0+\dots+j_k=n-k \\ j_0 \geq 0, \dots, j_k \geq 0}} M_k(t) M_{j_0}(s) \cdots M_{j_k}(s).$$

Después de estas observaciones ya estamos listos para definir los cumulantes monótonos.

Definición 3.4.4. Sea $h_n = h_n^a$ el coeficiente lineal del polinomio $M_n(t)$ (el coeficiente de N en $m_n^{N \cdot a}$). Decimos que h_n es el n -ésimo cumulante monótono de a .

Proposición 3.4.5. Los cumulantes monótonos satisfacen las propiedades 2 y 3 de los cumulantes generalizados.

Demostración. Tomemos a, a_1, a_2, \dots variables aleatorias independientes e idénticamente distribuidas. Es fácil ver que

$$m_n^{N \cdot (\lambda a)} = m_n^{\lambda a_1 + \dots + \lambda a_N} = \lambda^n m_n^{a_1 + \dots + a_N} = \lambda^n m_n^{N \cdot a}.$$

Por lo tanto el coeficiente lineal del primer polinomio será λ^n veces el coeficiente lineal del segundo polinomio, (vistos como polinomios en N) y se sigue la propiedad 2.

Por la asociatividad de la independencia monótona tenemos que $N \cdot (M \cdot a) = M \cdot a_1 + \dots + M \cdot a_N = (NM) \cdot a$ y por lo tanto si nos fijamos en los coeficientes lineales de $m_n^{N \cdot (M \cdot a)} = m_n^{(NM) \cdot a}$ (vistos como polinomios sobre N), concluimos que $h_n^{M \cdot a} = M h_n^a$, que es precisamente la propiedad 3. \square

Para corroborar la propiedad 1 debemos expresar los momentos en términos de los cumulantes, para ello utilizaremos el siguiente lema.

Lema 3.4.6. Se cumplen las siguientes ecuaciones:

$$\frac{dM_0(t)}{dt} = 0,$$

$$\frac{dM_n(t)}{dt} = \sum_{k=1}^n k h_{n-k+1} M_{k-1}(t) \quad \text{para } n \geq 1,$$

con condiciones iniciales $M_0(0) = 1$ y $M_n(0) = 0$ para $n \geq 1$.

Demostración. Por definición, tenemos que $M_i(s) = sh_i + s^2(\dots)$, y la ecuación (3.1) nos dice que

$$M_n(t+s) - M_n(t) = M_n(s) + \sum_{k=1}^{n-1} \sum_{\substack{j_0+\dots+j_k=n-k \\ j_0 \geq 0, \dots, j_k \geq 0}} M_k(t)M_{j_0}(s) \cdots M_{j_k}(s).$$

Ahora sólo hace falta comparar el coeficiente de s en la ecuación anterior. Sabemos que

$$\frac{dM_n(t)}{dt} = \lim_{s \rightarrow 0} \frac{M_n(t+s) - M_n(t)}{s}$$

que es precisamente el coeficiente lineal de s en el polinomio $M_n(t+s) - M_n(t)$. Por otro lado, para que sólo quede una s en la suma derecha, necesitamos que sólo uno de j_0, \dots, j_k sea distinto de cero, entonces tendremos que $j_i = n-k$ y esto nos aportará una h_{n-k+1} en los $k+1$ posibles casos (qué j_i es distinta de cero). Eso nos da que el coeficiente lineal del lado derecho es

$$h_n + \sum_{k=1}^{n-1} (k+1)h_{n-k}M_k(t).$$

Haciendo el cambio de variable $k = k+1$ terminamos. □

El siguiente teorema nos proporciona una fórmula que expresa a los momentos en términos de los cumulantes.

Teorema 3.4.7. Sean $(m_n)_{n \geq 1}$ y $(h_n)_{n \geq 1}$ las sucesiones de momentos y cumulantes monótonos de una variable aleatoria, respectivamente. Entonces tenemos que

$$m_n = \sum_{k=1}^n \sum_{1=i_0 < i_1 < \dots < i_k = n+1} \frac{1}{k!} \prod_{l=1}^k i_{l-1} h_{i_l - i_{l-1}}.$$

Demostración. Si integramos las ecuaciones del lema anterior, obtenemos que

$$M_0(t) = 1,$$

$$M_n(t) = \sum_{k=1}^n kh_{n-k+1} \int_0^t M_{k-1}(s) ds \quad \text{para } n \geq 1.$$

Entonces tenemos que

$$\begin{aligned}
M_n(t) &= \sum_{k_1=1}^n k_1 h_{n-k_1+1} \int_0^t M_{k_1-1}(t_1) dt_1 \\
&= \sum_{k_1=1}^n \sum_{k_2=1}^{k_1-1} k_1 k_2 h_{n-k_1+1} h_{k_1-k_2} \int_0^t \int_0^{t_1} M_{k_2-1}(t_2) dt_2 dt_1 \\
&= \sum_{k_1=1}^n \sum_{k_2=1}^{k_1-1} \sum_{k_3=1}^{k_2-1} k_1 k_2 k_3 h_{n-k_1+1} h_{k_1-k_2} h_{k_2-k_3} \int_0^t \int_0^{t_1} \int_0^{t_2} M_{k_3-1}(t_3) dt_3 dt_2 dt_1 \\
&\vdots \\
&= \sum_{k_1=1}^n \cdots \sum_{k_{j-1}=1}^{k_{j-2}-1} \prod_{l=1}^j k_l h_{k_{l-1}-k_l} \int_0^t \cdots \int_0^{t_{j-1}} M_0(t_j) dt_j \cdots dt_1 \\
&= \sum_{k_1=1}^n \cdots \sum_{k_{j-1}=1}^{k_{j-2}-1} \prod_{l=1}^j k_l h_{k_{l-1}-k_l} \frac{t^j}{j!} \\
&= \sum_{j=1}^n \sum_{1=i_0 < i_1 < \cdots < i_j = n+1} \frac{t^j}{j!} \prod_{l=1}^j i_{l-1} h_{i_l - i_{l-1}},
\end{aligned}$$

donde $i_l := k_{n-l}$. La ecuación que buscamos es cuando tomamos $t = 1$. \square

Corolario 3.4.8. Los cumulantes monótonos $h_n = h_n^a$ satisfacen la propiedad 1.

3.4.1. Particiones monótonas

Ahora veamos que podemos escribir la última fórmula de la sección anterior de una manera un poco más bonita utilizando una forma especial de particiones.

- Definición 3.4.9.**
1. Dados dos bloques V, W de una partición, decimos que V es un bloque interno de W (o de forma equivalente, decimos que V está dentro de W , que W es un bloque exterior de V , o que W cubre a V) si existen $i, j \in W$ tales que $i < k < j$ para cada $k \in V$.
 2. La profundidad de un bloque V de una partición que no se cruza es el número de bloques (incluyendo a V) que cubren al bloque V . La profundidad de una partición que no se cruza es la mayor profundidad entre todos sus bloques.
 3. Una partición ordenada es una pareja (π, λ) compuesta por una partición π y un orden lineal λ en sus bloques. Podemos ver a una partición ordenada como una sucesión $(\pi, \lambda) = (V_1, \dots, V_k)$ dónde $V_i \leq_\lambda V_j$ si y

sólo si $i \leq j$. También podemos escribirla como $(\pi, \lambda) := (V_1 < V_2 < \dots < V_k)$. Denotaremos por $\mathcal{LP}(n)$ a las particiones ordenadas, y de manera análoga, denotaremos por $\mathcal{LNC}(n)$ y $\mathcal{LI}(n)$ al conjunto de las particiones ordenadas (π, λ) con la restricción de que $\pi \in \mathcal{NC}(n)$ ó $\pi \in \mathcal{I}(n)$, respectivamente.

4. Una partición monótona es una partición ordenada (π, λ) con $\pi \in \mathcal{NC}(n)$ tal que para $V, W \in \pi$, $V \geq_\lambda W$ si V es un bloque interior de W . Denotamos por $\mathcal{M}(n)$ al conjunto de particiones monótonas.

Observación 3.4.10. Podemos observar que $|\mathcal{LP}(n)| = \sum_{k=1}^n k! |\{\pi \in \mathcal{P}(n) : |\pi| = k\}|$. Y con esta idea podemos reescribir la fórmula entre momentos y cumulantes clásica de la siguiente manera:

$$m_n = \sum_{\pi \in \mathcal{P}(n)} c_\pi = \sum_{k=1}^n c_\pi \frac{k!}{k!} |\{\pi \in \mathcal{P}(n) : |\pi| = k\}| = \sum_{(\pi, \lambda) \in \mathcal{LP}(n)} \frac{c_\pi}{|\pi|!}.$$

De manera análoga podemos reescribir las fórmulas para el caso booleano y libre.

Resulta que esta forma de ver a las fórmulas entre momentos y cumulantes es precisamente la que se puede generalizar a el caso monótono.

Teorema 3.4.11. Sean h_n los cumulantes monótonos de una medida de probabilidad con momentos finitos de todos los órdenes. Entonces se cumple que

$$m_n = \sum_{(\pi, \lambda) \in \mathcal{M}(n)} \frac{h_\pi}{|\pi|!}.$$

Demostración. Tomemos k y n enteros positivos fijos y sean n_1, \dots, n_k enteros positivos tales que $n_1 + \dots + n_k = n$. Vamos a contar cuántas particiones monótonas cumplen que $|V_j| = n_j$ para $1 \leq j \leq k$. Es importante observar que por definición de partición monótona V_k no tiene ningún bloque en su interior y por lo tanto debe ser un intervalo, como queremos que $|V_k| = n_k$ tenemos $n - n_k + 1$ opciones para elegir el bloque V_k . Ahora nos olvidamos de ese bloque y nos fijamos en $\pi \setminus V$, por la Observación 2.5.11 sabemos que $\pi \setminus V \in \mathcal{NC}(n - n_k)$, y queremos que tenga $k - 1$ bloques y que sea monótona, (no importa el orden que le demos porque sus bloques no estarán relacionados con V_k). Entonces podemos fijarnos en V_{k-1} , de nuevo debe ser un intervalo de tamaño n_{k-1} en $\pi \setminus V \in \mathcal{NC}(n - n_k)$ y lo podemos elegir de $n - n_k - n_{k-1} + 1$ maneras. Luego borramos el bloque V_{k-1} y repetimos el proceso hasta terminar con todos los bloques. Es sencillo ver que cada proceso nos dará como resultado una partición monótona distinta tal que $|V_j| = n_j$ para $1 \leq j \leq k$. Por lo tanto, las formas posibles de elegir los bloques son $(n - n_k + 1)(n - n_k - n_{k-1} + 1) \dots (n - \sum_{j=1}^k n_j + 1) = \prod_{m=1}^k (n - \sum_{j=m}^k n_j + 1)$.

Por lo tanto, podemos separar la suma sobre las particiones monótonas dependiendo de la cantidad de bloques, k , que tengan y el tamaño de sus bloques, n_1, \dots, n_k , de esta forma obtenemos que

$$\begin{aligned} \sum_{(\pi, \lambda) \in \mathcal{M}(n)} \frac{h_\pi}{|\pi|!} &= \sum_{k=1}^n \sum_{\substack{n_1 + \dots + n_k = n \\ n_1 \geq 1, \dots, n_k \geq 1}} \frac{h_{n_1} \cdots h_{n_k}}{k!} \left(\prod_{m=1}^k (n - \sum_{j=m}^k n_j + 1) \right) \\ &= \sum_{k=1}^n \sum_{1=i_0 < i_1 < \dots < i_k = n+1} \frac{h_{i_1 - i_0} \cdots h_{i_k - i_{k-1}}}{k!} \prod_{l=1}^k i_{l-1} \end{aligned}$$

donde escribimos $i_m := n - \sum_{j=m+1}^k n_j + 1$. La última expresión es igual a m_n por el Teorema 3.4.7. \square

Siguiendo un proceso bastante similar al que seguimos a lo largo de esta sección para definir los cumulante monótonos univariados y relacionarlos con los momentos, se pueden definir los cumulantes monótonos multivariados y relacionarlos con los momentos mixtos, para ver los detalles el lector puede consultar [HS11a]. La fórmula multivariada es la siguiente.

Teorema 3.4.12. Tenemos la siguiente fórmula para expresar los momentos en términos de los cumulantes monótonos multivariados $(h_\pi)_{\pi \in \mathcal{NC}}$:

$$m_n(a_1, \dots, a_n) = \sum_{(\pi, \lambda) \in \mathcal{M}(n)} \frac{h_\pi[a_1, \dots, a_n]}{|\pi|!}.$$

3.5. Relación de cumulantes con transformadas

El objetivo de esta sección es relacionar nuestros cumulantes con las transformadas que definimos en el Capítulo 1, como las transformadas se definieron para una sola variable, sólo ocuparemos los cumulantes univariados. Primero demostraremos tres sencillos lemas, que nos simplificarán las cuentas, cada uno de los lemas corresponde a un tipo de independencia. Los enunciados son bastante parecidos y las demostraciones son análogas.

3.5.1. Fórmulas que relacionan recursiones

Lema 3.5.1 (Clásico). Sean $(a_m)_{m \in \mathbb{N}}$ y $(b_m)_{m \in \mathbb{N}}$ dos sucesiones de reales y definimos $a_0 = 1$. Entonces las siguientes fórmulas son equivalentes:

1.

$$a_k = \sum_{i=1}^k \binom{k-1}{i-1} b_i a_{k-i} \quad \forall k \in \mathbb{N}$$

2.

$$\ln \left(\sum_{j=0}^{\infty} \frac{a_j}{j!} s^j \right) = \sum_{i=1}^{\infty} \frac{b_i}{i!} s^i$$

3.

$$a_k = \sum_{\pi \in \mathcal{P}(k)} b_{\pi} \quad \forall k \in \mathbb{N}$$

Demostración. Veamos que las dos últimas fórmulas cumplen la misma recursión del primer inciso.

Para ello fijamos las b 's y nos tomamos $a''_0 = 1$ y a''_1, \dots, a''_m tales que

$$\ln \left(1 + \sum_{j=1}^{\infty} \frac{a''_j}{j!} s^j \right) = \sum_{i=1}^{\infty} \frac{b_i}{i!} s^i.$$

Si derivamos de ambos lados obtenemos que

$$\frac{\sum_{j=1}^{\infty} \frac{a''_j}{(j-1)!} s^{j-1}}{\sum_{j=0}^{\infty} \frac{a''_j}{j!} s^j} = \sum_{i=1}^{\infty} \frac{b_i}{(i-1)!} s^{i-1}.$$

Y si pasamos multiplicando concluimos que

$$\sum_{k=1}^{\infty} \frac{a''_k}{(k-1)!} s^{k-1} = \left(\sum_{i=1}^{\infty} \frac{b_i}{(i-1)!} s^{i-1} \right) \left(\sum_{j=0}^{\infty} \frac{a''_j}{j!} s^j \right) = \sum_{k=1}^{\infty} \left(\sum_{i+j=k} \frac{b_i}{(i-1)!} \frac{a''_j}{j!} \right) s^{k-1}.$$

Igualando cada coeficiente observamos que:

$$\frac{a''_k}{(k-1)!} = \sum_{i=1}^k \frac{b_i}{(i-1)!} \frac{a''_{k-i}}{(k-i)!} \quad \forall k \in \mathbb{N}.$$

Si multiplicamos por $(k-1)!$ concluimos que los coeficientes a''_i cumplen la fórmula recursiva:

$$a''_k = \sum_{i=1}^k \binom{k-1}{i-1} b_i a''_{k-i} \quad \forall k \in \mathbb{N}.$$

Por otro lado, si definimos

$$a'_0 = 1$$

$$a'_k = \sum_{\pi \in \mathcal{P}(k)} b_{\pi} \quad \forall k \in \mathbb{N}$$

Entonces podemos separar la suma sobre todas las particiones, dependiendo del tamaño del bloque V que contiene al elemento 1. Es sencillo ver que

hay $\binom{k-1}{i-1}$ bloques de tamaño i que contienen al elemento 1, para cada uno de ellos, aparecerá un b_i que podemos factorizar y lo que queda después de factorizar es la suma sobre todas las particiones de los $k-i$ elementos restantes. Es decir:

$$\sum_{\pi \in \mathcal{P}(k)} b_\pi = \sum_{i=1}^k \sum_{\substack{\pi \in \mathcal{P}(k) \\ |V|=i}} b_i b_{\pi \setminus V} = \sum_{i=1}^k b_i \sum_{\substack{\pi \in \mathcal{P}(k) \\ |V|=i}} b_{\pi \setminus V} = \sum_{i=1}^k b_i \binom{k-1}{i-1} \sum_{\sigma \in \mathcal{P}(k-i)} b_\sigma.$$

Entonces tenemos que

$$a'_k = \sum_{\pi \in \mathcal{P}(k)} b_\pi = \sum_{i=1}^k b_i \binom{k-1}{i-1} \sum_{\sigma \in \mathcal{P}(k-i)} b_\sigma = \sum_{i=1}^k b_i \binom{k-1}{i-1} a'_{k-i}.$$

Como los coeficientes a_k , a'_k y a''_k cumplen la misma recursión entonces $a_k = a'_k = a''_k$ para toda $k \in \mathbb{N}$ y concluimos que las tres fórmulas son equivalentes. \square

Lema 3.5.2 (Booleano). Sean $(a_m)_{m \in \mathbb{N}}$ y $(b_m)_{m \in \mathbb{N}}$ dos sucesiones de reales y definimos $a_0 = 1$. Entonces las siguientes fórmulas son equivalentes:

1.

$$a_k = \sum_{i=1}^k b_i a_{k-i} \quad \forall k \in \mathbb{N}$$

2.

$$1 - \frac{1}{\sum_{j=0}^{\infty} a_j s^j} = \sum_{i=1}^{\infty} b_i s^i$$

3.

$$a_k = \sum_{\pi \in \mathcal{I}(k)} b_\pi \quad \forall k \in \mathbb{N}$$

Demostración. Fijamos las b 's y nos tomamos $a''_0 = 1$ y a''_1, \dots, a''_m tales que

$$1 - \frac{1}{\sum_{j=0}^{\infty} a''_j s^j} = \sum_{i=1}^{\infty} b_i s^i.$$

Si multiplicamos de ambos lados por $\sum_{j=0}^{\infty} a''_j s^j$ obtenemos que

$$\frac{\sum_{j=1}^{\infty} \frac{a''_j}{(j-1)!} s^{j-1}}{\sum_{j=0}^{\infty} \frac{a''_j}{j!} s^j} = \sum_{i=1}^{\infty} \frac{b_i}{(i-1)!} s^{i-1}.$$

Y si pasamos multiplicando concluimos que

$$\sum_{j=0}^{\infty} a_j'' s^j - 1 = \left(\sum_{i=1}^{\infty} b_i s^i \right) \left(\sum_{j=0}^{\infty} a_j'' s^j \right) = \sum_{k=1}^{\infty} \left(\sum_{i+j=k} b_i a_j'' \right) s^k.$$

Igualando cada coeficiente concluimos que los coeficientes a_i'' cumplen la fórmula recursiva:

$$a_k'' = \sum_{i=1}^k b_i a_{k-i}'' \quad \forall k \in \mathbb{N}.$$

Por otro lado, si definimos

$$a_0' = 1,$$

$$a_k' = \sum_{\pi \in \mathcal{I}(k)} b_{\pi} \quad \forall k \in \mathbb{N},$$

entonces podemos separar la suma sobre las particiones de intervalos, dependiendo del tamaño del bloque V que contiene al elemento 1. Es sencillo ver que sólo hay un bloque (intervalo) de tamaño i que contienen al elemento 1, en este caso aparecerá un b_i que podemos factorizar y lo que queda después de factorizar es la suma sobre todas las particiones de intervalos de los $k-i$ elementos restantes. Es decir:

$$\sum_{\pi \in \mathcal{I}(k)} b_{\pi} = \sum_{i=1}^k \left(\sum_{\substack{\pi \in \mathcal{I}(k) \\ |V|=i}} b_i b_{\pi \setminus V} \right) = \sum_{i=1}^k b_i \left(\sum_{\substack{\pi \in \mathcal{I}(k) \\ |V|=i}} b_{\pi \setminus V} \right) = \sum_{i=1}^k b_i \left(\sum_{\sigma \in \mathcal{I}(k-i)} b_{\sigma} \right).$$

Entonces tenemos que

$$a_k' = \sum_{\pi \in \mathcal{I}(k)} b_{\pi} = \sum_{i=1}^k b_i \left(\sum_{\sigma \in \mathcal{I}(k-i)} b_{\sigma} \right) = \sum_{i=1}^k b_i a_{k-i}'.$$

Como los coeficientes a_k , a_k' y a_k'' cumplen la misma recursión entonces $a_k = a_k' = a_k''$ para toda $k \in \mathbb{N}$ y concluimos que las tres fórmulas son equivalentes. \square

Lema 3.5.3 (Libre). Sean $(a_m)_{m \in \mathbb{N}}$ y $(b_m)_{m \in \mathbb{N}}$ dos sucesiones de reales y definimos $a_0 = 1$. Entonces las siguientes fórmulas son equivalentes:

1.

$$a_k = \sum_{j=1}^k b_j \sum_{\substack{i_1, \dots, i_j \in \{0, 1, \dots, k-j\} \\ i_1 + \dots + i_j = k-j}} a_{i_1} \cdots a_{i_j} \quad \forall k \in \mathbb{N}$$

2.

$$\sum_{k=0}^{\infty} a_k s^k = 1 + \sum_{j=1}^{\infty} b_j \left(s \sum_{i=0}^{\infty} a_i s^i \right)^j$$

3.

$$a_k = \sum_{\pi \in \mathcal{NC}(k)} b_{\pi} \quad \forall k \in \mathbb{N}$$

Demostración. Fijamos las b 's y nos tomamos $a'_0 = 1$ y a'_1, \dots, a'_m tales que

$$\sum_{k=0}^{\infty} a'_k s^k = 1 + \sum_{j=1}^{\infty} b_j \left(s \sum_{i=0}^{\infty} a'_i s^i \right)^j.$$

Cancelamos el término lineal y desarrollamos el lado derecho:

$$\begin{aligned} \sum_{k=1}^{\infty} a'_k s^k &= \sum_{k=1}^{\infty} b_i s^i \left(\sum_{j=0}^{\infty} a'_j s^j \right)^i \\ &= \sum_{k=1}^{\infty} b_i s^i \sum_{j=1}^k \sum_{\substack{i_1, \dots, i_j \in \{0, 1, \dots, k-j\} \\ i_1 + \dots + i_j = k-j}} (a'_{i_1} s^{i_1}) \cdots (a'_{i_j} s^{i_j}) \\ &= \sum_{k=1}^{\infty} s^k \sum_{j=1}^k b_j \sum_{\substack{i_1, \dots, i_j \in \{0, 1, \dots, k-j\} \\ i_1 + \dots + i_j = k-j}} a'_{i_1} \cdots a'_{i_j}. \end{aligned}$$

Igualando cada coeficiente concluimos que los coeficientes a'_i cumplen la fórmula recursiva:

$$a'_k = \sum_{j=1}^k b_j \sum_{\substack{i_1, \dots, i_j \in \{0, 1, \dots, k-j\} \\ i_1 + \dots + i_j = k-j}} a'_{i_1} \cdots a'_{i_j} \quad \forall k \in \mathbb{N}.$$

Por otro lado, si definimos

$$\begin{aligned} a''_0 &= 1 \\ a''_k &= \sum_{\pi \in \mathcal{NC}(k)} b_{\pi} \quad \forall k \in \mathbb{N} \end{aligned}$$

Entonces podemos separar la suma sobre las particiones π que no se cruzan, dependiendo del tamaño, j , del bloque V de π que contiene al elemento 1. Es decir,

$$a''_k = \sum_{j=1}^k \sum_{|V|=j} \sum_{\substack{\pi \in \mathcal{NC}(k) \\ V \in \pi}} b_{\pi}.$$

Supongamos que $V = \{v_1, \dots, v_j\}$ donde $1 = v_1 < \dots < v_j$. Entonces, por la condición de que π no se cruza tendremos que

$$\pi = V \cup \pi_1 \cup \dots \cup \pi_j,$$

donde π_r es una partición que no se cruza de $\{v_r + 1, v_r + 2, \dots, v_{r+1} - 1\}$ (tomamos $v_{j+1} := k$). Denotemos $i_r := v_{r+1} - v_r - 1$, entonces podemos identificar a π_r con un elemento de $NC(i_r)$, (si $i_r = 0$ podemos olvidarnos de la partición π_r). Por lo tanto tenemos que:

$$\begin{aligned} \sum_{\substack{\pi \in NC(k) \\ V \in \pi}} b_\pi &= \sum_{\substack{\pi = V \cup \pi_1 \cup \dots \cup \pi_j \\ \pi_r \in NC(i_r)}} b_j b_{\pi_1} \cdots b_{\pi_j} \\ &= b_j \left(\sum_{\pi_1 \in NC(i_1)} b_{\pi_1} \right) \cdots \left(\sum_{\pi_j \in NC(i_j)} b_{\pi_j} \right) \\ &= b_j a''_{i_1} \cdots a''_{i_j}. \end{aligned}$$

Además, es sencillo observar que hay una biyección entre los bloques V de π de tamaño j que contienen al elemento 1 y las j -tuplas $i_1, \dots, i_j \in \{0, 1, \dots, k-j\}$ tales que $i_1 + \dots + i_j = k-j$. Por lo tanto las a''_i cumplen la fórmula recursiva:

$$a''_n = \sum_{j=1}^k b_j \sum_{\substack{i_1, \dots, i_j \in \{0, 1, \dots, k-j\} \\ i_1 + \dots + i_j = k-j}} a''_{i_1} \cdots a''_{i_j}.$$

Como los coeficientes a_k , a'_k y a''_k cumplen la misma recursión entonces $a_k = a'_k = a''_k$ para toda $k \in \mathbb{N}$ y concluimos que las tres fórmulas son equivalentes. \square

3.5.2. Relación con las transformadas

Definición 3.5.4. Sean $(m_n)_{n \geq 1}$, $(b_n)_{n \geq 1}$, $(c_n)_{n \geq 1}$ y $(r_n)_{n \geq 1}$ los momentos y los cumulantes booleanos, clásicos y libres, respectivamente, de una medida de probabilidad μ con momentos finitos de todos los órdenes. Definimos las siguientes series formales de potencias:

$$M_\mu(z) := 1 + \sum_{n=1}^{\infty} m_n z^n.$$

$$M_\mu^*(z) := 1 + \sum_{n=1}^{\infty} \frac{m_n}{n!} z^n.$$

$$B_\mu(z) := \sum_{n=0}^{\infty} b_{n+1} z^n.$$

$$C_\mu(z) := \sum_{n=0}^{\infty} \frac{c_{n+1}}{n!} z^n.$$

$$R_\mu(z) := \sum_{n=0}^{\infty} r_{n+1} z^n.$$

Observación 3.5.5. Es importante observar que dada una medida de probabilidad μ (con soporte compacto), su serie de momentos M_μ y la transformada de Cauchy G_μ cumplen la relación

$$G_\mu(z) = \frac{1}{z} M_\mu\left(\frac{1}{z}\right), \quad (3.3)$$

la cuál se sigue directamente de la definición anterior y la Observación 1.4.7. Esta relación será importante, ya que las transformadas del Capítulo 1 las definimos en términos de la transformada de Cauchy.

Ahora veamos la siguiente proposición que se sigue de los tres lemas de la sección anterior.

Proposición 3.5.6. Dada una medida de probabilidad μ , tenemos las siguientes relaciones:

1.

$$\ln(M_\mu^*(z)) = zC_\mu(z). \quad (3.4)$$

2.

$$1 - \frac{1}{M_\mu(z)} = zB_\mu(z). \quad (3.5)$$

3.

$$M_\mu(z) = 1 + zM_\mu(z)R_\mu[zM_\mu(z)]. \quad (3.6)$$

Demostración. Sustituimos $a_n = m_n$ y $b_n = c_n$ en el Lema 3.5.1, $a_n = m_n$ y $b_n = b_n$ en el Lema 3.5.2, y $a_n = m_n$ y $b_n = r_n$ en el Lema 3.5.3. Es claro que las fórmulas entre momentos y cumulantes son precisamente las fórmulas en el inciso 3 de cada uno de los lemas. Por lo tanto, se cumplen las fórmulas del inciso 2 en cada uno de los lemas. Y sustituyendo el nombre que le dimos a cada serie formal de potencias, llegamos a las fórmulas deseadas. \square

Teorema 3.5.7. Dada una medida de probabilidad μ , tenemos las siguientes igualdades:

1.

$$\ln \mathcal{F}_\mu(z) = -izC_\mu(-iz) = \sum_{n=0}^{\infty} \frac{c_{n+1}}{n!} (-iz)^{n+1}.$$

2.

$$\mathcal{B}_\mu(z) = B_\mu(z) = \sum_{n=0}^{\infty} b_{n+1} z^n.$$

3.

$$\mathcal{R}_\mu(z) = R_\mu(z) = \sum_{n=0}^{\infty} r_{n+1} z^n.$$

Demostración. 1. Por la Observación 1.4.4 sabemos que $\mathcal{F}_\mu(z) = -izM_\mu^*(-iz)$. Y utilizando la fórmula (3.4) concluimos que $\ln \mathcal{F}_\mu(z) = -izC_\mu(-iz)$.

2. De la ecuación (3.3) podemos deducir que $zM_\mu(z) = G_\mu(\frac{1}{z})$, si lo sustituimos en la ecuación (3.5) dividida entre z concluimos que

$$B_\mu(z) = \frac{1}{z} - \frac{1}{zM_\mu(z)} = \frac{1}{z} - \frac{1}{G_\mu(\frac{1}{z})} = \mathcal{B}_\mu(z).$$

3. Utilizando las fórmulas (3.6) y (3.3) podemos observar que como series formales de potencias tenemos las siguientes igualdades:

$$zG(z) = M(\frac{1}{z}) = 1 + \frac{1}{z}M(\frac{1}{z})R\left[\frac{1}{z}M(\frac{1}{z})\right] = 1 + G(z)R[G(z)].$$

Si dividimos entre $G(z)$ obtenemos que

$$z = \frac{1}{G(z)} + R[G(z)].$$

Lo anterior nos dice que $G(z)$ y $R(z) + \frac{1}{z}$ son inversas. Por lo tanto, concluimos que

$$R(z) = G_\mu^{-1}(z) - \frac{1}{z} = \mathcal{R}_\mu(z).$$

□

Ya estamos listos para dar una demostración del Teorema 1.4.10 que dejamos pendiente en el Capítulo 1.

Demostración. Tomemos x y y variables aleatorias autoadjuntas en algún espacio de probabilidad C^* tales que x tiene distribución μ y y tiene distribución ν .

Si x y y son libres, entonces por el teorema anterior y la linealidad de los cumulantes libres concluimos que

$$\mathcal{R}_{\mu \boxplus \nu}(z) = \sum_{n=0}^{\infty} r_{n+1}^{x+y} z^n = \sum_{n=0}^{\infty} r_{n+1}^x z^n + \sum_{n=0}^{\infty} r_{n+1}^y z^n = \mathcal{R}_\mu(z) + \mathcal{R}_\nu(z).$$

Si x y y son independientes en el sentido booleano, entonces por el teorema anterior y la linealidad de los cumulantes booleanos concluimos que

$$\mathcal{B}_{\mu \boxplus \nu}(z) = \sum_{n=0}^{\infty} b_{n+1}^{x+y} z^n = \sum_{n=0}^{\infty} b_{n+1}^x z^n + \sum_{n=0}^{\infty} b_{n+1}^y z^n = \mathcal{B}_\mu(z) + \mathcal{B}_\nu(z).$$

Caso monótono. De la definición de la transformada F , la ecuación que queremos demostrar es equivalente a

$$G_{\mu \triangleright \nu}(z) = G_\mu\left(\frac{1}{G_\nu(z)}\right).$$

Por la ecuación (3.3), eso se transforma en ver que

$$\frac{1}{z} M_{\mu \triangleright \nu}\left(\frac{1}{z}\right) = \frac{1}{z} M_\nu\left(\frac{1}{z}\right) M_\mu\left(\frac{1}{z} M_\nu\left(\frac{1}{z}\right)\right).$$

Si multiplicamos por z y cambiamos el $\frac{1}{z}$ por z nos queda demostrar que

$$M_{\mu \triangleright \nu}(z) = M_\nu(z) M_\mu(z M_\nu(z)).$$

De la definición de M (y pensando que $m_0 = 1$), esto es lo mismo que ver que

$$\sum_{n=0}^{\infty} m_n^{x+y} z^n = M_\nu(z) \left(\sum_{k=0}^{\infty} m_k^x z^k (M_\nu(z))^k \right).$$

Si del lado derecho metemos a $M_\nu(z)$ y lo escribimos como serie tenemos que

$$\sum_{n=0}^{\infty} m_n^{x+y} z^n = \sum_{k=0}^{\infty} m_k^x z^k \left(\sum_{j=0}^{\infty} m_j^y z^j \right)^{k+1}.$$

Si nos fijamos en el coeficiente de z^n del lado derecho, tenemos que para cada k necesitamos encontrar $k+1$ enteros no negativos j_0, \dots, j_k que sumados nos den $n-k$ esto para que al multiplicarlo con el z^k completemos z^n . En cada uno de esos nos aportará al coeficiente el sumando $m_k^x m_{j_0}^y \cdots m_{j_k}^y$. Para llegar a la igualdad entre series que tenemos arriba vamos a ver la igualdad coeficiente a coeficiente. Es decir, queremos demostrar que para toda $n \geq 0$ se cumple que

$$m_n^{x+y} = \sum_{k=0}^n \sum_{\substack{j_0 + \dots + j_k = n-k \\ j_0 \geq 0, \dots, j_k \geq 0}} m_k^x m_{j_0}^y \cdots m_{j_k}^y$$

cuando x y y son monótonamente independientes. Pero esto es precisamente lo que nos dice el Teorema 3.4.1. \square

Observación 3.5.8. 1. Siguiendo la misma línea de la demostración que dimos para los primeros dos incisos, podemos dar una demostración para el Teorema 1.4.5:

$$\begin{aligned}
\log \mathcal{F}_{\mu*\nu}(z) &= \sum_{n=0}^{\infty} \frac{c_{n+1}^{x+y}}{n!} (-iz)^{n+1} \\
&= \sum_{n=0}^{\infty} \frac{c_{n+1}^x}{n!} (-iz)^{n+1} + \sum_{n=0}^{\infty} \frac{c_{n+1}^y}{n!} (-iz)^{n+1} \\
&= \log \mathcal{F}_{\mu}(z) + \log \mathcal{F}_{\nu}(z).
\end{aligned}$$

2. A pesar de que en la última demostración, pareciera que el caso monótono es más complicado, esto sólo se debe a que no habíamos hablado nada sobre la transformada F , mientras que para las otras transformadas ya habíamos demostrado varios lemas. Si nos ponemos a analizarlo con cuidado, descubriremos que el caso más sencillo es el monótono ya que sólo utiliza unas cuantas equivalencias y el Teorema 3.4.1 que se demuestra expandiendo de forma cuidadosa. De hecho, si observamos con detalle, para el caso monótono ni si quiera tuvimos que recurrir a los cumulantes. Recordemos que los cumulantes monótonos no cumplen la propiedad 3' y es por eso que no tenemos una proposición del estilo *la suma de las transformadas es la transformada de la convolución aditiva*, como ocurre en los otros tres casos. Por esta razón, el caso monótono nos pide ver que la convolución de transformadas nos da la transformada de la convolución aditiva y esta proposición no hace uso en ningún momento de los cumulantes.

3.6. Demostración del Teorema del Límite Central

Ya que desarrollamos toda la teoría de cumulantes, estamos listos para demostrar su poder, dando una demostración sencilla del TLC para cada una de las independencias.

Demostración. Denotemos por y_N las variables que nos interesan,

$$y_N := \frac{a_1 + \dots + a_N}{\sqrt{N}}.$$

Por las propiedades 2 y 3 de los cumulantes generalizados es sencillo observar que $\kappa_1(y_N) = 0$, $\kappa_2(y_N) = \sigma^2$ y que

$$\kappa_n^{y_N} = N^{-\frac{n-2}{2}} \kappa_n^{a_1} \rightarrow 0$$

cuando $N \rightarrow \infty$, para todo $n \geq 3$. Y por la propiedad 1 sabemos que los momentos $m_n^{y_N}$ convergen a los momentos m_n caracterizados por tener los

cumulantes $\kappa_1 = 0$, $\kappa_2 = \sigma^2$ y $\kappa_n = 0$ para todo $n \geq 3$. Ahora, como todos los cumulantes se anulan excepto el segundo será sencillo usar la fórmula entre momentos y cumulantes en cada independencia para ver qué variables aleatorias tienen esos cumulantes:

1. Clásico. De la fórmula entre momentos y cumulantes obtenemos que

$$m_n = \sum_{\pi \in \mathcal{P}(n)} c_\pi = \sum_{\pi \in \mathcal{P}_2(n)} (\sigma^2)^{n/2} = \begin{cases} \sigma^n (n-1)(n-3) \cdots 3 \cdot 1 & \text{si } n \text{ es par} \\ 0 & \text{si } n \text{ es impar.} \end{cases}$$

Y ya sabemos que esos son precisamente los momentos de una variable Gaussiana.

2. Libre. De la fórmula entre momentos y cumulantes obtenemos que

$$m_n = \sum_{\pi \in \mathcal{NC}(n)} r_\pi = \sum_{\pi \in \mathcal{NC}_2(n)} (\sigma^2)^{n/2} = \begin{cases} \sigma^n C_{n/2} & \text{si } n \text{ es par} \\ 0 & \text{si } n \text{ es impar,} \end{cases}$$

Y ya sabemos que esos son precisamente los momentos de una variable semicircular.

3. Booleano. De la fórmula entre momentos y cumulantes obtenemos que

$$m_n = \sum_{\pi \in \mathcal{I}(n)} b_\pi = \sum_{\pi \in \mathcal{I}_2(n)} (\sigma^2)^{n/2} = \begin{cases} \sigma^n & \text{si } n \text{ es par} \\ 0 & \text{si } n \text{ es impar.} \end{cases}$$

Y ya sabemos que esos son precisamente los momentos de una variable Bernoulli.

4. Monótono. Del Teorema 3.4.7 obtenemos que

$$\begin{aligned} m_n &= \sum_{k=1}^n \sum_{1=i_0 < i_1 < \cdots < i_k=n+1} \frac{1}{k!} \prod_{l=1}^k i_{l-1} h_{i_l - i_{l-1}} \\ &= \sum_{i_0=1, i_2=3, i_3=5, \dots, i_k=n+1} \frac{1}{(n/2)!} \prod_{l=1}^{n/2} (2l-1) \sigma^2 \\ &= \frac{n!}{(n/2)! 2^n (n/2)!} (\sigma^2)^{n/2} \\ &= \begin{cases} \left(\frac{\sigma}{2}\right)^n \binom{n}{n/2} & \text{si } n \text{ es par} \\ 0 & \text{si } n \text{ es impar.} \end{cases} \end{aligned}$$

Y ya sabemos que esos son precisamente los momentos de la distribución de la ley del arco seno.

□

Capítulo 4

Probabilidad Finita

El objetivo principal de este capítulo es abordar de forma combinatoria el concepto de libertad finita propuesta por Marcus en [Mar15]. Primero, presentaremos algunas herramientas necesarias para definir este concepto. Posteriormente, presentaremos algunos de los resultados de Marcus [Mar15]. Por último, presentaremos la aportación principal de esta tesis: definimos una noción de cumulantes finitos que linealizan la convolución aditiva simétrica y demostramos algunos de los teoremas del artículo de Marcus desde un enfoque combinatorio. Además, encontramos una fórmula que relaciona los momentos con los cumulantes finitos la cual permite entender mejor la relación entre cumulantes finitos y cumulantes libres.

4.1. Preliminares

En esta sección presentamos las herramientas necesarias para estudiar la convolución polinomial o finita. Estas herramientas son la Transformada de Laplace, la Transformada de Legendre, las Normas L^p y el Determinante de Fuglede-Kadison. Nuestra exposición se basa en el artículo de Marcus [Mar15].

4.1.1. Transformada de Laplace

Definición 4.1.1. Dada una función $f(x)$ definida para todo $x \geq 0$, la transformada de Laplace de f para el parámetro s , se define como

$$\mathcal{L}\{f\}(s) = \int_0^{\infty} e^{-xs} f(x) dx = \lim_{a \rightarrow \infty} \int_0^a e^{-xs} f(x) dx$$

donde estamos suponiendo que el límite de la derecha converge, en caso contrario diremos que f no tiene transformada de Laplace en s .

La transformada de Laplace es muy útil en combinatoria por que nos ayuda a convertir funciones generadoras exponenciales en funciones genera-

doras ordinarias (y viceversa). Sólo utilizaremos las siguientes dos propiedades sencillas de la transformada de Laplace.

Lema 4.1.2. 1. Si $k \in \mathbb{N}$, entonces la transformada de Laplace de x^k es:

$$\mathcal{L}\{x^k\}(s) = \frac{k!}{s^{k+1}}.$$

2. La transformada de Laplace es lineal, es decir, si f y g son funciones con transformada de Laplace en s y $\alpha \in \mathbb{R}$, entonces tenemos que

$$\mathcal{L}\{\alpha f + g\}(s) = \alpha \mathcal{L}\{f\}(s) + \mathcal{L}\{g\}(s).$$

Demostración. 1. Procedemos por inducción sobre k . Para la base $k = 0$ claramente $\mathcal{L}\{x^0\}(s) = \int_0^\infty e^{-xs} dx = \frac{1}{s}$. Para $k > 0$ utilizamos integración por partes con $v(x) = \frac{1}{-s}e^{-xs}$ y $u(x) = x^k$ y la hipótesis de inducción para ver que:

$$\begin{aligned} \mathcal{L}\{x^k\}(s) &= \int_0^\infty e^{-xs} x^k dx \\ &= \frac{1}{-s} e^{-xs} x^k \Big|_0^\infty - \int_0^\infty \frac{k}{-s} e^{-xs} x^{k-1} dx \\ &= 0 + \frac{k}{s} \mathcal{L}\{x^{k-1}\}(s) \\ &= \frac{k(k-1)!}{s s^k} \\ &= \frac{k!}{s^{k+1}}. \end{aligned}$$

Con esto concluimos la inducción y obtenemos la fórmula deseada.

2. Para este inciso sólo hace falta utilizar la definición y recordar que la integral es lineal:

$$\begin{aligned} \mathcal{L}\{\alpha f + g\}(s) &= \int_0^\infty e^{-xs} (\alpha f(x) + g(x)) dx \\ &= \lim_{a \rightarrow \infty} \int_0^a e^{-xs} (\alpha f(x) + g(x)) dx \\ &= \alpha \lim_{a \rightarrow \infty} \int_0^a e^{-xs} f(x) dx + \lim_{a \rightarrow \infty} \int_0^a e^{-xs} g(x) dx \\ &= \alpha \int_0^\infty e^{-xs} f(x) dx + \int_0^\infty e^{-xs} g(x) dx \\ &= \alpha \mathcal{L}\{f\}(s) + \mathcal{L}\{g\}(s). \end{aligned}$$

□

4.1.2. Transformada de Legendre

Definición 4.1.3. Sea f una función que es convexa en el intervalo $X \subseteq \mathbb{R}$. Y sea

$$X^* = \left\{ x^* \in \mathbb{R} : \sup_{x \in X} \{xx^* - f(x)\} < \infty \right\}.$$

La transformada de Legendre se define como la función $f^* : X^* \rightarrow \mathbb{R}$ tal que

$$f^*(s) = \sup_{x \in X} \{xs - f(x)\}.$$

En el caso en el que f es diferenciable tenemos la siguiente relación.

Lema 4.1.4. Si f es una función estrictamente convexa en X y diferenciable en un punto $z \in X$. Entonces $f'(z) \in X^*$ y

$$f^*(f'(z)) = zf'(z) - f(z).$$

Demostración. Tomemos $z \in X$, como f es diferenciable en z y estrictamente convexa, entonces satisface la desigualdad $f(x) \geq f(z) + (x-z)f'(z)$ para todo $x \in X$, además se da la igualdad si y sólo si $x = z$. Si acomodamos la ecuación obtenemos que $xf'(z) - f(x) \leq zf'(z) - f(z)$, y como la igualdad se da cuando $x = z$, entonces tenemos que

$$\sup_{x \in X} \{xf'(z) - f(x)\} = zf'(z) - f(z) < \infty.$$

Por definición esto significa que $f'(z) \in X^*$ y $f^*(f'(z)) = zf'(z) - f(z)$. \square

Corolario 4.1.5. Si f satisface las condiciones del lema anterior y además tiene segunda derivada. Entonces

$$(f')^{-1}(z) = (f^*)'(z) \text{ y}$$

$$f''((f^*)'(z)) = \frac{1}{(f^*)''(z)}.$$

Demostración. Si derivamos la ecuación del teorema anterior, obtenemos que

$$(f^*)'(f'(z))f''(z) = zf''(z) + f'(z) - f'(z) = zf''(z).$$

Como $f''(z) \neq 0$ por que f es estrictamente convexa, concluimos que f' es la inversa de $(f^*)'$, y por lo tanto obtenemos la primera ecuación. Esta ecuación implica que $f'((f^*)'(z)) = z$ y si derivamos, obtenemos la ecuación deseada. \square

4.1.3. Normas L^p

Definición 4.1.6. Dado un espacio de medida (X, μ) y un real $p > 0$, definimos la norma L^p de una función f como

$$\|f\|_p = \left(\int |f|^p d\mu \right)^{\frac{1}{p}}.$$

Para $p = \infty$, definimos

$$\|f\|_\infty = \lim_{p \rightarrow \infty} \|f\|_p = \inf \{a \geq 0 : \mu(\{x : |f(x)| > a\}) = 0\}.$$

Un resultado simple de la teoría de espacios L^p es:

Lema 4.1.7. Si μ es absolutamente continua con respecto a la medida de Lebesgue y f es continua entonces

$$\|f\|_\infty = \sup_{x \in X} \{|f(x)|\}.$$

Demostración. Tomemos una función f continua fija. Escribiremos $s := \sup_{x \in X} \{|f(x)|\}$. Sea E un conjunto medible de X con $\mu(E) < \infty$, y tal que f se anula en el complemento de E . Si $\mu(E) = 0$, entonces $s = \|f\|_p = 0$ y no hay nada que demostrar. Si $\mu(E) > 0$, tenemos que

$$\|f\|_p = \left(\int |f|^p d\mu \right)^{\frac{1}{p}} \leq \left(\int s^p d\mu \right)^{\frac{1}{p}} \leq s \mu(E)^{\frac{1}{p}}.$$

Como $\mu(E)^{\frac{1}{p}} \rightarrow 1$ cuando $p \rightarrow \infty$, entonces $\lim_{p \rightarrow \infty} \|f\|_p \leq s$. Por otro lado, dado $\epsilon > 0$, tenemos que

$$\mu(\{x : |f(x)| > s - \epsilon\}) \geq \delta, \quad \text{para alguna } \delta > 0.$$

Como esto pasa para una ϵ arbitraria entonces obtenemos que

$$\|f\|_\infty = \inf \{a \geq 0 : \mu(\{x : |f(x)| > a\}) = 0\} \geq s.$$

Por lo tanto, como

$$s \leq \|f\|_\infty = \lim_{p \rightarrow \infty} \|f\|_p \leq s,$$

concluimos que $\|f\|_\infty = s = \sup_{x \in X} \{|f(x)|\}$. □

La siguiente observación relaciona la transformada de Legendre con los espacios L^p .

Corolario 4.1.8. Sea $X \subset \mathbb{R}$ y μ una medida que es absolutamente continua con respecto a la medida de Lebesgue. Entonces para cualquier función continua $f : X \rightarrow \mathbb{R}$, tenemos que

$$f^*(s) = \ln \left\| e^{xs - f(x)} \right\|_\infty$$

para todo $s \in X^*$ (donde la transformada de Legendre y la norma L^∞ se toman sobre el espacio X).

Demostración. Como f es real, f^* también es real. Por lo tanto podemos escribir $f^*(s)$ como

$$\begin{aligned}
f^*(s) &= \ln(\exp(f^*(s))) \\
&= \ln \left(\exp \left(\sup_{x \in X} \{xs - f(x)\} \right) \right) \\
&= \ln \left(\sup_{x \in X} \{ \exp(xs - f(x)) \} \right) \\
&= \ln \left(\sup_{x \in X} \{ | \exp(xs - f(x)) | \} \right) \\
&= \ln \left\| e^{xs - f(x)} \right\|_{\infty}.
\end{aligned}$$

□

4.1.4. Determinante de Fuglede-Kadison

El determinante de un operador invertible fue introducido por primera vez en 1952 por Fuglede y Kadison [FK52]. La noción generaliza al determinante usual y se puede considerar para cualquier operador en un álgebra de von Neumann finita (\mathcal{N}, τ) con traza fiel normal. Dado un elemento normal invertible, $x \in \mathcal{N}$, el determinante de Fuglede-Kadison se define como

$$\begin{aligned}
\Delta_{\tau}^{FK} : GL_1(\mathcal{N}) &\rightarrow \mathbb{R}_+ \\
x &\mapsto \exp \left(\tau(\log(x^*x)^{\frac{1}{2}}) \right).
\end{aligned}$$

Algunas de las propiedades básicas de este determinante son las siguientes:

$$\begin{aligned}
\Delta_{\tau}^{FK}(e^y) &= |e^{\tau(y)}| \quad \text{para todo } y \in \mathcal{N}, \\
\Delta_{\tau}^{FK}(x) &= \Delta_{\tau}^{FK} \left((x^*x)^{\frac{1}{2}} \right) \quad \text{para todo } x \in GL_1(\mathcal{N}), \\
\Delta_{\tau}^{FK}(xy) &= \Delta_{\tau}^{FK}(x)\Delta_{\tau}^{FK}(y) \quad \text{para todo } x, y \in \mathcal{N}.
\end{aligned}$$

Para el caso en el que tomamos $\mathcal{N} = M_d(\mathbb{C})$ donde $\tau(x) = \frac{1}{n} \sum_{j=1}^d x_{j,j}$ es la traza normalizada y \det es el determinante usual, entonces tendremos que

$$\Delta_{\tau}^{FK}(A) = \det[A]^{\frac{1}{d}}.$$

donde A una matriz de $d \times d$ positiva definida.

Fuglede y Kadison demostraron que para una sucesión de matrices positivas definidas A_1, A_2, \dots que convergen adecuadamente a un operador límite a , el determinante normalizado converge a un límite bien definido y el funcional límite tiene muchas de las propiedades usuales que se esperarían en un determinante. La única propiedad no trivial que utilizaremos en este trabajo es que Δ_{τ}^{FK} está bien definida y es el límite de sus valores en una sucesión de matrices adecuada. En lo que resta del trabajo omitiremos el superíndice y subíndice para facilitar la lectura, es decir denotaremos al determinante de Fuglede-Kadison simplemente como $\Delta(\cdot)$

4.2. Convolución de polinomios

En esta sección presentaremos la convolución aditiva simétrica que será la análoga finita de las convoluciones aditivas presentadas en el Capítulo 1.

4.2.1. Convolución aditiva simétrica

La convolución aditiva simétrica que presentaremos se define y utiliza en [MSS15a] para dar cotas sobre las raíces de ciertos polinomios.

Definición 4.2.1 (Convolución aditiva simétrica). Sean A y B matrices reales simétricas de $m \times m$, donde $p(x) = \det[xI - A]$ y $q(x) = \det[xI - B]$. La convolución aditiva simétrica de p y q se define como

$$[p \boxplus_m q](x) = \mathbb{E}_Q \{ \det[xI - A - QBQ^T] \}$$

donde la esperanza se toma sobre las matrices ortonormales Q distribuidas uniformemente (usando la medida de Haar).

Observación 4.2.2. Si todas las raíces de p y de q son reales, entonces todas las raíces de $p \boxplus_m q$ son reales, debido a una teoría mucho más general desarrollada en [BB09].

En [MSS15a], los autores demuestran que la convolución aditiva simétrica se puede ver de la siguiente manera:

Teorema 4.2.3. Si $p(x) = \sum_{i=0}^m x^{m-i} (-1)^i a_i^p$ y $q(x) = \sum_{i=0}^m x^{m-i} (-1)^i a_i^q$ entonces tenemos que

$$[p \boxplus_m q](x) = \sum_{k=0}^m x^{m-k} (-1)^k \sum_{i+j=k} \frac{(m-i)!(m-j)!}{m!(m-i-j)!} a_i^p a_j^q.$$

Demostración. Denotaremos por $c_k(A)$ al coeficiente de $(-1)^k x^{d-k}$ del polinomio característico de una matriz, A , de dimensión d . Es importante observar que $c_k(A)$ es la k -ésima función simétrica elemental de los valores propios de A . Para subconjuntos $S, T \subseteq \{1, \dots, d\}$, escribiremos $A(S, T)$ para la submatriz de A indexada por las filas en S y las columnas en T . Escribiremos $A(S, :)$ para la submatriz que contiene las columnas de S . Observemos que

$$c_k(A) = \sum_{|S|=k} \det(A(S, S)).$$

Para demostrar el teorema veremos la igualdad para cada coeficiente. Es decir, veremos que si A y B son matrices simétricas de $d \times d$, entonces tenemos que

$$\mathbb{E}_Q c_k(A + QBQ^T) = \sum_{i+j=k} \frac{(d-i)!(d-j)!}{d!(d-k)!} c_i(A) c_j(B)$$

donde Q es una matriz ortonormal aleatoria.

Si escribimos $A = UCU^T$ y $B = VDV^T$ donde las matrices C y D son diagonales y las matrices U y V son ortonormales, entonces podemos ver que

$$\mathbb{E}_Q c_k(A + QBQ^T) = \mathbb{E}_Q c_k(C + U^T QVDV^T Q^T U) = \mathbb{E}_Q c_k(C + QDQ^T),$$

ya que $U^T QV$ es una matriz ortonormal. Debido al razonamiento anterior, podemos suponer que A y B son matrices diagonales. Entonces podemos ver que

$$\begin{aligned} \mathbb{E}_Q c_k(A + QBQ^T) &= \mathbb{E}_Q c_k \left(\sum_{i \leq d} a_i e_i e_i^T + \sum_{i \leq d} b_i (Qe_i)(Qe_i)^T \right) \\ &= \mathbb{E}_Q \sum_{S, T: |S|+|T|=k} c_k \left(\sum_{i \in S} a_i e_i e_i^T + \sum_{i \in T} b_i (Qe_i)(Qe_i)^T \right) \\ &= \mathbb{E}_Q \sum_{|S|+|T|=k} c_k (M_{ST}(A(S, S) \oplus B(T, T)) M_{ST}^T) \\ &= \mathbb{E}_Q \sum_{|S|+|T|=k} \det(M_{ST}^T M_{ST}(A(S, S) \oplus B(T, T))) \\ &= \sum_{|S|+|T|=k} \det A(S, S) \det B(T, T) \cdot (\mathbb{E}_Q \det(M_{ST}^T M_{ST})) \end{aligned}$$

donde $M_{ST} = [I(:, S)|(QI)(:, T)]$ es una matriz de $d \times k$ con columnas $\{e_i\}_{i \in S}$ y $\{Qe_i\}_{i \in T}$. Ahora podemos ver que la esperanza $\mathbb{E}_Q \det(M_{ST}^T M_{ST})$ depende solamente de $|S| := i$ y $|T| := j$. Si fijamos S y T y quitamos los subíndices para simplificar la notación, podemos ver que

$$\begin{aligned} \mathbb{E}_Q \det(M^T M) &= \mathbb{E}_Q c_i(I(:, S)I(:, S)^T) c_j(I(:, \bar{S})^T QI(:, T)I(:, T)^T Q^T I(:, \bar{S})) \\ &= \mathbb{E}_Q \mathbb{E}_\Pi(1) \cdot c_j(I(:, \bar{S})^T \Pi QI(:, T)I(:, T)^T Q^T \Pi^T I(:, \bar{S})), \end{aligned}$$

donde Π es una matriz aleatoria de permutación. Como $\Pi Q \stackrel{\text{dist}}{=} Q$, entonces lo anterior es igual a

$$\mathbb{E}_Q \mathbb{E}_R c_j(I(:, R)^T Q(:, T)Q(:, T)^T I(:, R)),$$

donde R es un subconjunto aleatorio uniforme de $[d]$ de tamaño $|\bar{S}| = d - i$. Ahora, si escribimos $P = Q(:, T)Q(:, T)^T$, tendremos que lo anterior es igual a

$$\mathbb{E}_Q \frac{1}{\binom{d}{d-i}} \sum_{|R|=d-i} c_j(P(R, R)) = \mathbb{E}_Q \frac{1}{\binom{d}{d-i}} \sum_{|R|=d-i} \sum_{W \subset R, |W|=j} c_j(P(W, W)).$$

Como cada W de tamaño $j = k - i \leq d - i$ aparece $\binom{d-j}{d-i-j}$ veces, entonces lo anterior es igual a

$$\mathbb{E}_Q \frac{1}{\binom{d}{d-i}} \sum_{W \subset [d], |W|=j} \binom{d-j}{d-i-j} c_j(P(W, W)) = \mathbb{E}_Q \frac{\binom{d-j}{d-i-j}}{\binom{d}{d-i}} c_j(P).$$

Como P es una proyección de rango j la anterior se reduce a

$$\frac{\binom{d-j}{d-i-j}}{\binom{d}{d-i}} \cdot 1 = \frac{(d-i)!(d-j)!}{d!(d-k)!}.$$

Por lo tanto podemos concluir que

$$\begin{aligned} \mathbb{E}_Q c_k(A + QBQ^T) &= \sum_{|S|+|T|=k} \det A(S, S) \det B(T, T) \cdot (\mathbb{E}_Q \det(M_{ST}^T M_{ST})) \\ &= \sum_{|S|+|T|=k} \det A(S, S) \det B(T, T) \frac{(d-|S|)!(d-|T|)!}{d!(d-k)!} \\ &= \sum_{i+j=k} \frac{(d-i)!(d-j)!}{d!(d-k)!} \sum_{|S|=i, |T|=j} \det A(S, S) \det B(T, T) \\ &= \sum_{i+j=k} \frac{(d-i)!(d-j)!}{d!(d-k)!} c_i(A) c_j(B), \end{aligned}$$

que es lo que queríamos. \square

4.2.2. Transformada U

En el artículo de libertad finita [Mar15], la transformada U juega un papel importante para formular los lemas y teoremas de una manera más elegante. Más adelante veremos que podemos olvidarnos de la transformada U y abordar los cálculos desde un enfoque combinatorio, que resultará un poco más sencillo porque ya contamos con una plataforma similar que vimos en los capítulos anteriores. El objetivo de esta sección es presentar la herramienta de la transformada U tal y como se presenta en el artículo [Mar15], para que posteriormente podamos comparar ambos métodos.

Notación 4.2.4. Dado un multiconjunto S , escribiremos $|S|$ para denotar al número de elementos en el multiconjunto (con multiplicidad). Por ejemplo $|\{1, 2, i, i\}| = 4$. Abusando de la notación, trataremos a los multiconjuntos como si fueran variables aleatorias. En cualquier caso, supondremos que la variable aleatoria está uniformemente distribuida en los elementos del multiconjunto. Por ejemplo, pensaremos que

$$\frac{1}{|S|} \sum_{s_i \in S} f(s_i) \quad \text{y} \quad \mathbb{E}\{f(s)\}$$

representan lo mismo.

Lema 4.2.5. Sea S un multiconjunto finito de los números complejos con $|S| = m$. Entonces existe un único multiconjunto T de números complejos tal que $|T| = m$ y tal que

$$\mathbb{E}\{\ln(x - s)^m\} = \ln \mathbb{E}\{(x - T)^m\}$$

para todo x (donde \ln es considerado como el valor principal del logaritmo complejo).

Demostración. Si aplicamos la función exponencial de ambos lados y simplificamos, llegamos a la proposición equivalente:

$$\prod_{s_i \in S} (x - s_i) = \sum_{k=0}^m \binom{m}{k} x^{m-k} (-1)^k \mathbb{E}\{T^k\}.$$

Observamos que ambos lados de la ecuación son polinomios mónicos, así que sólo hace falta mostrar que la igualdad se da para cada uno de los otros m coeficientes. Cada igualdad se puede ver como tener una restricción en $\mathbb{E}\{T^k\}$ para $1 \leq k \leq m$. Usando las identidades de Newton, esto es equivalente a tener restricciones en las primeras m funciones simétricas elementales de los elementos de T . Sin embargo, esto es equivalente a tener m soluciones de un polinomio de grado m , el cuál existe (y es único) sobre los números complejos. \square

Definición 4.2.6. Dado un multiconjunto S , al multiconjunto T que satisface las propiedades del Lema 4.2.5 le llamaremos la transformada U de S .

La expresión que será más útil al utilizar la transformada U es la siguiente:

$$\prod_{s_i \in S} (x - s_i) = \mathbb{E}\{(x - T)^m\}.$$

Es decir, cualquier polinomio mónico de grado m se puede escribir como combinación lineal de m polinomios, cada uno de los cuales tiene una sola raíz de multiplicidad m . En particular, si S es el multiconjunto de raíces (con multiplicidades) del polinomio p de grado m , entonces tenemos que $p(x) = \mathbb{E}\{(x - T)^m\}$, donde T es la transformada U de S . La propiedad de que $|T| = |S|$ es importante como lo veremos en el siguiente lema.

Lema 4.2.7. Sea S un multiconjunto de números reales y sea T su transformada U. Entonces

$$\mathbb{E}\{f(s)\} \in \mathbb{R}$$

para cualquier función f que es analítica en el soporte de T .

Demostración. Como todos los elementos de S son reales, entonces todos los coeficientes de

$$\prod_{s_i \in S} (x - s_i) = \sum_{k=0}^m \binom{m}{k} x^{m-k} (-1)^k \mathbb{E}\{T^k\}$$

son reales. Ahora, si consideramos el polinomio

$$q(x) = \prod_{t_i \in T} (x - t_i) = \sum_{i=0}^m q_i x^{m-i}.$$

Por las identidades de Newton, los coeficientes q_i se pueden expresar en función de los primeros m momentos (que acabamos de ver que son reales) y por lo tanto, son reales. Sea A_T una matriz real simétrica cuyos valores propios son los elementos de T . El teorema de Cayley-Hamilton nos dice que (como ecuación de matrices) $q(A_T) = 0$. Entonces

$$0 = \text{tr} [A_T q(A_T)] = \sum_i q_i \text{tr}[A_T^{m-i+1}]$$

es una expresión para $\text{tr} [A_T^{m+1}]$ como combinación lineal de $\text{tr}[A_T^k]$ para $k \leq m$ con coeficientes reales (y por lo tanto es real). Si procedemos inductivamente, tenemos que $\text{tr}[A_T^k] = \mathbb{E}\{T^k\} \in \mathbb{R}$ para todo k y por lo tanto lo mismo será cierto para la función analítica. \square

Observemos que si $|T|$ es más grande que $|S|$, entonces no podemos garantizar lo anterior. Para ver esto, tomamos $k > 0$ y podemos usar la misma lógica que en el Lema 4.2.5 para encontrar un multiconjunto W con $|W| = m + k$ cuyos primeros m momentos sean iguales a los de μ , pero que después el momento $m + 1$ sea complejo.

Aunque queramos tratar a nuestros multiconjuntos como distribuciones, debemos tener cuidado cuando vamos en la dirección contraria. Diremos que una distribución μ es m -realizable si existe un multiconjunto S_μ , con $|S_\mu| = m$, tal que la distribución uniforme de S_μ nos de las mismas probabilidades que μ . Diremos que el multiconjunto S_μ es su m -realización. En particular, la transformada U siempre se debe definir en términos de la realización de una distribución y no en la distribución tal cual. Es fácil observar que si una distribución es m -realizable, también será km -realizable para cualquier entero positivo k . Sin embargo, estas realizaciones serán muy diferentes.

La utilidad de la transformada U radica en su habilidad para transformar convoluciones de polinomios en independencia clásica. Esto se muestra en el siguiente lema:

Lema 4.2.8. Sean p y q polinomios de grado m con S y T las transformadas U de sus raíces. Como podemos tratar a S y T como independientes, entonces tenemos que

$$[p \boxplus_m q](x) = \mathbb{E}\{(x - S - T)^m\}.$$

Demostración. Sea $p(x) = \sum_{i=0}^m x^{m-i}(-1)^i p_i$. Como observamos anteriormente, tenemos que

$$p(x) = \mathbb{E}\{(x - S)^m\} = \sum_{k=0}^m x^{m-k}(-1)^k \binom{m}{k} \mathbb{E}\{S^k\}$$

y por lo tanto $p_i = \binom{m}{i} \mathbb{E}\{S^i\}$ (y análogamente $q_j = \binom{m}{j} \mathbb{E}\{T^j\}$). Por el Teorema 4.2.3 tenemos que

$$\begin{aligned} [p \boxplus_m q](x) &= \sum_{k=0}^m x^{m-k}(-1)^k \sum_{i+j=k} \frac{(m-i)!(m-j)!}{(m-k)!m!} p_i q_j \\ &= \sum_{k=0}^m x^{m-k}(-1)^k \sum_{i+j=k} \frac{(m-i)!(m-j)!}{(m-k)!m!} \binom{m}{i} \binom{m}{j} \mathbb{E}\{S^i\} \mathbb{E}\{T^j\} \\ &= \sum_{k=0}^m x^{m-k}(-1)^k \sum_{i+j=k} \binom{m}{i, j} \mathbb{E}\{S^i T^j\} \\ &= \mathbb{E}\{(x - S - T)^m\}. \end{aligned}$$

□

Como un corolario directo del Lema 4.2.8, obtenemos que la transformada U es invariante bajo traslaciones o multiplicación por un escalar.

Corolario 4.2.9. Sea T la transformada U del multiconjunto S . Entonces las transformadas U de

$$\{s + k : s \in S\} \quad \text{y} \quad \{ks : s \in S\}$$

son

$$\{t + k : t \in T\} \quad \text{y} \quad \{kt : t \in T\}$$

respectivamente.

Demostración. Sólo hay que utilizar el Lema 4.2.8 con $q = (x - k)^m$. □

4.3. Transformada \mathcal{R} m -finita

En esta sección, primero definiremos para cada m una nueva serie formal de potencias, a la que denotaremos como *transformada \mathcal{R} m -finita*. Posteriormente mostraremos la relación que hay entre estas series y la transformada \mathcal{R} que definimos en el Capítulo 1, con esto quedará claro el por qué

las denotamos con ese nombre. Por último veremos la relación entre esta transformada y la convolución de polinomios que introdujimos al principio de este capítulo. De esta forma mostraremos la relación entre la probabilidad libre y la convolución de polinomios.

4.3.1. Definición y propiedades básicas

Definición 4.3.1. Sea A un operador real simétrico cuya distribución espectral, μ_A , tiene soporte compacto. Para un entero m , definimos la serie de potencias

$$\mathcal{K}_{\mu_A}^m(s) = -\frac{\partial}{\partial s} \ln \|e^{-xs} \Delta(xI - A)\|_m$$

donde el dominio de integración es (ρ_A, ∞) . Diremos que $\mathcal{K}_{\mu_A}^m(s)$ es la *transformada \mathcal{K} m -finita de μ_A* . Y definimos la serie de potencias

$$\mathcal{R}_{\mu_A}^m(s) = \mathcal{K}_{\mu_A}^m(s) - \mathcal{K}_{\mu_0}^m(s)$$

donde μ_0 es la distribución constante 0. Diremos que $\mathcal{R}_{\mu_A}^m(s)$ es la *transformada \mathcal{R} m -finita de μ_A* .

Observación 4.3.2. 1. La transformada \mathcal{K}^m es la análoga de la inversa de la transformada de Cauchy en probabilidad libre. Además, a partir de la definición, es claro que es invariante bajo transformaciones unitarias de A .

2. Es importante observar que

$$\mathcal{K}_{\mu_0}^m(s) = \left(1 + \frac{1}{m}\right) \frac{1}{s}.$$

La demostración de este resultado es bastante sencilla, pero la escribiremos formalmente como un lema porque sirve para entender cómo se relacionan los conceptos involucrados en la definición de las transformadas finitas.

3. La transformada \mathcal{R}^m es la análoga a la transformada \mathcal{R} . En la probabilidad libre a la inversa de la transformada de Cauchy le restamos el término $\frac{1}{s}$ para obtener la transformada \mathcal{R} , la versión finita de lo anterior es restarle el término $\left(1 + \frac{1}{m}\right) \frac{1}{s}$ a la transformada \mathcal{K}^m para obtener la transformada \mathcal{R}^m .

Lema 4.3.3. Sea μ_0 la distribución constante 0. Para cualquier entero m se cumple que

$$\mathcal{K}_{\mu_0}^m(s) = \left(1 + \frac{1}{m}\right) \frac{1}{s}.$$

Demostración. Primero calcularemos la norma del determinante utilizando la transformada de Legendre:

$$\begin{aligned}
\|e^{-xs} \Delta(xI - 0)\|_m &= \|e^{-xs} (\det[xI - 0])^{\frac{1}{m}}\|_m = \|e^{-xs} x\|_m \\
&= \left(\int_0^\infty e^{-mxs} x^m dx \right)^{\frac{1}{m}} \\
&= (\mathcal{L}\{x^m\}(ms))^{\frac{1}{m}} \\
&= \left(\frac{m!}{(ms)^{m+1}} \right)^{\frac{1}{m}} \\
&= \frac{(m!)^{\frac{1}{m}}}{m^{\frac{m+1}{m}} s^{\frac{m+1}{m}}}.
\end{aligned}$$

Para concluir solo hay que aplicarle logaritmo a lo que obtuvimos y luego derivar:

$$\begin{aligned}
-\frac{\partial}{\partial s} \ln \|e^{-xs} \Delta(xI - 0)\|_m &= -\frac{\partial}{\partial s} \left(\ln \left(\frac{(m!)^{\frac{1}{m}}}{m^{\frac{m+1}{m}}} \right) - \frac{m+1}{m} \ln(s) \right) \\
&= \left(\frac{m+1}{m} \right) \frac{1}{s}.
\end{aligned}$$

□

4.3.2. Relación con probabilidad libre

El siguiente teorema es bastante importante porque muestra la conexión entre las transformadas finitas y la probabilidad libre. Para su demostración utilizaremos varias de las cosas que vimos en los preliminares de este capítulo.

Teorema 4.3.4 (Lema 4.6, [Mar15]). Sea A un operador Hermitiano cuya distribución espectral, μ_A , tiene soporte compacto. Entonces

$$\lim_{m \rightarrow \infty} \mathcal{K}_{\mu_A}^m(s) = \mathcal{G}_{\mu_A}^{-1}(s)$$

en todos los puntos $s \in (\rho_A, \infty)$.

Demostración. Primero definimos la función $f : (\rho_A, \infty) \rightarrow \mathbb{R}$ de tal manera que $f(x) = -\ln \Delta(xI - A)$. Es sencillo observar que f está bien definida y es continua en (ρ_A, ∞) . Además, si derivamos obtenemos que

$$f'(x) = -\text{tr}[(xI - A)^{-1}] = -\mathcal{G}_{\mu_A}(x).$$

Si volvemos a derivar obtenemos que

$$f''(x) = \text{tr}[(xI - A)^{-2}] > 0.$$

Por lo tanto f es estrictamente convexa y por el Corolario 4.1.5 tenemos que

$$(f^*)'(s) = (f')^{-1}(s) = (-\mathcal{G}_{\mu_A}(x))^{-1}(s) = \mathcal{G}_{\mu_A}^{-1}(-s).$$

Sustituyendo lo anterior en el Corolario 4.1.8 obtenemos que

$$\mathcal{G}_{\mu_A}^{-1}(s) = (f^*)'(-s) = -\frac{\partial}{\partial s} \ln \|e^{-xs} e^{-f(x)}\|_{\infty} = -\frac{\partial}{\partial s} \ln \|e^{-xs} \Delta(xI - A)\|_{\infty}.$$

Por el Lema 4.1.7 concluimos que

$$\begin{aligned} \lim_{m \rightarrow \infty} \mathcal{K}_{\mu_A}^m(s) &= \lim_{m \rightarrow \infty} \left(-\frac{\partial}{\partial s} \ln \|e^{-xs} \Delta(xI - A)\|_m \right) \\ &= -\frac{\partial}{\partial s} \ln \|e^{-xs} \Delta(xI - A)\|_{\infty} \\ &= \mathcal{G}_{\mu_A}^{-1}(s). \end{aligned}$$

□

4.3.3. Relación con la convolución de polinomios

Ya que mostramos la relación de la transformada finita con la probabilidad libre, ahora sólo queda ver la relación de la transformada finita con la convolución de polinomios. Para poder relacionar estas series de potencias que acabamos de definir con los polinomios necesitaremos prescindir de muchos de los coeficientes de la serie de potencias y quedarnos sólo con los primeros coeficientes, por ello, antes de seguir adelante introduciremos un tipo de notación bastante intuitiva.

Notación 4.3.5. Sea $f(x) = \sum_{i=0}^{\infty} a_i x^i$ un serie formal de potencias. Para un entero no negativo k , escribiremos $f \bmod [x^k]$ para denotar al polinomio formado por los primeros k coeficientes, es decir

$$f \bmod [x^k] = \sum_{i=0}^{k-1} a_i x^i.$$

Y escribiremos

$$f \equiv g \pmod{[x^k]}$$

si $f - g$ es el polinomio 0.

Observación 4.3.6. Varias propiedades se pueden derivar fácilmente de la definición. Por ejemplo, si a , b , c , d y h son series formales de potencias, entonces el sistema

$$\begin{aligned} a(x) &\equiv b(x) \pmod{[x^k]}, \\ c(x) &\equiv d(x) \pmod{[x^k]}, \end{aligned}$$

implica que

$$a(x) + c(x) \equiv b(x) + d(x) \pmod{[x^k]},$$

y también implica que

$$a(x)c(x) \equiv b(x)d(x) \pmod{[x^k]}.$$

Además, combinando las dos anteriores podemos concluir que

$$h(a(x)) \equiv h(b(x)) \pmod{[x^k]}.$$

Si además sabemos que h es invertible, entonces utilizando la inversa en el inciso anterior concluimos que:

$$a(x) \equiv b(x) \pmod{[x^k]} \iff h(a(x)) \equiv h(b(x)) \pmod{[x^k]}.$$

Ahora sí estamos listos para comenzar, el objetivo principal de esta sección será vincular la transformada \mathcal{R}^m que acabamos de definir en 4.3.1, con la convolución aditiva simétrica que definimos en 4.2.1. Para ello, seguiremos la misma idea de demostración que utiliza Adam Marcus en su artículo, lo primero que haremos es escribir $\mathcal{R}_{\mu_A}^m(s)$ en términos de la transformada U de A . La idea en los siguientes dos lemas es hacer los mismos cálculos que hicimos para la distribución constante 0 en el Lema 4.3.3 pero para cualquier distribución.

Lema 4.3.7 (Lema 4.3, [Mar15]). Si A es una matriz real simétrica de $m \times m$, entonces

$$\frac{\|e^{-xs}\Delta(xI - A)\|_m^m}{\|e^{-xs}\Delta(xI - 0)\|_m^m} \equiv \mathbb{E}\{e^{-msT_A}\} \pmod{[s^{m+1}]}$$

donde T_A es la transformada U de $\lambda(A)$.

Demostración. Para una matriz real simétrica de $m \times m$, A , tenemos la simplificación

$$\Delta(xI - A)^m = \det[xI - A] = \mathbb{E}\{(x - T_A)^m\},$$

donde T_A es la transformada U de $\lambda(A)$. Entonces

$$\begin{aligned}
\frac{\|e^{-xs}\Delta(xI - A)\|_m^m}{\|e^{-xs}\Delta(xI - 0)\|_m^m} &= \frac{(ms)^{m+1}}{m!} \|e^{-xs}\Delta(xI - A)\|_m^m \\
&= \frac{(ms)^{m+1}}{m!} \int_{\rho_A}^{\infty} e^{-mxs} \Delta(xI - A)^m dx \\
&= \frac{(ms)^{m+1}}{m!} \int_{\rho_A}^{\infty} e^{-(ms)x} \mathbb{E}\{(x - T_A)^m\} dx \\
&= \frac{(ms)^{m+1}}{m!} e^{-ms\rho_A} \int_0^{\infty} e^{-(ms)y} \mathbb{E}\{(y + \rho_A - T_A)^m\} dy \\
&= \mathbb{E}\left\{\frac{(ms)^{m+1}}{m!e^{ms\rho_A}} \mathcal{L}\{(y + \rho_A - T_A)^m\}(ms)\right\} \\
&= \mathbb{E}\left\{\frac{(ms)^{m+1}}{m!e^{ms\rho_A}} \sum_{i=0}^m \binom{m}{i} (\rho_A - T_A)^{m-i} \mathcal{L}\{x^i\}(ms)\right\} \\
&= \mathbb{E}\left\{\frac{(ms)^{m+1}}{m!e^{ms\rho_A}} \sum_{i=0}^m \frac{m!}{i!(m-i)!} (\rho_A - T_A)^{m-i} \frac{i!}{(ms)^{i+1}}\right\} \\
&= \mathbb{E}\left\{\frac{1}{e^{ms\rho_A}} \sum_{i=0}^m \frac{(ms)^{m-i} (\rho_A - T_A)^{m-i}}{(m-i)!}\right\} \\
&\equiv \mathbb{E}\left\{\frac{1}{e^{ms\rho_A}} e^{ms(\rho_A - T_A)}\right\} \pmod{[s^{m+1}]} \\
&\equiv \mathbb{E}\{e^{-msT_A}\} \pmod{[s^{m+1}]},
\end{aligned}$$

que es lo que queríamos demostrar. \square

Corolario 4.3.8 (Corolario 4.4, [Mar15]). Si A es una matriz real simétrica de $m \times m$, entonces

$$\mathcal{R}_{\mu_A}^m(s) \equiv -\frac{1}{m} \frac{\partial}{\partial s} \ln \mathbb{E}\{e^{-msT_A}\} \pmod{[s^m]}$$

donde T_A es la transformada U de $\lambda(A)$.

Demostración. Si al lema anterior le aplicamos la función logaritmo de ambos lados, que es invertible, por la Observación 4.3.6 tendremos que

$$\ln\left(\frac{\|e^{-xs}\Delta(xI - A)\|_m^m}{\|e^{-xs}\Delta(xI - 0)\|_m^m}\right) \equiv \ln \mathbb{E}\{e^{-msT_A}\} \pmod{[s^{m+1}]}.$$

Como estas series de potencias coinciden es los primeros $m + 1$ coeficientes, sus derivadas coinciden en los primeros m coeficientes. Con esto y de la

definición de la transformada \mathcal{R} obtenemos que:

$$\begin{aligned}
\mathcal{R}_{\mu_A}^m(s) &= \mathcal{K}_{\mu_A}^m(s) - \mathcal{K}_{\mu_0}^m(s) \\
&= -\frac{\partial}{\partial s} \ln \|e^{-xs} \delta(xI - A)\|_m + \frac{\partial}{\partial s} \ln \|e^{-xs} \delta(xI - 0)\|_m \\
&= -\frac{1}{m} \frac{\partial}{\partial s} \ln \left(\frac{\|e^{-xs} \Delta(xI - A)\|_m^m}{\|e^{-xs} \Delta(xI - 0)\|_m^m} \right) \\
&\equiv -\frac{1}{m} \frac{\partial}{\partial s} \ln \mathbb{E}\{e^{-msT_A}\} \quad \text{mod } [s^m].
\end{aligned}$$

□

Ya que relacionamos a $\mathcal{R}_{\mu_A}^m(s)$ con la transformada U de A , utilizaremos el Lema 4.2.8 para relacionar la convolución aditiva simétrica con la transformada U.

Lema 4.3.9 (Lema 4.5, [Mar15]). Sean A y B matrices reales simétricas de $m \times m$. Las siguientes proposiciones son equivalentes:

1.

$$\mathcal{R}_{\mu_A}^m(s) + \mathcal{R}_{\mu_B}^m(s) = \mathcal{R}_{\mu_{A+B}}^m(s) \quad \text{mod } [s^m].$$

2.

$$\det[xI - A] \boxplus_m \det[xI - B] = \det[xI - A - B].$$

Demostración. Sean T_A , T_B y T_{A+B} las transformadas U de $\lambda(A)$, $\lambda(B)$ y $\lambda(A+B)$, respectivamente. Por el Lema 4.2.8, sabemos que la proposición del segundo inciso es equivalente a

$$\mathbb{E}\{(x - T_A - T_B)^m\} = \mathbb{E}\{(x - T_{A+B})^m\}.$$

Esto último ocurre si y sólo si los primeros m momentos de $T_A + T_B$ y T_{A+B} son iguales. Esto a su vez es equivalente a que

$$\mathbb{E}\{e^{-ms(T_A+T_B)}\} = \mathbb{E}\{e^{-msT_{A+B}}\} \quad \text{mod } [s^{m+1}]$$

y como podemos asumir que T_A y T_B son independientes, lo último es cierto si y sólo si

$$\mathbb{E}\{e^{-msT_A}\} \mathbb{E}\{e^{-msT_B}\} = \mathbb{E}\{e^{-msT_{A+B}}\} \quad \text{mod } [s^{m+1}].$$

Como \ln es invertible, por la Observación 4.3.6 tenemos que lo anterior es equivalente a

$$f_A(s) + f_B(s) \equiv f_{A+B}(s) \quad \text{mod } [s^{m+1}]$$

donde $f_X(s) = -\frac{1}{m} \ln \mathbb{E}\{e^{-msT_X}\}$. Por lo tanto, por el Corolario 4.3.8, sólo hace falta mostrar que la ecuación anterior es equivalente a

$$\frac{\partial}{\partial s} f_A(s) + \frac{\partial}{\partial s} f_B(s) \equiv \frac{\partial}{\partial s} f_{A+B}(s) \quad \text{mod } [s^m].$$

La primer implicación es obvia, (ya habíamos mencionado que si dos series coinciden en los primeros $m+1$ coeficientes, entonces sus derivadas coinciden en los primeros m). Para la segunda implicación sólo hace falta mostrar que $f_A(0) + f_B(0) = f_{A+B}(0)$. Sin embargo, lo anterior es sencillo ya que por definición tenemos que $f_X(0) = 0$ para todo X . \square

Observemos que en el último lema utilizamos fuertemente que A y B son matrices de $m \times m$. Esto es porque los primeros m momentos caracterizan completamente la distribución, lo cual es falso para distribuciones más generales.

4.4. Cumulantes finitos

Hasta ahora, lo único que hemos hecho en este capítulo es repasar los puntos clave del artículo [Mar15]. En esta sección introduciremos el concepto de cumulante finito y demostraremos algunos de los teoremas del artículo (los cuales presentamos en la sección anterior) pero desde un enfoque combinatorio, tomando como base las propiedades que cumplen los cumulantes para los principales cuatro tipos de independencia.

Los resultados principales de esta sección son parte de un trabajo de investigación que será mandado a publicación [AP16].

4.4.1. Definición y notación

Definición 4.4.1. A los primeros d coeficientes de $\mathcal{R}_{\mu_A}^d(s)$ los llamaremos cumulantes d -finitos y los denotaremos por κ_i^A . Es decir, $\kappa_1^A, \kappa_2^A, \dots, \kappa_d^A$ son tales que:

$$\mathcal{R}_{\mu_A}^d(s) \equiv \sum_{j=0}^{d-1} \kappa_{j+1}^A s^j \pmod{[s^d]}.$$

Observación 4.4.2. Es importante recalcar que estos cumulantes dependen del entero d . Es decir, para cada $d \in \mathbb{N}$ estamos definiendo d cumulantes distintos. La razón para sólo definir los primeros d y no hacerlo para todos los coeficientes de $\mathcal{R}_{\mu_A}^d(s)$, es que la relación de la transformada con la convolución de polinomios sólo depende de los primeros d coeficientes.

Notación 4.4.3. 1. Dada A , una matriz real simétrica de $d \times d$, denotaremos por $p_A(x) := \det[xI - A]$ al polinomio característico de A . Además, si r_1, r_2, \dots, r_d son los valores propios de A . Entonces definimos:

$$a_0^A := 1.$$

$$a_k^A := \sum_{sym} r_1 r_2 \cdots r_k \quad \text{para } k = 1, \dots, d.$$

2. Escribiremos

$$(d)_n := \frac{d!}{(d-n)!} = d(d-1) \cdots (d-n+1) \quad \text{para } 0 \leq n \leq d.$$

Y denotaremos por $(d)_\pi$ a la extensión multiplicativa de las $(d)_n$.

Observación 4.4.4. 1. De la definición anterior podemos obtener esta importante relación:

$$p_A(x) = (x - r_1) \cdots (x - r_d) = \sum_{i=0}^d x^{d-i} (-1)^i a_i^A.$$

2. Observemos que en el Lema 4.2.3 tenemos que $a_k = a_k^A$ y $b_k = a_k^B$. Por lo tanto podemos reescribirlo de la siguiente manera:

$$a_k^{p_A \boxplus p_B} = \sum_{i+j=k} \frac{(d-i)!(d-j)!}{d!(d-i-j)!} a_i^A a_j^B = (d)_k \sum_{i+j=k} \frac{a_i^A}{(d)_i} \frac{a_j^B}{(d)_j},$$

para $k = 1, 2, \dots, d$.

Con la notación que acabamos de introducir, podemos reescribir el Corolario 4.3.8 (Corolario 4.4, [Mar15]) de la siguiente manera:

Lema 4.4.5. Si A es una matriz real simétrica de $d \times d$, entonces

$$\mathcal{R}_{\mu_A}^d(s) \equiv -\frac{1}{d} \frac{\partial}{\partial s} \ln \left(\sum_{i=0}^d \frac{(-d)^i a_i^A}{(d)_i} s^i \right) \quad \text{mod } [s^d].$$

Demostración. Por la definición de transformada U y la notación que acabamos de introducir tenemos que

$$\sum_{i=0}^d x^{d-i} (-1)^i a_i^A = (x - r_1) \cdots (x - r_d) = \sum_{i=0}^d x^{d-i} (-1)^i \binom{d}{i} \mathbb{E}\{T^i\}.$$

Por lo tanto tenemos que $a_i^A = \frac{(d)_i}{i!} \mathbb{E}\{T^i\}$ para $i = 1, \dots, d$. Entonces mod $[x^{d+1}]$ observamos que:

$$\mathbb{E}\left\{e^{-dsT_A}\right\} \equiv \mathbb{E}\left\{\sum_{i=0}^d \frac{(-ds)^i T_A^i}{i!}\right\} \equiv \sum_{i=0}^d (-d)^i s^i \frac{\mathbb{E}\{T_A^i\}}{i!} \equiv \sum_{i=0}^d (-d)^i s^i \frac{a_i^A}{(d)_i}.$$

Esta última igualdad es justo lo que estamos reescribiendo, si la sustituimos en el Corolario 4.3.8, obtenemos esta nueva versión. \square

4.4.2. Los cumulantes finitos son lineales

Primero utilizaremos lo que aprendimos en el Capítulo 2 sobre álgebras de incidencia, para relacionar a los cumulantes finitos, κ^A , de una matriz A con los coeficientes, a^A , de su polinomio característico.

Lema 4.4.6 (Relación entre cumulantes y coeficientes). Sea A una matriz de $d \times d$. Sean κ_n^A , para $n = 1, \dots, d$, los cumulantes finitos definidos a partir de A y sean a_n^A , para $n = 0, 1, \dots, d$, los coeficientes del polinomio característico p_A (ver Notación 4.4.3). Entonces se cumplen las siguientes dos fórmulas (una inversa de la otra):

$$a_n^A = \frac{(d)_n}{(-d)^n n!} \sum_{\pi \in \mathcal{P}(n)} (-d)^{|\pi|} \kappa_\pi^A (n-1)!_\pi \quad \text{para } n = 1, \dots, d, \quad (4.1)$$

donde $(n-1)!_\pi := \prod_{V \in \pi} (|V| - 1)!$.

$$\kappa_n^A = \frac{(-d)^n}{d(n-1)!} \sum_{\pi \in \mathcal{P}(n)} \frac{(-1)^{|\pi|} n!_\pi a_\pi^A (|\pi| - 1)!}{(d)_\pi} \quad \text{para } n = 1, \dots, d, \quad (4.2)$$

donde $n!_\pi := \prod_{V \in \pi} (|V|)!$.

Demostración. Por el Lema 4.4.5 sabemos que:

$$\sum_{j=1}^d \kappa_j^A s^{j+1} \equiv \mathcal{R}_{\mu_A}^d(s) \equiv -\frac{1}{d} \frac{\partial}{\partial s} \ln \left(\sum_{i=0}^d \frac{(-1)^i a_i^A d^i}{(d)_i} s^i \right) \quad \text{mod } [s^d].$$

A partir de lo anterior se nos ocurre definir

$$a_i = \frac{(-d)^i a_i^A}{(d)_i} i! \quad \text{para } i = 0, 1, \dots, d \text{ y}$$

$$b_j = -d \kappa_j^A (j-1)! \quad \text{para } j = 1, \dots, d.$$

Observemos que $a_0^A = 1$, además si multiplicamos la ecuación anterior por $-d$ y sustituimos los valores que acabamos de definir llegamos a que

$$\frac{\partial}{\partial s} \ln \left(\sum_{i=0}^d \frac{a_i}{i!} s^i \right) = \sum_{j=1}^d \frac{b_j}{(j-1)!} s^{j-1} \quad \text{mod } [s^d].$$

Por el Lema 3.5.1 sabemos que esta fórmula es equivalente a la siguiente:

$$a_n = \sum_{\pi \in \mathcal{P}(n)} b_\pi \quad \text{para } n = 1, \dots, d.$$

También podemos utilizar la fórmula de inversión de Möbius en la ecuación anterior para obtener la ecuación inversa:

$$b_n = \sum_{\pi \in \mathcal{P}(n)} a_\pi (-1)^{|\pi|-1} (|\pi| - 1)! \quad \text{para } n = 1, \dots, d.$$

Sustituyendo los valores de las a_n y b_n en las dos ecuaciones anteriores obtenemos que

$$\frac{(-1)^n a_n^A d^n}{(d)_n} n! = \sum_{\pi \in \mathcal{P}(n)} (-d)^{|\pi|} \kappa_\pi^A (n-1)!_\pi \quad \text{para } n = 1, \dots, d$$

y que

$$-d \kappa_n^A (n-1)! = \sum_{\pi \in \mathcal{P}(n)} \frac{(-d)^n a_\pi^A n!_\pi}{(d)_\pi} (-1)^{|\pi|-1} (|\pi| - 1)! \quad \text{para } n = 1, \dots, d.$$

Si despejamos y simplificamos las fórmulas anteriores obtenemos las fórmulas deseadas. \square

Utilizaremos la fórmula anterior para ver que los cumulantes finitos se comportan de manera lineal respecto a la convolución de funciones. Aunque aún no hemos definido una independencia de manera formal, cabe mencionar que esta sería una de las propiedades deseables de los cumulantes.

Proposición 4.4.7. Sean A y B matrices reales simétricas de $d \times d$. Entonces tenemos que

$$\kappa_k^{pA \boxplus d p B} = \kappa_k^A + \kappa_k^B \quad \text{para } k = 1, \dots, d.$$

Demostración. Por la reformulación del Lema 4.2.3 (Observación 4.4.4) sabemos que:

$$\frac{a_k^{pA \boxplus d p B}}{(d)_k} = \sum_{i+j=k} \frac{a_i^A}{(d)_i} \frac{a_j^B}{(d)_j}.$$

Si sustituimos $\frac{a_i^A}{(d)_i}$ y $\frac{a_j^B}{(d)_j}$ utilizando la ecuación (4.1) del lema anterior, el lado derecho de la ecuación de arriba nos queda igual a

$$\begin{aligned} & \sum_{i+j=k} \left(\frac{1}{(-d)^i i!} \sum_{\pi \in \mathcal{P}(i)} (-d)^{|\pi|} (n-1)!_\pi \kappa_\pi^A \right) \left(\frac{1}{(-d)^j j!} \sum_{\pi \in \mathcal{P}(j)} (-d)^{|\pi|} (n-1)!_\pi \kappa_\pi^B \right) \\ &= \frac{1}{(-d)^k k!} \sum_{i+j=k} \binom{k}{i} \left(\sum_{\pi \in \mathcal{P}(i)} (-d)^{|\pi|} (n-1)!_\pi \kappa_\pi^A \right) \left(\sum_{\pi \in \mathcal{P}(j)} (-d)^{|\pi|} (n-1)!_\pi \kappa_\pi^B \right). \end{aligned}$$

Observemos que $\binom{k}{i}$ son las formas de partir a $[k]$ en dos subconjuntos ordenados, el primero con i elementos y el segundo con j elementos. Por

lo tanto, podemos tomar la suma para $i = 0, \dots, k$ sobre los subconjuntos $S \subset [k]$ y luego fijarnos en $\mathcal{P}(S) \cong \mathcal{P}(i)$ y $\mathcal{P}(S^c) \cong \mathcal{P}(j)$, donde $S^c = [k] - S$. De esta forma, obtenemos que:

$$\begin{aligned}
& \sum_{i+j=k} \binom{k}{i} \left(\sum_{\pi \in \mathcal{P}(i)} (-d)^{|\pi|} (n-1)!_{\pi} \kappa_{\pi}^A \right) \left(\sum_{\pi \in \mathcal{P}(j)} (-d)^{|\pi|} (n-1)!_{\pi} \kappa_{\pi}^B \right) \\
&= \sum_{S \subset [k]} \left(\sum_{\pi_1 \in \mathcal{P}(S)} (-d)^{|\pi_1|} (n-1)!_{\pi_1} \kappa_{\pi_1}^A \right) \left(\sum_{\pi_2 \in \mathcal{P}(S^c)} (-d)^{|\pi_2|} (n-1)!_{\pi_2} \kappa_{\pi_2}^B \right) \\
&= \sum_{\pi_1 \cup \pi_2 \in \mathcal{P}(k)} \left((-d)^{|\pi_1|} (n-1)!_{\pi_1} \kappa_{\pi_1}^A \right) \left((-d)^{|\pi_2|} (n-1)!_{\pi_2} \kappa_{\pi_2}^B \right).
\end{aligned}$$

La última suma la podemos reescribir, si primero nos tomamos las particiones $\pi \in \mathcal{P}(k)$ y para cada partición nos fijamos en las particiones π_1 y π_2 cuya unión es π . De esta manera, la última fórmula es igual a

$$\begin{aligned}
&= \sum_{\pi \in \mathcal{P}(k)} \sum_{\pi_1 \cup \pi_2 = \pi} \left((-d)^{|\pi_1|} (n-1)!_{\pi_1} \kappa_{\pi_1}^A \right) \left((-d)^{|\pi_2|} (n-1)!_{\pi_2} \kappa_{\pi_2}^B \right) \\
&= \sum_{\pi \in \mathcal{P}(k)} \sum_{\pi_1 \cup \pi_2 = \pi} (-d)^{|\pi_1| + |\pi_2|} (n-1)!_{\pi_1} (n-1)!_{\pi_2} \kappa_{\pi_1}^A \kappa_{\pi_2}^B \\
&= \sum_{\pi \in \mathcal{P}(k)} \sum_{\pi_1 \cup \pi_2 = \pi} (-d)^{|\pi|} (n-1)!_{\pi} \kappa_{\pi_1}^A \kappa_{\pi_2}^B \\
&= \sum_{\pi \in \mathcal{P}(k)} (-d)^{|\pi|} (n-1)!_{\pi} \sum_{\pi_1 \cup \pi_2 = \pi} \kappa_{\pi_1}^A \kappa_{\pi_2}^B \\
&= \sum_{\pi \in \mathcal{P}(k)} (-d)^{|\pi|} (n-1)!_{\pi} (\kappa^A + \kappa^B)_{\pi}.
\end{aligned}$$

De esta manera, obtenemos que

$$a_k^{p_A \boxplus d p_B} = \frac{(d)_k}{(-d)^k k!} \sum_{\pi \in \mathcal{P}(k)} (-d)^{|\pi|} (\kappa^A + \kappa^B)_{\pi} (n-1)!_{\pi} \quad \text{para } k = 1, \dots, d.$$

Por otro lado, la ecuación (4.1) que escribe a los coeficientes en términos de cumulantes nos dice que:

$$a_k^{p_A \boxplus d p_B} = \frac{(d)_k}{(-d)^k k!} \sum_{\pi \in \mathcal{P}(k)} (-d)^{|\pi|} \kappa_{\pi}^{p_A \boxplus d p_B} (n-1)!_{\pi} \quad \text{para } k = 1, \dots, d.$$

Observemos que estas fórmulas son iguales excepto que en la primera tenemos $\kappa^A + \kappa^B$ y en la segunda tenemos $\kappa^{p_A \boxplus d p_B}$, como sabemos que esta fórmula tiene inversa entonces podemos concluir que son iguales, es decir:

$$\kappa_k^{p_A \boxplus d p_B} = \kappa_k^A + \kappa_k^B \quad \text{para } k = 1, \dots, d.$$

□

Teorema 4.4.8 (Linealidad). Sean A , B y C matrices reales simétricas de $d \times d$. Las siguientes proposiciones son equivalentes:

1.

$$\mathcal{R}_{\mu_A}^d(s) + \mathcal{R}_{\mu_B}^d(s) = \mathcal{R}_{\mu_C}^d(s) \pmod{[s^d]}.$$

2.

$$p_A \boxplus_d p_B = p_C.$$

Demostración. La igualdad de las series módulo s^d en la primera proposición ocurre si y sólo si los primeros d coeficientes son iguales, es decir:

$$\kappa_k^A + \kappa_k^B = \kappa_k^C \quad \text{para } k = 1, \dots, d.$$

Por el lema anterior esto ocurre si y sólo si

$$\kappa_k^{p_A \boxplus_d p_B} = \kappa_k^C \quad \text{para } k = 1, \dots, d.$$

Por las fórmulas que relacionan cumulantes con los coeficientes lo anterior sucede si y sólo si

$$a_k^{p_A \boxplus_d p_B} = a_k^C \quad \text{para } k = 1, \dots, d.$$

Y lo anterior es la igualdad coeficiente a coeficiente de lo que nos dice la segunda proposición. \square

Si el lema anterior lo utilizamos con $C = A + B$ podemos concluir como corolario el Lema 4.5, [Mar15].

Corolario 4.4.9 (Lema 4.5, [Mar15]). Sean A y B matrices reales simétricas de $m \times m$. Las siguientes proposiciones son equivalentes:

1.

$$\mathcal{R}_{\mu_A}^m(s) + \mathcal{R}_{\mu_B}^m(s) = \mathcal{R}_{\mu_{A+B}}^m(s) \pmod{[s^m]}.$$

2.

$$\det[xI - A] \boxplus_m \det[xI - B] = \det[xI - A - B].$$

4.4.3. Fórmulas entre momentos y cumulantes finitos

Ya que demostramos los teoremas desde un enfoque combinatorio, nos gustaría ir un poco más allá y dar una fórmula entre momentos y cumulantes finitos, para ello, primero debemos relacionar a los momentos con los coeficientes del polinomio característico.

Lema 4.4.10 (Relación entre momentos y coeficientes). Sea A una matriz de $d \times d$. Sean m_n^A , para $n = 1, \dots, d$, los momentos de la matriz A y sean a_n^A , para $n = 0, 1, \dots, d$, los coeficientes del polinomio característico p_A (ver Notación 4.4.3). Entonces se cumplen las siguientes dos fórmulas (una inversa de la otra):

$$a_n^A = \frac{(-1)^n}{n!} \sum_{\pi \in \mathcal{P}(n)} (-d)^{|\pi|} m_\pi^A (n-1)!_\pi \quad \text{para } n = 1, \dots, d \quad (4.3)$$

$$m_n^A = \frac{(-1)^n}{d(n-1)!} \sum_{\pi \in \mathcal{P}(n)} (-1)^{|\pi|} a_\pi^A i!_\pi (|\pi| - 1)! \quad \text{para } n = 1, \dots, d. \quad (4.4)$$

Demostración. Recordemos que $m_k^A = \frac{r_1^k + r_2^k + \dots + r_d^k}{d}$ y que $a_k^A := \sum_{sym} r_1 r_2 \cdots r_k$ para $k = 1, \dots, d$, entonces, si definimos

$$b'_k = d m_k^A = r_1^k + r_2^k + \dots + r_d^k \quad \text{para } k = 1, \dots, d$$

tendremos que las a^A y las b' cumplen la recursión dada por las Identidades de Newton:

$$k a_k^A = \sum_{i=1}^k (-1)^{i-1} a_{k-i}^A b'_i \quad \text{para } k = 1, \dots, d.$$

Si ahora tomamos

$$a_k = k! a_k^A \quad \text{para } k = 1, \dots, d \text{ y}$$

$$b_k = (-1)^k (k-1)! b'_k \quad \text{para } k = 1, \dots, d$$

y los sustituimos en la recursión anterior, entonces tendremos que

$$\frac{a_k}{(k-1)!} = \sum_{i=1}^k (-1)^{i-1} \frac{a_{k-i}}{(k-i)!} \frac{b_i}{(-1)^{i-1} (i-1)!} \quad \text{para } k = 1, \dots, d.$$

Si simplificamos obtenemos que

$$a_k = \sum_{i=1}^k \binom{k-1}{i-1} b_i a_{k-i} \quad \text{para } k = 1, \dots, d.$$

Y por el Lema 3.5.1 tenemos que

$$a_n = \sum_{\pi \in \mathcal{P}(n)} b_\pi \quad \text{para } n = 1, \dots, d.$$

Y podemos usar la fórmula de inversión de Möbius para obtener que

$$b_n = \sum_{\pi \in \mathcal{P}(n)} a_\pi (-1)^{|\pi|-1} (|\pi| - 1)! \quad \text{para } n = 1, \dots, d.$$

Por último hay que sustituir los valores de las a 's y las b 's y simplificar para obtener las fórmulas buscadas. \square

Ahora sí, ya que relacionamos los cumulantes finitos con las a 's y estas a su vez las relacionamos con los momentos, podemos concatenar ambas fórmulas para relacionar a los cumulantes finitos con los momentos.

Teorema 4.4.11. Sea A una matriz de $d \times d$. Sean m_n^A y κ_n^A , para $n = 1, \dots, d$, los momentos y cumulantes de la matriz A , respectivamente. Se cumplen la siguientes fórmulas (una inversa de la otra):

$$\kappa_n = \frac{d^{n-1}}{(n-1)!} \sum_{\pi \in \mathcal{P}(n)} \frac{(-1)^{|\pi|} (|\pi| - 1)!}{(d)_\pi} \sum_{\sigma \leq \pi} (-d)^{|\sigma|} m_\sigma (n-1)!_\sigma,$$

para $n = 1, \dots, d$ y

$$m_n = \frac{1}{d^{n+1} (n-1)!} \sum_{\pi \in \mathcal{P}(n)} (-1)^{|\pi|} (d)_\pi (|\pi| - 1)! \sum_{\sigma \leq \pi} (-d)^{|\sigma|} \kappa_\sigma (n-1)!_\sigma,$$

para $n = 1, \dots, d$.

Demostración. La fórmula (4.2) nos dice que

$$\begin{aligned} \kappa_n &= \frac{(-d)^n}{d(n-1)!} \sum_{\pi \in \mathcal{P}(n)} \frac{(-1)^{|\pi|} i!_\pi a_\pi (|\pi| - 1)!}{(d)_\pi} \\ &= \frac{(-d)^n}{d(n-1)!} \sum_{\pi \in \mathcal{P}(n)} \frac{(-1)^{|\pi|} (|\pi| - 1)!}{(d)_\pi} \prod_{i=1}^j |V_i|! a_{|V_i|} \end{aligned}$$

donde estamos suponiendo que $\pi = \{V_1, \dots, V_j\}$. El producto que aparece al final lo podemos escribir en términos de momentos si utilizamos la fórmula (4.3):

$$\begin{aligned} \prod_{i=1}^j |V_i|! a_{|V_i|} &= \prod_{i=1}^j (-1)^{|V_i|} \sum_{\sigma_i \in \mathcal{P}(|V_i|)} (-d)^{|\sigma_i|} m_{\sigma_i} (n-1)!_{\sigma_i} \\ &= (-1)^n \prod_{i=1}^j \sum_{\sigma_i \in \mathcal{P}(|V_i|)} (-d)^{|\sigma_i|} m_{\sigma_i} (n-1)!_{\sigma_i} \\ &= (-1)^n \prod_{i=1}^j \sum_{\sigma_i \in \mathcal{P}(|V_i|)} \prod_{V \in \sigma_i} (-d) m_{|V|} (|V| - 1)! \\ &= (-1)^n \sum_{\substack{\sigma = \sigma_1 + \dots + \sigma_j \\ \sigma_1 \in \mathcal{P}(|V_1|), \dots, \sigma_j \in \mathcal{P}(|V_j|)}} (-d)^{|\sigma|} m_\sigma (n-1)!_\sigma \\ &= (-1)^n \sum_{\sigma \leq \pi} (-d)^{|\sigma|} m_\sigma (n-1)!_\sigma. \end{aligned}$$

Por lo tanto concluimos que

$$\begin{aligned}\kappa_n &= \frac{(-d)^n}{d(n-1)!} \sum_{\pi \in \mathcal{P}(n)} \frac{(-1)^{|\pi|} (|\pi| - 1)!}{(d)_\pi} (-1)^n \sum_{\sigma \leq \pi} (-d)^{|\sigma|} m_\sigma (n-1)!_\sigma \\ &= \frac{d^{n-1}}{(n-1)!} \sum_{\pi \in \mathcal{P}(n)} \frac{(-1)^{|\pi|} (|\pi| - 1)!}{(d)_\pi} \sum_{\sigma \leq \pi} (-d)^{|\sigma|} m_\sigma (n-1)!_\sigma.\end{aligned}$$

Para la segunda fórmula debemos realizar un procedimiento análogo, sólo que ahora debemos unir la fórmula (4.4) con la fórmula (4.1). \square

Observación 4.4.12. Si en las fórmulas del lema anterior, primero sumamos sobre las particiones σ y luego sobre las particiones π podemos reescribir las fórmulas de la siguiente forma:

$$\kappa_n = \frac{d^{n-1}}{(n-1)!} \sum_{\sigma \in \mathcal{P}(n)} (-d)^{|\sigma|} m_\sigma (n-1)!_\sigma \sum_{\pi \geq \sigma} \frac{(-1)^{|\pi|} (|\pi| - 1)!}{(d)_\pi},$$

para $n = 1, \dots, d$ y

$$m_n = \frac{1}{d^{n+1} (n-1)!} \sum_{\sigma \in \mathcal{P}(n)} (-d)^{|\sigma|} \kappa_\sigma (n-1)!_\sigma \sum_{\pi \geq \sigma} (-1)^{|\pi|} (d)_\pi (|\pi| - 1)!,$$

para $n = 1, \dots, d$.

4.4.4. Análisis de las fórmulas obtenidas

Analizaremos la fórmula que escribe a los momentos en términos de los cumulantes, para ello será más útil la segunda fórmula obtenida. Antes de comenzar introduciremos la siguiente notación para simplificar las cuentas.

Definición 4.4.13. Para cada partición $\sigma \in P(n)$, donde $n \in \mathbb{N}$, denotaremos como P_σ al siguiente polinomio en d :

$$P_\sigma(d) := \sum_{\pi \geq \sigma} (-1)^{|\pi|} (d)_\pi (|\pi| - 1)!.$$

Teorema 4.4.14. Para todo $\sigma \in \mathcal{P}(n)$ tenemos que $P_\sigma(d)$ es un polinomio de grado $n + 1 - |\sigma|$ y su coeficiente principal es $\frac{(-1)^{|\sigma|} (n-1)! n_\sigma}{(n+1-|\sigma|)!}$.

Demostración. Procederemos por inducción sobre n . La base $n = 1$ es trivial. Supongamos que la afirmación es válida para todo $m < n$ y veamos que es válida para n . Digamos que $\sigma = \{V_0, V_1, \dots, V_{k-1}\}$, (esto quiere decir que $|\sigma| = k$). Supongamos que V_0 es el bloque de σ que contiene al elemento n y digamos que $|V_0| = j$. Podemos distinguir dos casos: $j = 1$ o $j > 1$.

Caso $j = 1$. Definimos $\sigma' \in \mathcal{P}(n-1)$ como $\sigma' = \{V_1, \dots, V_{k-1}\}$. Y a cada partición $\pi \geq \sigma'$ de la forma $\pi = \{U_1, \dots, U_r\}$ le asociaremos las particiones

$\pi_0, \pi_1, \dots, \pi_r \in \mathcal{P}(n)$, definidas de la siguiente forma: $\pi_0 = \{U_1, \dots, U_r, \{n\}\}$ y $\pi_i = \{U_1, \dots, U_{i-1}, U_i \cup \{n\}, U_{i+1}, \dots, U_r\}$ para $i = 1, \dots, r$.

Para las siguientes cuentas es importante recordar que $|\pi| = k$ y que $(d)_{\pi_i} = \prod_{j=1}^k \frac{d!}{(d-|V_j^i|)!}$ donde $\pi_i = \{V_1^i, \dots, V_k^i\}$ (hay que observar que $V_j^i = V_j$ si $j \neq i$ y $V_i^i = V_i \cup \{n+1\}$).

Si juntamos todas las pequeñas observaciones anteriores, obtenemos las siguientes igualdades:

$$\begin{aligned}
& \sum_{i=0}^{|\pi|} (-1)^{|\pi_i|} (d)_{\pi_i} (|\pi_i| - 1)! & (4.5) \\
& = (-1)^{|\pi_0|} (d)_{\pi_0} (|\pi_0| - 1)! + \sum_{i=1}^r (-1)^{|\pi_i|} (d)_{\pi_i} (|\pi_i| - 1)! \\
& = (-1)^{r+1} d (d)_{\pi} r! + \sum_{i=1}^r (-1)^r (d - |V_i|) (d)_{\pi} (r - 1)! \\
& = -d (-1)^r (d)_{\pi} r! + (-1)^{|\pi|} (d)_{\pi} (r - 1)! \sum_{i=1}^r (d - |V_i|) \\
& = (-1)^r (d)_{\pi} (r - 1)! (-dr) + (-1)^{|\pi|} (d)_{\pi} (r - 1)! (dr - (n - 1)) \\
& = (-1)^r (d)_{\pi} (r - 1)! (-dr + dr - (n - 1)) \\
& = (-1)^{|\pi|} (d)_{\pi} (|\pi| - 1)! (-n + 1).
\end{aligned}$$

Por otro lado, es importante observar que si $\pi \geq \sigma'$, entonces $\pi_0, \dots, \pi_r \geq \sigma$. Además, cada $\pi' \geq \sigma$ está asociado a exactamente un $\pi \geq \sigma'$. Por lo tanto tenemos que:

$$\begin{aligned}
P_{\sigma}(d) &= \sum_{\pi' \geq \sigma} (-1)^{|\pi'|} (d)_{\pi'} (|\pi'| - 1)! \\
&= \sum_{\pi \geq \sigma'} \left(\sum_{i=0}^{|\pi|} (-1)^{|\pi_i|} (d)_{\pi_i} (|\pi_i| - 1)! \right) \\
&= \sum_{\pi \geq \sigma'} \left((-1)^{|\pi|} (d)_{\pi} (|\pi| - 1)! (-n + 1) \right) \\
&= -(n - 1) \sum_{\pi \geq \sigma'} (-1)^{|\pi|} (d)_{\pi} (|\pi| - 1)! \\
&= -(n - 1) P_{\sigma'}(d).
\end{aligned}$$

Por la hipótesis de inducción sabemos que $P_{\sigma'}(d)$ es un polinomio de grado $(n - 1) + 1 - |\sigma'| = (n - 1) + 1 - (k - 1) = n + 1 - k$ y su coeficiente principal es $\frac{(-1)^{|\sigma'|} ((n-1)-1)! n_{\sigma'}}{((n-1)+1-|\sigma'|)!} = \frac{(-1)^{k-1} (n-2)! n_{\sigma}}{(n+1-k)!}$. Por lo tanto, concluimos que $P_{\sigma}(d)$ es un polinomio de grado $n + 1 - k = n + 1 - |\sigma|$ y su coeficiente

principal es

$$-(n-1) \frac{(-1)^{k-1} (n-2)! n_\sigma}{(n+1-k)!} = \frac{(-1)^{|\sigma|} (n-1)! n_\sigma}{(n+1-|\sigma|)!},$$

que es lo que queríamos demostrar.

Caso $j > 1$. A cada elemento $\pi \in \mathcal{P}(n)$ de la forma $\pi = \{U_0, U_1, \dots, U_{r-1}\}$ (donde U_0 es el bloque de π que contiene al elemento n) le asociamos el elemento $\pi' = \{U_0 \setminus \{n\}, U_1, \dots, U_{r-1}\} \in \mathcal{P}(n-1)$. Digamos que $U_\pi = U_0 \setminus \{n\}$. Es sencillo observar que $\pi \geq \sigma$ si y sólo si $\pi' \geq \sigma'$. Además, para todo $\pi \geq \sigma$ tenemos que $|\pi| = r = |\pi'|$ y que $(d)_\pi = (d - |U_\pi|)(d)_{\pi'}$. Por lo tanto tenemos que

$$\begin{aligned} P_\sigma(d) &= \sum_{\pi \geq \sigma} (-1)^{|\pi|} (d)_\pi (|\pi| - 1)! \\ &= \sum_{\pi \geq \sigma} (-1)^{|\pi'|} (d)_{\pi'} (d - |U_\pi|) (|\pi'| - 1)! \\ &= d \sum_{\pi' \geq \sigma'} (-1)^{|\pi'|} (d)_{\pi'} (|\pi'| - 1)! - \sum_{\pi' \geq \sigma'} (-1)^{|\pi'|} (d)_{\pi'} (|\pi'| - 1)! |U_\pi| \\ &= dP_{\sigma'}(d) - \sum_{\pi' \geq \sigma'} (-1)^{|\pi'|} (d)_{\pi'} (|\pi'| - 1)! |U_\pi|. \end{aligned}$$

Para poder aplicar inducción nos hace falta modificar un poco más la segunda suma, ya que aparece un factor $|U_\pi|$ que no aparece en la hipótesis de inducción. Lo que haremos será separarla en varias sumas que si conocemos. Primero recordemos que por definición $\sigma' = \{V_0 \setminus \{n\}, V_1, \dots, V_k\}$. Llamemos $V := V_0 \setminus \{n\}$. Definimos las particiones $\sigma_1, \dots, \sigma_{k-1} \in \mathcal{P}(n-1)$ de la siguiente forma:

$$\sigma_i = \{V_1, \dots, V_{i-1}, V_i \cup V, V_{i+1}, \dots, V_{k-1}\},$$

para $i = 1, \dots, k-1$. La razón de definir estas particiones es la siguiente. Si tomamos un $\pi' \geq \sigma'$ entonces sabemos que $V \subset U_\pi$ y por lo tanto debe ser de la forma $U_\pi = V \cup (\cup_{i \in I_\pi} V_i)$ donde $I_\pi \subset \{1, \dots, k-1\}$. Es sencillo observar que $i \in I_\pi$ si y sólo si $\pi' \geq \sigma_i$. Por lo tanto tendremos que:

$$|U_\pi| = |V| + \sum_{i \in I_\pi} |V_i| = |V| + \sum_{i=1}^{k-1} 1_{\pi' \geq \sigma_i} |V_i|.$$

Utilizando lo anterior podemos separar esa suma de la siguiente manera:

$$\begin{aligned}
& \sum_{\pi' \geq \sigma'} (-1)^{|\pi'|} (d)_{\pi'} (|\pi'| - 1)! |U_{\pi}| \\
&= \sum_{\pi' \geq \sigma'} (-1)^{|\pi'|} (d)_{\pi'} (|\pi'| - 1)! \left(|V| + \sum_{i=1}^{k-1} 1_{\pi' \geq \sigma_i} |V_i| \right) \\
&= |V| \sum_{\pi' \geq \sigma'} (-1)^{|\pi'|} (d)_{\pi'} (|\pi'| - 1)! + \sum_{\pi' \geq \sigma'} \left(\sum_{i=1}^{k-1} 1_{\pi' \geq \sigma_i} |V_i| (-1)^{|\pi'|} (d)_{\pi'} (|\pi'| - 1)! \right) \\
&= |V| P_{\sigma'}(d) + \sum_{i=1}^{k-1} |V_i| \left(\sum_{\pi' \geq \sigma'} 1_{\pi' \geq \sigma_i} (-1)^{|\pi'|} (d)_{\pi'} (|\pi'| - 1)! \right) \\
&= |V| P_{\sigma'}(d) + \sum_{i=1}^{k-1} |V_i| \sum_{\pi' \geq \sigma_i} (-1)^{|\pi'|} (d)_{\pi'} (|\pi'| - 1)! \\
&= |V| P_{\sigma'}(d) + \sum_{i=1}^{k-1} |V_i| P_{\sigma_i}(d).
\end{aligned}$$

Si sustituimos esta ecuación en la que ya teníamos anteriormente podemos concluir que:

$$P_{\sigma}(d) = dP_{\sigma'}(d) - \left(|V| P_{\sigma'}(d) + \sum_{i=1}^{k-1} |V_i| P_{\sigma_i}(d) \right) = (d - |V|) P_{\sigma'}(d) - \sum_{i=1}^{k-1} |V_i| P_{\sigma_i}(d).$$

Por hipótesis de inducción sabemos que $P_{\sigma'}(d)$ es un polinomio en d de grado $(n-1) + 1 - |\sigma'| = n - k$ y su coeficiente principal es

$$\frac{(-1)^{|\sigma'|} ((n-1) - 1)! n_{\sigma'}}{((n-1) + 1 - |\sigma'|)!} = \frac{(-1)^k (n-2)! \binom{j-1}{j}}{((n-1) + 1 - k)!} = \frac{(-1)^k (n-2)! n_{\sigma}}{(n-k)! j} (j-1).$$

Por lo tanto, al multiplicarlo por $d - |V|$ obtenemos un polinomio de grado $n + 1 - |\sigma|$ con el mismo coeficiente principal.

También por hipótesis de inducción sabemos que para $i = 1, \dots, k-1$ se cumple que $P_{\sigma_i}(d)$ es un polinomio de grado $(n-1) + 1 - (|\sigma_i|) = (n-1) + 1 - (k-1) = n + 1 - k$ con coeficiente principal

$$\frac{(-1)^{|\sigma_i|} ((n-1) - 1)! n_{\sigma_i}}{((n-1) + 1 - (|\sigma_i|))!} = \frac{(-1)^{k-1} (n-2)! \binom{j-1+|V_i|}{j|V_i|}}{(n+1-k)!}.$$

Por lo tanto, $P_{\sigma}(d)$ es un polinomio de grado $n + 1 - |\sigma|$ y con coeficiente

principal:

$$\begin{aligned}
& \sum_{i=1}^{k-1} |V_i| \left(\frac{(-1)^{k-1} (n-2)! \binom{n_\sigma}{j|V_i|}^{j-1+|V_i|}}{(n+1-k)!} \right) \\
&= \frac{(-1)^{k-1} (n-2)! n_\sigma}{(n+1-k)! j} \left(\sum_{i=1}^{k-1} (j-1+|V_i|) \right) \\
&= \frac{(-1)^{k-1} (n-2)! n_\sigma}{(n+1-k)! j} ((j-1)(k-1) + (n-j)).
\end{aligned}$$

Con lo anterior podemos concluir que $P_\sigma(d)$ es un polinomio de grado $n+1-|\sigma|$ y con coeficiente principal:

$$\begin{aligned}
& \frac{(-1)^k (n-2)! n_\sigma}{(n-k)! j} (j-1) - \frac{(-1)^{k-1} (n-2)! n_\sigma}{(n+1-k)! j} ((j-1)(k-1) + (n-j)) \\
&= \frac{(-1)^k (n-2)! n_\sigma}{(n+1-k)! j} (n+1-k)(j-1) + \frac{(-1)^k (n-2)! n_\sigma}{(n+1-k)! j} ((j-1)(k-1) + (n-j)) \\
&= \frac{(-1)^k (n-2)! n_\sigma}{(n+1-k)! j} ((n+1-k)(j-1) + (j-1)(k-1) + (n-j)) \\
&= \frac{(-1)^k (n-2)! n_\sigma}{(n+1-k)! j} (n(j-1) + (n-j)) \\
&= \frac{(-1)^k (n-2)! n_\sigma}{(n+1-k)! j} (j(n-1)) \\
&= \frac{(-1)^{|\sigma|} (n-1)! n_\sigma}{(n+1-|\sigma|)!},
\end{aligned}$$

que es lo que queríamos demostrar. \square

Teorema 4.4.15. Los cumulantes finitos tienden a los cumulantes libres. Es decir:

$$\lim_{d \rightarrow \infty} \kappa_\sigma^{(d)} = r_\sigma$$

para toda $\sigma \in \mathcal{P}$.

Demostración. Primero debemos modificar la fórmula de momentos en términos de cumulantes finitos utilizando la notación que introdujimos:

$$\begin{aligned}
m_n &= \frac{1}{d^{n+1} (n-1)!} \sum_{\sigma \in \mathcal{P}(n)} (-d)^{|\sigma|} \kappa_\sigma (n-1)!_\sigma P_\sigma(d) \\
&= \sum_{\sigma \in \mathcal{P}(n)} \frac{\kappa_\sigma}{d^{n+1-|\sigma|}} \frac{(-1)^{|\sigma|} (n-1)!_\sigma}{(n-1)!} P_\sigma(d)
\end{aligned} \tag{4.6}$$

Sea $Q_\sigma(d) := \frac{(n+1-|\sigma|)!}{(-1)^{|\sigma|}(n-1)!n_\sigma} P_\sigma(d)$ un polinomio en d . Por el lema anterior sabemos que $Q_\sigma(d)$ es un polinomio mónico de grado $n+1-|\sigma|$. Además

$$\begin{aligned} \frac{(-1)^{|\sigma|}(n-1)!_\sigma}{(n-1)!} P_\sigma(d) &= \frac{(-1)^{|\sigma|}(n-1)!_\sigma}{(n-1)!} \frac{(-1)^{|\sigma|}(n-1)!n_\sigma}{(n+1-|\sigma|)!} Q_\sigma(d) \\ &= \frac{n!_\sigma}{(n+1-|\sigma|)!} Q_\sigma(d). \end{aligned}$$

Entonces tenemos que:

$$\begin{aligned} m_n &= \sum_{\sigma \in \mathcal{P}(n)} \frac{\kappa_\sigma}{d^{n+1-|\sigma|}} \frac{n!_\sigma}{(n+1-|\sigma|)!} Q_\sigma(d) \\ &= \sum_{\sigma \in \mathcal{P}(n)} \frac{n!_\sigma}{(n+1-|\sigma|)!} \frac{Q_\sigma(d)}{d^{n+1-|\sigma|}} \kappa_\sigma \\ &= \sum_{\sigma \in \mathcal{NC}(n)} \frac{Q_\sigma(d)}{d^{n+1-|\sigma|}} \kappa_\sigma. \end{aligned}$$

Donde la última igualdad se obtiene al aplicar el Lema 2.5.4 para las particiones que no se cruzan y para todas las particiones. Como $Q_\sigma(d)$ es un polinomio mónico de grado $n+1-|\sigma|$, despejando $\kappa_n^{(d)}$ concluimos que:

$$\begin{aligned} \lim_{d \rightarrow \infty} \kappa_n^{(d)} &= \lim_{d \rightarrow \infty} \frac{d^n}{(d)_n} \left(m_n - \sum_{\substack{\sigma \in \mathcal{NC}(n) \\ \sigma \neq 1_n}} \frac{Q_\sigma(d)}{d^{n+1-|\sigma|}} \kappa_\sigma^{(d)} \right) \\ &= m_n - \sum_{\substack{\sigma \in \mathcal{NC}(n) \\ \sigma \neq 1_n}} \lim_{d \rightarrow \infty} \frac{Q_\sigma(d)}{d^{n+1-|\sigma|}} \kappa_\sigma^{(d)} \\ &= m_n - \sum_{\substack{\sigma \in \mathcal{NC}(n) \\ \sigma \neq 1_n}} \lim_{d \rightarrow \infty} \kappa_\sigma^{(d)} \\ &= m_n - \sum_{\substack{\sigma \in \mathcal{NC}(n) \\ \sigma \neq 1_n}} r_\sigma \\ &= r_n. \end{aligned}$$

Lo anterior prueba el resultado para toda $n \in \mathbb{N}$. Finalmente, por la multiplicatividad de los cumulantes se sigue el resultado para cualquier partición. \square

Bibliografía

- [AP16] Octavio Arizmendi and Daniel Perales. Cumulants for finite convolution. Preprint (2016).
- [Ari08] Octavio Arizmendi. *Divisibilidad Infinita Libre de Medidas de Probabilidad*. PhD thesis, Tesis de Licenciatura en Matemáticas, Universidad de Guanajuato, 2008.
- [BB09] Julius Borcea and Petter Brändén. The lee-yang and pólya-schur programs. ii. theory of stable polynomials and applications. *Communications on Pure and Applied Mathematics*, 62(12):1595–1631, 2009.
- [DRS72] Peter Doubilet, Gian-Carlo Rota, and Richard Stanley. On the foundations of combinatorial theory (vi): The idea of generating function. In *Sixth Berkeley symposium on mathematical statistics and probability*, volume 2, pages 267–318, 1972.
- [FK52] Bent Fuglede and Richard V Kadison. Determinant theory in finite factors. *Annals of Mathematics*, pages 520–530, 1952.
- [HS11a] Takahiro Hasebe and Hayato Saigo. Joint cumulants for natural independence. *Electron. Commun. Probab*, 16:491–506, 2011.
- [HS11b] Takahiro Hasebe and Hayato Saigo. The monotone cumulants. In *Annales de l'institut Henri Poincaré (B)*, volume 47, pages 1160–1170, 2011.
- [Leh04] Franz Lehner. Cumulants in noncommutative probability theory i. noncommutative exchangeability systems. *Mathematische Zeitschrift*, 248(1):67–100, 2004.
- [Mar15] Adam W Marcus. Polynomial convolutions and (finite) free probability. 2015.
- [MSS15a] Adam Marcus, Daniel A Spielman, and Nikhil Srivastava. Finite free convolutions of polynomials. *arXiv preprint arXiv:1504.00350*, 2015.

- [MSS15b] Adam Marcus, Daniel A Spielman, and Nikhil Srivastava. Interlacing families i: Bipartite ramanujan graphs of all degrees. volume 128, pages 307–325, 2015.
- [MSS15c] Adam Marcus, Daniel A Spielman, and Nikhil Srivastava. Interlacing families ii: Mixed characteristic polynomials and the kadison-singer problem. *Annals of Mathematics*, 128(1):327–350, 2015.
- [MSS15d] Adam W Marcus, Daniel A Spielman, and Nikhil Srivastava. Interlacing families iv: Bipartite ramanujan graphs of all sizes. In *Foundations of Computer Science (FOCS), 2015 IEEE 56th Annual Symposium on*, pages 1358–1377. IEEE, 2015.
- [Mur00] Naofumi Muraki. Monotonic convolution and monotonic lévy-hincin formula. *preprint*, 2000.
- [NS06] Alexandru Nica and Roland Speicher. *Lectures on the combinatorics of free probability*, volume 13. Cambridge University Press, 2006.
- [Sta12] Richard P Stanley. *Enumerative Combinatorics, Volume 1*. Cambridge University Press, 2012.
- [SW97] Roland Speicher and Reza Woroudi. Boolean convolution. *Fields Inst. Commun*, 12:267–279, 1997.
- [Voi85] Dan Voiculescu. Symmetries of some reduced free product c^* -algebras. In *Operator algebras and their connections with topology and ergodic theory*, pages 556–588. Springer, 1985.