



CIMAT

Centro de Investigación en Matemáticas, A.C.

FAMILIAS DE POLINOMIOS QUE SE ENTRELAZAN

T E S I S

Que para obtener el grado de

Maestro en Ciencias

con Orientación en

Probabilidad y Estadística

Presenta

Daniel Perales Anaya

Director de Tesis:

Dr. Octavio Arizmendi Echegaray

Guanajuato, Gto., 25 de mayo de 2018



CIMAT

Centro de Investigación en Matemáticas, A.C.

FAMILIAS DE POLINOMIOS QUE SE ENTRELAZAN

T E S I S

Que para obtener el grado de

Maestro en Ciencias

con Orientación en

Probabilidad y Estadística

Presenta

Daniel Perales Anaya

Director de Tesis:

Dr. Octavio Arizmendi Echegaray

Autorización de la versión final

Guanajuato, Gto., 25 de mayo de 2018

Para mi familia, Guillermo e Isabel,

Ana, Lulú, Raquel y Lorena.

Agradecimientos

Es complicado poder mencionar a la gran cantidad de personas que directa o indirectamente influyeron en la realización de este trabajo. Agradezco inmensamente a mis padres, Guillermo Perales e Isabel Anaya, por el gran apoyo y respaldo que siempre me han brindado, así como a mis hermanas Ana, Lulú, Raquel y Lorena. Agradezco también a Karina, quien a lo largo de estos dos años siempre me ha motivado a dar lo mejor de mí en todos mis deberes y a la vez siempre sabe cómo distraerme del trabajo y relajarme cuando me siento abrumado.

Agradezco a todos los entrenadores, participantes, compañeros y amigos de la Olimpiada Mexicana de Matemáticas donde descubrí mi amor por las matemáticas, en especial quisiera mencionar a Ian, Jorge, Fernandez y Lalo que siempre me apoyan y me motivan a seguir adelante en este viaje a través de las matemáticas. A la Dra. Radmila Bulajich y al Dr. Rogelio Valdez les agradezco por su gran apoyo durante mi participación en la Olimpiada y a la Dra. Elizabeth Romo Medrano por el apoyo durante mi educación básica.

Quiero agradecer a mi roomie Melida por toda su ayuda en diversos problemas derivan de vivir sólo y lejos de casa. Agradezco también a todos mis compañeros de la maestría Sonny, Alberto, Mafer, Lemus, Daniel, Chapingo, Javier, Adrián, Camilo, Santi, Sebas, Nigel, Melina y Hermes. La colaboración y discusión que mantuve con ustedes fue fundamental para afrontar todos los proyectos y tareas con los que sufrimos estos dos años. Sin duda todos ustedes hicieron que mi estancia en Guanajuato fuera más placentera. Agradezco a todos los profesores que a lo largo de estos dos años lograron transmitirme un poco de su experiencia, en especial al Dr. Víctor Pérez-Abreu, la importancia que le da a sus alumnos y su dedicación al dar clase siempre me motivaron a aprender.

Le agradezco a los sinodales, Dr. Víctor Pérez-Abreu y Dr. Carlos Vargas, por su disposición para revisar la tesis y sus valiosos comentarios que ayudaron a pulir este trabajo. Agradezco al CIMAT por su gran labor en el desarrollo de las matemáticas en México y por apoyarme para asistir a dos seminarios internacionales, uno de ellos involucrado directamente con este trabajo. También quiero agradecer a todo el personal de CIMAT cuya labor diaria creó el ambiente perfecto para aprovechar al máximo mi tiempo en este centro. Agradezco a CONACYT por otorgarme la beca nacional durante los casi dos años de maestría, sin este apoyo habría sido muy complicado costear los gastos que implica un posgrado.

Por último, agradezco especialmente a mi asesor de tesis, Dr. Octavio Arizmendi Echegaray, por todas sus sugerencias y aportaciones para esta tesis, así como por su exigencia para siempre hacer el mejor trabajo posible. Gracias también por el tiempo que le dedico a nuestros trabajos de investigación, por los consejos a lo largo de la maestría y por el inmenso apoyo, interés y tiempo que le dedicó a ayudarme con mi solicitud al doctorado.

Índice general

1. Familias de Polinomios que se Entrelazan y Estabilidad Real	3
1.1. Familias de Polinomios que se Entrelazan	3
1.2. Estabilidad Real	10
1.2.1. Definición y ejemplos	11
1.2.2. Operaciones que preservan estabilidad	12
1.2.3. Construcción de familias que se entrelazan	15
2. Aplicaciones	21
2.1. Gráficas de Ramanujan Bipartitas de Todos los Grados . . .	21
2.1.1. Introducción	22
2.1.2. Preliminares	24
2.1.3. Resultado principal	29
2.2. Solución del Problema de Kadison-Singer	31
2.2.1. Introducción	31
2.2.2. Resultado principal	32
2.2.3. Demostración del resultado principal	35
2.3. Invertibilidad Restringida	39
2.3.1. Introducción	39
2.3.2. El caso isotrópico con remplazo	41
2.3.3. El caso no isotrópico y la 4-norma de Schatten	43
2.3.4. El caso isotrópico sin remplazo	46
3. Convoluciones de Polinomios y Gráficas de Ramanujan	51
3.1. Convoluciones de Polinomios	51
3.1.1. Fórmulas de Convolución	53
3.1.2. Transformadas y Cotas	60
3.2. Gráficas de Ramanujan bipartitas de todos los tamaños . . .	68
3.2.1. Entrelazado para Permutaciones	69
3.2.2. Cuadratura	74

Introducción

La técnica de familias de polinomios que se entrelazan fue introducida en el año 2013 por Marcus, Spielman y Srivastava con el artículo en proceso “Interlacing Families I” [MSS13] y una versión preliminar de “Interlacing Families II” que subieron al arXiv. Posteriormente estos dos artículos fueron publicados en la prestigiosa revista *Annals of Mathematics* [MSS15b], [MSS15c], y se publicó el artículo “Interlacing Families IV” [MSS15d] junto con un artículo sobre convolución de polinomios [MSS15a] en donde se desarrollan herramientas utilizadas en el artículo anterior. Por último, a finales de 2017 se subió al arXiv el artículo que completa la serie “Interlacing Families III” [MSS17]. En los cuatro artículos de la serie, los autores utilizan la técnica de familias de polinomios que se entrelazan para obtener resultados importantes, entre ellos destacan los dos primeros artículos de la serie. En el primero se demuestra la existencia de gráficas de Ramanujan de todos los órdenes, un resultado importante en Teoría de Gráficas. Y en el segundo resuelven el problema de Kadison-Singer, un resultado que tiene aplicaciones en muchas ramas importantes de las matemáticas. Debido a su relevancia, en 2014 los mismos autores elaboraron un artículo [MSS14] en el cual hacen un resumen de la técnica de familias de polinomios que se entrelazan aplicada a polinomios característicos y cómo la utilizaron.

La novedosa herramienta de familias que se entrelazan ha probado ser muy útil. Debido a que es un trabajo muy reciente, es posible que sea útil en la solución de otros problemas complicados. Por esta razón, es muy importante poder comprender a profundidad las razones por las que esta herramienta funciona.

El principal aporte de este trabajo es precisamente profundizar en la técnica de familias de polinomios que se entrelazan y presentar de manera ordenada las ideas importantes de los cuatro artículos de la serie “Interlacing Families” desarrollada por Marcus, Spielman y Srivastava. Para lograr esto, dedicamos un capítulo a las ideas comunes y similares de estos cuatro artículos y posteriormente dedicamos dos capítulos a especializarnos en cada una de las diversas áreas donde se ha aplicado la técnica. Se espera que con este trabajo, las siguientes generaciones puedan asimilar más rápido la técnica de familias de polinomios que se entrelazan y sean capaces de aplicarla sistemáticamente en la solución de otros problemas.

En el primer capítulo presentaremos el método de familias de polinomios que se entrelazan, que es la idea novedosa detrás de las aplicaciones del capítulo 2 y 3. Para poder utilizar este método, es necesario demostrar que ciertas familias cumplen con esta propiedad y para ello utilizaremos la teoría de estabilidad real, que es el contenido de la segunda sección del primer capítulo.

En el segundo capítulo hablaremos de los tres primeros artículos de la serie de polinomios que se entrelazan, está dividido en tres secciones, una por cada artículo. En cada sección repasaremos los preliminares necesarios para presentar los resultados principales y esbozaremos la mayoría de las demostraciones. A grandes rasgos, la idea detrás de las demostraciones siempre es reducir el problema original al problema de analizar cierto polinomio y luego acotar sus raíces. En el primer artículo [MSS15b] se demuestra la existencia de familias infinitas de gráficas de Ramanujan regulares (e irregulares) bipartitas para cada grado mayor a 2. En la demostración se utiliza una variante de la conjetura de Bilu y Linial sobre la existencia de 2-levantamientos adecuados para cada gráfica. En el segundo artículo [MSS15c] se demuestran dos conjeturas, las cuales ya se sabía que implicaban una solución positiva del problema de Kadison-Singer. La primera es la conjetura de Weaver KS_2 , y la segunda es una formulación de Casazza, Edidin, Kalra y Paulsen [CEKP07] de la conjetura de adoquinado para la cual se obtienen cotas explícitas. En el tercer artículo de la serie [MSS17], primero se deriva una prueba simple del principio de invertibilidad restringida de Bourgain y Tzafriri, y posteriormente se mejora la cota de dos maneras posibles.

Por último, en el capítulo 3 primero hablaremos sobre ciertas convoluciones de polinomios y repasaremos las cotas obtenidas en [MSS15a] para estas convoluciones. Con las herramientas desarrolladas en ese artículo seremos capaces de demostrar la existencia de multigráficas de Ramanujan, regulares y bipartitas para cada grado y cada cantidad de vértices par mayor a 2. Esto último es el contenido del cuarto artículo de la serie [MSS15d].

Capítulo 1

Familias de Polinomios que se Entrelazan y Estabilidad Real

En este capítulo presentaremos el método de familias de polinomios que se entrelazan, que será la idea importante detrás de las aplicaciones del siguiente capítulo. Para poder utilizar este método es necesario demostrar que ciertas familias cumplen con esta propiedad y para ello utilizaremos la teoría de estabilidad real, la cual presentaremos en la segunda sección.

1.1. Familias de Polinomios que se Entrelazan

En esta sección veremos el método de familias de polinomios que se entrelazan, este método le da nombre a la serie de cuatro artículos publicada por Marcus, Spielman y Srivastava que se mencionaron en la Introducción. La idea importante es bastante sencilla y es por eso que podemos presentarla desde un principio sin necesidad de ningún preliminar. A pesar de ser muy sencilla, es fundamental para resolver algunos problemas importantes, los cuales veremos en el siguiente capítulo.

Primero daremos la definición de polinomios con entrelazado común y una propiedad que satisfacen. Posteriormente, veremos un teorema que nos permite encontrar de forma sistemática familias de polinomios que se entrelazan. Por último, generalizaremos esta propiedad para familias de polinomios que se entrelazan.

Notación 1.1.1. Dado un polinomio f con raíces reales, denotaremos por $\lambda_k(f)$ a la k -ésima raíz más grande de f . Usaremos $\lambda_{\max}(f)$ y $\lambda_{\min}(f)$ para referirnos a la mayor y menor raíz de f , respectivamente.

Definición 1.1.2 (Entrelazado y entrelazado común). Para $\alpha_1, \dots, \alpha_{n-1}, \beta_1, \dots, \beta_n \in \mathbb{R}$ decimos que el polinomio $f(x) = \prod_{i=1}^{n-1} (x - \alpha_i)$ *entrelaza* al

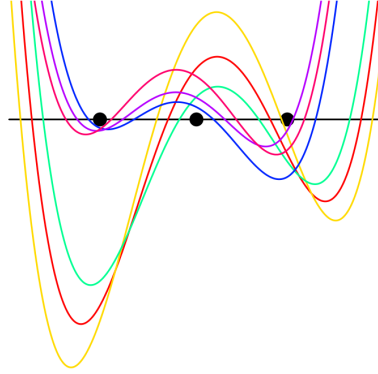


Figura 1.1: Polinomios con un entrelazado común. Todos entrelazan al polinomio con raíces en los puntos negros.

polinomio $g(x) = \prod_{i=1}^n (x - \beta_i)$ si

$$\beta_n \leq \alpha_{n-1} \leq \beta_{n-1} \leq \cdots \leq \alpha_1 \leq \beta_1.$$

La misma definición aplica si omitimos la raíz β_n (g de grado $n-1$). Decimos que $f_1, \dots, f_m$ tienen un *entrelazado común* si existe un polinomio g que entrelace a cada uno de ellos.

Teorema 1.1.3. Supongamos que $f_1, \dots, f_m$ son polinomios de grado n con todas sus raíces reales y coeficiente principal positivo, y sea f una combinación convexa de ellos, es decir, $f = \alpha_1 f_1 + \cdots + \alpha_m f_m$ para $\alpha_1, \dots, \alpha_m \geq 0$. Denotemos por

$$L_k = \min_{1 \leq j \leq m} \lambda_k(f_j) \quad y \quad U_k = \max_{1 \leq j \leq m} \lambda_k(f_j)$$

al mínimo y máximo, respectivamente, de las raíces k -ésimas de los polinomios. Si $f_1, \dots, f_m$ tienen un entrelazado común, entonces

$$L_k \leq \lambda_k(f) \leq U_k, \quad \forall k = 1, \dots, n.$$

En particular, par cada $k = 1, \dots, n$, existen $i, j \in \{1, \dots, m\}$ tales que

$$\lambda_k(f_i) \leq \lambda_k(f) \leq \lambda_k(f_j).$$

Demostración. Observemos que si $\alpha_j = 0$ para alguna j , el polinomio f_j no afectará en la suma y podemos reducir el problema al caso con $m-1$ polinomios. Por lo tanto, podemos suponer que $\alpha_1, \dots, \alpha_m > 0$. Además, si los polinomios $f_1, \dots, f_m$ tienen una raíz en común, entonces podemos

tomar $f_0(x)$ el máximo común divisor de los polinomios y considerar los siguientes polinomios

$$\hat{f}_i(x) = \frac{f_i(x)}{f_0(x)}.$$

De esta forma los polinomios $\hat{f}_1, \dots, \hat{f}_m$ no tendrán raíces en común y siguen cumpliendo con tener un entrelazado común, por lo que podemos probar el resultado para estos nuevos polinomios y al final multiplicar por $f_0(x)$, esto último le agrega las mismas raíces a los polinomios y a f . Entonces, podemos suponer que los polinomios no tienen raíces comunes. Ahora, como los polinomios $f_1, \dots, f_m$ tienen un entrelazado común, entonces tenemos que

$$L_m \leq U_m \leq L_{m-1} \leq U_{m-1} \leq \dots \leq L_1 \leq U_1.$$

Del hecho de que los polinomios tienen coeficiente principal positivo, obtenemos que para $j = 1, \dots, n$,

$$f_j(U_k) \geq 0, \quad f_j(L_k) \leq 0, \quad \text{si } k \text{ es impar,}$$

$$f_j(U_k) \leq 0, \quad f_j(L_k) \geq 0, \quad \text{si } k \text{ es par.}$$

Y como $\alpha_1, \dots, \alpha_m > 0$ entonces

$$f(U_k) \geq 0, \quad f(L_k) \leq 0, \quad \text{si } k \text{ es impar,}$$

$$f(U_k) \leq 0, \quad f(L_k) \geq 0, \quad \text{si } k \text{ es par.}$$

Por el teorema del valor intermedio concluimos que f tiene raíces en los intervalos (L_k, U_k) para todo $k = 1, \dots, n$. Y como a lo más tiene n raíces, entonces esas son todas sus raíces y concluimos que $L_k \leq \lambda_k(f) \leq U_k$. $\square$

Observación 1.1.4. 1. Como f tiene la libertad de ser cualquier combinación convexa podemos elegirlo de diversas formas.

- a) Podemos tomar simplemente la suma $f = f_1 + \dots + f_m$, es decir, $\alpha_1 = \dots = \alpha_m = 1$.
- b) Podemos tomar cualquier distribución de probabilidad μ con soporte en $[m]$ y tomar los $\alpha_k = \mathbb{P}(\mu = k)$ para $k = 1, \dots, m$, (que es lo mismo que poner la restricción extra de que $\alpha_1 + \dots + \alpha_m = 1$). De esta forma, tendríamos que f es el polinomio esperado, ponderando respecto a la distribución μ , es decir,

$$f = \alpha_1 f_1 + \dots + \alpha_m f_m = \mathbb{E}_\mu[f_\mu].$$

- c) Un caso particular del anterior es tomar la distribución uniforme (que de hecho es tomar la suma y normalizarla). Este caso será bastante útil en el resto de este trabajo.

2. Si los polinomios no tienen un entrelazado común, entonces la suma podría ni si quiera tener raíces reales. Por ejemplo, podemos considerar $(x+1)(x+2) + (x-1)(x-2)$.
3. Aún si la suma de los polinomios tiene raíces reales, la conclusión del teorema anterior podría ser falsa (en el caso de la raíz más grande) si el intervalo de las raíces más grandes se traslapa con el intervalo de las segundas raíces más grandes. Por ejemplo, las raíces de la suma de $(x+5)(x-9)(x-10)$ con $(x+6)(x-1)(x-8)$ son aproximadamente -5.3, 6.4 y 7.4 y por lo tanto no se cumple el teorema.
4. La condición de entrelazado común es un poco fuerte y se podría suavizar si en la conclusión sólo nos importa la raíz k -ésima para algún k fijo. Entonces podríamos enunciar una versión suavizada de la siguiente forma.

Sean $f_1, \dots, f_m$ polinomios con todas sus raíces reales y coeficiente principal positivo tales que hay un intervalo que contiene a todas las k -ésimas raíces y no contiene ningún otra raíz, es decir, existe un intervalo I_k , tal que $\lambda_k(f_j) \in I_k$ para todo j y $\lambda_i(f_j) \notin I_k$ para $i \neq k$ y todo j . Entonces, se cumple que $L_k \leq \lambda_k(\mathbb{E}[f_\mu]) \leq U_k$.

Al pedir que los polinomios tengan un entrelazado común, estamos pidiendo que el intervalo I_k exista para todo k y por lo tanto tenemos la conclusión para todo k . La demostración de esta variante es exactamente igual, sólo que en este caso, basta con considerar sólo el intervalo (L_k, U_k) .

Si nos fijamos cuidadosamente, la definición de entrelazado común nos pide que las raíces de los polinomios involucrados tengan un orden bastante específico. Entonces, para corroborar que en verdad se cumple esta propiedad necesitaríamos conocer a la perfección la posición de las raíces de todos los polinomios, lo cual difícilmente ocurre en los polinomios con los que usualmente nos encontramos en distintos problemas. Por lo tanto, nos gustaría tener alguna forma sencilla y sistemática de encontrar polinomios con entrelazado común, que es lo que haremos a continuación. Observemos que el lema anterior implica que toda combinación convexa de polinomios con entrelazado común tiene raíces reales. En el siguiente lema veremos un con-verso del anterior en ese sentido, ya que nos dice que si toda combinación convexa tiene raíces reales, entonces los polinomios tienen un entrelazado común. Este resultado ya ha aparecido varias veces en la literatura, entre ellas en el Teorema 2.1 de [Ded92], Teorema 2' de [Fel80] y Teorema 3.6 de [CS07]. Aquí presentaremos la versión que aparece en la mayoría de los artículos de Marcus, Spielman y Srivastava junto con su demostración.

Lema 1.1.5. Sean $f_1, \dots, f_m$ polinomios de grado n con coeficiente principal positivo. Entonces, $f_1, \dots, f_m$ tienen un entrelazado común si y sólo si $\sum_{i=1}^m \alpha_i f_i$ tiene todas sus raíces reales para cualesquiera $\alpha_1, \dots, \alpha_m \geq 0$.

Demostración. La ida se sigue directamente del Teorema 1.1.3. Para el regreso primero veamos que si f_i y f_j tienen un entrelazado común para cada $1 \leq i < j \leq m$, entonces $f_1, \dots, f_m$ tienen un entrelazado común. Lo anterior se puede ver por contradicción, primero observemos que si $f_1, \dots, f_m$ no tienen un entrelazado común, entonces para algún $k = 1, \dots, n-1$ los intervalos I_k e I_{k-1} se intersectan, donde I_k es el menor intervalo que contiene a las raíces k -ésimas de los polinomios ($\lambda_k(f_j) \in I_k$ para todo j). Lo cual a su vez nos dice que existen $1 \leq i, j \leq m$ con $i \neq j$ tales que $\lambda_k(f_j) < \lambda_{k-1}(f_i)$, contradiciendo el hecho de que f_i y f_j tienen un entrelazado común para cada $1 \leq i < j \leq m$. Lo anterior nos dice que el hecho de que m polinomios tengan la propiedad de entrelazado común es equivalente a que cada pareja formada por 2 de los m polinomios tenga la propiedad de entrelazado común. Por lo tanto, basta probar el lema para el caso de dos polinomios, que por comodidad llamaremos f_0 y f_1 . Para $t \in [0, 1]$ definimos

$$f_t = (1-t)f_0 + tf_1.$$

Sin pérdida de generalidad podemos suponer que f_0 y f_1 no tienen raíces en común (de lo contrario, podemos quitarles la raíz común, aplicar el procedimiento a continuación y al final agregarles la raíz). Observemos que por continuidad, cuando t varía de 0 a 1, las raíces de f_t definen n curvas continuas, $C_1, \dots, C_n$, cada una empieza en una raíz de f_0 y termina en una raíz de f_1 . Como las combinaciones convexas tienen raíces reales entonces todas las curvas deben estar contenidas en la recta real. Por último, si alguna de las curvas contiene en su interior (digamos en $s \in (0, 1)$) a alguna raíz (digamos r) de f_0 o f_1 (digamos f_0), entonces esto se traduce en que $f_s(r) = 0$ y $f_0(r) = 0$. De lo anterior tenemos que $0 = f_s(r) = (1-s)f_0(r) + sf_1(r) = sf_1(r)$, lo cual nos dice que r también es raíz de f_1 y contradice el hecho de que no tenían raíces en común. Por lo tanto, cada una de las n curvas define un intervalo cerrado que contiene exactamente una raíz de f_0 y una raíz de f_1 (los extremos de la curva), entonces los intervalos formados por estos pares de raíces no se intersectan y concluimos que f_0 y f_1 tienen un entrelazado común. $\square$

Ya que contamos con esta caracterización de polinomios con entrelazado común, podemos concentrarnos en ver que las combinaciones convexas de polinomios tienen raíces reales. Al saber lo anterior, el Lema 1.1.5 nos dice que los polinomios tienen un entrelazado común que a su vez, gracias al Teorema 1.1.3, nos dice dónde se encuentran las raíces de las combinaciones convexas. Para ver que las combinaciones convexas tienen raíces reales utilizaremos una teoría que generaliza esta noción, conocida como estabilidad real, que será el contenido de la siguiente sección.

Para finalizar esta sección generalizaremos el Teorema 1.1.3 para que sea más fácil utilizar la técnica en las aplicaciones de los próximos capítulos. La razón por la cual el Teorema 1.1.3 por sí sólo no es de mucha utilidad, es que nos pide encontrar conjuntos de polinomios donde todos tienen un entrelazado común. Sin embargo, varios de los conjuntos de polinomios con los que nos encontraremos en los próximos capítulos, cumplen que sólo ciertos subconjuntos de ellos tienen un entrelazado común. A pesar de no tener un entrelazado común, estos conjuntos cuentan con una muy buena estructura, la cual nos permite aplicar el Teorema 1.1.3 de forma iterativa. Este tipo de conjuntos se puede definir de manera más simple utilizando una estructura de árbol, como lo haremos a continuación y que es la forma en que se presenta en [MSS17]. Otra forma de presentar esto es utilizando sucesiones, como en [MSS15b] y [MSS15c], la cual analizaremos en la observación que le sigue a la definición.

Definición 1.1.6 (Familias que se entrelazan). Una *familia que se entrelaza* consiste de un árbol finito con raíz, T , donde a cada vértice v le asociamos un polinomio con raíces reales, $f_v(x) \in \mathbb{R}[x]$, con dos propiedades:

1. Todo polinomio $f_v(x)$ correspondiente a un vértice, que no es una hoja, es una combinación convexa de los polinomios correspondientes a sus hijos.
2. Para todos los vértices $v_1, v_2 \in T$ con un padre común, todas las combinaciones convexas de $f_{v_1}(x)$ y $f_{v_2}(x)$ tienen raíces reales.

Decimos que un conjunto de polinomios es una *familia que se entrelaza* si podemos identificar los polinomios con las hojas de un árbol de este tipo.

Observación 1.1.7. La definición anterior es la forma más general en que podemos definir a las familias de polinomios que se entrelazan para que sigan cumpliendo las propiedades que nos interesan. Para las aplicaciones, utilizaremos árboles bastante particulares, como familias de polinomios en las que los vértices del árbol a distancia k los indexaremos por sucesiones $(s_1, \dots, s_k) \in S_1 \times \dots \times S_k$ para $k \leq m$ donde $S_1, \dots, S_m$ son conjuntos finitos. Los hijos de un vértice $(s_1, \dots, s_k)$ serán los indexados por $(s_1, \dots, s_k, t)$ para $t \in S_{k+1}$. La raíz del árbol será la sucesión vacía y la denotaremos por $\emptyset$. Las hojas del árbol serán los vértices de la forma $(s_1, \dots, s_m) \in S_1 \times \dots \times S_m$, es decir, los que estén a distancia m . Por último, los vértices que no son hojas, serán combinaciones convexas específicas de sus hijos, por ejemplo, la suma o el promedio de ellos. La elección de la combinación convexa dependerá del resultado que queramos probar, en particular, de cómo nos gustaría que fuera $f_\emptyset$.

En la Figura 1.2 se muestran 8 polinomios de grado 4 que a pesar de no tener un entrelazado común (observar que la tercer raíz del azul es menor a la

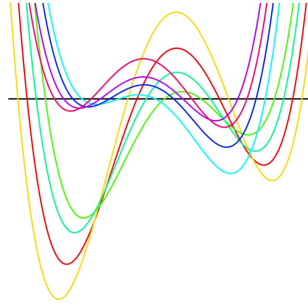


Figura 1.2: Familia de polinomios que se entrelazan sin entrelazado común.

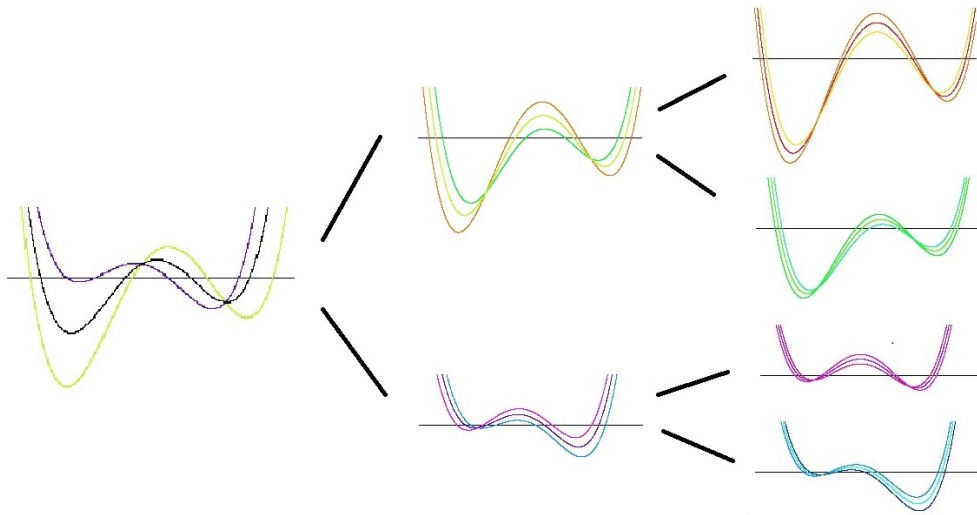


Figura 1.3: Árbol de entrelazado.

segunda raíz del verde), forman una familia de polinomios que se entrelaza. Para ver de forma más clara esto, a la derecha de la Figura 1.3 hemos dividido los 8 polinomios en 4 parejas, en cada pareja los polinomios se entrelazan y hemos sobrepuesto el promedio de ambos polinomios (cuyas raíces ya sabemos que están entre las raíces de los otros). En medio se encuentran los 4 polinomios promedio anteriores divididos en dos parejas, en cada pareja los polinomios se entrelazan y de nuevo hemos sobrepuesto el promedio de cada pareja de polinomios (cuyas raíces ya sabemos que están entre las raíces de los otros). Hasta la izquierda se encuentran estos dos últimos promedios, que también se entrelazan y sobrepusimos el promedio de ellos (que resulta ser el promedio de los 8 polinomios originales).

Ya que contamos con esta definición, podemos observar que el Teorema 1.1.3 aplica para el caso particular en el que el árbol tiene altura 1. Además,

de la definición de familia entrelazada, podemos observar que cualquier subárbol de T sigue cumpliendo ambas propiedades. Por lo tanto, si tomamos un vértice v del árbol que no es una hoja, y consideramos el subárbol con raíz en v , tendremos que los polinomios en las hojas de este árbol (que son un subconjunto de las anteriores) forma una familia que se entrelaza. Por lo tanto, podemos generalizar la propiedad que nos da el Teorema 1.1.3, si lo aplicamos iterativamente en el árbol.

Teorema 1.1.8. Sea f una familia que se entrelaza de polinomios de grado d con raíz $f_\emptyset$ y hojas $\{f_l\}_{l \in L}$. Entonces, para todos los índices $1 \leq j \leq n$, existen hojas a y b tales que

$$\lambda_j(f_a) \geq \lambda_j(f_\emptyset) \geq \lambda_j(f_b).$$

Demostración. Procedemos por inducción sobre la altura m del árbol. El caso base $m = 1$ es precisamente el Teorema 1.1.3. Ahora, fijemos un $1 \leq j \leq m$, y supongamos que ya sabemos el resultado para todo árbol con altura $k < m$ y queremos verlo para uno de altura m , primero nos fijamos en la raíz $f_\emptyset$. Por el Teorema 1.1.3 sabemos que algún hijo f_v cumple que

$$\lambda_j(f_v) \geq \lambda_j(f_\emptyset).$$

Como el subárbol con raíz f_s tiene altura menor a m , por inducción sabemos que tiene una hoja a tal que

$$\lambda_j(f_{v_a}) \geq \lambda_j(f_{v_s}) \geq \lambda_j(f_\emptyset),$$

lo cual concluye la inducción. Esto lo podemos hacer exactamente igual para todo j y de manera análoga cambiando los $\geq$ por $\leq$, con lo cual tenemos el resultado. $\square$

1.2. Estabilidad Real

El objetivo de esta sección es presentar las herramientas necesarias para demostrar que ciertas combinaciones convexas de polinomios cumplen con tener raíces reales, lo cual nos permitirá utilizar los teoremas de la sección anterior para acotar sus raíces. La forma en que procederemos será generalizar la propiedad de tener raíces reales a un concepto análogo en polinomios multivariados, conocido como estabilidad real. Encontraremos ciertos polinomios multivariados que tienen esta propiedad y algunas operaciones que la preservan. De esta manera, la receta general será encontrar un polinomio real estable multivariado y aplicarle algunas operaciones que preservan la propiedad hasta obtener un polinomio univariado, en cuyo caso la propiedad de ser real estable es equivalente a la de tener raíces reales. Primero presentaremos la teoría de estabilidad real y luego utilizaremos la receta anterior para encontrar familias de polinomios que se entrelazan.

1.2.1. Definición y ejemplos

La definición principal que utilizaremos de estabilidad real es la Definición 2.1 de [BB08].

Definición 1.2.1 (Estabilidad y Estabilidad Real). Un polinomio multivariado $f \in \mathbb{C}[z_1, \dots, z_n]$ es *estable* si no tiene raíces con todas sus coordenadas estrictamente en el plano superior, es decir, si

$$\operatorname{Im}(z_i) > 0 \quad \forall 1 \leq i \leq n \quad \Rightarrow \quad f(z_1, \dots, z_n) \neq 0.$$

Decimos que f es *real estable* si tiene coeficientes reales ($f \in \mathbb{R}[z_1, \dots, z_n]$). Denotaremos por $\mathcal{H}_n(\mathbb{C})$ y $\mathcal{H}_n(\mathbb{R})$ al conjunto de polinomios en n variables estables y reales estables respectivamente.

La principal razón de utilizar la definición anterior es que directamente podemos ver que un polinomio real estable univariado tiene todas sus raíces reales.

Observación 1.2.2. Si $f \in \mathcal{H}_1(\mathbb{R})$ esto quiere decir que f tiene coeficientes reales y no tiene raíces en el semiplano superior. Como f tiene coeficientes reales, es un hecho conocido que si $z \in \mathbb{C}$ es raíz compleja, entonces $\bar{z}$ también lo es, por lo tanto f tampoco tiene raíces en el semiplano inferior y concluimos que f tiene todas sus raíces reales. Es decir, todo polinomio real estable univariado tiene todas sus raíces reales.

A continuación, presentamos una forma equivalente de definir estabilidad, que resulta útil cuando se quieren encontrar polinomios multivariados real estables.

Observación 1.2.3. f es estable (respectivamente, real estable) si y sólo si para todo $\alpha \in \mathbb{R}^n$ y $v \in \mathbb{R}_+^n$ (donde $\mathbb{R}_+ = (0, \infty)$ es el conjunto de reales positivos) el polinomio univariado $f(\alpha + vt)$ es estable (respectivamente, real estable).

Ahora, presentamos algunos ejemplos sencillos de polinomios estables.

- Ejemplo 1.2.4.**
1. Los monomios $f(z_1, \dots, z_n) = \prod_{i=1}^n z_i$ son real estables, ya que al estar en el semiplano superior, los z_i tienen norma estrictamente positiva y su producto no puede ser 0.
 2. Los polinomios lineales $f(z_1, \dots, z_n) = \sum_{i=1}^n a_i z_i$ con coeficientes reales a_i del mismo signo son real estables. Al pedir que los a_i tengan el mismo signo forzamos a que la suma sea una combinación convexa de complejos en el semiplano superior o en el semiplano inferior (según sea el caso) lo cual no puede ser 0.
 3. El polinomio $f(z_1, z_2) = 1 - z_1 z_2$ es real estable, ya que el producto de dos números en el semiplano superior no puede ser positivo.

A continuación, presentamos un ejemplo bastante importante de un tipo de polinomio multivariado real estable. Este tipo de polinomios nos será bastante útil en lo que resta del trabajo debido a que tras aplicarles operaciones que preservan estabilidad real obtendremos polinomios univariados real estables que ya sabemos tienen raíces reales. La forma de plantearlo y su demostración provienen de la Proposición 2.4 de [BB08].

Proposición 1.2.5. Si $A_1, \dots, A_n$ son matrices positivas semidefinidas de $m \times m$ y B es una matriz hermitiana de $m \times m$, entonces

$$f(z_1, \dots, z_n) = \det \left(\sum_{i=1}^n z_i A_i + B \right)$$

es real estable (o el polinomio idénticamente cero).

Demostración. Supongamos primero que las matrices $A_1, \dots, A_n$ son positivas definidas. Sea $z(t) = \alpha + \lambda t$ con $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$, $\lambda = (\lambda_1, \dots, \lambda_n) \in \mathbb{R}_+^n$ y $t \in \mathbb{C}$. Observemos que $P := \sum_{j=1}^n \lambda_j A_j$ es positiva definida y por lo tanto es invertible y tiene raíz cuadrada. Entonces, si tomamos $H := B + \sum_{j=1}^n \alpha_j A_j$, que es una matriz hermitiana, tenemos que

$$\begin{aligned} f(z(t)) &= \det \left(\sum_{j=1}^n \alpha_j A_j + \sum_{j=1}^n \lambda_j t A_j + B \right) \\ &= \det(tP + H) = \det(P) \det(tI + P^{-\frac{1}{2}} H P^{-\frac{1}{2}}). \end{aligned}$$

Por lo tanto, $f(z(t))$ es un polinomio en t que es un múltiplo constante del polinomio característico de una matriz hermitiana y por lo tanto tiene todas sus raíces reales. Por la Observación 1.2.3 concluimos que f es real estable. Para el caso general de matrices positivas semidefinidas, podemos tomar el límite de matrices positivas definidas y usar el hecho de que el límite en cada restricción univariada debe tener raíces reales o ser el polinomio cero. $\square$

1.2.2. Operaciones que preservan estabilidad

Ahora veremos algunas de las operaciones que preservan estabilidad y estabilidad real, antes de eso recordemos el Teorema de Hurwitz que será utilizado para la demostración de una de ellas. La forma en que lo enunciamos es la misma que el Teorema 2.3 de [BB08]. También se pueden revisar el Teorema 2.5 y Corolario 2.6 de [Con78].

Teorema 1.2.6 (Teorema de Hurwitz). Sea D un dominio (conjunto abierto y conexo) en $\mathbb{C}^n$ y supongamos que $\{f_k\}_{k \in \mathbb{N}}$ es una sucesión de funciones analíticas que no se anulan en D y convergen a f uniformemente en subconjuntos compactos de D . Entonces f no se anula en D (o es idénticamente 0).

Las operaciones más simples que preservan estabilidad las recopilamos en el siguiente teorema.

Teorema 1.2.7. Las siguientes transformaciones lineales en $\mathbb{C}[z_1, \dots, z_n]$ preservan estabilidad, es decir, mandan polinomios estables en polinomios estables (o el polinomio 0).

1. **Permutación.** $f(z_1, \dots, z_n) \mapsto f(z_{\sigma(1)}, \dots, z_{\sigma(n)})$ para alguna permutación $\sigma : [n] \rightarrow [n]$.
2. **Escalar.** $f(z_1, \dots, z_n) \mapsto f(az_1, \dots, z_n)$ para $a > 0$.
3. **Diagonalización.** $f(z_1, \dots, z_n) \mapsto f(z_2, z_2, \dots, z_n) \in \mathbb{C}[z_2, \dots, z_n]$.
4. **Inversión.** $f(z_1, \dots, z_n) \mapsto z_1^d f(-\frac{1}{z_1}, z_2, \dots, z_n)$ donde $d = \deg_i(f)$ es el grado de z_1 en f .
5. **Evaluación.** $f(z_1, \dots, z_n) \mapsto f(a, z_2, \dots, z_n) \in \mathbb{C}[z_2, \dots, z_n]$ donde $a \in \mathcal{H} \cup \mathbb{R}$.
6. **Diferenciación.** $f(z_1, \dots, z_n) \mapsto \frac{\partial}{\partial z_1} f(z_1, \dots, z_n)$.

Además, dado que todas estas operaciones preservan los coeficientes reales entonces también preservan estabilidad real. Sólo es necesario recordar que al evaluar debemos hacerlo en $a \in \mathbb{R}$. Observemos también que utilizando la propiedad de permutación deducimos directamente que las demás propiedades se valen para todas las entradas, no sólo la primera.

Demostración.

1. Si $f(z_{\sigma(1)}, \dots, z_{\sigma(n)})$ no es estable, entonces existen entradas tales que $f(y_{\sigma(1)}, \dots, y_{\sigma(n)}) = 0$ con $\text{Im}(y_{\sigma(i)}) > 0$ para $1 \leq i \leq n$. Tomando $z_i = y_{\sigma(i)}$ tendremos que $f(z_1, \dots, z_n) = 0$ con $\text{Im}(z_i) > 0$ para $1 \leq i \leq n$. Entonces, por contrapositiva concluimos que si $f(z_1, \dots, z_n)$ es estable, entonces $f(z_{\sigma(1)}, \dots, z_{\sigma(n)})$ también.
2. Para $a > 0$ es claro que $\text{Im}(z_1) > 0$ si y sólo si $\text{Im}(az_1) > 0$ por lo que $f(z_1, \dots, z_n)$ se anula en $\mathcal{H}$ sí y sólo si $f(az_1, \dots, z_n)$ se anula en $\mathcal{H}$ y concluimos que $f(z_1, \dots, z_n)$ es real estable si y sólo si $f(az_1, \dots, z_n)$ es real estable.
3. Si $f(z_2, z_2, \dots, z_n) = 0$ con $\text{Im}(z_{\sigma(i)}) > 0$ para $2 \leq i \leq n$, entonces tomando $z_1 = z_2$ nos da que $f(z_1, \dots, z_n) = 0$ con $\text{Im}(z_i) > 0$ para $1 \leq i \leq n$, y por contrapositiva concluimos que diagonalizar preserva estabilidad.
4. Observemos que $z \mapsto -\frac{1}{z}$ preserva el semiplano superior, y por lo tanto, $f(z_1, \dots, z_n)$ se anula en $\mathcal{H}$ sí y sólo si $z_1^d f(-\frac{1}{z_1}, z_2, \dots, z_n)$ se anula en $\mathcal{H}$, con lo que concluimos que la inversión preserva estabilidad.

5. Para la evaluación, el caso en el que $a \in \mathcal{H}$ se sigue de la definición. Para el caso $a \in \mathbb{R}$ tomamos los polinomios $f_k(z_2, \dots, z_n) = f(a + \frac{1}{k}i, z_2, \dots, z_n)$. De la Observación 1.2.3, como f_k es estable entonces para todo $\alpha \in \mathbb{R}^{n-1}$ y $v \in \mathbb{R}_+^{n-1}$ el polinomio univariado $f_k(\alpha + vt)$ es estable. Si fijamos $\alpha \in \mathbb{R}^{n-1}$ y $v \in \mathbb{R}_+^{n-1}$ esto nos dice que $f_j(\alpha + vt)$ no se anulan en $\mathcal{H}$ y por el Teorema de Hurwitz con $D = \mathcal{H}$ obtenemos que el límite de los polinomios $f(\alpha + vt)$ tampoco se anula. De nuevo por la Observación 1.2.3 concluimos que $f(a, z_2, \dots, z_n)$ es real estable.
6. Por último, la diferenciación se sigue del Teorema de Gauss-Lucas que dice que si $f \in \mathbb{C}[z]$, las raíces de $f'(z)$ están en la envolvente convexa de las raíces de $f(z)$. Para demostrar este teorema podemos tomar $\lambda_1, \dots, \lambda_d \in \mathbb{C}$ las raíces de f y suponer sin pérdida de generalidad que f y f' no tienen raíces en común. Si z es raíz de f' obtenemos que

$$0 = \frac{f'(z)}{f(z)} = \sum_{i=1}^d \frac{1}{z - \lambda_i} = \sum_{i=1}^d \frac{\overline{z - \lambda_i}}{|z - \lambda_i|^2}.$$

Si tomamos el conjugado obtenemos que

$$0 = \sum_{i=1}^d \frac{z - \lambda_i}{|z - \lambda_i|^2} = z \sum_{i=1}^d \frac{1}{|z - \lambda_i|^2} - \sum_{i=1}^d \frac{\lambda_i}{|z - \lambda_i|^2},$$

de donde concluimos que

$$z = \frac{\sum_{i=1}^d \frac{|z - \lambda_i|^{-2}}{\sum_{i=1}^d |z - \lambda_i|^{-2}} \lambda_i}{1}$$

es una combinación convexa de las raíces de $f(z)$.

□

Ahora veamos un operador que preserve estabilidad real que nos será muy útil. A pesar de no ser tan elemental como los anteriores, sigue siendo sencillo.

Lema 1.2.8. Si $f(z_1, \dots, z_m)$ es estable (real estable), entonces también lo es

$$(1 - \partial_{z_i})f \quad \forall i.$$

Demostración. La demostración es similar a la propiedad de diferenciación. Podemos tomar $z_2, \dots, z_m$ números complejos fijos con parte imaginaria positiva. Entonces, el polinomio univariado $q(z_1) = p(z_1, z_2, \dots, z_m)$ es estable. Ahora, tomemos $\lambda_1, \dots, \lambda_d \in \mathbb{C}$ las raíces de q y supongamos sin pérdida de

generalidad que q y $q - q'$ no tienen raíces en común. Si z es raíz de $q - q'$, entonces obtenemos que $q(z) = q'(z)$ y por lo tanto

$$1 = \frac{q'(z)}{q(z)} = \sum_{i=1}^d \frac{1}{z - \lambda_i} = \sum_{i=1}^d \frac{\overline{z - \lambda_i}}{|z - \lambda_i|^2}.$$

Si denotamos por $S = \sum_{i=1}^d |z - \lambda_i|^{-2}$ y tomamos el conjugado de la expresión anterior, obtenemos que

$$\frac{1}{S} \sum_{i=1}^d |z - \lambda_i|^{-2} = 1 = \sum_{i=1}^d \frac{z - \lambda_i}{|z - \lambda_i|^2} = zS - \sum_{i=1}^d \lambda_i |z - \lambda_i|^{-2},$$

y concluimos que

$$z = \frac{1}{S} \sum_{i=1}^d |z - \lambda_i|^{-2} \left(\lambda_i + \frac{1}{S} \right)$$

es una combinación convexa de las raíces de q recorridas $\frac{1}{S}$ a la derecha. Como q no tenía raíces en $\mathcal{H}$, entonces $q - q'$ tampoco y por lo tanto, $(1 - \partial_{z_1})f$ es estable. El caso para $i \neq 1$ se sigue de la propiedad de permutación y el caso real estable se sigue de ver que la operación preserva los coeficientes reales. $\square$

En general, existe un teorema que caracteriza los operadores lineales que preservan estabilidad (y estabilidad real), para ello se puede consultar el Teorema 1.3 de [BB10], a continuación presentamos una versión simplificada, sobretodo en notación, de este teorema. Sea $\mathbb{C}_k[z_1, \dots, z_n]$ el espacio vectorial de polinomios complejos en $z_1, \dots, z_n$, donde cada variable tiene grado a lo más k . Diremos que una transformación lineal es no degenerada si su rango tiene dimensión al menos 2.

Teorema 1.2.9. Un operador lineal no degenerado $T : \mathcal{C}_k[z_1, \dots, z_n] \rightarrow \mathcal{C}[z_1, \dots, z_n]$ preserva estabilidad real si y sólo si el polinomio $2n$ -variado

$$G_T(z_1, \dots, z_n, w_1, \dots, w_n) := T \left[(z_1 + w_1)^k \dots (z_n + w_n)^k \right]$$

es estable, donde el operador T sólo actúa en las variables z .

1.2.3. Construcción de familias que se entrelazan

Ahora utilizaremos la teoría de estabilidad real para encontrar una familia de polinomios que se entrelazan bastante útil en las aplicaciones del siguiente capítulo. Empezaremos por ver unos lemas sencillos de álgebra lineal.

Lema 1.2.10. Si A es una matriz invertible y u, v son vectores, entonces $\det(A + uv^*) = \det(A) \det(1 + v^* A^{-1} u)$.

Demostración. Basta observar que

$$\begin{aligned}\det(I + uv^*) &= \det \left(\begin{pmatrix} I & 0 \\ v^* & 1 \end{pmatrix} \begin{pmatrix} I + uv^* & u \\ 0 & 1 \end{pmatrix} \begin{pmatrix} I & 0 \\ -v^* & 1 \end{pmatrix} \right) \\ &= \det \begin{pmatrix} I & u \\ 0 & 1 + v^*u \end{pmatrix} = 1 + v^*u,\end{aligned}$$

y por lo tanto

$$\det(A + uv^*) = \det(A) \det(I + (A^{-1}u)v^*) = \det(A)(1 + v^T(A^{-1}u)).$$

□

El lema anterior nos ayuda a demostrar el siguiente teorema que es bastante conocido, para ver su demostración referimos al lector al capítulo 6 de [Mey00] o a la sección 15.8 de [Har97].

Teorema 1.2.11 (Fórmula de Jacobi). Para una matriz cuadrada invertible A y otra matriz B de las mismas dimensiones,

$$\partial_x \det[xA + B] = \det[xA + B] \text{Tr}[A(xA + B)^{-1}].$$

En particular si evaluamos en $x = 0$ obtenemos que

$$\partial_x \det(xA + B)|_{x=0} = \det(B) \text{Tr}(AB^{-1}).$$

Por último, utilizamos el teorema anterior para derivar el siguiente lema que nos será bastante útil.

Lema 1.2.12. Para toda matriz cuadrada A y vector aleatorio v tenemos que

$$\mathbb{E} \det(A - vv^*) = (1 - \partial_t) \det(A + t\mathbb{E}vv^*)|_{t=0}.$$

Demostración. Primero supongamos que A es invertible. Por el lema del determinante de matrices, tenemos que

$$\begin{aligned}\mathbb{E} \det(A - vv^*) &= \mathbb{E} \det(A)(1 - v^*A^{-1}v) = \mathbb{E} \det(A)(1 - \text{Tr}(A^{-1}vv^*)) \\ &= \det(A)(1 - \mathbb{E} \text{Tr}(A^{-1}vv^*)) = \det(A) - \det(A) \text{Tr}(A^{-1}\mathbb{E}vv^*).\end{aligned}$$

Por otro lado, del teorema anterior tenemos que

$$(1 - \partial_t) \det(A + t\mathbb{E}vv^*) = \det(A + t\mathbb{E}vv^*) - \det(A) \text{Tr}(A^{-1}\mathbb{E}vv^*),$$

y tomando $t = 0$ terminamos.

Si A no es invertible, podemos tomar una sucesión de matrices invertibles que tienda a A y por continuidad (ya que la expresión a la que queremos llegar es un polinomio en las entradas de la matriz), concluimos que el teorema también se vale para A . □

A continuación presentamos la definición de polinomio característico mixto, cuyo nombre está inspirado en el discriminante mixto, al cual se asemeja en la definición a pesar de ser bastante distinto en su estructura.

Definición 1.2.13 (Polinomio característico mixto). Dadas $A_1, \dots, A_m$, matrices de tamaño $n \times n$, definimos su polinomio característico mixto como

$$\mu[A_1, \dots, A_m](x) = \left(\prod_{i=1}^m (1 - \partial_{z_i}) \right) \det \left(xI + \sum_{i=1}^m z_i A_i \right) \Big|_{z_1=\dots=z_m=0}.$$

Utilizando inductivamente el lema anterior podemos obtener que el polinomio característico esperado de una suma de matrices aleatorias de rango 1 es una función de la covarianza de las matrices.

Teorema 1.2.14. Sean $v_1, v_2, \dots, v_m$ vectores columna aleatorios e independientes en $\mathbb{C}^d$. Para cada i sea $A_i = \mathbb{E} v_i v_i^*$. Entonces

$$\mathbb{E} \chi \left[\sum_{i=1}^m v_i v_i^* \right] (x) = \mu[A_1, \dots, A_m](x).$$

Demostración. La idea de la demostración es probar por inducción sobre el número de vectores, k , que involucramos. Incluso podemos tomar cualquier matriz M en lugar de xI , es decir, veremos que para toda matriz M ,

$$\mathbb{E} \det \left(M - \sum_{i=1}^k v_i v_i^* \right) (x) = \left(\prod_{i=1}^k (1 - \partial_{z_i}) \right) \det \left(M + \sum_{i=1}^k z_i A_i \right) \Big|_{z_1=\dots=z_k=0}.$$

El caso base $k = 0$ es directo ya que

$$\mathbb{E} \det(M)(x) = \det(M).$$

Para el paso inductivo supongamos que el resultado se cumple para $j < k$ y veamos que se vale para k . Para ello basta hacer una cadena de igualdades donde utilizamos la independencia de los vectores, el lema anterior,

linealidad y la hipótesis de inducción:

$$\begin{aligned}
\mathbb{E} \det \left(M - \sum_{i=1}^k v_i v_i^* \right) &= \mathbb{E}_{v_1, \dots, v_{k-1}} \mathbb{E}_{v_k} \det \left(M - \sum_{i=1}^{k-1} v_i v_i^* - v_k v_k^* \right) \\
&= \mathbb{E}_{v_1, \dots, v_{k-1}} (1 - \partial_{z_k}) \det \left(M - \sum_{i=1}^{k-1} v_i v_i^* + z_k A_k \right) \Big|_{z_k=0} \\
&= (1 - \partial_{z_k}) \mathbb{E}_{v_1, \dots, v_{k-1}} \det \left((M + z_k A_k) - \sum_{i=1}^{k-1} v_i v_i^* \right) \Big|_{z_k=0} \\
&= (1 - \partial_{z_k}) \left(\prod_{i=1}^{k-1} (1 - \partial_{z_i}) \right) \det \left(M + z_k A_k + \sum_{i=1}^k z_i A_i \right) \Big|_{z_1=\dots=z_{k-1}=0} \Big|_{z_k=0} \\
&= \left(\prod_{i=1}^k (1 - \partial_{z_i}) \right) \det \left(M + \sum_{i=1}^k z_i A_i \right) \Big|_{z_1=\dots=z_k=0},
\end{aligned}$$

con lo cual terminamos la inducción y el caso $k = m$ nos da el resultado deseado. $\square$

Ahora, de la Proposición 1.2.5 y el Lema 1.2.8 se sigue el siguiente resultado.

Corolario 1.2.15. El polinomio característico mixto de matrices semidefinidas positivas tiene todas sus raíces reales.

Demostración. Primero empezamos con el polinomio multivariado

$$\det \left(xI + \sum_{i=1}^m z_i A_i \right),$$

el cual sabemos que es real estable gracias a la Proposición 1.2.5. Luego, el Lema 1.2.8 nos dice que el operador $1 - \partial$ preserva estabilidad real, y por lo tanto,

$$\left(\prod_{i=1}^m (1 - \partial_{z_i}) \right) \det \left(xI + \sum_{i=1}^m z_i A_i \right)$$

también es real estable. Por último, la propiedad de evaluación nos dice que evaluar $z_1 = \dots = z_m = 0$ preserva estabilidad real, por lo que obtenemos que

$$\mu[A_1, \dots, A_m](x) = \left(\prod_{i=1}^m (1 - \partial_{z_i}) \right) \det \left(xI + \sum_{i=1}^m z_i A_i \right) \Big|_{z_1=\dots=z_m=0},$$

es real estable, y como es un polinomio univariado, por la Observación 1.2.2 nos dice que tiene raíces reales. $\square$

Por último, usaremos lo anterior para ver que cualquier sucesión de vectores aleatorios independientes con soporte finito, $v_1, \dots, v_m$ definen una familia de polinomios que se entrelaza. Este resultado es bastante útil para dos de nuestras aplicaciones.

Teorema 1.2.16. Para $v_1, \dots, v_m$ vectores aleatorios independientes, sea l_i el tamaño del soporte, $\{w_{i,1}, \dots, w_{i,l_i}\}$, de v_i , y supongamos que toma cada valor con probabilidad $p_{i,1}, \dots, p_{i,l_i}$, respectivamente. Para $j_1 \in [l_1], \dots, j_m \in [l_m]$, definamos

$$q_{j_1, \dots, j_m} = \left(\prod_{i=1}^m p_{i,j_i} \right) \chi \left[\sum_{i=1}^m w_{i,j_i} w_{i,j_i}^* \right] (x).$$

Para $1 \leq k \leq m$ y $j_1 \in [l_1], \dots, j_k \in [l_k]$, definimos

$$q_{j_1, \dots, j_k} = \left(\prod_{i=1}^k p_{i,j_i} \right) \mathbb{E}_{v_{k+1}, \dots, v_m} \chi \left[\sum_{i=1}^k w_{i,j_i} w_{i,j_i}^* + \sum_{i=k+1}^m v_i v_i^* \right] (x).$$

Por último, definimos

$$q_\emptyset = \mathbb{E}_{v_1, \dots, v_m} \chi \left[\sum_{i=1}^m v_i v_i^* \right] (x).$$

Entonces los polinomios $q_{j_1, \dots, j_m}$ forman una familia que se entrelaza, en el sentido de la Observación 1.1.7.

Demostración. Si queremos ver que los polinomios $q_{j_1, \dots, j_m}$ forman una familia que se entrelaza, necesitamos demostrar que para cualquier asignación parcial $j_1, \dots, j_k$ (posiblemente vacía) se cumple que los polinomios

$$\{q_{j_1, \dots, j_k, t}(x)\}_{t=1, \dots, l_k}$$

tienen un entrelazado común. Por el Lema 1.1.5 sabemos que basta probar que para cualesquiera reales $\lambda_1, \dots, \lambda_{l_{k+1}}$ que sumen 1, el polinomio

$$\sum_{t=1}^{l_{k+1}} \lambda_t q_{j_1, \dots, j_k, t}(x)$$

tiene raíces reales. Para ello, podemos tomar la variable aleatoria u que toma el valor $w_{k+1,t}$ con probabilidad λ_t para $t = 1, \dots, l_{k+1}$. De esta forma tenemos que

$$\sum_{t=1}^{l_{k+1}} \lambda_t q_{j_1, \dots, j_k, t}(x) = \left(\prod_{i=1}^k p_{i,j_i} \right) \mathbb{E}_{u, v_{k+2}, \dots, v_m} \chi \left[\sum_{i=1}^k w_{i,j_i} w_{i,j_i}^* + uu^* + \sum_{i=k+2}^m v_i v_i^* \right] (x)$$

$$= \left(\prod_{i=1}^k p_{i,j_i} \right) \mu \left[w_{1,j_1} w_{1,j_1}^*, \dots, w_{k,j_k} w_{k,j_k}^*, \sum_{t=1}^{l_{k+1}} \lambda_t w_{k+1,t} w_{k+1,t}^*, \mathbb{E} v_{k+2} v_{k+2}^*, \dots, \mathbb{E} v_m v_m^* \right],$$

que es múltiplo de un polinomio característico mixto y por lo tanto tiene raíces reales gracias al Corolario 1.2.15. $\square$

Capítulo 2

Aplicaciones

En este capítulo estudiaremos los tres primeros artículos de la serie de polinomios que se entrelazan. El capítulo está dividido en tres secciones, una por cada artículo, en cada sección primero repasaremos los preliminares necesarios para presentar los resultados principales y posteriormente veremos las demostraciones. Estas demostraciones siguen dos pasos esenciales:

1. Reducir el problema original al problema de analizar el polinomio característico esperado.
2. Calcular el polinomio característico esperado y acotar sus raíces.

Para el primer paso utilizaremos las ideas del capítulo anterior. Primero veremos que los polinomios forman una familia de polinomios que se entrelazan utilizando la teoría de estabilidad real y luego aplicaremos la técnica de familias de polinomios que se entrelazan. El segundo paso varía dependiendo de cada artículo, en algunos casos utilizaremos cotas con las que ya se contaba previamente y en otros utilizaremos herramientas novedosas como una generalización multivariada del argumento barrera, el cual fue introducido por Batson, Spielman y Srivastava en [BSS12].

2.1. Gráficas de Ramanujan Bipartitas de Todos los Grados

Esta sección está basada en el primer artículo de la serie [MSS15b]. En este artículo se demuestra la existencia de familias infinitas de gráficas de Ramanujan regulares bipartitas para cada grado mayor a 2. Para ello se utiliza una variante de la conjetura de Bilu y Linial sobre la existencia de 2-levantamientos adecuados para cada gráfica. Además, se demuestra la existencia de familias infinitas de gráficas de Ramanujan “irregulares” bipartitas.

Comenzaremos con una introducción y motivación rápida, para luego plantear el problema y la solución utilizando familias de polinomios que se entrelazan.

2.1.1. Introducción

Notación 2.1.1. Dada G una gráfica con n vértices, denotaremos por A_G (o simplemente A si no hay ambigüedad) a su matriz de adyacencia, por $\chi(A_G)$ a su polinomio característico y por $\lambda_1(A_G) \geq \lambda_2(A_G) \geq \dots \geq \lambda_n(A_G)$ a sus valores propios (omitiremos el paréntesis si no hay ambigüedad).

- Observación 2.1.2.**
1. Si G es d -regular, entonces $A\mathbf{1} = d\mathbf{1}$ donde $\mathbf{1}$ es el vector de n unos y es sencillo ver que entonces $\lambda_1 = d$ (este es un valor propio trivial).
 2. Si G es bipartita, entonces sus valores propios son simétricos alrededor del cero. Si además es d -regular, entonces $\lambda_n = -d$ (este sería otro valor propio trivial).
 3. G es conexa si y sólo si $\lambda_2 < d$.

Se dice que una gráfica finita es un *expansor* si es dispersa y tiene fuertes propiedades de conexidad, las cuales se cuantifican dependiendo de los vértices o aristas en ella. Intuitivamente, una gráfica es un buen expansor si cada subconjunto de vértices que no es “muy grande” tiene una vecindad “grande”. En términos de sus valores propios, una gráfica d -regular, G , es un buen expansor si los valores absolutos de sus valores propios no triviales son pequeños respecto a d . El estudio de gráficas que son buenos expansores tiene varias aplicaciones a teoría de la complejidad, el diseño de redes computacionales robustas, y a la teoría de códigos de corrección de errores, es por ello que es importante encontrar buenos expansores. En este trabajo sólo trataremos informalmente a las gráficas expansoras ya que sólo se usan como motivación, al lector interesado en las definiciones formales y sus aplicaciones en ciencias de la computación lo referimos a [HLW06].

Ejemplo 2.1.3. Intuitivamente es claro por qué las gráficas completas son buenos expansores, ya que tienen muchas aristas. En efecto, se puede ver que K_d y $K_{d,d}$, la gráfica completa de d vértices y la gráfica bipartita completa con 2 componentes de d vértices, cada una tiene todos sus valores propios no triviales iguales a -1 y 0 , los cuales son pequeños respecto a d (cuando d crece) y por lo tanto son buenos expansores.

Para las aplicaciones que mencionamos nos gustaría buscar gráficas que tengan menos aristas y sigan siendo buenos expansores. Por ejemplo, nos gustaría construir infinitas familias de gráficas d -regulares con n vértices, en donde d sea fijo mientras que $n \rightarrow \infty$.

En 1886, Alon y Boppana vieron que el intervalo más pequeño en el cual podían caer los valores propios no triviales tenía que ser $[-2\sqrt{d-1}, 2\sqrt{d-1}]$. A continuación enunciamos el teorema formalmente, para más información sobre este resultado referimos al lector a [Alo86] y [Nil91].

Teorema 2.1.4. Para cada entero $d \geq 3$ y cada $\varepsilon > 0$, existe un n_0 tal que toda gráfica d -regular G con más de n_0 vértices tiene un valor propio no trivial que es mayor a $2\sqrt{d-1} - \varepsilon$.

Esto nos da una idea de cuáles pueden ser los mejores expansores a los que aspiramos para un d fijo y motiva la siguiente definición.

Definición 2.1.5 (Gráficas de Ramanujan). Una gráfica d -regular G es de Ramanujan si todos sus valores propios no triviales tienen valor absoluto a lo más $2\sqrt{d-1}$.

De acuerdo a lo que vimos antes para buenos expansores, este tipo de gráficas serían los mejores expansores posibles y en ello radica su importancia. A continuación presentamos un breve resumen de la historia sobre gráficas de Ramanujan:

- En 1988, Margulis [Mar88] y Lubotzky-Phillips-Sarnak [LPS88] vieron que existen infinitas sucesiones de gráficas Ramanujan para $d = p + 1$ con p primo. Para estas demostraciones se utilizaron gráficas de Cayley y análisis basados en teoría de números. En otros trabajos, [Chi92], [JL97], [Mor94] y [Piz90], se encontraron construcciones de gráficas de Ramanujan para $d = p^k + 1$ con p primo y $k \geq 1$.
- En 1994, en el Problema 10.7.3 de [Lub10], Lubotzky planteó la pregunta de si existen familias infinitas de gráficas de Ramanujan para cada grado mayor a 2.
- En 2003, Friedman [Fri03a] demostró que casi toda gráfica d -regular casi cumple la cota para ser Ramanujan. Es decir, para todo $\varepsilon > 0$, el valor absoluto de todos los valores propios no triviales de casi toda matriz d -regular suficientemente grande son a lo más $2\sqrt{d-1} + \varepsilon$.
- En 2006, Bilu y Linial [BL06] conjeturan que se pueden construir familias haciendo 2-levantamientos de gráficas, los cuales definiremos más adelante.
- En 2012, en el caso irregular, Puder [Pud15], demostró que con alta probabilidad un levantamiento de orden mayor de una gráfica G tiene nuevos valores propios acotados en valor absoluto por $\sqrt{3}\rho$, donde ρ es el radio espectral de la cubierta universal de G , la cual definiremos más adelante.

- En 2013, Marcus, Spielman y Srivastava [MSS13] demostraron que existen infinitas familias de gráficas Ramanujan bipartitas para cada $d \geq 3$. Además, demostraron la existencia de familias infinitas de gráficas de Ramanujan “irregulares” bipartitas. **Estos serán los resultados que presentaremos en esta sección.**
- En 2015, Marcus, Spielman y Srivastava [MSS15d] probaron que para cada n (par) y para cada $d \geq 3$, existe una multigráfica de Ramanujan bipartita d -regular. Este resultado lo presentaremos en el siguiente capítulo.

2.1.2. Preliminares

2-levantamientos

La estrategia que Bilu y Linial propusieron en [BL06] para construir familias infinitas de gráficas de Ramanujan era encontrar una forma de duplicar el número de vértices sin cambiar el grado ni agregar valores propios con valor absoluto mayor a $2\sqrt{d-1}$. Una operación que cumple con duplicar el número de vértices sin cambiar el grado es un 2-levantamiento.

Definición 2.1.6. Dada una gráfica $G = (V, E)$ con $V = \{v_1, \dots, v_n\}$, un 2-levantamiento de G es una gráfica $\hat{G} = (\hat{V}, \hat{E})$ con $\hat{V} = \{w_1, u_1, w_2, u_2, \dots, w_n, u_n\}$, es decir, construimos dos vértices w_i, u_i por cada vértice $v_i \in V$. A este par de vértices se le conoce como la *fibra* del vértice original. A cada arista en G le corresponden dos aristas en $\hat{G}$. Si (v_i, v_j) es una arista en E , consideramos las fibras $\{w_i, u_i\}$ y $\{w_j, u_j\}$ de v_i y v_j , respectivamente, y en $\hat{E}$ ponemos alguno de los dos pares de aristas siguientes, ya sea

$$\{(w_i, w_j), (u_i, u_j)\} \quad \text{o} \quad \{(w_i, u_j), (u_i, w_j)\}.$$

Observación 2.1.7. 1. La definición de 2-levantamiento es equivalente a decir que existe una cubierta $2 : 1$ de $\hat{G}$ a G . Donde una cubierta es un homomorfismo de gráficas $\pi : \hat{V} \rightarrow V$ que manda de forma biyectiva la estrella (conjunto de vértices vecinos) de cada vértice $\hat{v} \in \hat{V}$ a la estrella de $\pi(\hat{v})$. Si hay una cubierta (de cualquier orden) de $\hat{G}$ a G , decimos que $\hat{G}$ es un levantamiento de G y que G es un cociente de $\hat{G}$.

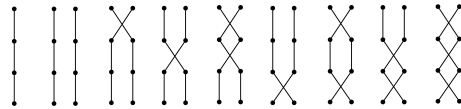


Figura 2.1: A la izquierda se encuentra un camino de 3 aristas, a la derecha sus $2^3 = 8$ posibles 2-levantamientos.

2. En total tendremos 2^m posibles 2-levantamientos de G , donde m es la cantidad de aristas de G .
3. Si en $\hat{E}$ sólo ponemos parejas del primer tipo $\{(w_i, w_j), (u_i, u_j)\}$, obtendremos dos copias disjuntas de la gráfica original.
4. Si en $\hat{E}$ sólo ponemos parejas del segundo tipo $\{(w_i, u_j), (u_i, w_j)\}$, obtendremos la doble-cubierta de G .
5. Si $\{w_i, u_i\}$ es la fibra de v_i entonces $\text{grad}(w_i) = \text{grad}(u_i) = \text{grad}(v_i)$. Por lo tanto, cualquier 2-levantamiento, $\hat{G}$, de una gráfica d -regular, G , también será d -regular.
6. Si G es bipartita, entonces cualquier 2-levantamiento, $\hat{G}$, es bipartito (basta meter en cada parte las fibras de los vértices que originalmente estaban en esa parte).
7. A cada 2-levantamiento, le podemos asignar un signado $s : E \rightarrow \{\pm 1\}$ de las aristas de G . Para ello tomamos $s(v_i, v_j) = 1$ si en $\hat{G}$ aparece la primera pareja $\{(w_i, w_j), (u_i, u_j)\}$, y tomamos $s(v_i, v_j) = -1$ si en $\hat{G}$ aparece la segunda pareja $\{(w_i, u_j), (u_i, w_j)\}$. Al tener esto podemos definir la matriz de adyacencia signada A_s como la misma que la matriz de adyacencia A de G , excepto que ahora, en la entrada correspondiente al vértice (v_i, v_j) escribimos $s(v_i, v_j)$, es decir, dejamos un 1 si aparece la primer pareja, y ponemos un -1 si aparece la segunda.
8. Lo interesante de asociar esta matriz signada a cada levantamiento es que los valores propios de un 2-levantamiento de G son la unión, con multiplicidades, de los valores propios de su matriz de adyacencia, A , y los valores propios de la matriz de adyacencia signada A_s asociada al levantamiento, (ver Lema 3.1 de [BL06]). Imitando la notación de Friedman [Fri03b], diremos que los valores propios de A son los valores propios viejos y que los valores propios de A_s son los valores propios nuevos.

Bilu y Linial demostraron que toda gráfica de grado máximo d tiene un signado, s , tal que los valores propios de A_s tienen valor absoluto a lo más $O(\sqrt{d \log^3 d})$ y conjeturaron que toda gráfica d -regular tiene un signado, s , tal que los valores propios de A_s tienen valor absoluto a lo más $2\sqrt{d-1}$.

Observación 2.1.8. Observemos que si esta conjetura es cierta, entonces podemos obtener una familia infinita de gráficas de Ramanujan d -regulares. Para ello, podemos tomar una gráfica G que sea d -regular y Ramanujan, y construir una que cumpla lo mismo y tenga el doble de vértices. Esta gráfica será precisamente el 2-levantamiento asociado al signado s que cumple la conjetura. Observemos que este 2-levantamiento tiene el doble de vértices, es d -regular y Ramanujan. Las primeras dos propiedades se siguen de la

definición y observaciones anteriores, la última se sigue de que los valores propios del 2-levantamiento son los de A y A_s , los de A_s están acotados en valor absoluto por $2\sqrt{d-1}$ gracias a la conjetura, y los valores propios no triviales de A están acotados en valor absoluto por $2\sqrt{d-1}$ por que G es Ramanujan. De esta forma, si empezamos con K_{d+1} , que sabemos es d -regular y Ramanujan, podemos utilizar la conjetura para obtener una familia infinita de gráficas de Ramanujan d -regulares. Hay que recalcar que los tamaños de las gráficas en esta familia sólo serían de la forma $2^j(d+1)$.

Utilizando la técnica de polinomios que se entrelazan podremos demostrar que toda gráfica d -regular tiene un signado, s , tal que los valores propios de A_s son a lo más $2\sqrt{d-1}$. La diferencia con la conjetura original es que no se tiene una cota por abajo (o del valor absoluto), sólo por arriba, esto tiene sentido ya que si recordamos los resultados para familias de polinomios que se entrelazan, sólo eramos capaces de controlar una raíz a la vez. Sin embargo, gracias al hecho de que los valores propios de gráficas bipartitas son simétricos respecto a 0, si acotamos sus valores propios por arriba obtenemos automáticamente una cota por abajo.

Observación 2.1.9. Como los 2-levantamientos de bipartitas siguen siendo bipartitas, podemos hacer exactamente el mismo proceso (pero empezando con $K_{d,d}$, que sabemos es d -regular, Ramanujan y bipartita), para obtener una familia infinita de gráficas de Ramanujan, bipartitas y d -regulares.

Cubierta universal

Ahora presentaremos los preliminares necesarios para demostrar la existencia de familias infinitas de gráficas de Ramanujan “irregulares” bipartitas.

Definición 2.1.10. Para $c, d \geq 2$ decimos que una gráfica bipartita es (c, d) -biregular si todos los vértice de un lado de la bipartición tienen grado c y todos los vértices del otro lado tienen grado d . La matriz de adyacencia de una gráfica (c, d) -biregular siempre tiene valores propios $\pm\sqrt{cd}$, a los cuales llamaremos valores propios triviales.

Fen y Li [FL96] (ver también [LS96]) probaron una generalización de la cota de Alon y Boppana que aplica para gráficas (c, d) -biregulares: *Para $\varepsilon > 0$, cualquier gráfica (c, d) -biregular suficientemente grande tiene un valor propio no trivial mayor a $\sqrt{c-1} + \sqrt{d-1} - \varepsilon$.* Lo anterior, motiva a decir que una gráfica (c, d) -biregular es Ramanujan si todos sus valores propios no triviales tienen valor absoluto menor o igual a $\sqrt{c-1} + \sqrt{d-1}$. También utilizaremos la técnica de familias de polinomios que se entrelazan para demostrar la existencia de familias infinitas de gráficas de Ramanujan (c, d) -biregulares para $c, d \geq 3$.

Tanto la cota de Alon y Boppana como la de Feng y Li se pueden ver como casos particulares de un fenómeno más general. Para entenderlo necesitaremos introducir el concepto de cubierta universal.

- Definición 2.1.11.** 1. Una *caminata* en una gráfica es una sucesión de vértices $(v_0, \dots, v_n)$ tal que cualquier pareja consecutiva (v_i, v_{i+1}) es una arista de la gráfica.
2. Una caminata es *simple* si no repite vértices.
3. Una caminata es de *no retroceso* si $v_{i-1} \neq v_{i+1}$ para todo i .
4. Decimos que una caminata w' es *continuación* de una caminata w si se obtiene añadiendo un vértice, digamos $w = (v_0, \dots, v_n)$ y $w' = (v_0, \dots, v_n, v_{n+1})$ para algún v_{n+1} .
5. La *cubierta universal con raíz v_0 de una gráfica G* es la gráfica T cuyos vértices son todas las caminatas de no retroceso en G que empiezan en v_0 . Dos vértices w, w' de T son adyacentes si uno es continuación del otro. Se puede ver que la cubierta universal es única salvo isomorfismo y es independiente de la elección de la raíz v_0 . También se sabe que la cubierta universal es el único árbol infinito tal que cualquier levantamiento conexo de G es cociente de T , (ver capítulo 6 de [HLW06]). *Nota: en caso de que G sea un árbol, el mismo sería su cubierta universal.*
6. Denotaremos por A_T a la matriz de adyacencia de una cubierta universal T , se puede ver que es una matriz simétrica infinita.
7. El radio espectral $\rho(T)$ de una cubierta universal T se define como

$$\rho(T) = \sup_{\|x\|_2=1} \|A_T x\|_2,$$

donde $\|x\|_2^2 := \sum_{i=1}^{\infty} x(i)^2$ cuando la serie converge.

8. El radio espectral de un árbol finito es la norma de su matriz de adyacencia.

La cota que generaliza a Alon-Boppana y Feng-Li es de Greenberg [Gre95] (ver también [Cio06]) .

Teorema 2.1.12. Para todo $\varepsilon > 0$ y toda familia infinita de gráficas con la misma cubierta universal, T , cualquier gráfica suficientemente grande tiene un valor propio no trivial mayor a $\rho(T) - \varepsilon$. Donde al igual que antes, el mayor valor propio (y el menor en el caso bipartito) se considera trivial.

El teorema anterior motiva la definición en general de una gráfica de Ramanujan (ver [Gre95] o Definición 6.7 de [HLW06]).

Definición 2.1.13. Una gráfica arbitraria *es de Ramanujan* si todos sus valores propios no triviales tienen valor absoluto menor o igual al radio espectral de su cubierta universal.

Observación 2.1.14. Como la cubierta universal de una gráfica d -regular es el árbol d -regular infinito, que se sabe tiene radio espectral $2\sqrt{d-1}$ (ver [GM88]) entonces el teorema anterior generaliza la cota de Alon y Boppana. Por otro lado, como la cubierta universal de una gráfica (c, d) -biregular es el árbol (c, d) -biregular infinito en el que los grados de los vértices alternan entre c y d entre cada nivel y como este árbol tiene radio espectral $\sqrt{c-1} + \sqrt{d-1}$ (ver [LS96]) entonces el teorema anterior generaliza la cota de Feng y Li.

Previo al artículo de Marcus, Spielman y Srivastava se habían realizado varios estudios de levantamientos aleatorios de gráficas, algunos resultados sobre el espectro se pueden encontrar en [ABG10], [AL06], [ALM02], [LR05], [LP10], [LSV11].

El polinomio de emparejamientos

En esta sección introduciremos los resultados necesarios para analizar el polinomio $f_\emptyset$ que obtendremos de ver a los polinomios característicos de las matrices signadas como una familia de polinomios que se entrelazan. Resulta que este será el polinomio de emparejamientos que ya había sido estudiado antes.

Empezaremos con la siguiente definición de Heilmann y Lieb [HL72].

Definición 2.1.15 (Polinomio de Emparejamientos). 1. Un *emparejamiento* en una gráfica es un subconjunto de sus aristas, donde no hay dos que compartan un vértice. Denotaremos por m_i al número de emparejamientos con i aristas de una gráfica G (con $m_0 = 1$).

2. Definimos el *polinomio de emparejamientos* de una gráfica G con n vértices como

$$\mu_G(x) = \sum_{i \geq 0} x^{n-2i} (-1)^i m_i.$$

El siguiente teorema de Godsil y Gutman [GG81] nos dice que el polinomio característico esperado es el polinomio de emparejamientos.

Teorema 2.1.16. Dada una gráfica G , fijemos un orden en sus vértices. Para $s \in \{\pm 1\}^m$ denotemos por A_s a la matriz de adyacencia signada correspondiente a s y denotemos por $f_s(x) = \det(xI - A_s)$ al polinomio característico de A_s . Entonces se cumple que

$$\mathbb{E}_{s \in \{\pm 1\}^m} [f_s(x)] = \mu_G(x),$$

donde suponemos que s se distribuye uniformemente en $\{\pm 1\}^m$.

Para ver una demostración simple de este teorema se puede consultar el Apéndice A de [MSS15b]. A continuación, utilizamos el concepto de árbol de caminos de Godsil [God81] y [God93] para obtener cotas de las raíces del polinomio de emparejamientos.

Definición 2.1.17. Dada una gráfica G y un vértice u , el *árbol de caminos* $P(G, u)$ contiene un vértice por cada caminata simple en G que comienza en u . Dos vértices en $P(G, u)$ son adyacentes si la caminata simple que corresponde a uno de ellos es la continuación de la caminata simple que le corresponde al otro.

Teorema 2.1.18. Sea $P(G, u)$ el árbol de caminos de G . Entonces el polinomio de emparejamientos de G divide al polinomio característico de la matriz de adyacencia de $P(G, u)$. En particular, todas las raíces de $\mu_G(x)$ son reales y tienen valor absoluto a lo más $\rho(P(G, u))$.

Del teorema anterior podemos derivar el siguiente lema.

Lema 2.1.19. Sea G una gráfica y T su cubierta universal. Entonces las raíces de $\mu_G(x)$ están acotadas en valor absoluto por $\rho(T)$.

Observación 2.1.20. El caso particular del Teorema 2.1.18, donde vemos que las raíces de $\mu_G(x)$ son reales también aparece como el Teorema 4.2 de [HL72]. El Lema 2.1.19 en el caso de una gráfica G con grado máximo d nos dice que las raíces de $\mu_G(x)$ tienen valor absoluto menor o igual a $2\sqrt{d-1}$, este resultado también se puede consultar en el Teorema 4.3 de [HL72].

2.1.3. Resultado principal

Con los preliminares que acabamos de presentar, ya estamos listos para obtener el resultado principal.

Teorema 2.1.21. Los polinomios $\{f_s\}_{s \in \{\pm 1\}^m}$ forman una familia que se entrelaza.

Demostración. Nos gustaría utilizar el Teorema 1.2.16, para poder aplicarlo diremos cuáles son los vectores columna aleatorios $v_1, \dots, v_m$. Definimos v_i como el vector aleatorio que toma los valores $v_{i,1} = e_u + e_v$ y $v_{i,-1} = e_u - e_v$ con probabilidad $\frac{1}{2}$ cada uno, donde suponemos que la arista i de G tiene como extremos a los vértices u y v , y denotemos por $e_u = (0, \dots, 0, 1, 0, \dots, 0)^T$ al vector unitario en dirección u (un 1 en la posición u y $n-1$ ceros). Entonces en el caso en que la gráfica sea d -regular tendremos que

$$A_s + Id = \sum_{i=1}^m v_{i,s_i} v_{i,s_i}^T,$$

donde $s = (s_1, \dots, s_m)$. En el caso en que la gráfica no sea d -regular, podemos suponer que d es el grado máximo y agregar los vectores $v_{m+1}, v_{m+2}, \dots, v_{m+n}$ que serán deterministas y serán de la forma $v_{m+u} = \sqrt{d - \text{grad}(u)} e_u$ para que de esta forma tengamos que

$$A_s + Id = \sum_{i=1}^m v_{i,s_i} v_{i,s_i}^T + \sum_{i=1}^n v_{m+i} v_{m+i}^T.$$

Continuaremos con el caso d -regular porque es más simple la notación, pero el otro caso es análogo. Entonces por el Teorema 1.2.16, sabemos que los polinomios

$$\begin{aligned} q_{s_1, \dots, s_m}(x) &= \frac{1}{2^m} \chi \left[\sum_{i=1}^m v_{i, s_i} v_{i, s_i}^T \right] \\ &= \frac{1}{2^m} \det(xI - A_s - Id) = \frac{1}{2^m} \det(I(x-d) - A_s) = \frac{f_s(x-d)}{2^m}, \end{aligned}$$

forman una familia que se entrelaza. Como multiplicar por 2^m no modifica las raíces y evaluar en $x+d$ en lugar de en x sólo recorre las raíces d unidades a lo largo del eje x , concluimos que los polinomios $\{f_s\}_{s \in \{\pm 1\}^m}$ forman una familia que se entrelaza. $\square$

Teorema 2.1.22. Sea G una gráfica con matriz de adyacencia A y cubierta universal T . Entonces existe un signado s de A tal que todos los valores propios de A_s son a lo más $\rho(T)$. En particular, si G es d -regular, entonces existe un signado s tal que los valores propios de A_s son a lo más $2\sqrt{d-1}$.

Demostración. Por el teorema anterior sabemos que $\{f_s\}_{s \in \{\pm 1\}^m}$ forman una familia que se entrelaza. Por el Teorema 2.1.16 sabemos que

$$f_\emptyset = \mathbb{E}_{s \in \{\pm 1\}^m} [f_s(x)] = \mu_G(x),$$

y por el Teorema 1.1.8 y el Lema 2.1.19 concluimos que existe un s tal que

$$\lambda_1(f_s) \leq \lambda_1(\mu_G) \leq \rho(T).$$

Para el caso regular observamos que la cubierta universal de una gráfica d -regular es el árbol infinito d -regular y este tiene radio espectral a lo más $2\sqrt{d-1}$. $\square$

Lema 2.1.23. Cualquier valor propio no trivial de una gráfica (c, d) -biregular completa es cero.

Demostración. La matriz de adyacencia de esta gráfica tiene rango 2, así que todos sus valores propios distintos de $-\sqrt{cd}$ y $\sqrt{cd}$ deben ser cero. $\square$

Teorema 2.1.24. Para cada $d \geq 3$, existe una sucesión infinita de gráficas d -regulares bipartitas de Ramanujan.

Demostración. El Teorema 2.1.22 prueba la conjetura de Bilu y Linial, mientras que en las Observaciones 2.1.8 y 2.1.9 ya vimos como demostrar este teorema a partir de la conjetura. $\square$

Teorema 2.1.25 (5.6 IF1). Para cada $c, d \geq 3$, existe una sucesión infinita de gráficas (c, d) -biregulares bipartitas de Ramanujan.

Demostración. La demostración de este teorema es análoga al anterior. El Lema 2.1.23 nos permite iniciar con la gráfica (c, d) -biregular completa y el Teorema 2.1.22 nos permite aplicar un 2-levantamiento de una gráfica (c, d) -biregular de Ramanujan y obtener una gráfica (c, d) -biregular con el doble de vértices (es fácil de la definición de 2-levantamiento) y que sigue siendo de Ramanujan (ya que el radio espectral de su cubierta universal es precisamente $\sqrt{c-1} + \sqrt{d-1}$). $\square$

2.2. Solución del Problema de Kadison-Singer

Esta sección está basada en el segundo artículo de la serie [MSS15c]. En este artículo, se utiliza el método de familias de polinomios que se entrelazan para demostrar dos conjeturas, las cuales ya se sabía que implicaban una solución positiva del problema de Kadison-Singer. La primera es la conjetura de Weaver KS_2 , y la segunda es una formulación de Casazza, Edidin, Kalra y Paulsen [CEKP07] de la conjetura de adoquinado para la cual se obtienen cotas explícitas.

2.2.1. Introducción

En 1959, Kadison y Singer [KS59] plantearon la siguiente pregunta fundamental.

Pregunta 2.2.1 (Problema de Kadison-Singer). ¿Todo estado puro en el álgebra (abeliana) de von Neumann, $\mathbb{D}$, de los operadores diagonales acotados en l_2 tiene una única extensión a un estado puro en $B(l_2)$, el álgebra de von Neumann de todos los operadores acotados en l_2 ?

Una respuesta positiva es equivalente a varias conjeturas (Ver [And79a], [And79b], [And81], [BT89], [CCLV05], [CT06], [Wea04]). Además, es una implicación directa de la conjetura de adoquinado con proyecciones de Ake-
mann y Anderson (Conjetura 7.1.3 de [AA91]). Previo al resultado de Marcus, Spielman y Srivastava, utilizando hipótesis más fuertes se habían encontrado soluciones parciales al problema en [BHKW88], [BT91], [BT89], [Pau11], [BD11], [Law10], [AAT12], [Pop14]. Para una discusión de la historia y otras conjeturas relacionadas ver [CFTW06].

La forma en que Marcus, Spielman y Srivastava resolvieron el problema es por medio de la demostración de la conjetura de Weaver KS_2 (ver Teorema 2 de [Wea04]) que dice lo siguiente:

Conjetura 2.2.2 (KS_2). Existen constantes universales $\eta \geq 2$ y $\theta > 0$ tales que se cumple lo siguiente. Sean $w_1, \dots, w_m \in \mathbb{C}^d$ tales que $\|w_i\| \leq 1$ para $i = 1, \dots, m$ y supongamos que

$$\sum_{i=1}^m |\langle u, w_i \rangle|^2 = \eta,$$

para todo vector unitario $u \in \mathbb{C}^d$. Entonces existe una partición S_1, S_2 de $\{1, \dots, m\}$ tal que

$$\sum_{i \in S_j} |\langle u, w_i \rangle|^2 \leq \eta - \theta,$$

para todo vector unitario $u \in \mathbb{C}^d$ y para $j \in \{1, 2\}$.

Se sabe que de esta conjetura se sigue directamente la conjetura de adoquinado con proyecciones (ver página 229 de [Wea04]). Por otro lado, veremos una prueba de la conjetura de adoquinado original de Anderson:

Conjetura 2.2.3 (Adoquinado de Anderson). Para todo $\varepsilon > 0$ existe un $r \in \mathbb{N}$ tal que para toda matriz autoadjunta compleja de $n \times n$ y con ceros en la diagonal, T , existen proyecciones diagonales $P_1, \dots, P_r$, con $\sum_{i=1}^r P_i = I$ tales que

$$\|P_i T P_i\| \leq \varepsilon \|T\| \quad \text{para } i = 1, \dots, r.$$

Una diferencia entre la conjetura de adoquinado y KS_2 es que la primera se puede extender fácilmente a operadores infinitos con un sencillo argumento de compactación (ver [And79a]) y luego se sigue inmediatamente la solución al problema de Kadison-Singer tal como se menciona en el Lema 5 artículo original [KS59]. Por otro lado, la reducción del problema a la conjetura de adoquinado con proyecciones requiere teoría de operadores no elemental.

2.2.2. Resultado principal

La clave para obtener mejores cotas a las que se tenían anteriormente, y por ende resolver el problema, radica en utilizar el método probabilista para encontrar algo con probabilidad positiva en lugar de con alta probabilidad. El resultado principal de esta sección es el siguiente.

Teorema 2.2.4. Si $\varepsilon > 0$ y $v_1, \dots, v_m \in \mathbb{C}^d$ son vectores aleatorios independientes con soporte finito tales que

$$\sum_{i=1}^m \mathbb{E} v_i v_i^* = I_d, \quad \text{y} \quad \mathbb{E} \|v_i\|^2 \leq \varepsilon, \quad \forall i,$$

entonces

$$\mathbb{P} \left[\left\| \sum_{i=1}^m v_i v_i^* \right\| \leq (1 + \sqrt{\varepsilon})^2 \right] > 0$$

Si recordamos lo que vimos en el primer capítulo, ya deberíamos tener una idea de qué es lo que haremos para demostrar el resultado. Antes de pasar a la demostración veamos cómo el teorema anterior implica el siguiente corolario, que es una generalización de la Conjetura 2.2.2. De esta forma quedará claro por qué este es justo el teorema que necesitamos.

Corolario 2.2.5. Sea r un entero positivo y sean $u_1, \dots, u_m \in \mathbb{C}^d$ vectores tales que

$$\sum_{i=1}^m \mathbb{E} u_i u_i^* = I, \quad \text{y} \quad \mathbb{E} \|u_i\|^2 \leq \delta, \quad \forall i.$$

Entonces existe una partición $\{S_1, \dots, S_r\}$ de $[m]$ tal que

$$\left\| \sum_{i \in S_j} u_i u_i^* \right\| \leq \left(\frac{1}{\sqrt{r}} + \sqrt{\delta} \right)^2, \quad \text{para } j = 1, \dots, r.$$

Demostración. Lo que queremos hacer es encontrar una partición con r bloques donde cada bloque tenga norma pequeña, y contamos con un teorema que nos dice que podemos controlar la norma con probabilidad positiva, es decir, para alguna realización. Entonces, lo que haremos será tomar todas las posibles particiones y ver que con probabilidad positiva una de ellas nos sirve. Para poder codificar las particiones tomaremos vectores de tamaño rd , d por el tamaño del vector y r la cantidad de bloques de la partición.

Para cada $i \in [m]$ y $k \in [r]$, definimos $w_{i,k} \in \mathbb{C}^{rd}$ como la suma directa de r vectores en $\mathbb{C}^d$, de los cuales todos serán vectores de ceros, que denotaremos por 0_d , excepto el k -ésimo, en donde pondremos una copia de u_i , es decir,

$$w_{i,1} = \begin{pmatrix} u_i \\ 0_d \\ \vdots \\ 0_d \\ 0_d \end{pmatrix}, \quad w_{i,2} = \begin{pmatrix} 0_d \\ u_i \\ \vdots \\ 0_d \\ 0_d \end{pmatrix}, \dots, w_{i,r} = \begin{pmatrix} 0_d \\ 0_d \\ \vdots \\ 0_d \\ u_i \end{pmatrix}.$$

El vector $w_{i,k}$ representará el evento en que el vector u_i está en el bloque k . Supondremos que cada vector u_i tiene la misma probabilidad de estar en cualquier bloque, es decir, para $i \in [m]$, elegimos v_i como el vector aleatorio que toma los valores $\sqrt{r}w_{i,1}, \sqrt{r}w_{i,2}, \dots, \sqrt{r}w_{i,r}$ cada uno con probabilidad $\frac{1}{r}$. Observemos que el factor $\sqrt{r}$ es sólo para que al tomar $u_i u_i^*$ el r que obtengamos se cancele con el $\frac{1}{r}$ de la probabilidad. Ahora, observemos que

$$\mathbb{E} v_i v_i^* = \frac{1}{r} \sum_{k=1}^r r w_{i,k} w_{i,k}^* = \begin{pmatrix} u_i u_i^* & 0_{d \times d} & \dots & 0_{d \times d} \\ 0_{d \times d} & u_i u_i^* & \dots & 0_{d \times d} \\ \vdots & \vdots & \ddots & \vdots \\ 0_{d \times d} & 0_{d \times d} & \dots & u_i u_i^* \end{pmatrix},$$

por lo que

$$\sum_{i=1}^m \mathbb{E} v_i v_i^* = I_{rd},$$

y se cumple la primera condición del problema. Además, $\|v_i\|^2 = r \|u_i\|^2 \leq r\delta$. Ahora, al utilizar el Teorema 2.2.4 con $\varepsilon = r\delta$, obtenemos que existe

una realización de los v_i que cumple la cota, y basta tomar la partición de forma que $S_k = \{i \in [m] : v_i = w_{i,k}\}$ para $k = 1, \dots, r$, es decir, S_k será precisamente el conjunto de vectores que en la realización quedaron en el bloque k . Formalmente, la cota nos dice

$$\begin{aligned} (1 + \sqrt{r}\delta)^2 &\geq \left\| \sum_{i=1}^m v_i v_i^* \right\| = \left\| \sum_{k=1}^r \sum_{i \in S_k} (\sqrt{r} w_{i,k}) (\sqrt{r} w_{i,k})^* \right\| \\ &= r \left\| \sum_{k=1}^r \sum_{i \in S_k} w_{i,k} w_{i,k}^* \right\| \geq r \left\| \sum_{i \in S_k} w_{i,k} w_{i,k}^* \right\| = r \left\| \sum_{i \in S_k} u_i u_i^* \right\| \end{aligned}$$

y concluimos que

$$\left\| \sum_{i \in S_j} u_i u_i^* \right\| \leq \left(\frac{1}{\sqrt{r}} + \sqrt{\delta} \right)^2, \quad \text{para } j = 1, \dots, r.$$

□

Observación 2.2.6. Si tomamos $r = 2$ y $\delta = \frac{1}{18}$, esto implica la Conjetura 2.2.2 para $\eta = 18$ y $\theta = 2$.

Por otro lado, el Corolario 2.2.5 también implica la Conjetura 2.2.3 con $r = (6/\varepsilon)^4$. Para demostrarlo veremos primero una definición y un teorema. Siguiendo al notación de [CEKP07] tenemos la siguiente definición.

Definición 2.2.7. Decimos que una matriz cuadrada, T , puede ser (r, ε) -adoquinada, si hay proyecciones a las coordenadas, $P_1, \dots, P_r$ tales que

$$\sum_{i=1}^r P_i = 1 \quad \text{y} \quad \|P_i T P_i\| \leq \varepsilon \|T\|, \quad \forall i \in \{1, \dots, r\}.$$

Necesitaremos también el Teorema 3 de [CEKP07], que se demuestra fácilmente con álgebra lineal elemental.

Lema 2.2.8. Supongamos que existe una función $r : \mathbb{R}_+ \rightarrow \mathbb{N}$ tal que toda matriz de proyección, Q , de tamaño $2n \times 2n$ y con todas sus entradas de la diagonal iguales a $1/2$, puede ser $(r(\varepsilon), \frac{1+\varepsilon}{2})$ -adoquinada para todo $\varepsilon > 0$. Entonces toda matriz de $n \times n$, autoadjunta y con zeros en la diagonal, T , puede ser $(r^2(\varepsilon), \varepsilon)$ -adoquinada para todo $\varepsilon > 0$.

Ya con esto, podemos demostrar el siguiente resultado, el cual es una versión cuantitativa de la Conjetura 2.2.3

Teorema 2.2.9. Para todo $\varepsilon > 0$, toda matriz compleja, autoadjunta y con zeros en la diagonal, T , puede ser (r, ε) -adoquinada con $r = (\frac{6}{\varepsilon})^4$.

Demostración. Sea Q una matriz de proyección arbitraria de tamaño $2n \times 2n$ y cuyas entradas en la diagonal son todas $\frac{1}{2}$. Entonces, $Q = (u_i^* u_j)_{i,j \in [2n]}$ es la matriz de Gram de $2n$ vectores $u_1, \dots, u_{2n} \in \mathbb{C}^n$ con $\|u_i\|^2 = \frac{1}{2} = \delta$. Si aplicamos el Corolario 2.2.5 a estos vectores para cualquier r , obtenemos una partición $S_1, \dots, S_r$ de $[2n]$. Si tomamos P_k como la proyección en los índices de S_k , tenemos que para cada $k = 1, \dots, r$

$$\|P_k Q P_k\| = \|(u_i^* u_j)_{i,j \in S_k}\| = \left\| \sum_{i \in S_k} u_i u_i^* \right\| \leq \left(\frac{1}{\sqrt{r}} + \frac{1}{\sqrt{2}} \right)^2 < \frac{1}{2} + \frac{3}{\sqrt{r}}.$$

Por lo tanto, para $r = \frac{36}{\varepsilon}$, cualquier Q puede ser $(r, \frac{1+\varepsilon}{2})$ -adoquinada. Utilizando el Lema 2.2.8 terminamos. $\square$

Una buena pregunta es si se puede optimizar la dependencia de r en ε en este teorema. En [CEKP07] se demuestra que lo más a lo cual se podría mejorar es a $r = \frac{1}{\varepsilon^2}$.

2.2.3. Demostración del resultado principal

Para demostrar el Teorema 2.2.4, observemos que de las condiciones del teorema es sencillo ver que podemos encontrar una familia de polinomios que se entrelazan, y como en todas las aplicaciones, habrá que acotar la raíz más alta del polinomio característico esperado, el cual sabemos que será el polinomio característico mixto gracias al Teorema 1.2.14. Primero nos concentraremos en encontrar una cota para este polinomio y posteriormente daremos la demostración completa del Teorema 2.2.4. El teorema que queremos demostrar es el siguiente:

Teorema 2.2.10. Supongamos que $A_1, \dots, A_m$ son matrices positivas semidefinidas con $\sum_{i=1}^m A_i = I$ y $\text{Tr}(A_i) \leq \varepsilon$. Entonces,

$$\lambda_{\max}(\mu[A_1, \dots, A_m](x)) \leq (1 + \sqrt{\varepsilon})^2.$$

Para demostrar este teorema comenzaremos por expresar a μ de otra forma.

Lema 2.2.11. Supongamos que $A_1, \dots, A_m$ son matrices hermitianas positivas semidefinidas. Si $\sum_{i=1}^m A_i = I$, entonces

$$\mu[A_1, \dots, A_m](x) = \left(\prod_{i=1}^m (1 - \partial_{z_i}) \right) \det \left(\sum_{i=1}^m z_i A_i \right) \Big|_{z_1 = \dots = z_m = x}.$$

Demostración. Sabemos que para cualquier función diferenciable es lo mismo diferenciar y evaluar $z_i + x$ que primero evaluar y luego diferenciar, es decir,

$$\partial_{y_i}(f(y_i))|_{y_i=z_i+x} = \partial_{z_i}(f(z_i+x)).$$

Entonces, basta tomar $y_i = z_i + x$ para $i = 1, \dots, m$ y obtendremos que

$$\begin{aligned} & \left(\prod_{i=1}^m (1 - \partial_{z_i}) \right) \det \left(xI + \sum_{i=1}^m z_i A_i \right) \Big|_{z_1=\dots=z_m=0} \\ &= \left(\prod_{i=1}^m (1 - \partial_{z_i}) \right) \det \left(\sum_{i=1}^m z_i A_i \right) \Big|_{z_1=\dots=z_m=x}. \end{aligned}$$

□

El lema anterior nos dice que

$$\mu[A_1, \dots, A_m](x) = Q(x, x, \dots, x),$$

donde

$$Q(z_1, \dots, z_m) = \left(\prod_{i=1}^m (1 - \partial_{z_i}) \right) \det \left(\sum_{i=1}^m z_i A_i \right).$$

Ahora nos gustaría acotar las raíces de la nueva expresión. Para ello lo mejor será hacer una cota superior multivariada de las raíces de Q . Para entender bien a qué nos referimos con esto primero definiremos el conjunto Ab_p y las funciones barrera.

Definición 2.2.12. Sea $p(z_1, \dots, z_m)$ un polinomio multivariado. Decimos que $z \in \mathbb{R}^m$ *está por encima de las raíces de p* si

$$p(z + t) > 0, \quad \forall t = (t_1, \dots, t_m) \in \mathbb{R}^m, \quad t_i \geq 0,$$

es decir, si p es positivo para todo punto con todas sus coordenadas mayores o iguales a las de z . Denotaremos por Ab_p al conjunto de puntos que están arriba de las raíces de p .

Entonces, para demostrar el Teorema 2.2.10 basta ver que $((1+\sqrt{\varepsilon})^2, \dots, (1+\sqrt{\varepsilon})^2) \in Ab_Q$. Para lograr esto hay que utilizar inductivamente el argumento barrera. En particular, hay que construir el polinomio Q aplicando operaciones de la forma $(1 - \partial_{y_i})$, para poder tener un control de cómo varían las raíces se utilizan las siguientes funciones:

Definición 2.2.13 (Función barrera). Sea p un polinomio real estable y $z = (z_1, \dots, z_m) \in Ab_p$. La *función barrera de p en dirección i evaluada en z* es

$$\Phi_p^i(z) = \frac{\partial_{z_i} p(z)}{p(z)} = \partial_{z_i} \log p(z).$$

De forma equivalente podemos definir Φ_p^i como

$$\Phi_p^i(z) = \frac{q'_{z,i}(z_i)}{q_{z,i}(z_i)} = \sum_{j=1}^r \frac{1}{z_i - \lambda_j},$$

donde la restricción univariada $q_{z,i}(t) = p(z_1, \dots, z_{i-1}, t, z_{i+1}, \dots, z_m)$ tiene raíces $\lambda_1, \dots, \lambda_r$ que sabemos son reales por la evaluación en polinomios reales estables (ver Teorema 1.2.7). Observemos que los Φ_p^i son funciones de m variables.

La propiedad importante que ocuparemos de las funciones barrera es que arriba de las raíces del polinomio estas funciones son no crecientes y convexas en cada coordenada.

Lema 2.2.14. Supongamos que p es real estable y $z \in Ab_p$. Entonces para $1 \leq i, j \leq m$ y $\delta \geq 0$, se tiene

1. **Monotonidad**, $\Phi_p^i(z + \delta e_j) \leq \Phi_p^i(z)$.
2. **Convexidad**, $\Phi_p^i(z + \delta e_j) \leq \Phi_p^i(z) + \delta \cdot \partial_{z_j} \Phi_p^i(z + \delta e_j)$.

No demostraremos este lema ya que utiliza cosas más técnicas. Para la demostración original de Marcus, Spielman y Srivastava se puede consultar el Lema 5.7 de [MSS15c], ahí los autores utilizan una caracterización de los polinomios bivariados reales estables. Esta caracterización se encuentra en el Corolario 6.7 de [BB10], y se prueba utilizando una adaptación de resultado de [HV07] que hicieron en [LPR05]. También se puede ver que el lema anterior se sigue del Corolario 4.6 de [BGLS01]. Además, Terence Tao presentó una prueba más elemental en su blog [Tao13].

Ya que tenemos estas propiedades a la mano, sólo nos falta ver la siguiente cota. Este será el último paso importante antes de poder demostrar el Teorema 2.2.10.

Lema 2.2.15. Supongamos que $p(z_1, \dots, z_m)$ es real estable, $z \in Ab_p$, y $\delta > 0$ cumple que

$$\Phi_p^j(z) \leq 1 - \frac{1}{\delta}.$$

Entonces para $1 \leq i \leq m$ se cumple que

$$\Phi_{p - \partial_{z_j} p}^i(z + \delta e_j) \leq \Phi_p^i(z)$$

Para la demostración de este lema se puede consultar el Lema 5.7 de [MSS15c]. Ya que tenemos estas herramientas, estamos listos para demostrar el Teorema 2.2.10 y posteriormente el resultado principal del artículo.

Demostración. Sean $P(y_1, \dots, y_m) = \det(\sum_{i=1}^m y_i A_i)$ y $t = \sqrt{\varepsilon} + \varepsilon$. Como las matrices A_i son positivas semidefinidas y $\det(t \sum_{i=1}^m A_i) = \det(tI) > 0$, entonces $(t, \dots, t) \in Ab_P$. Por el Teorema 1.2.11 sabemos que

$$\Phi_P^i(y_1, \dots, y_m) = \frac{\partial_i P(y_1, \dots, y_m)}{P(y_1, \dots, y_m)} = Tr \left(\left(\sum_{i=1}^m y_i A_i \right)^{-1} A_i \right).$$

Por lo tanto,

$$\Phi_P^i(t, \dots, t) = \frac{\text{Tr}(A_i)}{t} \leq \frac{\varepsilon}{t} = \frac{\varepsilon}{\varepsilon + \sqrt{\varepsilon}}.$$

Tomamos $\phi = \frac{\varepsilon}{\varepsilon + \sqrt{\varepsilon}}$,

$$\delta = \frac{1}{1 - \phi} = 1 + \sqrt{\varepsilon},$$

y para $k = 1, \dots, m$, definimos

$$P_k(y_1, \dots, y_m) = \left(\prod_{i=1}^k (1 - \partial_{y_i}) \right) P(y_1, \dots, y_m).$$

Sea $x^0 = (t, \dots, t)$ y $x^k = (t + \delta, \dots, t + \delta, t, \dots, t)$ el vector con $t + \delta$ en las primeras k coordenadas y t en el resto. Inductivamente podemos ver que $x^k \in \text{Ab}_{P_k}$ y que para $1 \leq i \leq m$ se tiene $\Phi_{P_k}^i(x^k) \leq \phi$. El caso base es $k = 0$ y ya lo hicimos, ya que $x^0 = t1 \in \text{Ab}_P$ y $\Phi_P^i(t1) \leq \phi$. Ahora, suponemos que ya tenemos el resultado para $j \leq k$ y lo demostramos para $k+1$. Observemos que P_k es real estable y por hipótesis de inducción sabemos que $x^k \in \text{Ab}_{P_k}$, $\delta > 0$ y

$$\Phi_{P_k}^{k+1}(z) \leq \phi = 1 - \frac{1}{\delta}.$$

Entonces por el Lema 2.2.15 sabemos que para $1 \leq i \leq m$ se cumple que

$$\Phi_{P_{k+1}}^i(x^{k+1}) = \Phi_{(1 - \partial_{y_{k+1}})P_k}^i(x^k + \delta e_{k+1}) \leq \Phi_{P_k}^i(x^k) \leq \phi.$$

Además, como Φ es no creciente en cada coordenada y δe_{k+1} es un vector no negativo, tenemos que

$$\frac{\partial_{y_i} P_k(x^{k+1})}{P_k(x^{k+1})} = \phi_{P_k}^i(x^k + \delta e_{k+1}) \leq \phi_{P_k}^i(x^k) < 1,$$

y por lo tanto $P_{k+1}(x^{k+1}) = (P_k - \partial_{y_i} P_k)(x^{k+1}) > 0$. De lo anterior concluimos que $x^{k+1} \in \text{Ab}_{P_{k+1}}$ y terminamos la inducción. Entonces, tenemos que $x^m = (t + \delta, \dots, t + \delta) \in \text{Ab}_{P_m}$ y como

$$P_m(x, \dots, x) = Q(x, \dots, x) = \mu[A_1, \dots, A_m](x),$$

concluimos que la raíz más grande de este último polinomio es a lo más

$$t + \delta = 1 + \sqrt{\varepsilon} + \sqrt{\varepsilon} + \varepsilon = (1 + \sqrt{\varepsilon})^2.$$

□

Ahora sí, para demostrar el Teorema 2.2.4 basta usar lo que sabemos de familias de polinomios que se entrelazan y la cota en las raíces que acabamos de obtener en el Teorema 2.2.10.

Demostración. Tomamos $A_i = \mathbb{E}v_i e_i^*$ y observamos que para todo $1 \leq i \leq m$,

$$\text{Tr}(A_i) = \mathbb{E}\text{Tr}(v_i v_i^*) = \mathbb{E}v_i^* v_i = \mathbb{E}\|v_i\|^2 \leq \varepsilon.$$

El Teorema 1.2.14 nos dice que el polinomio característico esperado de $\sum_i v_i v_i^*$ es el polinomio característico mixto $\mu[A_1, \dots, A_m](x)$ y por el Teorema 2.2.10 sabemos que la raíz más grande de este polinomio es a lo más $(1 + \sqrt{\varepsilon})^2$. Por último, el Teorema 1.2.16 nos dice que los polinomios $q_{j_1, \dots, j_m}$ forman una familia que se entrelaza y el Teorema 1.1.8 implica que existen $j_1, \dots, j_m$, tales que para esa realización

$$\left\| \sum_{i=1}^m v_i v_i^* \right\| \leq \lambda_{\max} \left(\chi \left(\sum_{i=1}^m v_i v_i^* \right) \right) \leq \lambda_{\max}(\mu[A_1, \dots, A_m](x)) \leq (1 + \sqrt{\varepsilon})^2.$$

□

2.3. Invertibilidad Restringida

Esta sección esta basada en el tercer artículo de la serie [MSS17]. En este artículo se deriva una prueba simple del principio de invertibilidad restringida de Bourgain y Tzafriri y luego se mejora en dos formas. Se demuestra que el rango estable se puede remplazar por el rango estable de la 4-norma de Schatten y que existen mejores cotas cuando la cantidad de columnas en la matriz no excede por mucho la cantidad de filas. Empezaremos con una introducción y los preliminares necesarios para posteriormente plantear los resultados principales y demostrarlos utilizando la técnica de familias de polinomios que se entrelazan.

2.3.1. Introducción

Para $d \geq m$ y una matriz B de $d \times m$, sus valores singulares $\sigma_1(B) \geq \sigma_2(B) \geq \dots \geq \sigma_d(B)$ (omitiremos el paréntesis cuando no haya ambigüedad) se definen como la raíz cuadrada de los valores propios de B^*B . El menor de los valores singulares es una medida de qué tan lejos está B de ser una matriz singular y también se puede encontrar como

$$\sigma_{\min}(B) = \min_{x \neq 0} \frac{\|Bx\|}{\|x\|},$$

donde $\|x\|$ denota a la norma euclidiana del vector x . Por otro lado, el rango estable de B se define como

$$\text{srnk}(B) = \frac{\|B\|_2^2}{\|B\|_\infty^2} = \frac{\sum_{i=1}^d \sigma_i^2}{\sigma_1^2},$$

donde $\|B\|_p$ denota la p -norma de Schatten, que es la norma l_p de los valores singulares de B :

$$\|B\|_p = \sqrt[p]{\sum_{i=1}^d (\sigma_i(B))^p}.$$

Al caso particular de $p = 2$ se le conoce como la norma de Fobrenius o norma de Hilbert-Schmidt y suele denotarse como $\|\cdot\|_F$. En el caso particular de $p = \infty$ obtenemos la norma del operador que es igual al valor singular más grande de B y también se puede definir como

$$\|B\|_\infty = \max_{\|x\|=1} \|Bx\|.$$

El principio de invertibilidad restringida [BT87] establece que si una matriz B tiene rango estable alto, entonces debe contener una submatriz de columnas grande, B_S , de tamaño $d \times |S|$, cuyo menor valor singular sea grande. Donde $S \subset [m]$ y B_S es la submatriz de B con las columnas determinadas por S . Este resultado se ha generalizado en [Ver01], [SS12], [You14], [NY17] y desde entonces ha sido útil en teoría de espacios de Banach, minería de datos y más recientemente, en ciencia computacional teórica.

Partiremos del mejor resultado de este tipo que se tenía antes de utilizar familias de polinomios que se entrelazan, el cual se puede consultar en [SS12].

Teorema 2.3.1. Supongamos que B es una matriz de $d \times m$ y sea $k \leq \text{srnk}(B)$ un entero positivo. Entonces existe un subconjunto $S \subset [m]$ de tamaño k tal que

$$\sigma_{\min}(B_S)^2 \geq \left(1 - \sqrt{\frac{k}{\text{srnk}(B)}}\right)^2 \frac{\|B\|_F^2}{m}.$$

El caso **isotrópico** es cuando $BB^T = I_d$. En este caso tenemos que $\text{srnk}(B) = d$ y el lado derecho del teorema se convierte en

$$\left(1 - \sqrt{\frac{k}{d}}\right)^2 \frac{d}{m}.$$

A partir de los resultados de Krasikov [Kra06] se puede llegar a que el valor $(1 - \sqrt{k/d})^2$ es una buena cota asintótica de la menor raíz del polinomio de Laguerre asociado $L_k^{d-k}(x)$, después de un reescalamiento apropiado, lo cual motiva la demostración en el caso isotrópico. Por otro lado, $(1 - \sqrt{k/d})^2$ es el punto más pequeño del soporte de la distribución Marchenko-Pastur, [MP67]. A partir de este hecho, se puede derivar que la cota del teorema anterior es asintóticamente buena, para ello se puede consultar una entrada de Srivastava en el siguiente blog [Sri].

Cabe mencionar que una forma genérica de probar las cotas de este artículo se puede ver en el marco más general de convoluciones de polinomios, del cual hablaremos en el siguiente capítulo, estas desigualdades son asintóticamente justas como se ve en [Mar15] y a veces mejoran las cotas de esta sección.

2.3.2. El caso isotrópico con remplazo

En esta sección demostraremos el caso isotrópico del Teorema 2.3.1. Para demostrarlo elegiremos el conjunto S tomando k columnas de B aleatoriamente con remplazo y viendo que el polinomio característico esperado tiene las mismas raíces que el polinomio de Laguerre asociado, el cual tomaremos como en la definición 5.1.5 de [Sze39].

Definición 2.3.2. El polinomio de Laguerre asociado de grado n y parámetro α es

$$L_n^{(\alpha)} = e^x x^{-\alpha} \frac{1}{n!} \partial_x^n (e^{-x} x^{n+\alpha}) = \frac{x^{-\alpha}}{n!} (\partial_x - 1)^n x^{n+\alpha}.$$

Lo que utilizaremos de estos polinomios es una cota en la menor raíz, que es el Teorema 1 de [Kra06].

Teorema 2.3.3. Para $\alpha > -1$,

$$\lambda_k(L_k^{(\alpha)}(x)) \geq V^2 + 3V^{4/3}(U^2 - V^2)^{-1/3},$$

donde $V = \sqrt{k + \alpha + 1} - \sqrt{k}$ y $U = \sqrt{k + \alpha + 1} + \sqrt{k}$.

Ya con esta cota podemos demostrar el teorema en el caso isotrópico.

Teorema 2.3.4. Supongamos que B es una matriz de $d \times m$ tal que $BB^T = I_d$ y sea $k \leq \text{srnk}(B)$ un entero positivo. Entonces existe un subconjunto $S \subset [m]$ de tamaño k tal que

$$\sigma_{\min}(B_S)^2 \geq \left(1 - \sqrt{\frac{k}{\text{srnk}(B)}}\right)^2 \frac{\|B\|_F^2}{m}.$$

Demostración. Supongamos que las columnas de B son los vectores $u_1, \dots, u_m \in \mathbb{R}^d$. Observemos que para un conjunto S de tamaño $k < d$,

$$\sigma_{\min}(B_S)^2 = \lambda_k(B_S^T B_S) = \lambda_k(B_S B_S^T) = \lambda_k\left(\sum_{i \in S} u_i u_i^T\right),$$

es decir, podemos expresar al menor valor singular como la k -ésima raíz del polinomio característico de una suma de matrices de rango 1. Como ya hemos visto en aplicaciones anteriores esto lo podemos hacer con familias de

polinomios que se entrelazan. Sean $v_1, \dots, v_k$ vectores aleatorios todos con distribución uniforme en el conjunto $\{u_1, \dots, u_m\}$. Por el Teorema 1.2.16 sabemos que los polinomios

$$q_{j_1, \dots, j_k}(x) = \det \left[xI - \sum_{i=1}^k u_{j_i} u_{j_i}^T \right] \quad j_i \in [m],$$

forman una familia que se entrelaza, y por el Teorema 1.1.8, uno de esos polinomios cumple que

$$\lambda_k \left(\sum_{i=1}^k u_{j_i} u_{j_i}^T \right) = \lambda_k(q_{j_1, \dots, j_k}(x)) \geq \lambda_k(q_\emptyset),$$

donde $q_\emptyset = \mathbb{E}_{v_1, \dots, v_k} \chi \left[\sum_{i=1}^k v_i v_i^* \right]$. Es importante remarcar que el polinomio $q_{j_1, \dots, j_k}$ que cumple no puede tener $j_a = j_b$ para $a \neq b$ ya que esto implicaría que elegimos dos columnas iguales y que $\sigma_{\min}(B_S) = 0$. Así que a final de cuentas, elegir con remplazo no nos afecta en nada.

Entonces, ahora tenemos que acotar la raíz k -ésima de $q_\emptyset$. Primero observemos que podemos reescribir este polinomio como

$$q_\emptyset(x) = \left(1 - \frac{1}{m} \partial_x \right)^k x^d.$$

Para ello, veamos que la condición $BB^T = I_d$ es equivalente a $\sum_{i=1}^m u_i u_i^T = I_d$, por lo que para $j = 1, \dots, k$ tendremos

$$\mathbb{E} v_j v_j^T = \frac{1}{m} \sum_{i=1}^m u_i u_i^* = \frac{1}{m} I_d.$$

Por el Teorema 1.2.14 sabemos que

$$\begin{aligned} q_\emptyset(x) &= \left(\prod_{i=1}^k (1 - \partial_{z_i}) \right) \det \left(xI + \sum_{i=1}^k z_i A_i \right) \Big|_{z_1 = \dots = z_k = 0} \\ &= \left(\prod_{i=1}^k (1 - \partial_{z_i}) \right) \det \left(xI + \left(\frac{1}{m} \sum_{i=1}^m z_i \right) I \right) \Big|_{z_1 = \dots = z_k = 0} \\ &= \left(\prod_{i=1}^k (1 - \partial_{z_i}) \right) \left(x + \frac{1}{m} \sum_{i=1}^k z_i \right)^d \Big|_{z_1 = \dots = z_k = 0}. \end{aligned}$$

Si ahora expandimos y nos fijamos en todos los monomios resultantes, podemos ver que para que no se anulen al evaluar en 0, cada variable debe aparecer a lo más una vez, y cada monomio de la forma

$$\frac{1}{m^n} x^{d-n} z_{i_1} \dots z_{i_n} \quad \text{con } i_1 < \dots < i_n,$$

aportará un $x^{d-n}(\frac{-1}{m})^n$ al operarlo con $\prod_{i=1}^k(1 - \partial_{z_i})$, ya que se anulará en todas partes excepto con $(-1)^d \partial_{z_{i_1}} \dots \partial_{z_{i_n}}$. Como hay $\binom{k}{n}$ monomios de tamaño n concluimos que

$$q_\emptyset(x) = \sum_{n=0}^d x^{d-n} \frac{(-1)^n}{m^n} \binom{k}{n} = \left(1 - \frac{1}{m} \partial_x\right)^k x^d.$$

Por último, de la definición de polinomio de Laguerre asociado podemos concluir que

$$x^{-(d-k)} q_\emptyset(x) = (-1)^k \frac{k!}{m^k} L_k^{(d-k)}(mx),$$

que nos dice que la k -ésima raíz más grande de $q_\emptyset$ es la menor raíz (que también es la k -ésima) de $L_k^{(d-k)}(mx)$, para la cual ya tenemos una cota. Utilizando la cota del Teorema 2.3.3 con $\alpha = d - k$ y olvidándonos de la parte $3V^{4/3}(U^2 - V^2)^{-1/3}$ obtenemos que

$$\lambda_k(q_\emptyset(x)) = \lambda_k(L_k^{(d-k)}(mx)) = \frac{\lambda_k(L_k^{(d-k)}(x))}{m} \geq \frac{V^2}{m} > \frac{(\sqrt{d} - \sqrt{k})^2}{m}.$$

Juntando todo lo anterior concluimos que

$$\sigma_{\min}(B_S)^2 \geq \left(1 - \sqrt{\frac{k}{d}}\right)^2 \frac{d}{m} = \left(1 - \sqrt{\frac{k}{\text{srang}(B)}}\right)^2 \frac{\|B\|_F^2}{m}.$$

□

2.3.3. El caso no isotrópico y la 4-norma de Schatten

En esta sección extenderemos las técnicas de la sección anterior para ver que el Teorema 2.3.1 es cierto en el caso no isotrópico, en donde demostraremos una cota que reemplaza el rango estable con el rango estable de 4-norma de Schatten:

$$\text{srang}_4 = \frac{\|B\|_2^4}{\|B\|_4^4} = \frac{\left(\sum_{i=1}^d \sigma_i^2\right)^2}{\sum_{i=1}^d \sigma_i^4}.$$

Observación 2.3.5. Observemos que

$$\sigma_1^2 \sum_{i=1}^d \sigma_i^2 \geq \sum_{i=1}^d \sigma_i^4,$$

y esto implica que

$$\text{srang}_4(B) = \frac{\left(\sum_{i=1}^d \sigma_i^2\right)^2}{\sum_{i=1}^d \sigma_i^4} \geq \frac{\sum_{i=1}^d \sigma_i^2}{\sigma_1^2} = \text{srang}(B),$$

por lo que al tomar $srank_4$ en lugar de $srank$ mejoramos el Teorema 2.3.1, especialmente cuando B tiene muchos valores singulares moderadamente grandes.

En este caso para obtener las cotas deseadas utilizaremos el concepto de $\alpha \text{ mín}(p(x))$ que es la menor raíz de $p(x) + \alpha \partial_x p(x)$, es decir

$$\alpha \text{ mín}(p(x)) = \lambda_{\text{mín}}(p(x) + \alpha p'(x)).$$

También se puede definir esto en términos de la función barrera menor

$$\Phi_p(x) = -\frac{p'(x)}{p(x)} = \sum_{i=1}^d \frac{1}{\lambda_i - x}.$$

Observemos que como $\alpha \text{ mín} = \text{mín}\{z : \Phi_p(z) = \frac{1}{\alpha}\}$, entonces $\alpha \text{ mín}$ es menor que la menor raíz de $p(x)$. Además, para $\alpha > 0$, tenemos que $\alpha \text{ mín}(x^k) = -k\alpha$. Lo que utilizaremos de $\alpha \text{ mín}$ es que cumple la siguiente desigualdad, la cual nos dice que $\alpha \text{ mín}$ crece de forma controlada cuando se le aplica $1 - \lambda \partial_x$. Este lema es similar al Lema 3.4 de [BSS12] y puede verse como un análogo en el caso mínimo del Lema 2.2.15.

Lema 2.3.6. Si $p(x)$ es un polinomio con raíces reales y $\lambda > 0$, entonces $(1 - \lambda \partial_x)p(x)$ tiene raíces reales y

$$\alpha \text{ mín}((1 - \lambda \partial_x)p(x)) \geq \alpha \text{ mín}(p(x)) + \frac{1}{1/\lambda + 1/\alpha}.$$

Para la demostración de este lema puede consultarse el Lema 4.3 de [MSS17]. Por último, para plantear el resultado principal de esta sección será más útil trabajar con valores propios de $A = B^T B$ que los valores singulares de B . Para una matriz simétrica A , denotaremos por

$$\kappa_A = \frac{\text{Tr}(A)^2}{\text{Tr}(A^2)}.$$

Observemos que si $A = B^T B$, entonces $\kappa_A = srank_4(B)$. El resultado en el caso no isotrópico es el siguiente:

Teorema 2.3.7. Supongamos que $u_1, \dots, u_m$ son vectores con $\sum_{i=1}^m u_i u_i^T = A$. Entonces para todo entero $k \leq \kappa_A$, existe un conjunto $S \subset [m]$ de tamaño k con

$$\lambda_k \left(\sum_{i \in S} u_i u_i^T \right) \geq \left(1 - \sqrt{\frac{k}{\kappa_A}} \right)^2 \frac{\text{Tr}(A)}{m}.$$

Demostración. Para demostrarlo podemos utilizar exactamente la misma familia de polinomios que en el caso no isotrópico y con ello llegar a que existe un $q_{j_1, \dots, j_k}$ tal que

$$\lambda_k \left(\sum_{i=1}^k u_{j_i} u_{j_i}^T \right) = \lambda_k(q_{j_1, \dots, j_k}(x)) \geq \lambda_k q_\emptyset,$$

donde $q_\emptyset = \mathbb{E}_{v_1, \dots, v_k} \chi \left[\sum_{i=1}^k v_i v_i^* \right]$. La diferencia ahora es que este polinomio característico esperado no necesariamente será de Laguerre. Esto se debe a que ahora tenemos la condición $\mathbb{E}_{v_i} = \frac{1}{m} \sum_{i=1}^m u_i u_i^* = \frac{1}{m} A$ y en el caso anterior $A = I_d$. Sin embargo, igual podemos reescribir $q_\emptyset$ como

$$\mathbb{E} \det \left[xI - \sum_{i=1}^k v_i v_i^T \right] = x^{d-k} \prod_{i=1}^d (1 - \lambda_i \partial_x) x^k,$$

donde $\lambda_1, \dots, \lambda_d$ son los valores propios de $M =: \mathbb{E} v v^T = \frac{1}{m} A$. El procedimiento es similar al caso isotrópico, por el Teorema 1.2.14 sabemos que

$$\begin{aligned} q_\emptyset(x) &= \left(\prod_{i=1}^k (1 - \partial_{z_i}) \right) \det \left(xI + \sum_{i=1}^k z_i M \right) \Big|_{z_1 = \dots = z_k = 0} \\ &= \left(\prod_{i=1}^k (1 - \partial_{z_i}) \right) \det \left(xI + M \sum_{i=1}^k z_i \right) \Big|_{z_1 = \dots = z_k = 0} \\ &= \left(\prod_{i=1}^k (1 - \partial_{z_i}) \right) \sum_{n=0}^d x^{d-n} \left(\sum_{i=1}^k z_i \right)^n e_n(M) \Big|_{z_1 = \dots = z_k = 0}, \end{aligned}$$

donde $e_n(M)$ es la n -ésima función simétrica elemental de M y es igual

$$e_n(M) = \sum_{S \subset [d], |S|=n} \prod_{i \in S} \lambda_i.$$

De nuevo expandimos los $\left(\sum_{i=1}^k z_i \right)^n$ y nos fijamos en todos los monomios resultantes, podemos ver que para que no se anulen al evaluar en 0, cada variable debe aparecer a lo más una vez, y cada monomio de la forma $z_{i_1} \dots z_{i_n}$ con $i_1 < \dots < i_n$ aportará un $(-1)^n$ al operarlo con $\prod_{i=1}^m (1 - \partial_{z_i})$, ya que se anulará en todas partes excepto con $(-1)^d \partial_{z_{i_1}} \dots \partial_{z_{i_n}}$. Como hay $\frac{k!}{(k-n)!}$ monomios de tamaño n obtenemos que

$$q_\emptyset(x) = \sum_{n=0}^k x^{d-n} (-1)^n \frac{k!}{(k-n)!} e_n(M),$$

como $\partial_x^n x^k = x^{k-n} \frac{k!}{(k-n)!}$ para $n \leq k$ y es 0 para $n > k$ concluimos que

$$q_\emptyset(x) = x^{d-k} \sum_{n=0}^k \partial_x^n (-1)^n e_n(M) x^k = x^{d-k} \prod_{i=1}^d (1 - \lambda_i \partial_x) x^k.$$

Por último, necesitamos acotar la menor raíz de $\prod_{i=1}^d (1 - \lambda_i \partial_x) x^k$. Observemos que si multiplicamos a los vectores aleatorios v_i por una constante,

podemos suponer sin pérdida de generalidad que $Tr[M] = 1$, y con ello bastaría probar que

$$\lambda_k(p) \geq \left(1 - \sqrt{kTr[M^2]}\right)^2.$$

Aplicando el Lema 2.3.6 d veces para algún $\alpha > 0$ obtenemos que

$$\begin{aligned} \lambda_k(p) &\geq \lambda_k \left(\prod_{i=1}^d (1 - \lambda_i \partial_x) x^k \right) \geq \alpha \min \left(\prod_{i=1}^d (1 - \lambda_i \partial_x) x^k \right) \\ &\geq \alpha \min(x^k) + \sum_{i=1}^d \frac{1}{\lambda_i^{-1} + \alpha^{-1}} = -k\alpha + \sum_{i=1}^d \frac{1}{\lambda_i^{-1} + \alpha^{-1}} = -k\alpha + \sum_{i=1}^d \frac{1}{1 + \lambda_i \alpha^{-1}} \lambda_i. \end{aligned}$$

Como la función $y \rightarrow \frac{1}{1+y\alpha^{-1}}$ es convexa para $\alpha > 0$ y $\lambda_1 + \dots + \lambda_d = Tr[M] = 1$, la desigualdad de Jensen nos dice que

$$\sum_{i=1}^d \frac{1}{1 + \lambda_i \alpha^{-1}} \lambda_i \geq \frac{1}{1 + \left(\sum_{i=1}^d \lambda_i^2\right) \alpha^{-1}} = \frac{1}{1 + Tr(M^2) \alpha^{-1}}.$$

Por lo tanto,

$$\lambda_k(p) \geq -k\alpha + \frac{1}{1 + Tr(M^2) \alpha^{-1}},$$

para cualquier $\alpha > 0$. Si maximizamos esta expresión para $\alpha > 0$ obtenemos que

$$\lambda_k(p) \geq \left(1 - \sqrt{kTr[M^2]}\right)^2.$$

□

Observación 2.3.8. Vale la pena mencionar que en la Sección 4.1 de [MSS17] se presenta un algoritmo que corre en tiempo polinomial con el cual se puede obtener el subconjunto S del teorema anterior.

2.3.4. El caso isotrópico sin reemplazo

La cota del Teorema 2.3.1 en el caso isotrópico se puede mejorar si tomamos los vectores aleatoriamente sin reemplazo. Al hacer esto modificaremos nuestro árbol de la familia de polinomios que se entrelazan, ya que ahora no podemos tomar simplemente vectores aleatorios independientes, si no que tendremos que formar el árbol tomando en cuenta el subconjunto de vectores que tenemos. Esto también implica que se modificará el polinomio característico esperado, que ahora será un polinomios de Jacobi escalado, para el cual podemos derivar una nueva cota en las raíces más bajas. De esta forma se demuestra que existe un conjunto S de k columnas para el cual

$$\sigma_{min}(B_S)^2 \geq \frac{(\sqrt{d(m-k)} - \sqrt{k(m-d)})^2}{m^2}.$$

Como

$$\frac{(\sqrt{d(m-k)} - \sqrt{k(m-d)})^2}{m^2} \geq \left(1 + \frac{\sqrt{dk}}{m}\right) \left(1 - \sqrt{\frac{k}{d}}\right)^2 \frac{d}{m},$$

esto mejora el Teorema 2.3.1 en un factor constante cuando m es un múltiplo constante de d . La formulación formal del teorema es la siguiente.

Teorema 2.3.9. Supongamos que $u_1, \dots, u_m \in \mathbb{R}^n$ cumplen que $\sum_{i=1}^m u_i u_i^T = I$ y $k \leq d$ es un entero. Entonces existe un subconjunto $S \subset [m]$ de tamaño k tal que

$$\lambda_k \left(\sum_{i \in S} u_i u_i^T \right) \geq \frac{(\sqrt{d(m-k)} - \sqrt{k(m-d)})^2}{m^2}.$$

La demostración completa de este resultado requiere definir bien los vértices de nuestro árbol, encontrar una forma efectiva de escribir estos vértices en términos de las hojas y hacer bien las cuentas para encontrar el polinomio característico. Además hay que probar que en verdad es una familia de polinomios que se entrelaza, ya que en este caso no podemos utilizar directamente el Teorema 1.2.16, en su lugar habrá que probar directamente la propiedad 2 de la Definición 1.1.6. Dado que todos estos pasos requieren de hacer varias cuentas y prestar atención a muchos detalles, sólo presentaremos la idea general de la demostración, la solución detallada se puede encontrar en la Sección 5 de [MSS17].

Empezaremos por ver cómo formar el árbol para aplicar la técnica de familias de polinomios que se entrelazan. En este caso, las hojas del árbol corresponderán a subconjuntos de $[m]$ de tamaño k . La raíz corresponde al conjunto vacío $\emptyset$, y los otros vértices corresponden a subconjuntos de tamaño menor o igual a k , donde un conjunto es hijo de otro conjunto si contiene al otro y tiene exactamente un elemento más. Para cada $S \subset [m]$ definimos

$$p_S(x) = \chi \left(\sum_{i \in S} u_i u_i^T \right)$$

Las hojas serán los polinomios $p_S(x)$. Al vértice interno asociado con el conjunto T de tamaño menor a k , le asignamos el polinomio

$$f_T(x) = \mathbb{E}_{\substack{S \supset T \\ |S|=k}} p_S(x),$$

donde la esperanza se toma uniformemente sobre conjuntos S de tamaño k que contienen a T .

Para demostrar que es una familia de polinomios que se entrelaza hay que reescribir los polinomios f_T en términos de p_T .

Lema 2.3.10. Para cada subconjunto T de $[m]$ de tamaño $t \leq k$,

$$f_T(x) = \frac{1}{\binom{m-t}{k-t}} \sum_{\substack{S \supset T \\ |S|=k}} p_S(x) = \frac{(m-k)!}{(m-t)!} (x-1)^{-(m-d-k)} \partial_x^{k-t} (x-1)^{m-d-t} p_T(x).$$

En particular,

$$f_\emptyset(x) = \frac{(m-k)!}{m!} (x-1)^{-(m-d-k)} \partial_x^k (x-1)^{m-d} x^d.$$

Este lema se obtiene de aplicar de forma recursiva un lema mas simple.

Lema 2.3.11. Para cada subconjunto T de $[m]$ de tamaño t ,

$$\sum_{i \neq T} p_{T \cup \{i\}}(x) = (x-1)^{-(m-d-t-1)} \partial_x (x-1)^{m-d-t} p_T(x).$$

Posteriormente, para acotar las raíces de nuestro polinomio característico esperado $f_\emptyset(x)$, hacemos algo análogo a los dos casos anteriores. Primero cambiamos su k -ésima raíz por la menor raíz de un polinomio más simple.

Lema 2.3.12. La k -ésima raíz más grande de $f_\emptyset(x)$ es igual a la menor raíz del polinomio

$$\partial_x^d (x-1)^{m-k} x^k.$$

Y luego usamos d veces el siguiente lema, el cual demostraremos en el siguiente capítulo (ver Lema 3.1.32)

Lema 2.3.13. Para todo polinomio de raíces reales $p(x)$ de grado al menos dos y $\alpha > 0$,

$$\alpha \min(\partial_x p(x)) \geq \alpha \min(p(x)) + \alpha.$$

Con ello obtenemos la cota

Teorema 2.3.14. Para $m > d > k$,

$$\lambda_k(f_\emptyset(x)) > \frac{(\sqrt{d(m-k)} - \sqrt{k(m-d)})^2}{m^2}.$$

Con el siguiente lema, se puede mejorar la cota inferior en el teorema anterior a

$$\frac{\left(\sqrt{(d+1)(m-k)} - \sqrt{k(m-d-1)}\right)^2}{m^2}.$$

Lema 2.3.15. Para cada polinomio de raíces reales $p(x)$ y $\alpha > 0$,

$$\alpha \min(p) + \alpha \leq \lambda_{\min}(p).$$

Observación 2.3.16. La fórmula de Rodrigues para el polinomio de Jacobi de grado n con parámetros α y β que se presenta en (4.3.1) de [Sze39] es

$$P_n^{(\alpha, \beta)}(x) = \frac{(-1)^n}{2^n n!} (1-x)^{-\alpha} (1+x)^{-\beta} \partial_x^n (1-x)^{\alpha+n} (1+x)^{\beta+n}.$$

Podemos observar que estos polinomios están relacionados con el polinomios $f_\emptyset(x)$ de la demostración pasada de la siguiente forma:

$$P_k^{(m-d-k, d-k)}(2x-1) = \binom{m}{k} x^{-(d-k)} f_\emptyset(x).$$

Por lo tanto, cotas inferiores de la k -ésima raíz de $f_\emptyset$ se traducen en cotas de la menor raíz de los polinomios de Jacobi. Por lo pronto, ya existía una cota en [Kra06] para polinomios de Jacobi, la cual parece ser mejor que la obtenida en el Lema 2.3.15.

Capítulo 3

Convoluciones de Polinomios y Gráficas de Ramanujan

En este capítulo nos interesa entender principalmente los resultados del último artículo de la serie [MSS15d], los cuales analizaremos en la segunda sección. Para obtener estos resultados, es necesario hacer uso de unas cotas obtenidas en otro artículo de los mismos autores [MSS15a], es por eso que el objetivo de la primer sección será entender a grandes rasgos los resultados de este artículo. Las contribuciones importantes de estos artículos son identificar una nueva clase de polinomios característicos con raíces reales relacionados con gráficas aleatorias, y desarrollar nuevas herramientas tanto para establecer sus propiedades de entrelazado como para analizar la ubicación de sus raíces. Estos métodos son distintos a los del polinomio característico mixto y las cotas que se obtienen son mejores (por un factor de dos) a las que se obtienen con el argumento barrera, es por ello que decidimos presentar estos dos artículos en un capítulo distinto a los demás. Además, el polinomio característico esperado que consideran se puede calcular en tiempo polinomial, a diferencia de la mayoría de los otros.

3.1. Convoluciones de Polinomios

En esta sección presentaremos de manera concisa los resultados y las ideas principales detrás de las demostraciones del artículo [MSS15a]. A grandes rasgos, lo que nos gustaría analizar es qué pasa con los valores propios de dos matrices A y B , si rotamos una de ellas aleatoria y uniformemente, las sumamos, calculamos el valor esperado y nos fijamos en sus valores propios. El poder controlar cómo se comportan los valores propios nos permitirá acotar de manera precisa el espectro de ciertas matrices aleatorias que se pueden construir con estas operaciones y que nos permitirá obtener gráficas de Ramanujan. De manera análoga nos gustaría saber qué pasa si rotamos una de ellas aleatoria y uniformemente, y la multiplicamos con la otra, o si

rotamos ambas de manera independiente y luego las sumamos. Para poder trabajar con estas operaciones, será mucho más útil fijarse en los polinomios característicos p y q de las matrices A y B , respectivamente y tras realizar la operación, fijarnos en el polinomio característico. De esta forma, transformamos nuestras tres operaciones de matrices en 3 convoluciones de polinomios. Y ahora debemos entender la relación que hay entre las raíces del polinomio resultante y los dos que convolucionamos. Para entender esto, primero reescribiremos estas convoluciones en términos más sencillos, y con ello descubrimos que algunas de estas convoluciones ya se habían estudiado antes e incluso ya se tenían cotas, aunque no nos serán útiles. Posteriormente se mejoran las cotas inspirados en convoluciones que se tienen en probabilidad libre. Dado que presentar las demostraciones para las tres convoluciones es bastante extenso, nos concentraremos en explicar bien la primera de ellas y sólo haremos un esbozo de cómo obtener las otras dos; para ver los detalles de estas dos el lector puede consultar directamente [MSS15a].

En las siguientes definiciones denotaremos por $\chi_x(M) = \det(xI - M)$ al polinomio característico M evaluado en x . También denotaremos por $O(d)$ al grupo de las matrices ortogonales de dimensión d .

Definición 3.1.1 (Convolución Aditiva Simétrica). Sean $p(x) = \chi_x(A)$ y $q(x) = \chi_x(B)$ dos polinomios de raíces reales, para algunas matrices simétricas de $d \times d$, A y B . La convolución aditiva simétrica de p y q se define como

$$p(x) \boxplus_d q(x) = \mathbb{E}_Q \chi_x(A + QBQ^T),$$

donde la esperanza se toma sobre las matrices aleatorias ortogonales tomando la medida de Haar en $O(d)$.

Definición 3.1.2 (Convolución Multiplicativa Simétrica). Sean $p(x) = \chi_x(A)$ y $q(x) = \chi_x(B)$ dos polinomios de raíces reales, para algunas matrices simétricas semidefinidas positivas de $d \times d$, A y B . La convolución multiplicativa simétrica de p y q se define como

$$p(x) \boxtimes_d q(x) = \mathbb{E}_Q \chi_x(AQBQ^T),$$

donde Q se distribuye de acuerdo a la medida de Haar en $O(d)$.

Definición 3.1.3 (Convolución Aditiva Asimétrica). Sean $p(x) = \chi_x(AA^T)$ y $q(x) = \chi_x(BB^T)$ dos polinomios de raíces reales no negativas, para algunas matrices (no necesariamente simétricas) de $d \times d$, A y B . La convolución aditiva asimétrica de p y q se define como

$$p(x) \boxplus\boxplus_d q(x) = \mathbb{E}_{Q,R} \chi_x((A + QBR^T)(A + QBR^T)^T),$$

donde Q y R se distribuyen de acuerdo a la medida de Haar en $O(d)$.

3.1.1. Fórmulas de Convolución

En esta sección reescribiremos en términos más sencillos la convolución aditiva simétrica que acabamos de definir, y mencionamos algunos resultados que ya se tenían de esta convolución. Luego esbozaremos cómo reescribir las otras dos convoluciones.

Convolución Aditiva Simétrica

Teorema 3.1.4. Si $p(x) = \sum_{i=0}^d x^{d-i}(-1)^i a_i^p$ y $q(x) = \sum_{i=0}^d x^{d-i}(-1)^i a_i^q$ entonces tenemos que

$$[p \boxplus_d q](x) = \sum_{k=0}^m x^{d-k}(-1)^k \sum_{i+j=k} \frac{(d-i)!(d-j)!}{d!(d-i-j)!} a_i^p a_j^q.$$

Esto es equivalente a las identidades

$$[p \boxplus_d q](x) = \frac{1}{d!} \sum_{i=0}^d (d-i)! a_i^q p^{(i)}(x) = \frac{1}{d!} \sum_{i=0}^d p^{(i)}(x) q^{(d-i)}(y),$$

donde $p^{(i)}$ denota la i -ésima derivada de p .

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

$$\mathbb{E}_Q c_k(A + QBQ^T) = \mathbb{E}_Q c_k(C + U^T QV D V^T Q^T U) = \mathbb{E}_Q c_k(C + QDQ^T),$$

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. $\square$

Observación 3.1.5. Esta convolución fue estudiada por Walsh [Wal22], quien demostró algunos resultados. Uno de ellos es que si p y q son polinomios con raíces reales de grado d , entonces $p \boxplus_d q$ también es un polinomio con raíces reales de grado d . Además,

$$\lambda_1(p \boxplus_d q) \leq \lambda_1(p) + \lambda_1(q),$$

donde $\lambda_1(p)$ es la raíz más grande del polinomio p . También se puede ver [Mar49] y el Teorema 5.3.1 de [RS02].

Otra forma en la que se puede demostrar el teorema anterior es utilizando algo a lo que los autores denominan teorema de cuadratura, y que nos dice que tomar la esperanza respecto a matrices ortonormales es lo mismo que respecto a las matrices de permutación signadas.

Teorema 3.1.6 (Cuadratura simétrica aditiva). Para cualesquiera matrices cuadradas simétricas A y B , se cumple que

$$\mathbb{E}_Q \chi_x(A + QBQ^T) = \mathbb{E}_P \chi_x(A + PBP^T),$$

donde la primer esperanza es sobre las matrices ortonormales y la segunda sobre las matrices de permutación signadas.

El caso asimétrico se demuestra utilizando una idea similar a esta, así que las definiciones e ideas para este teorema las veremos en el caso asimétrico después de ver el multiplicativo.

Convolución Multiplicativa Simétrica

La convolución más sencilla de reescribir es la multiplicativa y queda de la siguiente manera.

Teorema 3.1.7. Si $p(x) = \sum_{i=0}^d x^{d-i}(-1)^i a_i^p$ y $q(x) = \sum_{i=0}^d x^{d-i}(-1)^i a_i^q$, y tienen puras raíces no negativas, entonces tenemos que

$$[p \boxtimes_d q](x) = \sum_{k=0}^d x^{d-k}(-1)^k \frac{a_k^p a_k^q}{\binom{d}{k}}.$$

Observación 3.1.8. Esta convolución ya había sido estudiada por Szegő [Sze22], y se sabía que si p y q son polinomios con raíces reales no negativas, entonces $p \boxtimes_d q$ también es un polinomio con raíces reales no negativas. Además,

$$\lambda_1(p \boxtimes_d q) \leq \lambda_1(p) \lambda_1(q).$$

Para demostrar el teorema anterior basta ver que se cumple la igualdad para cada coeficiente.

Teorema 3.1.9. Si A y B son matrices simétricas de $d \times d$, y Q es una matriz ortonormal aleatoria, entonces tenemos que

$$\mathbb{E}_Q e_k(AQBQ^T) = \frac{e_k(A)e_k(B)}{\binom{d}{k}}.$$

Para demostrar este hecho hay que seguir pasos análogos a los de la convolución aditiva simétrica. De hecho, ahora los pasos serán más sencillos. Primero hay que observar que basta demostrarlo para matrices diagonales, ya que la medida de Haar es invariante bajo multiplicación por matrices ortonormales. Para el caso de matrices diagonales hay que escribir el e_k como suma de determinantes de submatrices de tamaño k y usar el hecho de que la medida de Haar también es invariante bajo multiplicación por matrices de permutación aleatorias.

Convolución Aditiva Asimétrica

Ahora daremos una idea de cómo reescribir la convolución aditiva asimétrica.

Teorema 3.1.10. Si $p(x) = \sum_{i=0}^d x^{d-i}(-1)^i a_i^p$ y $q(x) = \sum_{i=0}^d x^{d-i}(-1)^i a_i^q$, y tienen puras raíces no negativas, entonces tenemos que

$$[p \boxplus_d q](x) = \sum_{k=0}^d x^{d-k}(-1)^k \sum_{i+j=k} \left(\frac{(d-i)!(d-j)!}{d!(d-i-j)!} \right)^2 a_i^p a_j^q.$$

Esto es equivalente a la identidad

$$[p \boxplus_d q](x) = \left(\frac{1}{d!} \right)^2 \sum_{i=0}^d ((d-k)!)^2 a_i^q (DxD)^i p(x).$$

Observemos que ambos lados de la ecuación son invariantes si multiplicamos (por la derecha o izquierda) a las matrices A y B por matrices ortonormales, por lo tanto si tomamos la descomposición en valores singulares podemos reducir el problema a demostrarlo para matrices diagonales. La demostración para el caso diagonal requiere dos pasos principales, primero el Teorema de Cuadratura que es pasar de matrices ortonormales aleatorias a matrices de permutación signadas aleatorias y posteriormente calcular explícitamente la esperanza sobre matrices de permutación signadas aleatorias.

Definición 3.1.11. Una matriz de permutación signada tiene todas sus entradas iguales a $-1, 0, 1$, con exactamente una entrada distinta de cero en cada fila y columna.

Observación 3.1.12. Una matriz de permutación signada uniformemente aleatoria se puede obtener multiplicando una matriz diagonal con entradas ± 1 uniformemente aleatoria con una matriz de permutación uniformemente aleatoria.

A continuación detallaremos un poco más como se dan estos dos pasos, utilizando los lemas que utilizan los autores, para una demostración detallada de estos lemas referimos al lector a la Sección 2.3 de [MSS15a].

Primer paso. De ortonormales a permutaciones.

Teorema 3.1.13 (Cuadratura asimétrica aditiva). Para cualesquiera matrices cuadradas A y B , se cumple que

$$\mathbb{E}_{P,S} \chi_x((A + PBS^T)(A + PBS^T)^T) = \mathbb{E}_{Q,R} \chi_x((A + QBR^T)(A + QBR^T)^T),$$

donde P, S son matrices de permutación signadas uniformemente aleatorias, y Q, R se distribuyen de acuerdo a la medida de Haar en $O(d)$.

Para demostrar el teorema anterior (y de manera análoga el caso simétrico) se necesita ver que la esperanza respecto a las permutaciones es invariante si multiplicamos a B por matrices ortonormales y esto se divide en dos pasos, primero hay que corroborarlo para las matrices ortonormales que actúan en subespacios de coordenadas bidimensionales y luego ver que estas matrices generan a todas las matrices ortonormales. A estas matrices de $d \times d$ las podemos denotar por $R_{st}(\theta)$, para $s \neq t$, y son las matrices cuyas entradas distintas de cero son

1. $R_{st}(\theta)[s, s] = R_{st}(\theta)[t, t] = \cos(\theta)$.
2. $R_{st}(\theta)[s, t] = -R_{st}(\theta)[t, s] = \sin(\theta)$.
3. $R_{st}(\theta)[k, k] = 1$ para $k \neq s, t$.

Ya con esta notación, partiremos del siguiente lema.

Lema 3.1.14. Sean A y B matrices de $d \times d$ y sea

$$\det(R_{st}(\theta)AR_{st}(\theta)^T + B) = \sum_k c_k e^{ik\theta}.$$

Entonces $c_k = 0$ para $|k| \geq 3$.

Con este lema podemos ver que si el polinomio característico esperado es invariante bajo un conjunto pequeño de estas rotaciones (para $\theta = 0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}$) entonces también es invariante bajo todas ellas (para todo θ):

Lema 3.1.15. Sean A, B matrices de $d \times d$ y sean $s, t \in [d]$ dos índices distintos. Sea P una matriz aleatoria con soporte en $O(d)$ tal que la distribución de P es la misma que la de $PR_{st}(j\frac{\pi}{2})$ para $j = 0, 1, 2, 3$. Entonces, para todo ángulo θ ,

$$\mathbb{E}_P \det(A + PBP^T) = \mathbb{E}_P \det(A + PR_{st}(\theta)BR_{st}(\theta)^T P^T).$$

Con este último lema podemos demostrar que la esperanza respecto a las permutaciones es invariante si multiplicamos a B por matrices ortonormales.

Lema 3.1.16. Para cualesquiera matrices A y B , y matrices ortonormales Q y R se cumple que

$$\mathbb{E}_{P,S} \chi_x((A + PQBR^T S^T)(A + PQBR^T S^T)^T) = \mathbb{E}_{P,S} \chi_x((A + PBS^T)(A + PBS^T)^T),$$

es decir, la esperanza no varía si tomamos QBR^T en lugar de B .

De el lema anterior se sigue el Teorema 3.1.13.

Segundo paso. Cálculo de la esperanza sobre permutaciones para matrices diagonales.

Para este paso bastará con hacer las cuentas para la esperanza sobre permutaciones, la demostración es un poco larga y requiere de varios pasitos no triviales.

Teorema 3.1.17. Para matrices diagonales A y B , se cumple que

$$\begin{aligned} & \mathbb{E}_{P,S} \chi_x((A + PBS^T)(A + PBS^T)^T) \\ &= \sum_{k=0}^d x^{d-k} (-1)^k \sum_{i+j=k} \left(\frac{(d-i)!(d-j)!}{d!(d-k)!} \right) e_i(AA^T) e_j(BB^T). \end{aligned}$$

Si combinamos este resultado con el anterior inmediatamente obtenemos el Teorema 3.1.10. Ya que contamos con esta nueva forma de expresar a la convolución aditiva asimétrica y utilizando la teoría de polinomios Hurtwitz estables podemos obtener el siguiente resultado.

Teorema 3.1.18. Si p y q son polinomios con raíces reales no negativas, entonces $p \boxplus_d q$ también es un polinomio con raíces reales no negativas.

La demostración de este teorema es el contenido de la sección 3 de [MSS15a]. A continuación daremos una breve idea de los resultados que se utilizan en la prueba. Primero hay que observar que el teorema se sigue directamente del siguiente.

Teorema 3.1.19. Si $p(x) = \sum_{i=0}^d x^{d-i} a_i^p$ y $q(x) = \sum_{i=0}^d x^{d-i} a_i^q$, y tienen puras raíces reales no positivas, entonces tenemos que

$$r(x) = \sum_{k=0}^d x^{d-k} \sum_{i+j=k} \left(\frac{(d-i)!(d-j)!}{d!(d-k)!} \right)^2 a_i^p a_j^q$$

también está en $\mathbb{P}^-(d)$, el conjunto de polinomio de grado d con todas sus raíces reales y no positivas.

La demostración del teorema anterior se basa en construir polinomios Hurwitz estables y ver que ello implica que cierto polinomio pertenezca a $\mathbb{P}^-$.

Definición 3.1.20. Un polinomio multivariado $p(z_1, \dots, z_m) \in \mathbb{R}[z_1, \dots, z_m]$ es Hurwitz estable si es idénticamente cero o cuando

$$\operatorname{Re}(z_i) > 0, \forall i \quad \Rightarrow \quad p(z_1, \dots, z_m) \neq 0.$$

Observemos que la definición de Hurwitz estable es bastante similar a la de real estable, sólo que ahora pedimos que no se anule en el semiplano con parte real positiva mientras que la estabilidad real nos lo pedía para el semiplano con parte imaginaria positiva. El siguiente lema nos dice que cierto tipo de polinomio Hurwitz estable pertenece a $\mathbb{P}^-$.

Lema 3.1.21. Si $r(x)$ es un polinomio tal que $h(x, y) = r(xy)$ es Hurwitz estable. Entonces $r \in \mathbb{P}^-$.

Las siguientes proposiciones son las que nos dicen cómo construir polinomios Hurwitz estables.

Resultado 3.1.22. Si $p(x) \in \mathbb{P}^-$, entonces el polinomio $f(x, y) = p(xy)$ es Hurwitz estable.

El siguiente es un resultado de polarización y su demostración es idéntica a la de la Proposición 3.4 de [BB09b].

Proposición 3.1.23 (Polarización). Sea

$$p(x, y) = \sum_{i=0}^d \sum_{j=0}^d c_{i,j} x^i y^j,$$

es un polinomio Hurwitz estable. Si para cada i , σ_i^x denota el i -ésimo polinomio elementario simétrico en las variables $x_1, \dots, x_d$, y σ_i^y el correspondiente en las variables $y_1, \dots, y_d$. Entonces el polinomio

$$P(x_1, \dots, x_d, y_1, \dots, y_d) = \sum_{i=0}^d \sum_{j=0}^d c_{i,j} \frac{\sigma_i^x \sigma_j^y}{\binom{d}{i} \binom{d}{j}}$$

es Hurwitz estable. Al polinomio P se le conoce como la polarización de p .

Por último, necesitamos la siguiente resultado de Lieb y Sokal [LS81] (Ver también el Teorema 8.4 de [BB09a]).

Proposición 3.1.24. Sean $P(z_1, \dots, z_d)$ y $Q(z_1, \dots, z_d)$ polinomios Hurwitz estables. Si D_i^z denota la derivada respecto a z_i , entonces

$$Q(D_1^z, \dots, D_d^z)P(z_1, \dots, z_d)$$

es Hurwitz estable.

Con los resultados anteriores es posible demostrar el Teorema 3.1.19 y posteriormente el Teorema 3.1.18.

3.1.2. Transformadas y Cotas

Antes de pasar a la parte de demostrar las cotas que nos interesan, daremos unos preliminares necesarios. Empecemos por observar que cada una de las tres convoluciones que estamos estudiando tiene un análogo en probabilidad libre, de hecho el nombre de cada una está inspirado en su contra parte libre. Recordemos también que en la teoría de probabilidad libre, cada una de estas tres convoluciones está equipada con una transformada natural de medidas de probabilidad. En esta subsección definiremos transformadas de polinomios análogas y las usaremos para acotar las raíces extremas de las convoluciones de polinomios.

Definición 3.1.25. 1. Veremos a un vector $(\lambda_1, \dots, \lambda_d)$ como una distribución discreta que toma cada valor con probabilidad $1/d$. Recordemos que la transformada de Cauchy-Hilbert-Stieltjes de tal distribución es

$$\mathcal{G}_\lambda(x) = \frac{1}{d} \sum_{i=1}^d \frac{1}{x - \lambda_i}.$$

2. Dado un polinomio p con raíces $\lambda_1, \dots, \lambda_d$, definimos $\mathcal{G}_p(x) = \mathcal{G}_\lambda(x)$.

3. La inversa de la transformada de Cauchy la definimos como

$$\mathcal{K}_p(w) = \max\{x : \mathcal{G}_p(x) = w\}.$$

4. La transformada R de Voiculescu está dada por $\mathcal{R}_\lambda(w) = \mathcal{K}_\lambda(w) - \frac{1}{w}$. Utilizaremos la misma notación para producir una transformación en polinomios,

$$\mathcal{R}_p(w) = \mathcal{K}_p(w) - \frac{1}{w}.$$

Observación 3.1.26. Observemos que para polinomios con raíces reales p , $\mathcal{K}_p(w)$ es el valor de x que es más grande que todas las raíces de p y tal que $\mathcal{G}_p(x) = w$. Además, como $\mathcal{G}_p(x) = \frac{1}{d} \frac{p'(x)}{p(x)}$,

$$\mathcal{G}_p(x) = w \quad \Leftrightarrow \quad p(x) - \frac{1}{wd} p'(x) = 0.$$

Esto nos dice que

$$\mathcal{K}_p(w) = \max \text{root}(U_{\frac{1}{wd}} p),$$

donde $U_\alpha p(x) = p(x) - \alpha Dp(x)$ y D es el operador diferencial.

Notación 3.1.27. Para simplificar la notación, denotaremos por

$$\max_\alpha(p) = \max \text{root}(U_\alpha p) = \mathcal{K}_p\left(\frac{1}{\alpha d}\right).$$

En probabilidad libre sabemos que si $\lambda \boxplus \mu$ es la convolución aditiva libre de dos distribuciones de probabilidad en los reales, λ y μ , entonces

$$\mathcal{R}_{\lambda \boxplus \mu}(w) = \mathcal{R}_\lambda(w) + \mathcal{R}_\mu(w).$$

Para la convolución aditiva finita, se obtiene una desigualdad análoga.

Las cotas que presentaremos a continuación se basan en el siguiente lema que nos permite juntar dos raíces del polinomio sin cambiar el valor de la transformada de Cauchy en un punto particular. Utilizando este lema varias veces podemos pasar de cualquier polinomio al caso de polinomios con una sola raíz, para los cuales las cotas son más sencillas.

Lema 3.1.28. Sea $\alpha > 0$, $d \geq 2$, y sea $p(x) \in \mathbb{P}(d)$ un polinomio con al menos dos raíces distintas. Digamos que $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d$ son todas las d raíces de p , donde lo anterior implica que $\lambda_1 > \lambda_k$ para algún k . Entonces existen reales μ y ρ tales que $p(x) = \hat{p}(x) + c\tilde{p}(x)$, donde $c > 0$ es una constante,

$$\tilde{p}(x) = (x-\mu)^2 \prod_{i \notin \{1,k\}} (x-\lambda_i) \in \mathbb{P}(d) \quad \text{y} \quad \hat{p}(x) = (x-\rho) \prod_{i \notin \{1,k\}} (x-\lambda_i) \in \mathbb{P}(d-1),$$

y además se cumple que

1. $\max_\alpha(\tilde{p}) = \max_\alpha(\hat{p}) = \max_\alpha(p)$.
2. $\lambda_1 > \mu > \lambda_k$.

3. $\rho > \lambda_1$. En particular, si $d \geq 3$, $\hat{p}$ tiene al menos dos raíces distintas.

Demostración. Lo que haremos será proponer μ , con lo que sabremos cómo es $\tilde{p}$, y esto a su vez nos determina qué polinomio debe ser $\hat{p}$, ρ y c . Posteriormente corroboraremos que se cumplen todas las propiedades. Denotamos por $t := \max_{\alpha}(p)$. La idea es elegir

$$\mu = t - \frac{2}{\frac{1}{t-\lambda_1} + \frac{1}{t-\lambda_k}},$$

de tal forma que

$$\frac{2}{t-\mu} = \frac{1}{t-\lambda_1} + \frac{1}{t-\lambda_k}.$$

Como queremos que $p = \tilde{p} + c\hat{p}$, entonces

$$\begin{aligned} c\hat{p}(x) &= p(x) - \tilde{p}(x) = [(x - \lambda_1)(x - \lambda_k) - (x - \mu)^2] \prod_{i \notin \{1,k\}} (x - \lambda_i) \\ &= [(2\mu - \lambda_1 - \lambda_k)x - (\mu^2 - \lambda_1 - \lambda_k)] \prod_{i \notin \{1,k\}} (x - \lambda_i), \end{aligned}$$

lo que nos dice que debemos tomar $c = 2\mu - \lambda_1 - \lambda_k$ y

$$\rho = \frac{\mu^2 - \lambda_1\lambda_k}{2\mu - \lambda_1 - \lambda_k} = \frac{\mu^2 - \lambda_1\lambda_k}{c}.$$

Observemos que en efecto $\hat{p}(x) \in \mathbb{P}(d-1)$.

De la definición de μ tenemos que $t - \mu$ es la media armónica de $t - \lambda_1$ y $t - \lambda_k$. Por un lado, esto nos dice que $t - \mu$ está entre $t - \lambda_1$ y $t - \lambda_k$ y por lo tanto μ está entre λ_1 y λ_k , con lo que obtenemos el inciso 2. Por otro lado, tenemos que $t - \mu$ es menor a la media aritmética, es decir,

$$t - \mu < \frac{t - \lambda_1 + t - \lambda_k}{2} = t - \frac{\lambda_1 + \lambda_k}{2}$$

que implica que $2\mu > \lambda_1 + \lambda_k$ y nos dice que c es positivo.

Para el inciso 3 observar que

$$\rho - \lambda_1 = \frac{\mu^2 - \lambda_1\lambda_k - c\lambda_1}{c} = \frac{\mu^2 - \lambda_1\lambda_k - 2\lambda_1\mu + \lambda_1^2 + \lambda\lambda_k}{c} = \frac{(\mu - \lambda_1)^2}{c} > 0.$$

Para el primer inciso, vemos que de la definición de μ tenemos que

$$\frac{D\tilde{p}(t)}{\tilde{p}(t)} = \frac{2}{t-\mu} + \sum_{i \notin \{1,k\}} \frac{1}{t-\lambda_i} = \frac{1}{t-\lambda_1} + \frac{1}{t-\lambda_k} + \sum_{i \notin \{1,k\}} \frac{1}{t-\lambda_i} = \frac{Dp(t)}{p(t)} = \frac{1}{\alpha},$$

de donde se sigue que $\max_{\alpha}(\tilde{p}) = \max_{\alpha}(p) = t$. Ahora, como U_{α} es lineal, entonces $c(U_{\alpha}\hat{p})(t) = (U_{\alpha}p)(t) - (U_{\alpha}\tilde{p})(t) = 0$, y por lo tanto t es raíz de $U_{\alpha}\hat{p}$. Como

$$t - \rho = \frac{ct - \mu^2 + \lambda_1\lambda_k}{c} = \frac{(\lambda_1\lambda_k - \lambda_1t - \lambda_2t + t^2) - (t^2 - 2\mu t + \mu^2)}{c}$$

$$= \frac{(\lambda_1 - t)(\lambda_k - t) - (t - \mu)^2}{c} > 0,$$

ya que $t \geq \lambda_1, \lambda_k$. Entonces $t > \rho$ y obtenemos que t es mayor que las raíces de $\hat{p}$, las cuales son ρ y λ_i con $i \notin \{1, k\}$. De esta forma concluimos que $t = \max_{\alpha}(\hat{p})$ y terminamos el inciso 1. $\square$

Convolución Aditiva Simétrica

En esta subsección nos dedicaremos a demostrar la siguiente cota para el caso aditivo simétrico.

Teorema 3.1.29. Para polinomios de grado d con raíces reales, p y q , se cumple que

$$\mathcal{R}_{p \boxplus_d q}(w) \leq \mathcal{R}_p(w) + \mathcal{R}_q(w).$$

Observación 3.1.30. 1. Por la Observación 3.1.26 esta desigualdad se puede reescribir como

$$\max_{\alpha}(p \boxplus_d q) + d\alpha \leq \max_{\alpha}(p) + \max_{\alpha}(q).$$

2. Observemos que esta cota es mejor que la de la Observación 3.1.5, especialmente cuando la mayoría de las raíces de p y q están lejos de sus raíces máximas.

Empezaremos por ver un lema que es consecuencia del Teorema 3.1.4 y que nos ayuda a lidiar con el caso en que p y q tienen distinto grado.

Lema 3.1.31. Para $p \in \mathbb{P}(d)$ y $q \in \mathbb{P}(k)$ con $k < d$, se cumple que

$$p \boxplus_d q = \frac{1}{d}(Dp) \boxplus_{d-1} q.$$

Demostración. Como $k < d$, entonces $q^{(d)}(y) = 0$ y por el Teorema 3.1.4 aplicado dos veces concluimos que

$$\begin{aligned} [p \boxplus_d q](x) &= \frac{1}{d!} \sum_{i=0}^d p^{(i)}(x) q^{(d-i)}(y) = \frac{1}{d!} \sum_{i=1}^d (p^{(1)})^{(i-1)}(x) q^{(d-i)}(y) \\ &\stackrel{(j=i-1)}{=} \frac{1}{d} \frac{1}{(d-1)!} \sum_{j=0}^{d-1} (Dp)^{(j)}(x) q^{(d-1-j)}(y) = \frac{1}{d} (Dp) \boxplus_{d-1} q. \end{aligned}$$

$\square$

La demostración del caso general es muy similar a la del caso particular en el que $p(x) \in \mathbb{P}(d)$ y $q(x) = x^{d-1}$, así que empezaremos por hacer este caso. Observemos que $\max_{\alpha}(x^{d-1}) = (d-1)\alpha$ y $p(x) \boxplus_d q(x) = Dp(x)$. Entonces lo que queremos probar es lo siguiente.

Lema 3.1.32. Para $\alpha \geq 0$, $d \geq 2$, y $p \in \mathbb{P}(d)$,

$$\max_{\alpha}(Dp) \leq \max_{\alpha}(p) - \alpha.$$

Demostración. Para un polinomio $p \in \mathbb{P}(d)$ definimos

$$\phi(p) = \max_{\alpha}(p) - \alpha - \max_{\alpha}(Dp).$$

Para cada $R > 0$, demostraremos por inducción sobre d que todos los polinomios $p \in \mathbb{P}(d)$ con todas sus raíces en $[-R, R]$ cumplen que $\phi(p) \geq 0$. Primero observemos que si p sólo tiene una raíz, es decir, si $p = (x - \lambda)^d$, entonces $\max_{\alpha}(p) = \lambda + d\alpha$ y $\max_{\alpha}(Dp) = \lambda + (d - 1)\alpha$, con lo que obtenemos la igualdad $\phi(p) = 0$. Por lo tanto, sólo tenemos que concentrarnos en los polinomios con al menos dos raíces distintas. Dado $R > 0$, sea p el polinomio con todas sus raíces en $[-R, R]$ que minimiza el valor de ϕ . Sabemos que existe este mínimo porque $[-R, R]^d$ es compacto. Supondremos que $\max_{\alpha}(Dp) > \max_{\alpha}(p) - \alpha$ para llegar a una contradicción. Por el Lema 3.1.28 podemos obtener polinomios $\tilde{p}$ y $\hat{p}$ tales que

$$\max_{\alpha}(\tilde{p}) - \alpha = \max_{\alpha}(\hat{p}) - \alpha = \max_{\alpha}(p) - \alpha.$$

Además, tenemos que $\tilde{p}$ tiene todas sus raíces en $[-R, R]$.

Ahora, en el caso base $d = 2$, tendremos que $\hat{p}$ tiene grado 1 y por lo tanto $U_{\alpha}D\hat{p}$ es igual al primer coeficiente de $\hat{p}$. Como este es positivo obtenemos que $(U_{\alpha}D\hat{p})(\max_{\alpha}(Dp)) > 0$. Como $\tilde{p} = p - c\hat{p}$ con c positivo, esto implica que $(U_{\alpha}D\tilde{p})(\max_{\alpha}(Dp)) < 0$. Entonces $\max_{\alpha}(D\tilde{p}) > \max_{\alpha}(Dp)$ y

$$\phi(\tilde{p}) = \max_{\alpha}(\tilde{p}) - \alpha - \max_{\alpha}(D\tilde{p}) < \max_{\alpha}(p) - \alpha - \max_{\alpha}(Dp) = \phi(p),$$

lo cual contradice el hecho de que p minimizaba ϕ .

Para $d \geq 3$, la hipótesis de inducción nos dice que $\max_{\alpha}(D\hat{p}) \leq \max_{\alpha}(Dp)$ y por lo tanto $(U_{\alpha}D\hat{p})(\max_{\alpha}(Dp)) > 0$, de manera análoga al caso base llegamos a una contradicción. Por lo tanto, concluimos que $\phi(p) \geq 0$ y como este es el valor mínimo obtenemos el resultado deseado. $\square$

Ahora sí estamos listos para demostrar el Teorema 3.1.29.

Demostración. Supongamos que $p \in \mathbb{P}(d)$ y $q \in \mathbb{P}(k)$. Procedemos por inducción sobre d . En el caso base de $d = 1$ observamos que $\max_{\alpha}(x - r) = r + \alpha$ y $(x - \lambda) \boxplus_1 (x - \mu) = (x - \lambda - \mu)$, por lo que se tiene la igualdad

$$\max_{\alpha}(p \boxplus_1 q) + \alpha = \max_{\alpha}(p) + \max_{\alpha}(q).$$

En el caso $d > k$, podemos usar el Lema 3.1.31, la hipótesis de inducción y el Lema 3.1.32 para terminar:

$$\max_{\alpha}(p \boxplus_d q) = \max_{\alpha}(Dp \boxplus_{d-1} q) \leq \max_{\alpha}(Dp) + \max_{\alpha}(q) - (d - 1)\alpha$$

$$\leq \max_{\alpha}(p) + \max_{\alpha}(q) - d\alpha,$$

En el caso $k = d$, fijemos un polinomio $q \in \mathbb{P}(d)$ y para $p \in \mathbb{P}(d)$ definimos

$$\phi(p) = \max_{\alpha}(p) + \max_{\alpha}(q) - d\alpha - \max_{\alpha}(p \boxplus_d q).$$

Para cada $R > 0$, demostraremos que todos los polinomios $p \in \mathbb{P}(d)$ con todas sus raíces en $[-R, R]$ cumplen que $\phi(p) \geq 0$. Primero observemos que si p sólo tiene una raíz, es decir, si $p = (x - \lambda)^d$, entonces $\max_{\alpha}(p) = \lambda + d\alpha$ y de la definición tenemos que $p \boxplus_d q = q(x - \lambda)$, por lo que $\max_{\alpha}(p \boxplus_d q) = \lambda + \max_{\alpha}(q)$, y entonces se obtiene la igualdad $\phi(p) = 0$. Por lo tanto, sólo tenemos que concentrarnos en los polinomios con al menos dos raíces distintas. Dado $R > 0$, sea p el polinomio con todas sus raíces en $[-R, R]$ que minimiza el valor de ϕ . Supondremos que

$$\max_{\alpha}(p \boxplus_d q) > \max_{\alpha}(p) + \max_{\alpha}(q) - d\alpha$$

para llegar a una contradicción. Por el Lema 3.1.28 podemos obtener polinomios $\tilde{p}$ y $\hat{p}$ tales que

$$\max_{\alpha}(\tilde{p}) + \max_{\alpha}(q) - d\alpha = \max_{\alpha}(\hat{p}) + \max_{\alpha}(q) - d\alpha = \max_{\alpha}(p) + \max_{\alpha}(q) - d\alpha.$$

Además, tenemos que $\tilde{p}$ tiene todas sus raíces en $[-R, R]$.

La hipótesis de inducción nos dice que

$$\max_{\alpha}(\hat{p} \boxplus_d q) \leq \max_{\alpha}(\hat{p}) + \max_{\alpha}(q) - d\alpha$$

y por lo tanto $(U_{\alpha}(\hat{p} \boxplus_d q))(\max_{\alpha}(p \boxplus_d q)) > 0$. Como $\tilde{p} = p - c\hat{p}$ con c positivo, esto implica que $(U_{\alpha}(\tilde{p} \boxplus_d q))(\max_{\alpha}(p \boxplus_d q)) < 0$. Entonces $\max_{\alpha}(\tilde{p} \boxplus_d q) > \max_{\alpha}(p \boxplus_d q)$ y

$$\phi(\tilde{p}) < \phi(p),$$

lo cual contradice el hecho de que p minimizaba ϕ . Por lo tanto, concluimos que $\phi(p) \geq 0$ y como este es el valor mínimo, terminamos la inducción. $\square$

Observación 3.1.33. Si se hacen las cuentas con mayor detenimiento es posible demostrar que la igualdad sólo se da cuando p o q son polinomios con una sola raíz.

Convolución Multiplicativa Simétrica

Para acotar las raíces de la convolución multiplicativa finita se utiliza una variante de la transformada S de Voiculescu. La cual definiremos como

$$\tilde{\mathcal{M}}_p(z) = z\mathcal{G}_p(z) - 1 = \frac{1}{d} \sum_{i=1}^d \sum_{j \geq 1} \left(\frac{\lambda_i}{z} \right)^j.$$

Para un polinomio p con raíces reales no negativas y $z > 0$, podemos observar que

$$z > \max \text{root}(p) \quad \Leftrightarrow \quad \tilde{\mathcal{M}}_p(z) < \infty.$$

Definimos la inversa de esta transformada, $\tilde{\mathcal{M}}_p^{(-1)}(w)$, como la mayor z tal que $\tilde{\mathcal{M}}_p(z) = w$, y definimos

$$\tilde{\mathcal{S}}_p(w) = \frac{w}{w+1} \tilde{\mathcal{M}}_p^{(-1)}(w).$$

Este es el recíproco de la transformada S usual. Con esta notación la cota que podemos obtener para la convolución multiplicativa es la siguiente

Teorema 3.1.34. Para polinomios con raíces reales no negativas, p y q , y para $w > 0$, se cumple que

$$\tilde{\mathcal{S}}_{p \boxtimes q}(w) = \tilde{\mathcal{S}}_p(w) \tilde{\mathcal{S}}_q(w).$$

La demostración de este teorema es similar a la del caso aditivo simétrico sólo que teniendo un poco más de cuidado con las cuentas, las cuales se vuelven un poco más complicadas debido a la necesidad de trabajar sólo con polinomios que tienen puras raíces positivas. Primero hacemos el caso para cuando p tiene sólo una raíz. Luego obtenemos un lema que nos ayuda a lidiar con el caso en que p y q tienen distinto grado, el cual ahora es consecuencia directa del Teorema 3.1.7 y en el que en lugar de tomar la derivada del polinomio con grado mayor, ahora tomamos la derivada polar respecto a 0 de este polinomio. En este caso necesitaremos reescribir la derivada polar y obtener algunas propiedades de este operador. Por último necesitamos un corolario del Lema 3.1.28, el cual nos encamina más a los polinomios específicos que deseamos construir en este caso. El lector interesado en todos los detalles puede consultar la Sección 4.2 de [MSS15a].

Convolución Aditiva Asimétrica

Para definir la transformada para la convolución aditiva asimétrica, necesitaremos de una operación que nos permita evaluar polinomios en x^2 en lugar de x .

Definición 3.1.35. Definimos $\mathbb{S}$ como la operación que convierte un polinomio en $p(x)$ en el polinomio $p(x^2)$. Observemos que si p tiene sólo raíces positivas reales, λ_i , entonces $\mathbb{S}p$ tiene raíces $\pm\sqrt{\lambda_i}$. Usaremos esta misma notación para cuando λ es una distribución de probabilidad con soporte en los reales no negativos, en este caso definimos $\mathbb{S}\lambda$ como la correspondiente distribución de probabilidad en $\pm\sqrt{\lambda_i}$.

Observación 3.1.36. En probabilidad libre se sabe que si λ y μ son distribuciones de probabilidad en los números positivos, y si $\boxplus$ es su convolución asimétrica libre, entonces

$$\mathcal{R}_{\mathbb{S}\lambda \boxplus \mathbb{S}\mu}(w) = \mathcal{R}_{\mathbb{S}\lambda}(w) + \mathcal{R}_{\mathbb{S}\mu}(w).$$

Inspirados en la igualdad anterior se puede obtener la siguiente cota para la convolución aditiva asimétrica.

Teorema 3.1.37. Para polinomios de grado d con raíces reales no negativas, p y q , se cumple que

$$\mathcal{R}_{\mathbb{S}(p \boxplus_d q)}(w) = \mathcal{R}_{\mathbb{S}p}(w) + \mathcal{R}_{\mathbb{S}q}(w).$$

Para demostrar este teorema primero podemos reescribirlo en términos de $\max_\alpha$ de la siguiente manera:

Teorema 3.1.38. Para $d \geq 1$ y $p, q \in \mathbb{P}^+(d)$, se cumple que para toda α ,

$$\max_\alpha(\mathbb{S}(p \boxplus_d q)) = \max_\alpha(\mathbb{S}p) + \max_\alpha(\mathbb{S}q) - 2\alpha d,$$

con igualdad sólo si alguno de los polinomios es x^d .

Con el siguiente lema, que es una especie de normalización, vemos que basta probar el teorema anterior para el caso particular $\alpha = 1$.

Lema 3.1.39. Para un polinomio con raíces reales, $p(x)$, se cumple que

$$\max_\alpha(p(x)) = \frac{1}{\alpha} \max \text{root}(U_1 p(\alpha x)) = \frac{1}{\alpha} \max_1(p(\alpha x)).$$

Como con las otras convoluciones, observamos que para lidiar con dos polinomios de distinto grado, basta primero aplicarle una operación al polinomio de mayor grado. En este caso la operación será DxD , también conocida como la derivada de Laguerre. Luego hay que reescribir la convolución aditiva simétrica en términos de esta derivada, ahora estos resultados se siguen del Teorema 3.1.10. Posteriormente utilizamos una consecuencia del Lema 3.1.28 para reducir el análisis a casos especiales. Por último se hacen los 3 casos especiales:

- $p(x) = (x - \lambda)^d$ y $q(x) = x^{d-1}$.
- $p(x) \in \mathbb{P}^+(d)$ y $q(x) = x^{d-1}$.
- $p(x) = (x - \lambda)^d$ y $q(x) = (x - \mu)^d$.

Cabe mencionar que la demostración del último caso no es nada sencilla y los autores dedican toda una subsección a esto. La idea es que hay que acotar la transformada de Cauchy de polinomios del tipo

$$q_d^{\lambda, \mu}(x) = (x - \lambda)^d \boxplus_d (x - \mu)^d,$$

para lo cual hay que relacionar esta transformada con polinomios de Chebyshev del segundo tipo, y esto requiere varios pasos y cotas un poco técnicas. El lector interesado en todos los detalles puede consultar las Secciones 4.3 y 4.4 de [MSS15a].

3.2. Gráficas de Ramanujan bipartitas de todos los tamaños

En esta sección analizaremos los resultados del último artículo de la serie [MSS15d]. Para ello será útil tener en mente lo que ya vimos sobre gráficas de Ramanujan en la primer sección del capítulo pasado. El teorema principal de este artículo es que la unión de m emparejamientos perfectos aleatorios de una bipartición de $2d$ vértices es Ramanujan con probabilidad positiva.

Teorema 3.2.1. Para $d \geq 3$, sean $P_1, \dots, P_d$ matrices de permutación independientes y uniformemente aleatorias de tamaño $n \times n$. Entonces, con probabilidad positiva, los valores propios no triviales de

$$A = \sum_{i=1}^d \begin{pmatrix} 0 & P_i \\ P_i^T & 0 \end{pmatrix}$$

son todos menores que $2\sqrt{d-1}$ en valor absoluto.

En este artículo también se demuestra una versión no bipartita.

Teorema 3.2.2. Para $d \geq 3$ y n par, sean $P_1, \dots, P_d$ matrices de permutación independientes y uniformemente aleatorias de tamaño $n \times n$. Sea M la matriz de adyacencia de cualquier emparejamiento perfecto fijo en d vértices. Entonces, con probabilidad positiva, los valores propios no triviales de

$$A = \sum_{i=1}^d P_i M P_i^T$$

son todos menores que $2\sqrt{d-1}$.

Observación 3.2.3. 1. Observemos que en el primer teorema, por cómo está construido A , esta será la matriz de adyacencia de una multigráfica bipartita y además es Ramanujan por que sus valores propios están acotados en valor absoluto por $2\sqrt{d-1}$. Mientras tanto, en el segundo teorema, por cómo está construido A , obtenemos una multigráfica d -regular. Sin embargo, no aseguramos que sea de Ramanujan debido a que no tenemos control sobre el menor valor propio.

2. Es importante recalcar que como en ambos teoremas tendremos uniones de emparejamientos independientes, las gráficas que se producen pueden tener multiaristas y por ende son multigráficas.
3. Observemos que si nos olvidamos del hecho de que son multigráficas, estos teoremas se pueden ver como una generalización del resultado principal de la Sección 2.1 (Ver Teorema 2.1.24). Esto es porque ahora seríamos capaces de construir gráficas de Ramanujan d -regulares en n

vértices para cualquier grado d y cualquier cantidad par de vertices n , mientras que en la Sección 2.1 sólo podíamos hacer esto para cualquier grado d y una infinidad específica de valores n que dependían del grado, los cuales eran de la forma $n = 2^k(d + 1)$ con $k \in \mathbb{N}$.

4. Hay que destacar que en este artículo no se presenta una forma de construir gráficas de Ramanujan en tiempo polinomial. Aunque esto puede ser de interés matemático, para las aplicaciones no es tan necesario ya que basta con generar gráficas casi Ramanujan y ya se cuenta con algoritmos eficientes para este fin.

La demostración de ambos teoremas sigue la misma estrategia y es similar a lo que hemos estado haciendo en las demás aplicaciones. En este trabajo sólo presentaremos el caso no bipartito, que es considerablemente más simple, y para el caso bipartito sólo daremos la idea general. La estrategia en ambos casos consiste de tres pasos importantes:

- Primero, se muestra que el polinomio característico esperado (de las gráficas aleatorias que nos importan) tiene raíces reales y proviene de familias de polinomios que se entrelazan, lo cual reduce el problema al de analizar las raíces de estos polinomios.
- Posteriormente, se deriva una fórmula cerrada para este polinomio característico esperado utilizando un teorema de cuadratura similar al de la sección anterior.
- Por último, se utilizan las cotas obtenidas en la sección anterior para acotar el espectro del polinomio característico esperado.

3.2.1. Entrelazado para Permutaciones

En esta subsección descompondremos las permutaciones aleatorias en transposiciones que actúan en dos vértices a la vez y mostraremos que estas transposiciones corresponden a transformaciones lineales que preservan las propiedades de raíces reales de ciertos polinomios multivariados. La parte importante será ver que si A y B son matrices simétricas, entonces el polinomio característico esperado de $A + PBP^T$ tiene raíces reales para una matriz de permutación P . Para ello necesitaremos de una clase de polinomios tipo determinante que serán el análogo a los polinomios característico mixtos de los capítulos pasados.

Empezaremos por definir qué es una transposición aleatoria y la clase especial de permutaciones aleatorias con las que trabajaremos, que son precisamente aquellas que se pueden obtener como composición de transposiciones aleatorias.

Definición 3.2.4 (Transposición aleatoria). Una transposición aleatoria es una matriz aleatoria que es igual a una transposición de dos índices a, b (fijos) con probabilidad α y que es igual a la identidad con probabilidad $1 - \alpha$, para algún $\alpha \in [0, 1]$.

Definición 3.2.5 (Realizable con transposiciones). Decimos que una matriz aleatoria p con soporte en las matrices de permutación, es realizable con transposiciones si existen transposiciones aleatorias $S_1, \dots, S_N$ tales que la distribución de P es la misma que la distribución del producto $S_N S_{N-1} \dots S_2 S_1$.

La importante de la última definición es que una matriz de permutación uniformemente aleatoria es realizable con transposiciones.

Lema 3.2.6. Sean P y S matrices de permutación de $n \times n$ uniformemente aleatorias. Entonces P y $P \oplus S = \begin{pmatrix} P & 0 \\ 0 & S \end{pmatrix}$ son realizables con transposiciones.

Demostración. Supongamos que S_{1k} es la transposición aleatoria que con probabilidad $\frac{1}{k}$ es la identidad y con probabilidad $\frac{k-1}{k}$ transpone e_1 con e_k . Definimos $M_2 = S_{12}$ y recursivamente para $k > 2$ definimos $M_k = M_{k-1} S_{1k} M_{k-1}$. Veamos por inducción que M_k permuta uniformemente las primeras k entradas (el resto se queda fijo porque ninguna transposición actúa en las demás entradas). El caso base $k = 2$ es trivial. Y si suponemos la hipótesis cierta para $k - 1$, tenemos dos posibles casos.

Con probabilidad $\frac{1}{k}$ tendremos que S_{1k} es la identidad, en este caso $M_k = M_{k-1} S_{1k} M_{k-1}$ deja fijo a k mientras que los demás índices se permutan de manera uniforme.

Con probabilidad $\frac{k-1}{k}$ transponemos e_1 con e_k , en este caso $M_k = M_{k-1} S_{1k} M_{k-1}$ manda a k en algún $j = 1, \dots, k-1$ con probabilidad $\frac{1}{k-1} \frac{k-1}{k} = \frac{1}{k}$, mientras que las entradas $\{1, \dots, k\} \setminus \{j\}$ se permutan uniformemente.

En ambos casos concluimos que M_k permuta uniformemente las primeras k entradas y por lo tanto M_n es una permutación aleatoria uniforme que es realizable con transposiciones.

Para $P \oplus S$ utilizamos el argumento anterior para realizar $P \oplus I$, luego para $I \oplus S$ y los multiplicamos. $\square$

Ahora presentaremos la noción de polinomios tipo determinante que nos será útil para encontrar polinomios univariados con raíces reales. Esto se puede ver como el análogo a estabilidad real que utilizamos en los capítulos anteriores.

Definición 3.2.7 (Polinomios tipo determinante). Un polinomio homogéneo $P(X_1, \dots, X_m)$ de grado d en las entradas de m matrices simétricas de $d \times d$, $X_1, \dots, X_m$, se llama tipo determinante si cumple las siguientes dos propiedades.

- **Hiperbolicidad respecto a $(I, I, \dots, I)$.** Las restricciones univariadas

$$q(t) = P(tI - A_1, \dots, tI - A_m)$$

tienen raíces reales para todas las matrices simétricas $A_1, \dots, A_m$. Para los interesados en aprender más sobre condición de hiperbolicidad los referimos a [Pem12].

- **Linealidad de rango 1.** Para un vector v , un índice $i \leq m$, y un real s , se cumple que

$$\begin{aligned} P(X_1, X_2, \dots, X_i + svv^T, \dots, X_m) \\ = P(X_1, X_2, \dots, X_m) + sD_{i, vv^T}P(X_1, X_2, \dots, X_m), \end{aligned}$$

donde denotamos por

$$D_{i, vv^T}P(X_1, X_2, \dots, X_m) = \left(\frac{\partial}{\partial s} P(X_1, X_2, \dots, X_i + svv^T, \dots, X_m) \right) \Big|_{s=0}$$

a la derivada direccional de P en dirección $(0, \dots, 0, vv^T, 0, \dots, 0)$, con vv^T en la posición i . Observemos que $D_{i, vv^T}P(X_1, X_2, \dots, X_m)$ es homogéneo de grado $d - 1$.

Ejemplo 3.2.8. Un ejemplo importante de polinomios tipo determinante es precisamente el determinante de una suma de matrices:

$$P(X_1, \dots, X_m) = \det(X_1 + \dots + X_m).$$

En efecto, la hiperbolicidad se sigue de que $P(tI - A_1, \dots, tI - A_m) = \det(mI - A_1 - \dots - A_m)$ es el polinomio característico mixto de una matriz simétrica, mientras que la linealidad de rango uno se sigue de la invarianza del determinante bajo cambio de bases y su linealidad respecto a las entradas de la matriz. Esto último también se puede hacer utilizando el Lema 1.2.10.

En la siguiente observación presentaremos tres resultados sencillos relacionados con los polinomios tipo determinante que nos serán útiles.

Observación 3.2.9. 1. **Los cambios de rango 1 se entrelazan.** Es decir, si $P(X_1, \dots, X_m)$ es del tipo determinante, entonces para cualquier entrada i , todo vector v y matrices simétricas $A_1, \dots, A_m$ tenemos que

$$P(tI - A_1, \dots, tI - A_m) \quad \text{y} \quad P(tI - A_1, \dots, tI - A_i - vv^T, \dots, tI - A_m)$$

se entrelazan.

En efecto, observemos que basta hacerlo para la primera entrada y en este caso, de la linealidad de rango 1 tenemos que

$$P(tI - A_1 - svv^T, tI - A_2, \dots, tI - A_m)$$

$$= P(tI - A_1, \dots, tI - A_m) - sD_{vv^T}P(tI - A_1, \dots, tI - A_m).$$

Como P es hiperbólico, si lo vemos como polinomio univariado en t , tiene raíces reales y por el Lema 1.1.5 sabemos que $D_{vv^T}P(tI - A_1, \dots, tI - A_m)$ entrelaza a $P(tI - A_1, \dots, tI - A_m)$ y de propiedades simples de entrelazado (se puede consultar [Fis06]) concluimos que $P(tI - A_1, \dots, tI - A_m)$ entrelaza a

$$\begin{aligned} & P(tI - A_1, \dots, tI - A_m) - D_{vv^T}P(tI - A_1, \dots, tI - A_m) \\ &= P(tI - A_1 - vv^T, tI - A_2, \dots, tI - A_m). \end{aligned}$$

2. **Las permutaciones preservan linealidad de rango 1.** Como el conjunto de matrices de rango 1 es invariante bajo conjugación por permutaciones, entonces si Π es una matriz de permutación y $P(X_1, \dots, X_m)$ es de rango 1 lineal, tendremos que $P(\Pi X_1 \Pi^T, X_2, \dots, X_m)$ también es de rango 1 lineal. Además, de la linealidad de D_{i, vv^T} se sigue que si P y Q son de rango 1 lineal, entonces $P + Q$ también es de rango 1 lineal.
3. **Si conjugamos una matriz simétrica por una transposición y la restamos de la matriz original obtenemos una matriz de rango 2 y traza 0.** En efecto, si A es simétrica con entradas a_{ij} y suponemos sin pérdida de generalidad que σ es una transposición de las dos primeras entradas, entonces tendremos que

$$A - \sigma A \sigma^T = \begin{pmatrix} a_{11} - a_{22} & a_{12} - a_{21} & a_{13} - a_{23} & a_{14} - a_{24} & \dots \\ a_{21} - a_{12} & a_{22} - a_{11} & a_{23} - a_{13} & a_{24} - a_{14} & \dots \\ a_{31} - a_{32} & a_{32} - a_{31} & 0 & 0 & \dots \\ a_{41} - a_{42} & a_{42} - a_{41} & 0 & \ddots & \\ \vdots & \vdots & \vdots & & \ddots \end{pmatrix}.$$

Entonces, es claro que tiene traza 0. Además, si $a_{11} - a_{22} = a_{12} - a_{21}$, las primeras dos filas son linealmente independientes y el resto son múltiplos de $(1, -1, 0, \dots, 0)$, mientras que si $a_{11} - a_{22} \neq a_{12} - a_{21}$, entonces la suma de las primeras dos filas es $(c, -c, 0, \dots, 0)$ con $c \neq 0$ y las demás son múltiplos escalares de esta. Por lo que concluimos que $A - \sigma A \sigma^T$ tiene rango 2.

Ahora utilizaremos lo que acabamos de aprender en un resultado que será bastante útil, el cual establece que las transposiciones aleatorias definen operadores lineales que preservan la propiedad de ser tipo determinante.

Lema 3.2.10 (Transposiciones aleatorias preservan el tipo determinante). Si $P(X_1, \dots, X_m)$ es del tipo determinante, entonces para cualquier $i = 1, \dots, m$ y transposición aleatoria S , el polinomio

$$\mathbb{E}_S P(X_1, \dots, S X_i S^T, \dots, X_m)$$

es del tipo determinante.

Demostración. Observemos que basta hacerlo para la primer entrada, ($i = 1$). Supongamos que S es igual a una transposición σ con probabilidad α e igual a la identidad con probabilidad $1 - \alpha$, entonces queremos ver que

$$\begin{aligned} Q(X_1, \dots, X_m) &= \mathbb{E}_S P(X_1, \dots, SX_i S^T, \dots, X_m) \\ &= (1 - \alpha)P(X_1, \dots, X_m) + \alpha P(\sigma X_1 \sigma^T, X_2, \dots, X_m) \end{aligned}$$

es hiperbólico y lineal de rango 1. La homogeneidad se sigue de que las transposiciones y combinaciones convexas preservan homogeneidad. El inciso 2 de la Observación 3.2.9 nos dice que la linealidad de rango 1 también se preserva. Para la hiperbolicidad debemos ver que cualquier restricción univariada en dirección $(I, \dots, I)$,

$$\begin{aligned} Q(tI - A_1, \dots, tI - A_m) &= (1 - \alpha)P(tI - A_1, \dots, tI - A_m) \\ &\quad + \alpha P(tI - \sigma A_1 \sigma^T, tI - A_2, \dots, tI - A_m), \end{aligned}$$

tiene raíces reales. Por el inciso 3 de la Observación 3.2.9 sabemos que $\sigma A_1 \sigma^T - A_1$ es de rango 2 y tiene traza 0, por lo que existen vectores a y b tales que

$$P(tI - \sigma A_1 \sigma^T, tI - A_2, \dots, tI - A_m) = P(tI - A_1 + bb^T - aa^T, tI - A_2, \dots, tI - A_m).$$

Por último, como P es del tipo determinante, por el inciso 1 de la Observación 3.2.9 tenemos que $P(tI - A_1 + bb^T, tI - A_2, \dots, tI - A_m)$ entrelaza tanto a

$$P(tI - A_1 + bb^T - aa^T, tI - A_2, \dots, tI - A_m)$$

como a

$$P(tI - A_1, tI - A_2, \dots, tI - A_m).$$

Esto nos dice que estos dos últimos tienen un entrelazado común y el Lema 1.1.5 nos permite concluir que su combinación convexa $Q(tI - A_1, \dots, tI - A_m)$ tiene raíces reales y terminamos. $\square$

El resultado principal de esta subsección es el siguiente y se obtiene de aplicar nN veces el lema que acabamos de probar. Este resultado establece que los polinomios característicos esperados de productos de transposiciones aleatorias siempre tienen raíces reales. Recordemos que estos polinomios cumplen la misma función que los polinomios característicos mixtos del capítulo pasado.

Teorema 3.2.11. Sean $A_1, \dots, A_m$ matrices simétricas de $d \times d$ y sean $\{S_{ij}\}_{i \leq m, j \leq N}$ transposiciones aleatorias independientes (no necesariamente idénticas). Entonces

$$\mathbb{E} \chi \left(\sum_{i=1}^m S_{i,N}^1 S_{i,N-1} \cdots S_{i1} A_i S_{i1}^T \cdots S_{iN}^T \right),$$

tiene raíces reales.

Demostración. Empezamos con $P(X_1, \dots, X_m) = \det(X_1 + \dots + X_m)$ que ya sabemos que es del tipo determinante y aplicamos nN veces el lema anterior para ver que cada transposición S_{ij} preserva el tipo determinante. Con ello obtenemos que

$$\mathbb{E}_{S_{1N}} \cdots \mathbb{E}_{S_{11}} \mathbb{E}_{S_{2N}} \cdots \mathbb{E}_{S_{n1}} \det \left(\sum_{i=1}^m S_{i,N}^1 S_{i,N-1} \cdots S_{i1} A_i S_{i1}^T \cdots S_{iN}^T \right),$$

es del tipo determinante y al tomar la restricción $X_i = \frac{t}{m}I - A_i$ terminamos. $\square$

Como consecuencia inmediata del teorema anterior y la teoría de familias de polinomios que se entrelazan obtenemos el siguiente resultado.

Teorema 3.2.12. Sean $A_1, \dots, A_d$ matrices simétricas de $n \times n$ y sean $P_1, \dots, P_d$ permutaciones aleatorias independientes realizables con transposiciones. Entonces para todo $k \leq n$, con probabilidad positiva se tiene

$$\lambda_k \left(\sum_{i=1}^d P_i A_i P_i^T \right) \leq \lambda_k \left(\mathbb{E}_\chi \left(\sum_{i=1}^d P_i A_i P_i^T \right) \right).$$

3.2.2. Cuadratura

En esta subsección derivaremos una fórmula cerrada para el polinomio característico esperado de la suma de permutaciones aleatorias. Primero hay que observar que los polinomios característicos esperados de permutaciones aleatorias se puede reemplazar por polinomios característicos esperados de matrices ortogonales aleatorias. Esto nos dice que los polinomios característicos no pueden distinguir entre el conjunto de matrices de permutación y el conjunto de matrices ortogonales. Esencialmente, esto sucede porque los determinantes son multilineales, lo cual causa ciertas restricciones para que tengan coeficientes de Fourier de grado muy bajo. Este teorema de cuadratura es bastante similar al que mencionamos en la sección anterior para la convolución aditiva simétrica y que esbozamos para el caso asimétrico. Debido a que es bastante técnico sólo daremos la idea general, referimos al lector a la Sección 4 de [MSS15d] para llenar los detalles.

El objetivo principal de este capítulo es demostrar el siguiente resultado

Teorema 3.2.13. Supongamos que A y B son matrices simétricas de $d \times d$ con $A\mathbf{1} = a\mathbf{1}$ y $B\mathbf{1} = b\mathbf{1}$, donde $\mathbf{1} = (1, \dots, 1)$. Sea $\chi_x(A) = (x - a)p(x)$ y $\chi_x(B) = (x - b)q(x)$. Entonces

$$\mathbb{E}_P \chi_x(A + PBP^T) = (x - (a + b))p(x) \boxplus_{d-1} q(x),$$

donde P es una permutación aleatoria uniforme.

Observemos que las matrices A, B, P tienen el vector propio $\mathbf{1}$, lo que nos dice que podemos tomar el cambio de base V que manda $\mathbf{1}$ al vector e_n , y digonalizarlas en bloques simultaneamente, es decir,

$$VAV^T = \begin{pmatrix} \hat{A} & 0 \\ 0 & a \end{pmatrix} = \hat{A} \oplus a, \quad VBV^T = \hat{B} \oplus b, \quad VPV^T = \hat{P} \oplus 1.$$

Como el determinante es invariante cuando cambiamos la base, tenemos que

$$\begin{aligned} \mathbb{E}_P \det(xI - A - PBP^T) &= \mathbb{E}_P \det(xI - VAV^T - (VPV^T)(VBV^T)(VP^TV^T)) \\ &= \mathbb{E}_{\hat{P}} \det(xI - \hat{A} \oplus a - (\hat{P} \oplus 1)(\hat{B} \oplus b)(\hat{P}^T \oplus 1)) \\ &= \mathbb{E}_{\hat{P}} \det((xI - \hat{A} - \hat{P}\hat{B}\hat{P}^T) \oplus (x - a - b)) \\ &= (x - a - b) \mathbb{E}_{\hat{P}} \det(xI - \hat{A} - \hat{P}\hat{B}\hat{P}^T), \end{aligned}$$

donde para una P , la transformación ortogonal $\hat{P}$ permuta a $\hat{e}_1, \dots, \hat{e}_n$, las proyecciones ortogonales a $\mathbf{1}$ de la base estándar de vectores $e_1, \dots, e_n$. Como estos son los vértices de un simplejo regular con d vértices en $\mathbb{R}^{d-1}$ centrado en el origen, podemos ver a los $\hat{P}$ como elementos del grupo simétrico de este simplejo, denotaremos a este grupo como A_{d-1} . Por otro lado, como $p(x) = \chi(\hat{A})$ y $q(x) = \chi(\hat{B})$, tenemos que $p(x) \boxplus_{d-1} q(x) = \mathbb{E}_Q \chi(\hat{A} - Q\hat{B}Q^T)$.

Por lo tanto el Teorema 3.2.13 se reduce a demostrar que cualesquiera matrices $\hat{A}$ y $\hat{B}$ de $(d-1) \times (d-1)$ cumplen que

$$\mathbb{E}_{\hat{P}} \chi(\hat{A} + \hat{P}\hat{B}\hat{P}^T) = \mathbb{E}_Q \chi(\hat{A} + Q\hat{B}Q^T),$$

donde Q es una matriz ortogonal aleatoria de Haar de tamaño $(d-1) \times (d-1)$ y P se distribuye uniformemente en A_{d-1} . Observemos que esto es muy parecido al teorema de cuadratura de la sección anterior, sólo que ahora P es un elemento en A_{d-1} en lugar de una permutación aleatoria signada. Como $\hat{A}$ es matriz simétrica, si tomamos $A = Ix - \hat{A}$ y $B = \hat{B}$, la igualdad anterior es equivalente a demostrar el siguiente teorema.

Teorema 3.2.14 (Teorema de cuadratura). Supongamos que A y B son matrices simétricas de $d \times d$, entonces

$$\mathbb{E}_{P \in A_d} \det(A + PBP^T) = \mathbb{E}_{Q \in O(d)} \det(A + QBQ^T),$$

donde A_d son las permutaciones.

Como ya mencionamos, la demostración de este teorema es un poco técnica y larga. Así que al igual que en la sección anterior sólo mencionaremos las ideas importantes. Los dos pasos importantes son:

- Demostrar que se tiene invarianza para A_d . Es decir, que el lado izquierdo es invariante si multiplicamos (del lado derecho) a P por matrices ortogonales.

- Ver que la Invarianza implica cuadratura.

Invarianza para A_d . La primer parte se resume en el siguiente teorema.

Teorema 3.2.15. Sean A y B matrices de $d \times d$, y sea

$$f_{A,B}(Q) = \det(A + QBQ^T).$$

Entonces, para todo $Q_0 \in O(d)$,

$$\mathbb{E}_{P \in A_d} f_{A,B}(P) = \mathbb{E}_{P \in A_d} f_{A,B}(PQ_0).$$

Para demostrar esta invarianza, lo dividimos en varios pasos.

1. Primero recordamos el Lema 3.1.14 que nos dice que los determinantes son de grado pequeño en rotaciones de rango 2.
2. Luego lo utilizamos para demostrar que se tiene Invarianza para A_2 , esto es similar al Teorema 3.1.16.
3. Posteriormente vemos que se tiene invarianza para $A_{i,j,k}$, que es el subgrupo de A_d que permuta las entradas i, j y k .
4. Por último se demuestra que $O_{i,j,k}$ genera a $O(d)$. Donde $O_{i,j,k}$ actúa en el subespacio lineal bidimensional paralelo al subespacio afín que generan los vértices i, j y k , y deja fijo al subespacio ortogonal a este.

Invarianza implica cuadratura. Esta parte es bastante más sencilla y se sigue del siguiente resultado más general.

Lema 3.2.16. Sea f una función de $O(d)$ a $\mathbb{R}$ y sea H un subgrupo finito de $O(d)$. Si

$$\mathbb{E}_{P \in H} f(P) = \mathbb{E}_{P \in H} f(PQ_0), \quad \forall Q_0 \in O(d),$$

entonces

$$\mathbb{E}_{P \in H} f(P) = \mathbb{E}_{Q \in O(d)} f(Q),$$

donde Q se distribuye de acuerdo a la medida de Haar y P es uniforme en H .

Para terminar esta sección, observemos que el Teorema 3.2.13 se puede generalizar fácilmente para suma de varias matrices, para ello basta aplicarlo $d - 1$ veces.

Teorema 3.2.17. Supongamos que $A_1, \dots, A_d$ son matrices simétricas de $n \times n$ con $A_i \mathbf{1} = a_i \mathbf{1}$ y $\chi_x(A_i) = (x - a_i)p_i(x)$. Entonces

$$\mathbb{E}_{P_1, \dots, P_d} \chi_x \left(\sum_{i=1}^d P_i A_i P_i^T \right) = \left(x - \sum_{i=1}^d a_i \right) p_1(x) \boxplus \dots \boxplus p_d(x),$$

donde $P_1, \dots, P_d$ son matrices de permutación aleatorias uniformes e independientes.

Ahora sí estamos listos para demostrar el Teorema 3.2.2.

Demostración. Sea M la matriz de adyacencia de un emparejamiento perfecto en n vértices, con n par. Por el Lema 3.2.6, la distribución uniforme en permutaciones es realizable con transposiciones aleatorias. Entonces, el Teorema 3.2.12 nos dice que con probabilidad positiva

$$\lambda_k \left(\sum_{i=1}^d P_i A_i P_i^T \right) \leq \lambda_k \left(\mathbb{E}_{\chi_x} \left(\sum_{i=1}^d P_i A_i P_i^T \right) \right).$$

Por el Teorema 3.2.17 sabemos que

$$\mathbb{E}_{\chi_x} \left(\sum_{i=1}^d P_i A_i P_i^T \right) = (x-d)(p(x) \boxplus_{n-1} \cdots \boxplus_{n-1} p(x)),$$

donde $\boxplus_{n-1}$ se aplica d veces y

$$p(x) = \frac{\chi_x(M)}{x-1} = (x-1)^{\frac{d}{2}-1} (x+1)^{\frac{d}{2}},$$

es el polinomio característico de un emparejamiento perfecto al que le quitamos la raíz trivial 1 (una de las múltiples que tiene). Por lo tanto, ahora debemos enfocarnos en acotar la raíz más grande de $p^{\boxplus d} = p \boxplus_{n-1} \cdots \boxplus_{n-1} p$, lo cual ya sabemos hacer gracias a la sección anterior. Observemos que la transformada de Cauchy de $p(x)$ es menor a la de $\chi(M)(x)$ para $x > 1$,

$$\mathcal{G}_p(x) = \frac{\frac{d}{2}-1}{d-1} \frac{1}{x-1} + \frac{\frac{d}{2}}{d-1} \frac{1}{x+1} < \frac{1}{2} \frac{1}{x-1} + \frac{1}{2} \frac{1}{x+1} = \frac{x}{x^2-1} = \mathcal{G}_{\chi(M)}(x).$$

Como ambas funciones son decrecientes en $x > 1$, obtenemos la siguiente cota para la transformada de Cauchy,

$$\mathcal{K}_p(w) < \mathcal{K}_{\chi(M)}(w) \quad \forall w > 0.$$

Si utilizamos $d-1$ veces la cota para la convolución aditiva simétrica obtenida en el Teorema 3.1.29 llegamos a que

$$\begin{aligned} \mathcal{K}_{p^{\boxplus d}} &\leq d \cdot \mathcal{K}_p(w) - \frac{d-1}{w} < d \cdot \mathcal{K}_{\chi(M)}(w) - \frac{d-1}{w} = dx - \frac{d-1}{w} \\ &= dx - \frac{(d-1)(x^2-1)}{x} = \frac{x^2+d-1}{x}, \end{aligned}$$

donde tomamos $\mathcal{K}_{\chi(M)}(w) = x$ y usamos que $w = \frac{x}{x^2-1}$. Por último, podemos ver que esta expresión se minimiza cuando $x = \sqrt{d-1}$ y se obtiene el valor $2\sqrt{d-1}$. Y como $\mathcal{K}_{p^{\boxplus d}}$ acota la raíz más grande de $p^{\boxplus d}$, concluimos que

existen matrices de permutación $P_1, \dots, P_d$ de tamaño $n \times n$, tales que los valores propios no triviales de

$$A = \sum_{i=1}^d P_i M P_i^T$$

son todos menores que $2\sqrt{d-1}$. Y con esto terminamos. □

La demostración del Teorema 3.2.1 es similar a la del Teorema 3.2.2, sólo que ahora utilizaremos la cota para la convolución aditiva asimétrica obtenida en el Teorema 3.1.37.

Bibliografía

- [AA91] Charles A Akemann and Joel Anderson. *Lyapunov theorems for operator algebras*, volume 458. American Mathematical Society, 1991.
- [AAT12] Charles Akemann, Joel Anderson, and Betül Tanbay. The kadison–singer problem for the direct sum of matrix algebras. *Positivity*, 16(1):53–66, 2012.
- [ABG10] Louigi Addario-Berry and Simon Griffiths. The spectrum of random lifts. *arXiv preprint arXiv:1012.4097*, 2010.
- [AL06] Alon Amit and Nathan Linial. Random lifts of graphs: edge expansion. *Combinatorics, Probability and Computing*, 15(3):317–332, 2006.
- [ALM02] Alon Amit, Nathan Linial, and Jiří Matoušek. Random lifts of graphs: independence and chromatic number. *Random Structures & Algorithms*, 20(1):1–22, 2002.
- [Alo86] Noga Alon. Eigenvalues and expanders. *Combinatorica*, 6(2):83–96, 1986.
- [And79a] Joel Anderson. Extensions, restrictions, and representations of states on C^* -algebras. *Transactions of the American Mathematical Society*, 249(2):303–329, 1979.
- [And79b] Joel Anderson. Extreme points in sets of positive linear maps on $B(H)$. *Journal of Functional Analysis*, 31(2):195–217, 1979.
- [And81] Joel Anderson. A conjecture concerning the pure states of $B(H)$ and a related theorem. In *Topics in modern operator theory*, pages 27–43. Springer, 1981.
- [BB08] Julius Borcea and Petter Brändén. Applications of stable polynomials to mixed determinants: Johnson’s conjectures, unimodality, and symmetrized fischer products. *Duke Mathematical Journal*, 143(2):205–223, 2008.

- [BB09a] Julius Borcea and Petter Brändén. The lee-yang and pólya-schur programs. I. linear operators preserving stability. *Inventiones Mathematicae*, 177(3):541, 2009.
- [BB09b] 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.
- [BB10] Julius Borcea and Petter Brändén. Multivariate pólya-schur classification problems in the weyl algebra. *Proceedings of the London Mathematical Society*, 101(1):73–104, 2010.
- [BD11] Anton Baranov and Konstantin Dyakonov. The feichtinger conjecture for reproducing kernels in model subspaces. *Journal of Geometric Analysis*, 21(2):276–287, 2011.
- [BGLS01] Heinz H Bauschke, Osman Guler, Adrian S Lewis, and Hristo S Sendov. Hyperbolic polynomials and convex analysis. *Canadian Journal of Mathematics*, 53(3):470–488, 2001.
- [BHKW88] Kenneth Berman, Herbert Halpern, Victor Kaftal, and Gary Weiss. Matrix norm inequalities and the relative dixmier property. *Integral Equations and Operator Theory*, 11(1):28–48, 1988.
- [BL06] Yonatan Bilu and Nathan Linial. Lifts, discrepancy and nearly optimal spectral gap. *Combinatorica*, 26(5):495–519, 2006.
- [BSS12] Joshua Batson, Daniel A Spielman, and Nikhil Srivastava. Twice-ramanujan sparsifiers. *SIAM Journal on Computing*, 41(6):1704–1721, 2012.
- [BT87] Jean Bourgain and Lior Tzafriri. Invertibility of ‘large’ submatrices with applications to the geometry of banach spaces and harmonic analysis. *Israel Journal of Mathematics*, 57(2):137–224, 1987.
- [BT89] J Bourgain and L Tzafriri. Restricted invertibility of matrices and applications. *Analysis at Urbana*, 2:1986–1987, 1989.
- [BT91] J Bourgain and L Tzafriri. On a problem of kadison and singer. *Journal Fur Die Reine Und Angewandte Mathematik*, 420:1–43, 1991.
- [CCLV05] Peter Casazza, Ole Christensen, Alexander Lindner, and Roman Vershynin. Frames and the feichtinger conjecture. *Proceedings of the American Mathematical Society*, 133(4):1025–1033, 2005.

- [CEKP07] Peter G Casazza, Dan Edidin, Deepti Kalra, and Vern I Paulsen. Projections and the kadison-singer problem. *Operators and Matrices*, 2007.
- [CFTW06] Peter G Casazza, Matthew Fickus, Janet C Tremain, and Eric Weber. The kadison-singer problem in mathematics and engineering: a detailed account. *Contemporary Mathematics*, 414:299, 2006.
- [Chi92] Patrick Chiu. Cubic ramanujan graphs. *Combinatorica*, 12(3):275–285, 1992.
- [Cio06] Sebastian M Cioabă. Eigenvalues of graphs and a simple proof of a theorem of greenberg. *Linear Algebra and its Applications*, 416(2-3):776–782, 2006.
- [Con78] J.B. Conway. *Functions of One Complex Variable I*. Functions of one complex variable. Springer Science & Business Media, 1978.
- [CS07] Maria Chudnovsky and Paul Seymour. The roots of the independence polynomial of a clawfree graph. *Journal of Combinatorial Theory, Series B*, 97(3):350–357, 2007.
- [CT06] Peter G Casazza and Janet Crandell Tremain. The kadison-singer problem in mathematics and engineering. *Proceedings of the National Academy of Sciences of the United States of America*, 103(7):2032–2039, 2006.
- [Ded92] Jean Pierre Dedieu. Obreschkoff’s theorem revisited: What convex sets are contained in the set of hyperbolic polynomials? *Journal of Pure and Applied Algebra*, 81(3):269–278, 1992.
- [Fel80] Harriet Fell. On the zeros of convex combinations of polynomials. *Pacific Journal of Mathematics*, 89(1):43–50, 1980.
- [Fis06] Steve Fisk. Polynomials, roots, and interlacing. *arXiv preprint math/0612833*, 2006.
- [FL96] Keqin Feng and Wen-Ch’ing Winnie Li. Spectra of hypergraphs and applications. *Journal of Number Theory*, 60(1):1–22, 1996.
- [Fri03a] Joel Friedman. A proof of alon’s second eigenvalue conjecture. In *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, pages 720–724. ACM, 2003.
- [Fri03b] Joel Friedman. Relative expanders or weakly relatively ramanujan graphs. *Duke Mathematical Journal*, 118(1):19–35, 2003.

- [GG81] Chris Godsil and Ivan Gutman. On the matching polynomial of a graph, in algebraic methods in graph theory. *Colloquia Mathematica Societatis János Bolyai*, 25:241–249, 01 1981.
- [GM88] Chris D Godsil and Bojan Mohar. Walk generating functions and spectral measures of infinite graphs. *Linear Algebra and its Applications*, 107:191–206, 1988.
- [God81] Christopher David Godsil. Matchings and walks in graphs. *Journal of Graph Theory*, 5(3):285–297, 1981.
- [God93] Chris Godsil. *Algebraic Combinatorics*, volume 6. New York: Chapman & Hall, 1993.
- [Gre95] Yoseph Greenberg. *On the spectrum of graphs and their universal covering*. PhD thesis, Hebrew University, 1995.
- [Har97] David A Harville. *Matrix algebra from a statistician’s perspective*, volume 1. Springer, 1997.
- [HL72] Ole J Heilmann and Elliott H Lieb. Theory of monomer-dimer systems. *Communications in Mathematical Physics*, 25(3):190–232, 1972.
- [HLW06] Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *Bulletin of the American Mathematical Society*, 43(4):439–561, 2006.
- [HV07] J William Helton and Victor Vinnikov. Linear matrix inequality representation of sets. *Communications on Pure and Applied Mathematics*, 60(5):654–674, 2007.
- [JL97] Bruce W Jordan and Ron Livné. Ramanujan local systems on graphs. *Topology*, 36(5):1007–1024, 1997.
- [Kra06] Ilia Krasikov. On extreme zeros of classical orthogonal polynomials. *Journal of Computational and Applied Mathematics*, 193(1):168–182, 2006.
- [KS59] Richard V Kadison and Isadore M Singer. Extensions of pure states. *American Journal of Mathematics*, 81(2):383–400, 1959.
- [Law10] Wayne Lawton. Minimal sequences and the kadison-singer problem. *Bulletin of the Malaysian Mathematical Sciences Society*, 33(2):169–176, 2010.
- [LP10] Nati Linial and Doron Puder. Word maps and spectra of random graph lifts. *Random Structures & Algorithms*, 37(1):100–135, 2010.

- [LPR05] Adrian Lewis, Pablo Parrilo, and Motakuri Ramana. The lax conjecture is true. *Proceedings of the American Mathematical Society*, 133(9):2495–2499, 2005.
- [LPS88] Alexander Lubotzky, Ralph Phillips, and Peter Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988.
- [LR05] Nathan Linial and Eyal Rozenman. Random lifts of graphs: perfect matchings. *Combinatorica*, 25(4):407–424, 2005.
- [LS81] Elliott H Lieb and Alan D Sokal. A general lee-yang theorem for one-component and multicomponent ferromagnets. *Communications in Mathematical Physics*, 80(2):153–179, 1981.
- [LS96] Wen-Ch'ing Winnie Li and Patrick Solé. Spectra of regular graphs and hypergraphs and orthogonal polynomials. *European Journal of Combinatorics*, 17(5):461–477, 1996.
- [LSV11] Eyal Lubetzky, Benny Sudakov, and Van Vu. Spectra of lifted ramanujan graphs. *Advances in Mathematics*, 227(4):1612–1645, 2011.
- [Lub10] Alex Lubotzky. *Discrete groups, expanding graphs and invariant measures*. Springer Science & Business Media, 2010.
- [Mar49] Morris Marden. *Geometry of polynomials*. Number 3. American Mathematical Society, 1949.
- [Mar88] Grigorii Aleksandrovich Margulis. Explicit group-theoretical constructions of combinatorial schemes and their application to the design of expanders and concentrators. *Problemy peredachi informatsii*, 24(1):51–60, 1988.
- [Mar15] Adam W Marcus. Polynomial convolutions and (finite) free probability. *Preprint*, 2015.
- [Mey00] Carl D Meyer. *Matrix analysis and applied linear algebra*, volume 71. SIAM, 2000.
- [Mor94] Moshe Morgenstern. Existence and explicit constructions of $q+1$ regular ramanujan graphs for every prime power q . *Journal of Combinatorial Theory, Series B*, 62(1):44–62, 1994.
- [MP67] Vladimir Alexandrovich Marchenko and Leonid Andreevich Pastur. Distribution of eigenvalues for some sets of random matrices. *Matematicheskii Sbornik*, 114(4):507–536, 1967.

- [MSS13] Adam Marcus, Daniel A Spielman, and Nikhil Srivastava. Interlacing families I: Bipartite ramanujan graphs of all degrees. In *2013 IEEE 54th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 529–537. IEEE, 2013.
- [MSS14] Adam W Marcus, Daniel A Spielman, and Nikhil Srivastava. Ramanujan graphs and the solution of the kadison-singer problem. *Proceedings of ICM*, III:375–386, 2014.
- [MSS15a] Adam Marcus, Daniel A Spielman, and Nikhil Srivastava. Finite free convolutions of polynomials. *arXiv preprint arXiv:1504.00350*, 2015.
- [MSS15b] Adam W Marcus, Daniel A Spielman, and Nikhil Srivastava. Interlacing families I: Bipartite ramanujan graphs of all degrees. *Annals of Mathematics*, 128(1):307–325, 2015.
- [MSS15c] Adam W 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 *2015 IEEE 56th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1358–1377. IEEE, 2015.
- [MSS17] Adam W Marcus, Daniel A Spielman, and Nikhil Srivastava. Interlacing families III: Sharper restricted invertibility estimates. *arXiv preprint arXiv:1712.07766*, 2017.
- [Nil91] Alon Nilli. On the second eigenvalue of a graph. *Discrete Mathematics*, 91(2):207–210, 1991.
- [NY17] Assaf Naor and Pierre Youssef. Restricted invertibility revisited. In *A Journey Through Discrete Mathematics*, pages 657–691. Springer, 2017.
- [Pau11] Vern Paulsen. Syndetic sets, paving and the feichtinger conjecture. *Proceedings of the American Mathematical Society*, 139(3):1115–1120, 2011.
- [Pem12] Robin Pemantle. Hyperbolicity and stable polynomials in combinatorics and probability. *arXiv preprint arXiv:1210.3231*, 2012.
- [Piz90] Arnold K Pizer. Ramanujan graphs and hecke operators. *Bulletin of the American Mathematical Society*, 23(1):127–137, 1990.

- [Pop14] Sorin Popa. A ii1 factor approach to the kadison–singer problem. *Communications in Mathematical Physics*, 332(1):379–414, 2014.
- [Pud15] Doron Puder. Expansion of random graphs: New proofs, new results. *Inventiones Mathematicae*, 201(3):845–908, 2015.
- [RS02] Qazi Ibadur Rahman and Gerhard Schmeisser. *Analytic theory of polynomials*. Number 26. Oxford University Press, 2002.
- [Sri] Nikhil Srivastava. Restricted invertibility by interlacing polynomials. <https://windowsontheory.org/2014/04/15/restricted-invertibility-by-interlacing-polynomials/>.
- [SS12] Daniel A Spielman and Nikhil Srivastava. An elementary proof of the restricted invertibility theorem. *Israel Journal of Mathematics*, 190(1):83–91, 2012.
- [Sze39] Gabor Szeg. *Orthogonal polynomials*, volume 23. American Mathematical Society, 1939.
- [Tao13] Terence Tao. Real stable polynomials and the kadison–singer problem. <https://terrytao.wordpress.com/2013/11/04/real-stable-polynomials-and-the-kadison-singer-problem/>, 2013.
- [Ver01] Roman Vershynin. John’s decompositions: selecting a large part. *Israel Journal of Mathematics*, 122(1):253–277, 2001.
- [Wal22] Joseph L Walsh. On the location of the roots of certain types of polynomials. *Transactions of the American Mathematical Society*, 24(3):163–180, 1922.
- [Wea04] Nik Weaver. The kadison–singer problem in discrepancy theory. *Discrete Mathematics*, 278(1-3):227–239, 2004.
- [You14] Pierre Youssef. Restricted invertibility and the banach–mazur distance to the cube. *Mathematika*, 60(1):201–218, 2014.