Universidad de Valladolid

Facultad de Ciencias

Trabajo Fin de Grado

Grado en Matemáticas

Familias de polinomios entrelazados.

La Conjetura de Kadison-Singer.

Autor: Álvaro Samperio Valdivieso

Tutor/es: Antonio Campillo López

Índice

Introducción

El TFG está dedicado a estudiar las propiedades conceptuales y técnicas de las familias entrelazadas de polinomios con coeficientes y raíces reales y a la reciente prueba de la Conjetura de Kadison-Singer, propuesta en 1959 por Richard Kadison (1925-2008) e Isadore Singer (1924) y resuelta positivamente por Adam Marcus, Daniel Spielman y Nikhil Srivastava en 2013. Los especialistas siempre habían dudado de que la conjetura fuese cierta, y se habían formulado nuevos enunciados equivalentes o relacionados con ella en términos elementales. Su solución se obtuvo utilizando también técnicas elementales, las de las familias de polinomios entrelazados.

La técnica de polinomios entrelazados permitió simultáneamente a Adam Marcus, Daniel Spielman y Nikhil Srivastava probar la existencia de grafos de Ramanujan regulares bipartitos de grado arbitrario y número de vértices arbitrariamente grande, resolviendo así parcialmente otro problema clásico de las matemáticas. El grado de un vértice de un grafo es el número de aristas que confluyen en él, el grafo es regular cuando todos los vértices tienen el mismo grado, que entonces se llama grado del grafo. Los grafos de Ramanujan regulares son aquellos que tienen una densidad óptima de distribución relativa de sus vértices, y se caracterizan por ser los que su función zeta (la función zeta de Ihara) satisface la hipótesis de Riemann. El problema de existencia de grafos de Ramanujan no bipartitos de grado arbitario y número de vértices arbitrariamente grande aún permanece abierto.

La conjetura de Kadison-Singer afirmaba que cada estado puro en la C*-álgebra operadores lineales diagonales sobre un espacio de Hilbert separable tiene una única extensión a la C*-álgebra de todos los operadores de dicho espacio. Se sabía que la conjetura era equivalente a otros enunciados sobre espacios de Hilbert de dimensión finita y, en particular, al formulado en 2004 por Weaver.

El TFG muestra la formulación completa y la prueba de la conjetura. El autor ha realizado a la vez el TFG para el Grado de Física con título "Formulación algebraica de la Mecánica Cuántica. La Conjetura de Kadison-Singer" conectado con este TFG de Matemáticas, y presentado a la misma convocatoria en el Grado de Física de la Facultad de Ciencias de la Universidad de Valladolid.

El primer capítulo está dedicado al estudio de las familias entrelazadas de polinomios y a demostrar sus propiedades principales. El segundo capítulo se dedica a la formulación física, en términos de análisis funcional, de la conjetura y a las reducciones sucesivas a otros problemas de dimensión finita que combinan probabilidad y álgebra lineal. El tercer capítulo se dedica a exponer la demostración de la conjetura, basada en la aplicación de la técnica de polinomios entrelazados para probar uno de los enunciados del segundo capítulo del que se conocía que implicaba la validez de la conjetura.

1. La Técnica básica. Familias de polinomios entrelazados.

En este primer capítulo abordaremos el concepto de familia entrelazada de polinomios, su caracterización y su utilidad.

Partiremos de la definición básica de entrelazado introducida por Marcus, Spielman y Srivastava en los artículos [5] y [6]. El primero de ellos ultiliza la técnica de polinomios entrelazados para probar que existen familias infinitas de grafos regulares bipartitos de Ramanujan de cualquier grado mayor que 2 con un número arbitrariamente grande de vértices, y el segundo, para probar la Conjetura de Kadison-Singer.

Después estudiaremos las consecuencias que se derivan de la propiedad de entrelazado. Algunos de los resultados están tomados de [7].

Por último, definiremos la propiedad de ser familia entrelazada, una propiedad relacionada con la de entrelazado que permite una aplicación con hipótesis más fáciles de comprobar en casos prácticos, de estos resultados.

La definición de familia entrelazada que establecemos es más general que la definición dada en estas fuentes anteriores, con el objetivo de que su posterior aplicación resulte ligeramente más sencilla.

Comenzaremos definiendo qué es un entrelazado.

Definición 1.1 *Decimos que un polinomio $g(x) = a \prod_{i=1}^{n-1}(x - \alpha_i)$ de grado n-1 entrelaza a un polinomio $f(x) = b \prod_{i=1}^{n}(x - \beta_i)$ de grado $n \geq 2$, ambos con todos sus coeficientes y raíces reales si*

$$\beta_1 \geq \alpha_1 \geq \beta_2 \geq \alpha_2 \geq ...\alpha_{n-1} \geq \beta_n$$

Decimos que los polinomios f_1, ..., f_m (de grado $n \geq 2$ y con todos sus coeficientes y raíces reales) tienen un entrelazado común si existe un polinomio que entrelaza a cada uno de ellos.

Las principales consecuencias de esta propiedad se derivan del siguiente teorema.

Teorema 1.2 *Sean f_1, ..., f_m polinomios de grado $n \geq 2$ de coeficientes y raíces reales con coeficiente principal positivo y entrelazado común.*

Llamamos:

$$m_k = min_{1 \leq j \leq m} \lambda_k(f_j) \qquad y \qquad M_k = max_{1 \leq j \leq m} \lambda_k(f_j)$$

al minimo y al máximo, respectivamente, de las raíces k-ésimas (ordenadas de mayor a menor) de los polinomios de la familia.

Sea $S = \rho_1 f_1 + ... + \rho_m f_m$ una combinación lineal de los polinomios con coeficientes reales no negativos, $\rho_1, ..., \rho_m \geq 0$ (un caso particular es una combinación convexa, es decir, con la condición adicional $\sum_{j=1}^{m} \rho_j = 1$).

Entonces, para todo $k \in \{1, ..., n\}$,

$$m_k \leq \lambda_k(S) \leq M_k.$$

Demostración: Comenzamos observando que la condición de entrelazado común da un orden a las raíces de tal manera que $M_1 \geq m_1 \geq ... \geq M_n \geq m_n$, puesto que un polinomio que entrelaza al conjunto debe tener sus raíces entre el mínimo de una raíz y el máximo de la siguiente.

Por otra parte, podemos reducir el problema al caso en que $\rho_j \neq 0$ para todo j = 1, ...,m; ya que si existe un coeficiente ρ_j nulo, el polinomio correspondiente (f_j) no afectará en la combinación lineal S.

Además, también podemos reducir el problema al caso en que no existen raíces comunes a todos los polinomios.

Si la k-ésima raíz es común a todos los polinomios, el mínimo y el máximo k-ésimo coinciden ($m_k = M_k$), y además ese punto es raíz de S, luego se cumple $m_k \leq \lambda_k(S) \leq M_k$.

Si los polinomios de partida $f_1, ..., f_m$ tienen alguna raíz en común, podemos dividir cada polinomio entre el máximo común divisor del conjunto y obtenemos, para todo k, los polinomios $h_k = \dfrac{f_k}{mcd(f_1, ..., f_m)}$, que no tienen raíces comunes y conservan el entrelazado común de los polinomios de partida. Una vez demostrado el teorema para estos polinomios h_k, al multiplicarlos por $mcd(f_1, ..., f_m)$ añadimos sólo las raíces comunes a $f_1, ..., f_m$; que también cumplen el resultado del teorema (como vimos en el párrafo anterior).

Ahora demostraremos el teorema en el caso de que los polinomios no tengan raíces comunes ni correspondan a coeficientes nulos.

Como los polinomios tienen coeficiente dominante positivo, cada f_j será positivo para x mayor que $\lambda_1(f_j)$. Esto es así, ya que si existiera un punto x_0 con $f_j(x_0) < 0$ y $x_0 > \lambda_1(f_j)$, como $f_j(x) \to +\infty$ cuando $x \to +\infty$ y los polinomios son funciones continuas, por el Teorema de Bolzano f_j tendría una raíz mayor que $\lambda_1(f_j)$, lo cual es absurdo.

Por tanto $f_j(M_1) \geq 0$ para todo j, y esto implica $S(M_1) \geq 0$.

Ahora, en el caso de que no haya ningún polinomio con raíces múltiples, la derivada de cada polinomio es distinta de 0 en cada raíz. Por tanto, las raíces de cada polinomio son los puntos en los que éste cambia de signo.

Esto hace que, para todo j, $f_j(m_1) \leq 0$, $f_j(M_2) \leq 0$, $f_j(m_2) \geq 0$, ..., y en general:

$$\begin{cases} f_j(M_k) \geq 0 \quad y \quad f_j(m_k) \leq 0 \quad si \quad k \quad es \quad impar \\ \\ f_j(M_k) \leq 0 \quad y \quad f_j(m_k) \geq 0 \quad si \quad k \quad es \quad par \end{cases} \tag{1}$$

Como todos los coeficientes ρ_j son positivos, también se tiene:

$$\begin{cases} S(M_k) \geq 0 \quad y \quad S(m_k) \leq 0 \quad si \quad k \quad es \quad impar \\ \\ S(M_k) \leq 0 \quad y \quad S(m_k) \geq 0 \quad si \quad k \quad es \quad par \end{cases} \tag{2}$$

En el caso de que exista algún polinomio con raíces múltiples, si la multiplicidad de la raíz es mayor o igual que 3, cualquier distribución de las raíces compatible con la existencia de entrelazado tendrá raíces comunes a todos los polinomios. Eliminándolas como se indicó anteriormente, reducimos el problema al caso de raíces con multiplicidad 2.

Sea $s \leq m$ tal que la raíz k-ésima de f_1, ..., f_s tiene multiplicidad 2. Entonces esa raíz es igual a m_k y a M_{k+1}. Además, para todo j = 1, ..., s; $f'_j(m_k) = 0$ y $f''_j(m_k) \neq 0$.

Por tanto, estos polinomios presentan un mínimo relativo en m_k si k es impar y un máximo relativo si k es par. Entonces, en cualquier caso, no cambian de signo entre M_k Y m_{k+1}, y se tiene:

$$\begin{cases} f_j(M_k) \geq 0 \quad y \quad f_j(m_{k+1}) \geq 0 \quad si \quad k \quad es \quad impar \\ \\ f_j(M_k) \leq 0 \quad y \quad f_j(m_{k+1}) \leq 0 \quad si \quad k \quad es \quad par \end{cases} \tag{3}$$

Para el resto de los polinomios (que no tienen raíces múltiples) se mantiene el resultado de la ecuación (1). Como k+1 es par cuando k es impar y viceversa, la suma S de los polinomios que cumplen (1) y los que cumplen (3), también cumple (2) para todo k en este caso de existencia de raíces múltiples.

Aplicando el Teorema de Bolzano a S en cada uno de los n intervalos $[m_k, M_k]$, S tiene al menos una raíz en cada uno de ellos. El grado de S es n (por ser combinación lineal de polinomios de grado n), luego por el Teorema fundamental del álgebra, tiene n raíces. Por tanto, tiene exactamente una raíz en cada uno de estos intervalos, es decir,

$$m_k \leq \lambda_k(S) \leq M_k.$$

para todo k = 1, ..., n.

□

El papel que juega la existencia de un entrelazado común en este resultado que da una localización muy precisa para las raíces de S es crucial, véase que sin un entrelazado común ni siquiera se podría garantizar que las raíces de S sean reales, como se puede ver en el siguiente ejemplo sencillo.

Ejemplo 1.3 *Los polinomios* $f_1 = x^2 + 3x$*, y* $f_2 = x^2 - 3x + 1$ *tienen todos sus coeficientes y raíces reales, pero la combinación lineal convexa* $S = \frac{1}{2}f_1 + \frac{1}{2}f_2 = x^2 + \frac{1}{2}$ *no tiene raíces reales.*

Observemos cómo en el Teorema 1.2 tenemos mucha libertad en la elección de los coeficientes ρ_j de la combinación lineal (siendo la única restricción que sean reales y no negativos). A continuación presentamos dos casos particulares de especial interés.

Corolario 1.4 *Sean* $f_1, \ldots, f_m$ *polinomios de grado* $n \geq 2$ *de coeficientes y raíces reales con coeficiente principal positivo y entrelazado común. Sea* $S = f_1 + \ldots + f_m$ *la suma de los polinomios. Entonces existe un índice* $j \in \{1, \ldots, m\}$ *tal que la mayor raíz de* f_j *es menor o igual que la mayor raíz de S.*

Demostración: El resultado es consecuencia directa del Teorema 1.2 en el caso de elección de todos los coeficientes ρ_j iguales a 1.

El resultado del teorema implica que la mayor raíz de S, $\lambda_1(S)$, está entre el mínimo y el máximo de las primeras raíces de los polinomios f_j:

$$m_1 \leq \lambda_1(S) \leq M_1.$$

Como $m_1 = min_{1 \leq j \leq m} \lambda_1(f_j) \leq \lambda_1(S)$, debe existir al menos un índice j tal que su mayor raíz sea menor o igual que la mayor raíz de la suma S. □

Corolario 1.5 *Sea f una variable aleatoria cuyo soporte es un conjunto finito de polinomios* $\{f_1, \ldots, f_m\}$ *de grado* $n \geq 2$ *de coeficientes y raíces reales con coeficiente principal positivo y entrelazado común. Denotamos por* $\mathsf{P}(f = f_j)$*, para* $j = 1, \ldots, m$*; a las probabilidades puntuales de la ley de probabilidad de f.*

Llamamos polinomio esperado a la esperanza matemática de esta variable:

$$(E)(f) = \mathsf{P}(f = f_1)f_1 + \mathsf{P}(f = f_2)f_2 + \ldots + \mathsf{P}(f = f_m)f_m$$

Entonces existe un polinomio f_j *tal que la mayor raíz de* f_j *es menor o igual que la mayor raíz del polinomio esperado.*

Exponemos una definición de los conceptos probabilísticos que aparecen en este teorema en la discusión del Teorema 2.10.

Demostración: En este caso, basta con aplicar el Teorema 1.2 en el caso particular en el que los coeficientes ρ_j de la combinación lineal son las probabilidades puntuales de los polinomios del soporte: $\rho_j = \mathsf{P}(f = f_j)$.

La combinación lineal que resulta de esta elección de los coeficientes es el polinomio esperado, y es una combinación convexa por la definición de probabilidad, ya que $\rho_j = \mathsf{P}(f = f_j) \geq 0$ para todo j y la suma de todas estas probabilidades puntuales es igual a 1.

Por el mismo razonamiento del corolario anterior, tenemos el resultado. $\square$

La hipótesis de existencia de entrelazado común en un conjunto de polinomios ofrece resultados muy importantes, pero exige un orden muy concreto en sus raíces.

Comprobar si un conjunto de polinomios cumple esta hipótesis, por tanto, exigiría conocer con precisión la posición de todas las raíces de todos ellos, lo cual es muy difícil o imposible en los casos de interés.

En su lugar, recurriremos a la siguiente caracterización de entrelazado común, que es más fácil de manejar.

Lema 1.6 *Sean f_1, ..., f_m polinomios de grado $n \geq 2$ de coeficientes y raíces reales con coeficiente principal positivo.*

Entonces f_1, ..., f_m tienen un entrelazado común si y sólo si $S = \rho_1 f_1 + ... + \rho_m f_m$ tiene todas sus raíces reales para cualquier combinación lineal de los polinomios con coeficientes reales no negativos.

Antes de de demostrar el lema, indicamos que la condición "cualquier combinación lineal de los polinomios con coeficientes reales no negativos" se puede sustituir por "cualquier combinación lineal convexa", dado que cualquier combinación lineal de los polinomios con coeficientes reales no negativos, S = $\sum_{j=1}^{m} \rho_j \mathrm{f}_j$ tiene todas sus raíces reales si y solo si las tiene la combinación lineal convexa $\frac{1}{\sum_{j=1}^{m} \rho_j}$ S.

Demostración: La implicación hacia la derecha es consecuencia directa del Teorema 1.2.

Probemos la otra implicación. Para empezar, veamos que la condición de poseer un entrelazado común es equivalente a que los polinomios del conjunto admitan un entrelazado común dos a dos.

La implicación hacia la derecha es trivial. Para la otra implicación, razonando por reducción al absurdo, si f_1, ..., f_m no tienen entrelazado común, entonces existe al menos un índice k $\in$ {1, ..., n-1} tal que la intersección de los intervalos [m_k, M_k] y [m_{k+1}, M_{k+1}] es no vacía.

Por tanto, deben existir i, j k $\in$ {1, ..., m} tal que $\lambda_k(f_j) > \lambda_{k+1}(f_i)$. Esto entra en contradiccion con el hecho de que los polinomios f_i y f_j deben tener entrelazado común si suponemos que los polinomios del conjunto tienen entrelazado común dos a dos.

Entonces para probar el lema, fijados dos índices i, j cualquiera, basta con probar el lema en el caso de dos polinomios f_i y f_j.

Además, al igual que en la demostración del Teorema 1.2, si los polinomios tienen raíces en común, podemos quitárselas dividiendo entre el máximo común divisor de los polinomios, resolver el problema y añadir las raíces comunes que quitamos, luego es suficiente con probar el caso en que los dos polinomios no tienen raíces comunes.

Para todo t $\in$ [0,1], definimos S_t = (1 - t)f_i + tf_j.

Por continuidad, cuando t varía de 0 a 1, las raíces de S_t describen n curvas continuas, cada una empezando en una raíz de f_i(en t = 0) y acabando en una raíz de f_j(en t = 1).

Si en el interior de alguna de las curvas (para t $\in$ (0,1)) hubiera alguna raíz x_0 de f_i, se tendría $S_t(x_0) = 0 = f_i(x_0)$. Y además, $S_t(x_0) = (1 - t)f_i(x_0) + tf_j(x_0) = tf_j(x_0) = 0$, con lo que x_0 sería raíz también de f_j, contradiciendo el hecho de que f_i y f_j no tienen raíces comunes.

Análogamente, si hubiera una raíz de f_j en el interior, tendría que ser raíz de f_i y llegaríamos a una contradicción.

Por lo tanto, las curvas que definen las raíces no tienen raíces de los polinomios en el interior.

Por otra parte, por hipótesis, las curvas son reales, pues están descritas por las raíces de S_t, que es una combinación convexa de f_i y f_j.

Entonces, cada una de las n curvas define un intervalo cerrado real con extremos en una raíz de f_i y otra de f_j, y la intersección de estos intervalos es vacía, por lo que f_i y f_j tienen entrelazado común. $\square$

Con este lema hemos encontrado una condición equivalente a la de existencia de entrelazado común que es una caracterización sencilla en casos de interés.

Sin embargo, la hipótesis de existencia de entrelazado común para todo el conjunto de polinomios es más exigente de lo necesario para obtener los resultados que estamos presentando.

En realidad, si el conjunto de polinomios de interés tiene una estructura determinada de tal manera que tengan un entrelazado común ciertos subconjuntos de él y ciertos conjuntos de combinaciones lineales de éstos con coeficientes reales no negativos, esto es suficiente para poder aplicar el Teorema 1.2 sucesivamente y obtener sus conclusiones en el todo el conjunto.

Esta estructura es a la que llamaremos de familia entrelazada.

Definición 1.7 *Sean C_1, ..., C_m conjuntos finitos, y para cada $(c_1, ..., c_m) \in C_1 \times$*

... × C_m, sea $f_{c_1,\dots,c_m}(x)$ un polinomio de grado $n \geq 2$ de coeficientes y raíces reales con coeficiente principal positivo.

Decimos que el conjunto de polinomios $\{f_{c_1,\dots,c_m}\}_{(c_1,\dots,c_m)\in C_1\times\dots\times C_m}$ es una familia entrelazada si se cumple lo siguiente:

Para cada elección de los subíndices $(c_1, \dots, c_m) \in C_1 \times \dots \times C_m$, existe una combinación lineal con coeficientes reales no negativos de los polinomios del conjunto $\left\{f_{c_1,\dots,c_{m-1},t}\right\}_{t\in C_m}$, que llamaremos $S_{c_1,\dots,c_{m-1}}$

Además, de manera recursiva, para todo $1 < k < m$ y todo $(c_1, \dots, c_{k-1}) \in C_1 \times \dots \times C_{k-1}$, existe una combinación lineal con coeficientes reales no negativos de los polinomios $\left\{S_{c_1,\dots,c_{k-1},t}\right\}_{t\in C_k}$ que llamaremos $S_{c_1,\dots,c_{k-1}}$.

Y además, existe una combinación lineal con coeficientes reales no negativos de los polinomios $\{S_t\}_{t\in C_1}$ que llamaremos $S_\emptyset$.

Estas elecciones de polinomios cumplen que para todo $k \in \{1, \dots, m\}$ y para todo $(c_1, \dots, c_k) \in C_1 \times \dots \times C_k$, los polinomios $\left\{S_{c_1,\dots,c_{k-1},t}\right\}_{t\in C_k}$ tienen entrelazado común.

Donde, en el caso $k = m$, definimos $S_{c_1,\dots,c_{m-1},t} \equiv f_{c_1,\dots,c_{m-1},t}$.

Observamos que los polinomios que hemos denominado con la letra S (con diferentes subíndices) forman una estructura que podemos identificar con un árbol, cuya raíz es $S_\emptyset$ y sus hojas son los polinomios del conjunto de partida.

Decimos que, para todo k, $S_{c_1,\dots,c_{k-1}}$ es el "padre" de los polinomios $\left\{S_{c_1,\dots,c_{k-1},t}\right\}_{t\in C_k}$ siendo cada polinomio los "hijos" del primero.

Con esta notación, cada padre es combinación lineal con coeficientes reales no negativos de sus hijos. A la hora de establecer la estructura de familia entrelazada, se tomará una combinación específica en cada formación de un padre a partir de sus hijos, desde las hojas hasta la raíz, que dependerá de la expresión que queramos que tenga la raíz como función de las hojas (polinomios de partida).

Después, para aplicar el Teorema 1.2 se hará recursivamente, desde la raíz hasta las hojas. Como esta aplicación recursiva en una familia será nuestra manera de utilizar la técnica de polinomios entrelazados más adelante, la justificamos en el siguiente teorema, que no es más que una version del Teorema 1.2 para familias entrelazadas.

Teorema 1.8 *Sean $C_1, \dots, C_m$ conjuntos finitos, y sea $\{f_{c_1,\dots,c_m}\}_{(c_1,\dots,c_m)\in C_1\times\dots\times C_m}$ una familia entrelazada de polinomios de grado $n \geq 2$ de coeficientes y raíces reales con coeficiente principal positivo, con raíz $S_\emptyset$.*

Entonces, para cada $j = 1, \dots, m$, existen al menos dos elecciones de los índices, $(c_1, \dots, c_m)^{min}$ y $(c_1, \dots, c_m)^{max}$ de manera que:

$$\lambda_j(f_{(c_1,\dots,c_m)^{min}}) \leq \lambda_j(S_\emptyset) \leq \lambda_j(f_{(c_1,\dots,c_m)^{max}}).$$

Demostración: Demostraremos el teorema por inducción sobre m. El caso m=1 es el Teorema 1.2, que ya fue probado.

Por otra parte, con un m >1 fijo, supongamos el resultado cierto para m-1 y veamos que se cumple para m. Veamos la desigualdad inferior (la otra es análoga).

Por ser $S_\emptyset$ combinación lineal con coeficientes reales no negativos de polinomios de $\{S_t\}_{t\in C_1}$, y tener éstos un entrelazado común, por el Teorema 1.2, para un j fijo, existe un índice $c_1 \in C_1$ tal que

$$\lambda_j(S_{c_1}) \leq \lambda_j(S_\emptyset)$$

Ahora, dado que la familia, por ser entrelazada, tiene estructura de árbol, podemos ver ese polinomio S_{c_1} como la raíz de un nuevo árbol, de una familia entrelazada de m-1 conjuntos, quedándonos con los polinomios que tienen c_1 como primer índice.

Por hipótesis, esta familia cumple el teorema, luego existe una asignación de índices $(c_2, ..., c_m) \in C_1 \times ... \times C_m$ tal que

$$\lambda_j(S_{c_1,...,c_m}) \leq \lambda_j(S_{c_1})$$

Luego, por la desigualdad anterior, hemos demostrado que existe una asignación de índices (c_1, ..., c_m) tal que:

$$\lambda_j(f_{c_1,...,c_m}) = \lambda_j(S_{c_1,...,c_m}) \leq \lambda_j(S_{c_1}) \leq \lambda_j(S_\emptyset)$$

con lo que queda probado el resultado. □

Por último, retomamos el asunto de la elección de las combinaciones lineales con coeficientes reales no negativos en cada paso para formar la estructura de árbol en una familia entrelazada.

La elección de la suma de polinomios en cada paso nos llevaría a que el polinomio raíz $S_\emptyset$ sea la suma de los polinomios del conjunto (las hojas), independientemente de la estructura del árbol y el tamaño de los conjuntos C_1, ..., C_m. Esto nos proporciona una versión del Corolario 1.3 para familias entrelazadas, que es utilizado en [9] y por Spielman, Marcus y Svatrisava en [5] para construir grafos bipartitos d-regulares de Ramanujan por un proceso de 2-recubrimientos.

En la prueba de la Conjetura de Kadison-Singer, que es el objetivo de este trabajo, utilizaremos una versión del Corolario 1.4 para familias entrelazadas, aunque la manera de construir el árbol para llegar a que la raíz sea la esperanza matemática de una cierta variable aleatoria y la motivación para hacerlo forman parte del contenido de los siguientes capítulos.

2. La formulación de la Conjetura. Reducciones y problemas equivalentes.

En esta sección plantearemos el problema de Kadison-Singer y hablaremos de su contexto histórico. Además estableceremos su equivalencia con diferentes problemas y demostraremos que su prueba se deduce de la de un teorema principal, cuya demostración detallamos en el siguiente apartado.

La Conjetura de Kadison-Singer tiene un origen relacionado con la mecánica cuántica. Realizamos un desarrollo extenso de la formulación algebraica de la mecánica cuántica y de su relación con este problema en el Trabajo de Fin de Grado del Grado en Física [8]. Aquí exponemos un resumen de las principales ideas.

En la descripción cuántica de un sistema total tenemos un conjunto de sistemas físicos, cada uno de los cuales asociamos con un estado, y un conjunto de instrumentos de medida que nos permiten determinar una propiedad de un sistema físico, cada uno de los cuales asociamos con un observable.

Sin entrar en detalles, estados y observables son clases de equivalencia de los conjuntos de sistemas físicos e instrumentos, respectivamente, que surgen al relacionar por un lado los sistemas y por otro lado los observables que son indistinguibles en el proceso de medida.

Un observable es un elemento autoadjunto de una C*-algebra, que debido a la construcción de Gelfand-Naimark-Segal (GNS) podemos ver como una C*-subálgebra de la C*-algebra B(H) de operadores lineales y acotados en un cierto espacio de Hilbert H.

A continuación definimos algunos conceptos de este último párrafo y otros conceptos relacionados, que utilizaremos a lo largo del capítulo.

Comenzamos definiendo conceptos como norma y producto escalar para llegar a la definición de espacio de Hilbert.

Un espacio normado es un espacio vectorial V sobre un cuerpo K con una norma, que es una aplicación $||\cdot||$: V $\longrightarrow \mathbb{R}$ que cumple:

- 1. $||x|| \geq 0$ para todo $x \in V$.
- 2. $||x|| = 0$ si y solo si $x = 0$.
- 3. $||\lambda x|| = |\lambda|\ ||x||$ para todo $x \in V$, y para todo $\lambda \in K$.
- 4. $||x + y|| \leq ||x|| + ||y||$ para todo $x, y \in V$ (desigualdad triangular).

Si, además, V es completo con la métrica definida por la norma, como $d(x, y) = ||x\text{-}y||$, se dice que V es un espacio de Banach.

En concreto, los espacios vectoriales con los que trabajamos aquí son complejos. El cuerpo K es **C** con el valor absoluto habitual.

Un producto escalar definido en un espacio vectorial V complejo es una aplicación $<\cdot,\cdot>$: V x V $\longrightarrow \mathbb{R}$ que define una métrica hermítica definida positiva, es decir, que cumple:

- 1. $<x + y, z> = <x, z> + <y, z>$ para todo x, y, z $\in$ V.
- 2. $<\lambda x, y> = \lambda <x, y>$ para todo x, y $\in$ V, para todo $\lambda \in$ **C**.
- 3. $<x, y> = \overline{<y, x>}$ para todo x, y $\in$ V (donde la línea horizontal superior es la conjugación compleja).
- 4. $<x, x> \geq 0$ para todo x $\in$; y además $<x, x> = 0$ si y solo si x = 0.

El producto escalar da lugar a la norma definida por $||x|| = <x, x>^{1/2}$. También cumple la desigualdad de Cauchy-Schwarz: $|<x, y>|^2 \leq ||x||\ ||y||$ para todo x, y $\in$ V. Un espacio H es de Hilbert si es un espacio con producto escalar que es espacio de Banach.

Si H es un espacio de Hilbert, los operadores lineales y continuos (para la topología dada por la norma) son exactamente aquellas aplicaciones lineales y continuas T: H $\longrightarrow$ H para las que existe una constante M >0 tal que $||T(x)|| \leq M\ ||x||$. Si T es un operador lineal y continuo en H, llamamos adjunto de T al único operador T^*, que también es continuo, tal que:

$$< T(x), y >=< x, T^*(y) >$$

para todo x, y $\in$ H. Decimos que T es autoadjunto si T = T^*.

El conjunto de operadores lineales y continuos de H se denota por B(H). Es otro espacio de Banach para la norma dada por $||T|| = \sup_{||x||=1} ||T(x)||$, y, de hecho, es un álgebra donde el producto está dado por la composición de operadores. En particular, B(H) es un álgebra de Banach, de acuerdo con la siguiente definición.

Ahora definimos varios conceptos relacionados con álgebras.

Un álgebra normada U es un álgebra sobre K con elemento unidad 1 que es también un espacio normado, y satisface las propiedades: $||AB|| \leq ||A||\ ||B||$ para todo A, B $\in$ U; y $||1|| = 1$. Si U es un espacio de Banach respecto de esta norma, decimos que es un álgebra de Banach.

Una C*-algebra U es un álgebra compleja de Banach con una involución (es decir, un automorfismo de álgebras de Banach cuyo cuadrado es la identidad) denotada por *, que es una aplicación de U en U, que cumple:

- 1. $(aS + bT)^* = \overline{a}S^* + \overline{b}T^*$
- 2. $(ST)^* = T^*S^*$
- 3. $(T^*)^* = T$

para todo S, T $\in$ U y para todo a, b $\in \mathbb{C}$. Además debe cumplir la siguiente condición:

- 4. $||T^*T|| = ||T||^2$

Llamamos a T* elemento adjunto de T, ya que en el caso particular de una C*-álgebra B(H), con H espacio de Hilbert, la involución es la operación de tomar el operador adjunto.

Dada una C*-álgebra U, decimos que S es una C*-subálgebra de U si es una subálgebra de U y contiene a los adjuntos de todos sus elementos.

Sea U un álgebra de Banach. Para A $\in$ U, definimos el espectro de A en U, $sp_U(A)$ como $sp_U(A) = \{\lambda \in \mathbb{C}$ tal que A - λ 1 no tiene inverso en U (por los dos lados)$\}$ Decimos que un operador A es positivo si $sp_U(A) \subset [0, \infty]$. Escribimos $A \geq 0$ si A es positivo. Escribimos, abusando del lenguaje, $A \geq B$ si A - B es positivo, con A, B $\in$ U. Los elementos A positivos, son aquellos para los que existe un elemento R $\in$ U tal que $A = R^*R$.

La medida de un observable A, en cualquier estado, da como resultado un elemento de su espectro. La definición de un observable como un operador autoadjunto asegura que el resultado de una medida siempre será real, dado que el espectro de un operador autoadjunto está contenido en $\mathbb{R}$.

Al contrario que en la mecánica clásica, el proceso de medida en la mecánica cuántica tiene una naturaleza probabilista, es decir, la medida de un observable A en un estado tiene una probabilidad asociada (que depende del estado) con soporte en $sp_U(A)$. La probabilidad de obtener al medir un resultado dentro de un subconjunto S de $sp_U(A)$ es la probabilidad asociada a S.

Asociamos cada estado con un funcional lineal continuo $\varphi : B(H) \longrightarrow \mathbb{R}$ en una C*-álgebra, que asigna a cada observable la esperanza matemática de la distribución de probabilidad, en su conjunto de posibles resultados de la medida, asociada a la medida del observable en cualquier sistema físico que esté en ese estado. Además debe ser positivo ($\varphi(A^*A) \geq 0$ para todo operador A) y debe estar normalizado ($\varphi(1) = 1$). Entendemos por estado a cualquier funcional lineal φ: B(H) $\longrightarrow \mathbb{R}$ positivo y normalizado.

Decimos que un estado φ es mezcla estadística de los estados $\varphi_1, ..., \varphi_n$ si es una combinación lineal convexa de éstos. Es decir, si $\rho_1\varphi_1 + ... + \rho_n\varphi_n$, con $\rho_1, ..., \rho_n \in \mathbb{R}$, con $\rho_i \geq 0$ para todo i; y $\sum_{i=1}^n \rho_i = 1$.

Llamamos estado puro a un elemento extremal del conjunto convexo de estados en una C*-algebra, es decir, φ es puro si para toda expresión de φ = tφ_1 + (1 - t)φ_2 como mezcla estadística de estados φ_1, φ_2, con t $\in$ [0,1], debe ser t = 0, t = 1, o φ_1 = φ_2,. Un estado puro se corresponde con un sistema físico "puro" en el sentido de no ser una mezcla estadística de sistemas más sencillos.

El conjunto de estados es convexo, es decir, cualquier mezcla estadística de estados del conjunto es un estado del conjunto. Además es compacto para una cierta topología llamada topología física, cuya definición no indicamos aquí. Estas dos propiedades garantizan la existencia de estados puros en nuestro sistema, por el Teorema de Klein-Milman.

Sea S una C*-subálgebra de una C*-álgebra U. Si φ es un estado en S, decimos que un estado Ψ en U es una extensión de φ si $\Psi = \varphi$ en S. La existencia de alguna extensión de cualquier estado en una subálgebra está garantizada por el Teorema de Hahn Banach, uno de los resultados fundamentales del análisis funcional, que exponemos a continuación.

Teorema de Hahn Banach: Sea V un espacio vectorial complejo, sea p una seminorma en V, esto es, una aplicación de p: V $\longrightarrow \mathbb{R}$ tal que

- 1. p(x) $\geq$ 0 para todo x $\in$ V.
- 2. p(λx) = $|\lambda|$ p(x) para todo x $\in$ V y para todo $\lambda \in \mathbb{C}$.
- 3. p(x + y) $\leq$ p(x) + p(y) para todo x, y $\in$ V (desigualdad triangular).

Sean M un subespacio vectorial de V y sea φ un funcional en M tal que $|\varphi(x)| \leq p(x)$ para todo x $\in$ M. Entonces, existe un funcional lineal Ψ en V tal que $|\Psi(x)| \leq p(x)$ para todo x $\in$ E, y Ψ = f en M.

Dada una C*-álgebra, los observables que conmutan son llamados compatibles y son los que pueden ser medidos a la vez.

Una C*-subálgebra abeliana maximal es una C*-subálgebra abeliana con la propiedad de que no existe ninguna C*-subálgebra abeliana que la contenga (excepto ella misma).

Si D es una C*-subálgebra abeliana maximal de B(H) con H de dimensión finita n, B(H) es equivalente al conjunto de matrices complejas de tamaño n x n, y D es equivalente a la C*-subálgebra de ésta consistente en el conjunto de matrices diagonales complejas de tamaño n x n.

Von Neumann demostró que si el espacio de Hilbert asociado es separable (como espacio topológico), y de dimensión infinita, cualquier C*-subálgebra abeliana maximal (respecto de la inclusión de conjuntos) debe ser equivalente a una de las tres opciones siguientes:

- 1. l^∞(N)$\subset B(l^2$(N)) (caso discreto)
- 2. L^∞(0,1)$\subset B(L^2$(0,1)) (caso continuo),
- 3. L^∞(0,1) $\oplus$ $l^\infty(\{1,...,n\})$ $\subset$ B(L^2(0,1) $\oplus$ $l^2(\{1,...,n\})$) (caso mixto).

Precisemos la estructura de los conjuntos que aparecen en cada caso.

En el caso discreto, $l^\infty(\mathbb{N}) \equiv l^\infty$ es el conjunto de las sucesiones de números naturales acotadas, y $l^2(\mathbb{N}) \equiv l^2$ es el espacio de Hilbert de las sucesiones de números naturales de cuadrado integrable, en el que el producto escalar de a, b $\in l^2$ a = (a(1), a(2), ...) y b = (b(1), b(2), ...) es $<$a, b $>=$ a(1)$\overline{b(1)}$ + a(2)$\overline{b(2)}$ + ...

En el caso continuo, L^∞(0,1) es el conjunto de las funciones reales acotadas en (0,1); L^2(0,1) es el espacio de Hilbert de las funciones reales de cuadrado integrable en (0,1), en el que el producto escalar de f, g $\in L^2$(0,1) es $\int_0^1 f\overline{g}$

Por último, en el caso mixto, $l^\infty(\mathbb{N})$ es el conjunto de las sucesiones de elementos contenidos en $\{1,...,n\}$), y $l^2(\mathbb{N})$ es el espacio de Hilbert de las sucesiones de números en $\{1,...,n\}$) de cuadrado integrable.

En el caso discreto, que es el de interés en este texto, la identificación de un elemento a = (a(1), a(2), ...) $\in l^\infty$ con un operador $\phi_a \in$ B(l^2) es clara: el operador actúa multiplicando la componente enésima de un elemento x $\in l^2$ por a(n), es decir ϕ_a = a(1)x(1) + a(2)x(2) + ... Con esta identificación está claro que los elementos de l^∞ conmutan entre sí, por conmutar el producto componente a componente.

A pesar de que no tratamos los casos continuo y mixto, la identificación es similar. Por ejemplo, a una función f acotada en (0,1) le corresponde un operador perteneciente a B(L^2(0,1)) que multiplica a cada función g de cuadrado integrable, siendo el producto fg también de cuadrado integrable, ya que f es acotada.

En 1959, Kadison y Singer [4] plantearon el siguiente problema:

Conjetura 2.1 (de Kadison-Singer) *¿Cada estado puro definido en una C*-subálgebra abeliana maximal de la C*-algebra B(l^2) de operadores acotados en l^2 admite una única extensión a un estado puro en B(l^2)?*

Este problema plantea la ambiguedad (o no) en la extensión de estados puros en el caso discreto. El problema análogo en los casos continuo y mixto fue resuelto con resultado negativo por Kadison y Singer. Sin embargo el problema en el caso discreto no fue resuelto hasta 2013 por Adam Marcus, Daniel Spielman y Nikhil Srivastava [6] con resultado positivo (a pesar de que Kadison y Singer creían que la respuesta sería negativa).

Antes de su demostración, diferentes autores habían probado que el problema de Kadison-Singer es equivalente a varios resultados que abarcan numerosos campos.

Weaver [11] introdujo la siguiente conjetura, llamada KS$_r$, cuya veracidad para algún r entero mayor o igual que 2 es equivalente a la veracidad de Conjetura de Kadison-Singer.

Conjetura 2.2 (KS$_r$, r $\geq$ 2) *Existen constantes universales $\eta \geq 2$ y $\theta > 0$ de manera que se tiene lo siguiente: Sea w_1, ..., $w_m \in \mathbb{C}^d$ cualquier elección de vectores cumpliendo $||w_i|| \leq 1$ para todo i y satisfaciendo*

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

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

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

para todo j y para todo vector unitario $u \in C^d$.

Nuestro objetivo es verificar el problema de Kadison-Singer demostrando la conjetura KS$_2$ de Weaver, como hicieron Spielman, Marcus y Srivastava en la prueba original [6].

Detallamos la demostración de la conjetura KS$_2$ entre el final de este capítulo y el capítulo siguiente (que consiste en la demostración del resultado fundamental de su prueba). Por el momento supondremos cierta la conjetura KS$_r$ para algún r entero mayor o igual que dos y veremos que esto implica el Teorema de Kadison-Singer (en el caso discreto).

Weaver [11] mostró que la veracidad de la conjetura KS$_r$ para algún r entero mayor o igual que dos es equivalente a la siguiente conjetura formulada por Anderson [1]. Nosotros sólo exponemos aquí la implicación hacia la derecha, como paso intermedio para llegar hasta la Conjetura de Kadison-Singer.

Conjetura 2.3 (de Pavimentación) *Para todo $\epsilon > 0$ existe un $r \in \mathbb{N}$ tal que para toda matriz T compleja autoadjunta, con todos sus elementos diagonales nulos y de tamaño n x n, existen proyecciones diagonales P_1,...,P_r cuya suma es la matriz identidad tal que*

$$||P_i T P_i|| \leq \epsilon ||T||$$

para todo i = 1, 2,..., r.

A continuación exponemos cuatro consideraciones sobre la notación que aparece en el teorema, que será de uso común en lo sucesivo.

La norma que utilizamos para un operador lineal y continuo T entre los espacios normados V y W es $||T|| = sup\{||T(x)||_{(W)}$ tal que $||x||_{(V)} \leq 1\} = sup\{||T(x)||_{(W)}$ tal que $||x||_{(V)} = 1\}$, donde con $||x||_{(V)}$ nos referimos a la norma que se considera en V (x $\in$ V) y con $||T(x)||_{(W)}$ nos referimos a la norma que se considera en W (T(x) $\in$ W). En un caso como el anterior en el que V y/o W sea de dimensión finita, consideramos que la norma en ese espacio es la norma euclidea sin indicarlo explícitamente.

Esta norma ya fue utilizada, en particular, para definir la estructura de álgebra de Banach de B(H) anteriormente.

En dimensión finita, como es el caso de este teorema, una proyección ortogonal es una matriz P de tamaño n x n tal que P = P^2 = P^*.

Una proyección diagonal es una proyección ortogonal cuyos elementos no diagonales son nulos y cuyos elementos diagonales son 0 ó 1.

Daremos una demostración del Teorema de Pavimentación probando el siguiente teorema en términos de proyecciones ortogonales en lugar de matrices autoadjuntas. La equivalencia entre los dos resultados fue establecida por Anderson y Akemann [1].

Teorema 2.4 *Existen constantes universales 0$\leq \epsilon \leq$1 y δ>0 de manera que se tiene lo siguiente: Para toda proyección $P = (p_{ij})$ de tamaño n x n cumpliendo max_i p_{ii} $\leq \delta$, existen proyecciones diagonales Q_1, ..., Q_r cuya suma es la matriz identidad tal que*

$$||Q_iPQ_i|| \leq 1 - \epsilon$$

para todo i = 1, 2,..., r.

Demostración: Suponemos que existen constantes r $\geq$ 2 entero, $\eta \geq 2$ y $\theta > 0$ de tal manera que cumplen la conjetura KS_r.

Veremos que el Teorema 2.4 se cumple entonces para las constantes $\epsilon = \dfrac{\theta}{\eta}$, $\delta = \dfrac{1}{\eta}$ y r = r. Vemos que, por ser $\eta \geq 0$, también $\delta \geq 0$, y además por ser $\eta - \theta \geq 0$, y ser ambas constantes positivas, se tiene 0$\leq \epsilon \leq$1

Sea P una proyección ortogonal con $\max_i p_{ii} \leq \dfrac{1}{\eta}$. Queremos demostrar que existen proyecciones diagonales Q_1, ..., Q_r cuya suma es la matriz identidad tal que

$$||Q_iPQ_i|| \leq 1 - \frac{\theta}{\eta}$$

Llamamos k al rango de P, es decir, k es la dimensión del subespacio vectorial imagen de P, $P(\mathbb{C}^n) \subset \mathbb{C}^n$.

Para todo i = 1, ..., n definimos los vectores $v_i = \sqrt{\eta}\ P(e_i) \in P(\mathbb{C}^n)$, donde $e_i \in \mathbb{C}^n$ es el vector i-ésimo de la base canónica, es decir aquel cuyas componentes son igual 0 a excepción de la componente i-ésima, que es igual a 1.

Estos vectores v_i; con i = 1, ..., n; cumplen las hipótesis de la conjetura KS_r

Por un lado, para todo i,

$$||v_i||^2 = \eta||P(e_i)||^2 = \eta < P(e_i), e_i > \leq \eta max_i p_{ii} \leq 1$$

Y por otro lado, para todo vector unitario u $\in$ P($\mathbb{C}^n$),

$$\sum_{i=1}^{n} |\langle u, v_i\rangle|^2 = \sum_{i=1}^{n} |\langle u, \sqrt{\eta}P(e_i)\rangle|^2 = \eta\sum_{i=1}^{n} |\langle P(u), e_i\rangle|^2 = \eta\sum_{i=1}^{n} |\langle u, e_i\rangle|^2 = \eta$$

Para pasar de la segunda igualdad a la tercera utilizamos el hecho de que P es autoadjunta por ser una proyección. Además, obtenemos la siguiente igualdad debido a que u es un vector de la imagen de P, y por tanto P(u) = u. La última igualdad es debida a que los vectores canónicos forman una base ortonormal y u es unitario.

Aplicando la Conjetura KS$_r$, existe una partición S$_1$, S$_2$, ... S$_r$ de $\{1, ...n\}$ tal que

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

para todo j = 1, ..., r; y para todo vector unitario u $\in$ P($\mathbb{C}^n$).

Para todo j = 1, ..., r, definimos la proyección diagonal Q$_j$ como:

$$Q_j(e_i) = \begin{cases} e_i & si \quad i \in S_j \\ \\ 0 & si \quad i \notin S_j \end{cases}$$

para todo i = 1, ..., n.

Ahora buscamos una cota de $||Q_jPQ_j||$ para todo j. Primero vemos que

$$||Q_jPQ_j|| = ||Q_jPPQ_j|| = ||Q_jP(Q_jP)^*|| = ||Q_jP||^2$$

Por otra parte,

$$\begin{aligned} ||Q_jP(u)|| &= sup\{||Q_jP(u)|| \ tal\ que\ ||u|| \leq 1\} = \\ &= sup\{||Q_jP(Pu)|| \ tal\ que\ ||u|| \leq 1\} = \\ &= sup\{||Q_jP(Pu)|| \ tal\ que\ ||Pu|| \leq 1\} \end{aligned}$$

ya que por ser P y (1-P) proyecciones ortogonales, $||u|| = ||Pu|| + ||(1-P)u||$, entonces si $||u|| \leq 1, ||Pu|| \leq 1||u|| \leq 1$. De la última igualdad deducimos que podemos calcular la norma de Q_jP como:

$$||Q_jP(u)|| = sup\{||Q_jP(u)|| \ tal\ que\ ||u|| = 1, u \in P(\mathbb{C}^n)\}$$

Sea u $\in$ P($\mathbb{C}^n$) unitario. Entonces,

$$||Q_jPQ_j(u)|| = ||Q_jP(u)||^2 = \sum_{i=1}^{n} |\langle Q_jP(u), e_i\rangle|^2 =$$

$$= \sum_{i=1}^{n} |\langle u, PQ_j(e_i)\rangle|^2 = \sum_{i\in S_j} |\langle u, P(e_i)\rangle|^2 = \frac{1}{\eta}\sum_{i\in S_j} |\langle u, v_i\rangle|^2 \leq 1 - \frac{\theta}{\eta}$$

Por tanto, para todo j, $||Q_jPQ_j|| \leq 1 - \frac{\theta}{\eta}$ y queda probado el resultado. $\square$

Nótese que, en el Teorema de Pavimentación, el número r de elementos de la partición es independiente de la dimensión n, lo que nos permitirá extender el resultado de matrices de tamaño n x n a operadores acotados en l^2, de dimensión infinita. A partir de esta información sobre elementos de $B(l^2)$ y otras conclusiones generales sobre estados en C*-álgebras, probaremos la Conjetura de Kadison-Singer. Algunos de los resultados que exponemos se pueden encontar en [10].

Previamente, veamos que el Teorema de Pavimentación sigue siendo cierto eliminando la restricción de que la matriz T debe ser autoadjunta.

Teorema 2.5 *Para todo ϵ >0 existe un $l \in \mathbb{N}$ tal que para toda matriz T compleja, con todos sus elementos diagonales nulos y de tamaño n x n, existen proyecciones diagonales $P_1,\ldots,P_r$ cuya suma es la matriz identidad tal que*

$$||P_iTP_i|| \leq \epsilon||T||$$

para todo i = 1, 2,..., l.

Demostración: Sea ϵ >0. Por el Teorema de Pavimentación, existe un $r \in \mathbb{N}$ tal que para toda matriz T compleja autoadjunta, con todos sus elementos diagonales nulos, con $||T|| \leq 1$ y de tamaño n x n, existen proyecciones diagonales P_1, ..., P_r cuya suma es la matriz identidad tal que

$$||P_iTP_i|| \leq \epsilon$$

para todo i = 1, 2,..., r.

Veamos que el Teorema 2.5 se cumple para la constante $l = r^2$.

Sea T una matriz compleja, con todos sus elementos diagonales nulos y de tamaño n x n. El caso en que T sea la matriz es trivial, lleva a la desigualdad $0 \leq 0$ para cualquier elección de l proyecciones cuya suma sea la matriz identidad.

Tratamos por lo tanto el caso $T \neq 0$. Definimos las siguientes matrices:

$$A = \frac{T + T^*}{2}$$

$$B = \frac{T - T^*}{2i}$$

Con esta definición, las matrices A y B son autoadjuntas, tienen todos sus elementos diagonales nulos y T = A + iB.

De la desigualdad triangular tomando normas en la definición de A y B, obtenemos $||A|| \leq ||T||$ y $||A|| \leq ||T||$.

Aplicando el Teorema de Pavimentación a las matrices $\frac{A}{||T||}$ y $\frac{A}{||T||}$, de norma menor o igual que 1, tenemos que existen proyecciones diagonales R_1,...,R_r cuya suma es la matriz identidad tal que

$$|||R_i \frac{A}{||T||} R_i|| \leq \frac{\epsilon}{2}$$

para todo i = 1, 2,..., r. Y también existen proyecciones diagonales S_1,...,S_r cuya suma es la matriz identidad tal que

$$|||R_i \frac{B}{||T||} R_i|| \leq \frac{\epsilon}{2}$$

para todo i = 1, 2,..., r.

Ahora definimos para cada uno de los l = m^2 pares (i, j) $\in$ {1, ..., r} x {1, ..., r}, la proyección diagonal P_{ij} = $R_i S_j$ = $S_j R_i$.

Estas proyecciones diagonales están bien definidas, ya que las matrices diagonales conmutan entre sí, y además el producto será una matriz diagonal de tamaño n x n con con elementos diagonales iguales a 0 ó 1 (resultado del producto 0 x 0, 0 x 1 ó 1 x 1, en cada caso). Veamos que el teorema se cumple para estas proyecciones.

La suma de las proyecciones P_{ij} es la matriz identidad:

$$\sum_{i,j} P_{ij} = \sum_{i,j} R_i S_j = \sum_{i=1}^{m} R_i \left(\sum_{j=1}^{m} S_j \right) = \sum_{i=1}^{m} R_i = 1$$

Por otro lado, para todo par (i, j), se tiene la siguiente desigualdad:

$$||P_i A P_i|| = ||S_i R_i A R_i S_i|| \leq ||R_i A R_i|| \leq \frac{\epsilon}{2} ||T||$$

donde la primera desigualdad es consecuencia de que la norma de una proyección ortogonal es menor o igual que 1, y obtuvimos la segunda desigualdad al aplicar el Teorema de Pavimentación a $\frac{A}{||T||}$.

Análogamente, obtenemos la desigualdad $|||P_i B P_i|| \leq \frac{\epsilon}{2} ||T||$.

Por último, aplicando la desigualdad triangular a la expresión T = A + iB, obtenemos el resultado buscado:

$$||P_i T P_i|| \leq ||P_i A P_i|| + ||P_i B P_i|| \leq \epsilon ||T||$$

$\square$

A continuación exponemos el teorema que extiende el resultado del Teorema de Pavimentación a dimensión infinita.

Las matrices n x n (pertenecientes a $M_n(\mathbb{C})$), serán remplazadas por operadores acotados T $\in$ $B(l_2)$ y las proyecciones diagonales serán reemplazadas por elementos de la C*-subálgebra Maximal Abeliana (CCOC) de $B(l_2)$, que se identifican con los elementos de l^∞, como vimos en la introducción, con la propiedad de que sus componentes (como elemento de l^∞) son 0 ó 1.

la condición sobre los elementos diagonales de un operador se expresará en términos de la función diag: $B(l_2) \longrightarrow l^\infty$, definida como: diag(T)(i) = $\langle T(e_i), e_i \rangle$, donde e_i es la sucesión con todos sus elementos nulos, salvo el elemento i-ésimo, que es igual a 1.

Con estas modificaciones, el teorema nos proporcionará un resultado muy potente que involucra a elementos de $B(l_2)$ y l^∞, espacios en los que se formula la Conjetura de Kadison-Singer, de manera que su prueba se deducirá casi inmediatamente de éste.

Teorema 2.6 *Para todo ϵ >0 existe un $l \in \mathbb{N}$ tal que para todo $T \in B(l_2)$ con diag(T) = 0, existen proyecciones diagonales P_1,...,P_l con $\sum_{i=1}^{l} P_i$ = 1 tal que*

$$||P_i T P_i|| \leq \epsilon ||T||$$

para todo i = 1, 2,..., l.

Demostración: Sea ϵ >0. Por el Teorema 2.5, existe un l $\in \mathbb{N}$ tal que para toda matriz A $\in$ $M_n(\mathbb{C})$, con todos sus elementos diagonales nulos, existen proyecciones diagonales R_1, ..., R_l $\in$ $M_n(\mathbb{C})$ con $\sum_{i=1}^{l} R_i = 1$ tal que

$$||R_i A R_i|| \leq \epsilon$$

para todo i = 1, 2,..., l.

Sea T $\in$ $B(l_2)$ con diag(T) = 0.

Para cada n $\in \mathbb{N}$ consideramos la función ϕ_n: $B(l_2) \longrightarrow M_n(\mathbb{C})$ que asigna a cada operador B $\in$ $B(l_2)$ la matriz formada por sus n x n primeros elementos de matriz en la base canónica ortonormal, es decir $(\phi_n(\mathrm{B}))_{ij} = \langle B(e_j), e_i \rangle$.

Trivialmente, vemos que, para todo n, diag(ϕ_n(T)) = 0 y $||\phi_n(\mathrm{T})||$ = 1. Por lo tanto, existen proyecciones diagonales $R_{n,1}$,...,$R_{n,l}$ $\in \mathbb{N}$ con $\sum_{i=1}^{l} R_i = 1$ tal que

$$||R_{n,i}\phi_n(T)R_{n,i}|| \leq \epsilon||\phi_n(T)|| \leq \epsilon||T||$$

para todo i = 1, 2,..., l.

Ahora veamos que existe una función estrictamente creciente f: $\mathbb{N} \longrightarrow \mathbb{N}$ y l sucesiones cuyos elementos son 0 ó 1, $\{y_i\}_{i=1}^{l} \subset$ $\{0, 1\}^{\mathbb{N}}$, de tal manera que para cada i fijo, y_i es el límite de la sucesión de sucesiones $\{x_{i,n}\}_{n\in\mathbb{N}} \subset \{0, 1\}^{\mathbb{N}}$, definida de la siguiente manera:

$$x_{i,n}(m) = \begin{cases} \left\langle R_{f(n),i}(e_m), e_m \right\rangle & si \quad m \leq f(n) \\ \\ 0 & si \quad m > f(n) \end{cases}$$

para cada i = 1, ..., l.

Estas sucesiones pertenecen a $\{0, 1\}^{\mathbb{N}}$, ya que los elementos de matriz diagonales de las proyecciones $R_{n,i}$ son 0 ó 1, por ser proyecciones diagonales.

Lo probamos para todo l $\in \mathbb{N}$ por inducción. El caso l = 1 se cumple para y_1 = (1, 1, 1, ...) y para la función identidad f = Id, ya que $R_{n,i}$ = 1, y con esta elección de f, la sucesión n-ésima $x_{i,n}$ tiene sus primeros n elementos igual a 1 y los demás nulos, por lo que converge a y_1.

Ahora suponemos cierto el caso l-1. Existe, por tanto, una función estrictamente creciente g: $\mathbb{N} \longrightarrow \mathbb{N}$ y l-1 sucesiones cuyos elementos son 0 ó 1, $\{y_i\}_{i=1}^{l-1} \subset \{0, 1\}^{\mathbb{N}}$, de tal manera que para cada i fijo, y_i es el límite de una sucesión $\{z_{i,n}\}_{n\in\mathbb{N}} \subset \{0, 1\}^{\mathbb{N}}$, definida de la siguiente manera:

$$z_{i,n}(m) = \begin{cases} \left\langle R_{g(n),i}(e_m), e_m \right\rangle & si \quad m \leq g(n) \\ \\ 0 & si \quad m > g(n) \end{cases}$$

para cada i = 1, ..., l-1.

Definimos $\{w_n\}_{n\in\mathbb{N}} \subset \{0, 1\}^{\mathbb{N}}$, como:

$$w_n(m) = \begin{cases} \left\langle R_{g(n),l}(e_m), e_m \right\rangle & si \quad m \leq g(n) \\ \\ 0 & si \quad m > g(n) \end{cases}$$

Como $\{0, 1\}^{\mathbb{N}}$ es producto del conjunto compacto $\{0, 1\}$, por el Teorema de Tychonoff es un espacio métrico compacto, luego la sucesión $\{w_n\}_{n\in\mathbb{N}} \subset \{0, 1\}^{\mathbb{N}}$ admite una subsucesión convergente $\{w_{n_k}\}_{k\in\mathbb{N}}$, cuyo límite denominamos $y_l \in \{0, 1\}^{\mathbb{N}}$.

Definimos la función h: $\mathbb{N} \longrightarrow \mathbb{N}$ como h(k) = n_k. Como h y g son estrictamente crecientes, su composición f = h $\circ$ g también lo es.

Vemos que la propiedad que estamos probando se cumple para la función f = h $\circ$ g y el conjunto de sucesiones $\{y_i\}_{i=1}^{l}$. En efecto, si para i = 1, ..., l-1; definimos $x_{i,n}$ = $z_{i,h(n)}$, $\{x_{i,n}\}_{n\in\mathbb{N}}$ es una subsucesión de $\{z_{i,h(n)}\}_{n\in\mathbb{N}}$ y por tanto converge su mismo límite y_i. Además se tiene que:

$$x_{i,n}(m) = \begin{cases} \left\langle R_{f(n),i}(e_m), e_m \right\rangle & si \quad m \leq f(n) \\ \\ 0 & si \quad m > f(n) \end{cases}$$

Para i=l, tomamos $x_{l,n} = w_{h(n)}$, que converge a y_i y cumple:

$$x_{l,n}(m) = \begin{cases} \left\langle R_{f(n),l}(e_m), e_m \right\rangle & si \quad m \leq f(n) \\ \\ 0 & si \quad m > f(n) \end{cases}$$

Hemos garantizado la existencia del conjunto $\{y_i\}_{i=1}^{l} \subset \{0, 1\}^{\mathbb{N}}$ con las propiedades que hemos mencionado. Si vemos cada vector y_i como un elemento de l^{∞}, l proyecciones diagonales, $\{P_i\}_{i=1}^{l} \subset l^{\infty}$, donde $P_i(m) = y_i(m)$ para todo i y para todo m. Veamos que el teorema se cumple para estas proyecciones diagonales.

Primeramente, comprobemos que $\sum_{i=1}^{l} P_i = 1$. Veamoslo para cada componente $\sum_{i=1}^{l} P_i(m)$ con $m \in \mathbb{N}$. Tenemos que para cada i fijo, $x_{i,n}(m)$ converge a $y_i(m)$. Al tomar valores $x_{i,n}(m)$ en el espacio discreto $\{0, 1\}$, esta convergencia exige que exista un número natural N_i tal que $x_{i,n}(m) = y_i(m)$ para todo $n \geq N_i$. Llamando $N = \max_i N_i$, se tiene:

$$\sum_{i=1}^{l} p_i(m) = \sum_{i=1}^{l} y_i(m) = \sum_{i=1}^{l} x_{i,N}(m) = \sum_{i=1}^{l} \left\langle R_{f(n),i}(e_m), e_m \right\rangle =$$

$$= \left\langle \left(\sum_{i=1}^{l} R_{f(n),i} \right)(e_m), e_m \right\rangle = \langle e_m, e_m \rangle = 1$$

Por último, veamos que $||P_i T P_i|| \leq \epsilon ||T||$. Para ello, utilizamos el siguiente resultado, que podemos encontrar en el Apéndice B.10 de [10].

Sea H un espacio de Hilbert con una base ortonormal $\{v_i\}_{i \in I}$. Sea A: H $\longrightarrow$ H un operador acotado y $\alpha > 0$ tal que $|\langle x, A(y) \rangle| \leq \alpha \, ||x|| \, ||y||$ para todo x, y $\in$ H con soporte finito (es decir, tal que $\{i \in I$ tal que $\langle x, v_i \rangle \neq 0\}$ y $\{i \in I$ tal que $\langle y, v_i \rangle \neq 0\}$ son conjuntos finitos). Entonces $||A|| \leq \alpha$.

En nuestro caso, un vector en l^2 tiene soporte finito si tiene un número finito de componentes. Sean u, v $\in l^2$ con soporte finito, y sea M tal que $u(n) = v(n) = 0$ para todo $n \geq M$. Busquemos una cota adecuada para $|\langle P_i T P_i(u), v \rangle|$ con la que poder aplicar el resultado anterior.

Para todo m entero con $1 \leq m \leq M$, existe un $N_m \in \mathbb{N}$ tal que $x_{i,n}(m) = y_i(m) = P_i(m)$ para todo $n \geq N_m$, debido a la convergencia de los $x_{i,n}$ definidos anteriormente. Llamamos $N = \max(M, \max_{1 \leq m \leq M} N_m)$.

Definimos la función c_N: $l^2 \longrightarrow \mathbb{N}$ que asigna a cada vector h de l^2 el vector que está formado por sus n primeras componentes, es decir, $c_N(h)(n) = h(n)$.

Para todo i = 1, 2, ..., l; se tiene:

$$\langle P_i T P_i(u), v \rangle = \langle T P_i(u), P_i(v) \rangle = \langle \phi_n(T) c_N(P_i(u)), c_N(P_i(v)) \rangle$$

dado que el soporte de $P_i(u)$ está contenido en el soporte de u (y se tiene la propiedad análoga para v).

Ahora, por la elección que hemos hecho del índice N, tenemos que $c_N(P_i(u)) = R_{N,i}(c_N(u))$, y $c_N(P_i(v)) = R_{N,i}(c_N(v))$. Sustituyendo en el último término de la igualdad superior:

$$\langle P_iTP_i(u), v\rangle = \langle TP_i(u), P_i(v)\rangle =$$
$$= \langle \phi_n(T)R_{N,i}(c_N(u)), R_{N,i}(c_N(v))\rangle = \langle R_{N,i}\phi_n(T)R_{N,i}(c_N(u)), (c_N(v))\rangle$$

Ahora, tomando el módulo del producto escalar:

$$|\,\langle P_iTP_i(u), v\rangle\,| = |\,\langle R_{N,i}\phi_n(T)R_{N,i}(c_N(u)), (c_N(v))\rangle\,| \leq$$
$$\leq ||R_{N,i}\phi_n(T)R_{N,i}(c_N(u))||||(c_N(v))|| \leq$$
$$\leq ||R_{N,i}\phi_n(T)R_{N,i}||||(c_N(u))||||(c_N(v))|| \leq \epsilon||T||||(c_N(u))||||(c_N(v))||$$

donde la primera desigualdad es la desigualdad de Cauchy-Schwarz de ese producto escalar, la segunda desigualdad es consecuencia de la definición de norma de un operador y la última desigualdad se obtiene aplicando la propiedad con la que definimos los $R_{N,i}$ al principio de la demostración del teorema.

Ahora, utilizando la propiedad anterior, tenemos que para todo i = 1, 2, ..., l; $||P_iTP_i|| \leq \epsilon\ ||T||$, y el teorema queda probado. □

En la demostración de la Conjetura de Kadison-Singer utilizamos, además del teorema anterior, una propiedad general sobre los estados puros en una C*-álgebra, que exponemos en el Teorema 2.9. Para la demostración del Teorema 2.9, necesitamos los dos lemas siguientes, también de carácter general en la teoría de álgebra de operadores.

Lema 2.7 *Sea U una C*-álgebra. Todo elemento $A \in U$ puede expresarse como la combinación de cuatro elementos positivos A_0, A_1, A_2, $A_3 \in U$ de la siguiente manera:*

$$A = \sum_{k=0}^{3} i^k A_k$$

Demostración: Sea A ∈ U. Es posible escribir A = B + iC, con B y C ∈ U autoadjuntos. En efecto, la ecuación se cumple para B = $\dfrac{A + A*}{2}$ y C = $\dfrac{A - A*}{2i}$, que son autoadjuntos.

Cada elemento autoadjunto Q ∈ U se puede descomponer como diferencia de dos elementos positivos, Q = Q^+ - Q^-, con Q^+, Q^- ∈ U positivos.

Una demostración precisa de este hecho se encuentra en la Proposición 4.2.3 de[3]. Sin entrar en detalles, el hecho de que Q sea autoadjunta permite definir su cálculo funcional, una herramienta que permite identificar las funciones continuas reales definidas en el espectro de Q con elementos de $\in$ U. La descomposición de Q (cuyo espectro es real) en dos elementos positivos (con espectro real no negativo) es análoga a la la descomposición de una función f real como resta de su parte positiva y la opuesta de su parte negativa (f = f^+ - f^-, con f^+(t) = max $\{t, 0\}$ y f^-(t) = max $\{-t, 0\}$).

Si aplicamos esta descomposición a B y C, tenemos A = B^+ - B^- + iC^+ - iC^-, lo que demuestra el lema. □

Lema 2.8 *Sea U una C*-álgebra con elemento unidad. Sea φ un estado puro en U. Entonces, para todo funcional lineal Ψ: U $\longrightarrow$ $\mathbb{N}$ con 0 $\leq \Psi \leq \varphi$, existe un t con 0 $\leq$ t $\leq$ 1 tal que se tiene Ψ = tφ.*

Demostración: Sea φ un estado puro en U y Ψ: U $\longrightarrow$ $\mathbb{N}$ con 0 $\leq \Psi \leq \varphi$. El elemento unidad de U, 1, es positivo, por lo que 0 $\leq \Psi(1) \leq \varphi(1) = 1$. Tratamos por separado los tres casos posibles, según el valor de $\Psi(1)$.

Comenzamos por el caso en que $\Psi(1) = 0$. Si A $\in$ U es positivo, $0 \leq \dfrac{A}{||A||} \leq 1$, luego 0 $\leq$ A $\leq$ ||A|| 1, Por la positividad de Ψ, tenemos:

$$0 \leq \Psi(A) \leq \Psi(||A||1) = ||A||\Psi(1) = 0$$

luego Ψ(A) = 0. Por el Lema 2.7 anterior, cualquier elemento del álgebra U se puede escribir como una combinación de elementos positivos. Por la linealidad de Ψ, Ψ es idéniticamente nulo en U, $\Psi = 0$.

El segundo caso es en el que $\Psi(1) = 1$. Como $0 \leq \Psi \leq \varphi$, tenemos φ - $\Psi \geq 0$. Además, (φ - Ψ) = 0, luego aplicando el argumento del caso anterior a la función φ - Ψ, φ - Ψ = 0, luego $\Psi = \varphi$.

Por último, estudiamos el caso en que 0 <$\Psi(1)$ <1. Podemos definir los funcionales $\Psi_1 = \dfrac{1}{1-\Psi(1)}$ (φ - Ψ) y $\Psi_2 = \dfrac{1}{\Psi(1)}$ Ψ, que son positivos por serlo Ψ y φ - Ψ, y que cumplen $\Psi_1(1) = \Psi_2(1) = 1$, luego son estados en U. Además,

$$(1 - \Psi(1))\Psi_1 + \Psi(1)\Psi_2 = \varphi - \Psi + \Psi = \varphi$$

luego φ es una combinación convexa de los estados $\Psi_1(1)$, $\Psi_2(1)$. Como φ es puro, es extremal en el conjunto de estados, por lo tanto esa combinación debe ser trivial. En este caso (1 - $\Psi(1)$) y $\Psi(1)$ son distintos de cero, por lo que $\Psi_1 = \Psi_2 = \varphi$, y de la definición de Ψ_2, resulta $\Psi = \Psi(1)\varphi$.

Los resultados obtenidos en los dos primeros casos también se pueden escribir como $\Psi = \Psi(1)\varphi$, por lo que el lema se cumple para la constante t = $\Psi(1)$. □

En el siguiente teorema obtenemos una propiedad muy fuerte para los estados puros en una C*-álgebra abeliana que será necesaria para garantizar que la extensión de un estado es única en la demostración de la Conjetura de Kadison-Singer, de manera que observamos el papel fundamental que juega en la conjetura la hipótesis de que el estado cuya extensión es única sea un estado puro.

Teorema 2.9 *Sea U una C*-álgebra abeliana con elemento unidad. Sea φ un estado puro en U. Entonces, para todo A, B $\in$ U, φ(AB) = φ(A) φ(B).*

Demostración: Sea φ un estado puro en U. Sean A, B $\in$ U.

Primeramente, tratamos el caso $0 \leq$ B ≤ 1. Sea C $\in$ U positivo. Dado que B, 1 - B y C son positivos, podemos escribiir B = D*D, (1 - B) = U*U y C = V*V, con D, U, V $\in$ U. Por un lado,

$$CB = V^*VD^*D = D^*V^*VD = (VD)^*VD \geq 0$$

y por otro lado,

$$C - CB = C(1 - B) = V^*VU^*U = U^*V^*VU = (VU)^*VU \geq 0$$

luego, $0 \leq$ CB $\leq$ C.

Ahora definimos el funcional Ψ: U $\longrightarrow \mathbb{C}$ como Ψ(Z) = φ(ZB) para todo Z $\in$ U. Como φ es positivo por ser un estado, y ZB ≥ 0 para todo Z ≥ 0, Ψ es positivo.

Además, si Z ≥ 0, también se tiene ZB ≥ 0, y

$$(\varphi - \Psi)(Z) = \varphi(Z) - \Psi(Z) = \varphi(Z) - \varphi(ZB) = \varphi(Z - ZB) \geq 0$$

con lo que φ - Ψ también es positivo.

Combinando los dos resultados, tenemos $0 \leq \Psi \leq \varphi$. Aplicando el Lema 2.8, existe t con $0 \leq$ t ≤ 1 tal que Ψ = tφ. Ahora es inmediato ver que:

$$\varphi(AB) = \Psi(A) = t\varphi(A) = t\varphi(1)\varphi(A) = \Psi(1)\varphi(A) = \varphi(B)\varphi(A) = \varphi(A)\varphi(B)$$

por tanto el teorema queda probado en este caso.

Para el caso general, sin la restricción $0 \leq$ B ≤ 1, utilizamos la descomposición de B en elementos positivos cuya existencia garantiza el Lema 2.7. Escribimos $B = \sum_{k=0}^{3} i^k ||B_k|| \dfrac{B_k}{||B_k||}$, con $B_k \geq 0$. Ahora, por la linealidad de φ y el hecho de que $0 \leq \dfrac{B_k}{||B_k||} \leq 1$:

$$\varphi(AB) = \varphi\left(A\sum_{k=0}^{3} i^k ||B_k|| \frac{B_k}{||B_k||}\right) = \sum_{k=0}^{3} i^k ||B_k|| \varphi\left(A\frac{B_k}{||B_k||}\right) =$$

$$= \sum_{k=0}^{3} i^k ||B_k|| \varphi(A)\varphi\left(\frac{B_k}{||B_k||}\right) = \varphi(A)\varphi\left(\sum_{k=0}^{3} i^k ||B_k|| \frac{B_k}{||B_k||}\right) = \varphi(A)\varphi(B)$$

aplicando el teorema en el caso anterior a A y $\frac{B_k}{||B_k||}$. $\square$

Mencionamos que el recíproco de los dos últimos resultados, Lema 2.7 y Teorema 2.8 también es cierto, aunque no los utilizamos en nuestro problema.

Ahora estamos en condiciones de demostrar el Teorema de Kadison-Singer.

Demostración:[de la Conjetura de Kadison-Singer (2.1)] Sea φ un estado puro en $l^\infty \subset B(l^2)$. Sea Ψ un estado en $B(l^2)$ que es una extensión de φ. La existencia de, al menos, una extensión de φ está garantizada por el Teorema de Hann Banach. Nuestro objetivo es encontrar una expresión explícita para la extensión Ψ.

Comencemos viendo que, para cualquier $T \in B(l^2)$ con $\text{diag}(T) = 0$, se tiene $\Psi(T) = 0$.

Sea $\epsilon > 0$. Por el Teorema 2.6, existe un $n \in \mathbb{N}$ tal que para todo $T \in B(l^2)$ con $\text{diag}(T) = 0$, existen proyecciones diagonales $P_1, ..., P_n \in l^\infty$ con $\sum_{i=1}^n P_i = 1$ tal que

$$||P_i T P_i|| \leq \epsilon$$

para todo $i = 1, 2, ..., n$.

Sea $T \in B(l^2)$ con $\text{diag}(T) = 0$, y sean $P_1, ..., P_n$ sus proyecciones diagonales que cumplen la desigualdad anterior.

Aplicando el Teorema 2.9 al estado puro φ, tenemos $\varphi(P_i) = \varphi(P_i^2) = \varphi(P_i)^2$ para todo i, donde $P_i{}^2 = P_i$ por ser P_i una proyección. Por tanto, $\varphi(P_i) = 0$ ó 1, para todo $i = 1, 2, ..., n$. Ahora bien, las suma de proyecciones P_i es la unidad de l^∞, y $\varphi(1) = 1$ por ser φ un estado. Entonces, existe un índice i_0 tal que $\varphi(P_{i_0}) = 1$ y $\varphi(P_i) = 0$ si $i \neq i_0$.

Como Ψ es una extensión de φ, también existe un índice i_0 tal que $\Psi(P_{i_0}) = 1$ y $\Psi(P_i) = 0$ si $i \neq i_0$.

Al ser Ψ un estado en $B(l^2)$, la función Ψ: $B(l^2)^2 \longrightarrow \mathbb{C}$ definida como $\Psi(X, Y) = \Psi(X^*Y)$ es un producto escalar en $B(l^2)$ (cumple las propiedades de la definicón de producto escalar que hemos dado en la introducción), por tanto cumple la desigualdad de Cauchy Schwartz: $|\Psi(X^*Y)|^2 \leq \Psi(X^*X)\Psi(Y^*Y)$ para todo $X, Y \in B(l^2)$. Por lo tanto,

$$|\Psi(P_i T P_j)|^2 \leq \Psi(P_i^* P_i)\Psi((TP_j)^* TP_j) = \Psi(P_i)\Psi((TP_j)^* TP_j)$$

para todo $i, j = 1, 2, ..., n$. En consecuencia, $\Psi(P_i\ T\ P_j) = 0$ si $i \neq i_0$. Asimismo, razonando análogamente sobre P_j, también $\Psi(P_i\ T\ P_j) = 0$ si $j \neq i_0$.

Entonces, introduciendo la suma de todas las proyecciones en $|\Psi(T)|$, tenemos:

$$|\Psi(T)| = |\Psi\left(\left(\sum_i P_i\right) T \left(\sum_j P_j\right)\right)| = |\sum_{i,j} \Psi(P_i T P_j)| =$$

$$= |\Psi(P_{i_0}TP_{i_0})| \leq ||P_{i_0}TP_{i_0}|| \leq \epsilon||T||$$

donde la penúltima desigualdad se da porque $||\Psi|| = 1$, al ser Ψ un estado.

Como esta desigualdad se tiene con ϵ para todo T $\in$ B(l^2) con diag(T) = 0, ϵ es una cota para la norma del funcional Ψ restringido al subespacio vectorial de B(l^2) formado por los elementos de B(l^2) con diag(T) = 0, $||\Psi|| \leq \epsilon$.

Como podemos aplicar el Teorema 2.6 (y por tanto este argumento), para cualquier $\epsilon \geq 0$, tenemos $||\Psi|| \leq \epsilon$ para todo $\epsilon \geq 0$, y por tanto $\Psi = 0$ en ese subespacio.

Ahora es sencillo ver que debe ser $\Psi = \varphi \circ$ diag en B(l^2). En efecto, si A $\in$ B(l^2), diag(A - diag(A)) = diag(A) - diag(A) = 0, luego aplicando el resultado anterior, Ψ(A - diag(A)) = 0, por lo que:

$$\Psi(A) = \Psi(diag(A)) = \varphi(diag(A))$$

ya que diag(A) $\in l^\infty$ por ser imagen de la función diag.

Hemos demostrado que si Ψ es una extensión de φ, entonces $\Psi = \varphi \circ$ diag, por lo que si existen dos funciones Ψ_1 y Ψ_2 que sean extensiones de φ, debe ser $\Psi_1 = \varphi \circ$ diag $= \Psi_1$.

Luego la extensión es única y la conjetura de Kadison-Singer queda probada con resultado positivo. $\square$

A lo largo de este capítulo, hemos demostrado que si la Conjetura KS_r es cierta para algún r ≥ 2, la conjetura de Kadison-Singer tiene resultado positivo.

En lo sucesivo nos centraremos en encontrar una prueba positiva de la conjetura KS_2, lo que, por lo tanto, es suficiente para que la Conjetura de Kadison-Singer quede demostrada.

Veremos a continuación que el teorema KS_2 se deduce del siguiente teorema, cuya demostración detallamos en el próximo capítulo:

Teorema 2.10 *Sea ϵ >0. Sean v_1, ..., v_m vectores aleatorios independientes en $\mathbb{C}^d$ con soporte finito tal que $\mathbf{E}||v_i|| \leq \epsilon$ para todo i y tal que*

$$\sum_{i=1}^{m} \mathbf{E} v_i v_i^* = I_d$$

Entonces

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

En este caso, denotamos por v_i^* el vector traspuesto conjugado de v_i.

Precisemos los conceptos probabilísticos que aparecen en la formulación de este teorema.

Un espacio probabilístico (Ω, σ, P) consta de un espacio muestral Ω cuyos elementos son los posibles resultados de un experimento aleatorio, una sigma-álgebra σ cuyos elementos son ciertos subconjuntos de Ω a los que es posible asignar una probabilidad y una probabilidad P:$\sigma \longrightarrow [0, 1]$ con $\mathsf{P}(\Omega) = 1$ y tal que para toda unión disjunta numerable de subconjuntos $\mathrm{A}_i \in \sigma$, $\cup_{i=1}^{\infty} \mathrm{A}_i$, se tiene $\mathrm{P}(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathrm{P}(\mathrm{A}_i)$.

Un vector aleatorio v en $\mathbb{C}^d$ con soporte finito es una función v:$\Omega \longrightarrow \mathbb{C}^d$ de tal manera que el conjunto de valores posibles de v, que es el conjunto imagen de v, $\mathrm{v}(\omega) \subset \mathbb{C}^d$, es finito; y si $\mathrm{B} \subset \mathrm{v}(\Omega)$, entonces su contraimagen por v está en σ, $\mathrm{v}^{-1}(\mathrm{B}) \in \sigma$.

Esta última propiedad nos permite definir la ley de probabilidad de v, que es una probabilidad $\mathsf{P} \circ \mathrm{v}^{-1}$ definida en el conjunto de partes de $\mathrm{v}(\Omega)$ como P $(\mathrm{v} \in \mathrm{B}) = \mathsf{P} \circ \mathrm{v}^{-1}$ (B).

Si v_1, ..., v_m vectores aleatorios en $\mathbb{C}^d$, el vector combinado dado por (v_1, ..., v_m) es un vector aleatorio en $(\mathbb{C}^d)^m$, por tanto tiene una ley de probabilidad asociada, que llamaremos ley de probabilidad conjunta de v_1, ..., v_m. Decimos que los vectores aleatorios v_1, ..., v_m son independientes si, eligiendo cualquier cantidad finita k de ellos, v_{j_k}, ..., v_{j_k}, su ley de probabilidad conjunta es el producto de sus leyes de probabilidad (individuales).

La esperanza matemática de un vector aleatorio v en $\mathbb{C}^d$ con soporte finito, que puede tomar valores z_1,...,$\mathrm{z}_n \in \mathbb{C}^d$, se define como $\mathbf{E}\mathrm{v} = \mathrm{z}_1\mathrm{P}(\mathrm{v} = \mathrm{z}_1) + ... + \mathrm{z}_n\mathrm{P}(\mathrm{v} = \mathrm{z}_n)$. La esperanza es lineal en el conjunto de vectores aleatorios.

La esperanza del vector conjunto (v_1, ..., v_m) de los vectores aleatorios v_1, ..., v_m es igual al producto de las esperanzas de todos los vectores si y solo si estos son independientes.

Es importante notar que, al tratarse de un soporte finito de probabilidad, la desigualdad que implica el teorema es equivalente a la existencia de al menos una asignación de v_1,...,v_m con probabilidad no nula, tal que $\|\sum_{i=1}^{m} v_i v_i^*\| \leq (1 + \sqrt{\epsilon})^2$.

De este resultado se sigue el siguiente corolario, en términos de proyecciones, del que se deduce trivialmente el Teorema KS_2:

Corolario 2.11 *Sea δ >0 y r un entero positivo. Sean u_1,...,$u_m \in C^d$ tal que $//u_i// \leq \delta$ para todo i y tal que*

$$\sum_{i=1}^{m} u_i u_i^* = I_d$$

Entonces existe una partición S_1,..., S_r de $\{1, ...m\}$ tal que

$$\left\| \sum_{i \in S_j} u_i u_i^* \right\| \leq \left(\frac{1}{\sqrt{r}} + \sqrt{\delta} \right)^2$$

para todo j = 1,..., r.

Demostración: Sean $\delta>0$, r un entero positivo y $u_1,...,u_m \in C^d$ cumpliendo las hipótesis del corolario. Para cada i $\in \{1,...m\}$ y cada k $\in \{1,...r\}$, definimos $w_{i,k} \in (\mathbb{C}^d)^r$ como el vector cuya componente k-ésima es u_i y el resto de sus componentes son el vector nulo en $\mathbb{C}^d$.

A partir de estos vectores, definimos los vectores aleatorios independientes $v_1,...,v_m \in (\mathbb{C}^d)^r$ tal que v_i toma los valores $\{\sqrt{r}w_{i,1},...\sqrt{r}w_{i,r}\}$ con distribución de probabilidad uniforme (esto es, cada valor con probabilidad $\frac{1}{r}$).

Calculamos para cualquier i, la esperanza de $v_i v_i^*$ es:

$$\mathbf{E}v_i v_i^* = \begin{pmatrix} u_i u_i^* & 0 & \cdots & 0 \\ 0 & u_i u_i^* & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_i u_i^* \end{pmatrix}$$

En este caso denotamos por 0 al elemento nulo $\mathbb{C}^{dxd}$ de resultante del producto de un vector nulo en $\mathbb{C}^d$ por su traspuesto.

Puesto que por hipótesis la suma de los $u_i u_i^*$ es la identidad, la suma de esperanzas resulta en la identidad

$$\mathbf{E}\left(\sum_{i=1}^{m} v_i v_i^*\right) = I_{rd}$$

Por otra parte, $||v_i|| = r||u_i|| \leq r\delta$ para todo i.

Estamos en las condiciones de aplicar el Teorema 2.10 a los vectores $v_1,...,v_m$ con $\epsilon = r\delta$ y dimensión rd. Existe, por tanto, una asignación de $v_1,...,v_m$ tal que:

$$(1+\sqrt{\epsilon})^2 \geq \left\|\sum_{i=1}^{m} v_i v_i^*\right\| = \left\|\sum_{k=1}^{r}\sum_{i:v_i=w_{i,k}} \sqrt{r}w_{i,k}\sqrt{r}w_{i,k}^*\right\|$$

Denominando $S_k = \{i : v_i = w_{i,k}\}$ para todo k = 1,..., r, tenemos que estos S_k son los elementos de la partición que cumplen la condición del corolario, ya que, por ser diagonal la matriz que aparece en la última igualdad, tenemos que:

$$\left\|\sum_{i\in S_j} u_i u_i^*\right\| = \left\|\sum_{i\in S_j} w_{i,k}w_{i,k}^*\right\| \leq \frac{1}{r}\left\|\sum_{k=1}^{r}\sum_{i:v_i=w_{i,k}} \sqrt{r}w_{i,k}\sqrt{r}w_{i,k}^*\right\| \leq \left(\frac{1}{\sqrt{r}}+\sqrt{\delta}\right)^2$$

para todo k = 1,..., r. □

Por último, terminamos esta sección utilizando este corolario para demostrar el Teorema KS_2.

Mencionamos que, a pesar de que aquí hemos propuesto la demostración del Teorema de Kadison Singer partiendo del Teorema 2.10, que implica el Corolario 2.11, que a su vez implica el Teorema KS_2, a partir del cual hemos visto que se puede obtener el Teorema de Pavimentación 2.3 y a partir de éste último obtener el Teorema de Kadison-Singer, es posible obtener el Teorema de Pavimentación 2.3 directamente como consecuencia del Corolario 2.11, tomando $r=(6/\epsilon)^4$, como se establece en [6].

Demostración: del Teorema KS_2 (2.2 en el caso particular $r = 2$)

Elegimos las constantes universales $\eta=18$ y $\theta=2$. Sea $w_1,...,w_m \in C^d$ cualquier elección de vectores cumpliendo las hipótesis del Teorema KS_2. Para todo i, definimos $u_i=w_i/\eta$. Además definimos las constantes $\delta = 1/18$ y $r = 2$. Entonces, $||u_i||\leq 1$ para todo i. Por otro lado

$$\sum_{i=1}^{m} |\left\langle u, u_i \right\rangle|^2 = 1$$

para todo vector unitario $u \in C^d$.

Eligiendo en la suma anterior como vector unitario $u_i/||u_i||$ para todo i, obtenemos que el conjunto de vectores $u_1,...,u_m$ debe ser ortogonal. Además debe ser una base ortogonal, ya que de existir algún vector unitario ortogonal a todos los u_i, no se cumpliría la suma. Además, como para todo vector unitario u,

$$||u|| = \sum_{i=1}^{m} |\left\langle u, u_i \right\rangle|^2 = 1$$

tenemos que el conjunto de vectores debe ser una base ortonormal. Por tanto:

$$\sum_{i=1}^{m} u_i u_i^* = I_d$$

Entonces los vectores $u_1,...,u_m$ satisfacen el Corolario 2.11 con constantes $\delta = 1/18$ y $r = 2$. Por tanto, existe una partición S_1,S_2 de $\{1, ...m\}$ tal que

$$\sum_{i \in S_j} |\left\langle u, w_i \right\rangle|^2 = \eta \left\| \sum_{i \in S_j} u_i u_i^* \right\| \leq \eta \left(\frac{1}{\sqrt{r}} + \sqrt{\delta} \right)^2 = 16 = 18 - 2 = \eta - \theta$$

para todo j y para todo vector unitario $u \in C^d$. □

3. El resultado principal y la prueba de la Conjetura de Kadison-Singer.

En este último capítulo exponemos una demostración detallada del Teorema 2.10, que por lo visto en el capítulo 2, es suficiente para completar la prueba positiva de la Conjetura de Kadison-Singer (2.1).

Esta demostración nos proporciona además un ejemplo de aplicación de la técnica de polinomios entrelazados , la cual mostramos en el capítulo 1.

Comenzaremos discutiendo el contenido del teorema y describiéndolo en términos de polinomios.

Para todo i = 1, ..., m, v_i es un vector aleatorio que toma valores en $\mathbb{C}^d$, por lo tanto $v_i v_i^*$ es una matriz compleja de tamaño d x d. Además es hermítica, es decir, es igual a su conjugada traspuesta; y es de rango 1, ya que la imagen de cualquier vector $z \in \mathbb{C}^d$ por $v_i v_i^*$ es proporcional a v_i, $(v_i v_i^*)z = (v_i^* z)v_i$. También es semidefinida positiva, es decir, $\overline{z} v_i v_i^* z = (v_i^* z)^* v_i^* z = |v_i^* z|^2 \geq 0$ para todo $z \in \mathbb{C}^d$.

Por lo tanto, $\sum_{i=1}^m v_i v_i^*$ es una matriz compleja de tamaño d x d hermítica, suma de matrices de rango 1. Por ser hermítica, las raíces de su polinomio caracteístico $\chi\left(\sum_{i=1}^m v_i v_i^*\right)(x) \equiv det\left(xI - \sum_{i=1}^m v_i v_i^*\right)$ (sus autovalores) son reales, y sus coeficientes también.

Esta matriz, como matriz de un operador T en $\mathbb{C}^d$, tiene una norma, $||T|| = sup\{||T(z)||$ tal que $||z|| \leq 1\}$, que es igual a la mayor raíz de su polinomio característico,

$$\left\|\sum_{i=1}^m v_i v_i^*\right\| = \lambda_1\left(\chi\left(\sum_{i=1}^m v_i v_i^*\right)\right)$$

Por otra parte, por ser v_1, ..., v_m vectores aleatorios en $\mathbb{C}^d$ con soporte finito, (v_1, ..., v_m) sólo puede tomar un número finito de valores, cada uno de los cuales da lugar a uno de los posibles valores que puede tomar el polinomio $\chi\left(\sum_{i=1}^m v_i v_i^*\right)$. Habrá por lo tanto un número finito de estos posibles valores. Esto nos permite ver al polinomio $\chi\left(\sum_{i=1}^m v_i v_i^*\right)$ como una variable aleatoria con soporte finito contenido en el conjunto de polinomios con coeficientes y raíces reales de grado d.

El teorema afirma, por tanto, que existe al menos una asignación del polinomio $\chi\left(\sum_{i=1}^m v_i v_i^*\right)$ con probabilidad no nula que cumple que su raíz más grande es menor o igual que $(1+\sqrt{\epsilon})^2$ (de tal manera que la asignación correspondiente de los vectores v_1, ..., v_m, a partir de los cuales se construye el polinomio, cumpla las hipótesis del teorema).

La idea principal de esta demostración es sortear la dificultad de tener que buscar una cota para las raíces de alguno de los valores (polinomios) que puede tomar la variable $\chi\left(\sum_{i=1}^m v_i v_i^*\right)$, de los cuales puede haber una cantidad finita, pero arbitrariamente grande; reduciendo el problema a demostrar que la esperanza de esta variable es un polinomio con todas sus raíces reales, buscar una cota superior adecuada para estas raíces y demostrar que los posibles valores que puede tomar la variable forman una familia entrelazada.

La esperanza matemática de la variable aleatoria $\chi\left(\sum_{i=1}^m v_i v_i^*\right)$, al ser esta de soporte finito, es una combinación lineal convexa de los posibles valores que puede tomar la variable. Demostraremos que estos posibles valores forman una familia de polinomios entrelazados, lo que nos permitirá aplicar una version del Corolario 1.5 para familias entrelazadas, que se deduce como un caso particular del Teorema 1.8, que garantizará la existencia de un polinomio cuya mayor raíz sea a lo sumo la mayor raíz de la esperanza. Utilizaremos que $(1+\sqrt{\epsilon})^2$ es una cota superior de esta última para completar la demostración.

La posibilidad de realizar esta reducción del problema es un ejemplo de la potencia de esta técnica de polinomios entrelazados, y fue la clave de la prueba original de la Conjetura de Kadison-Singer que realizaron Spielman, Marcus y Srivastava [6].

En la demostración de este teorema utilizaremos resultados que podemos encontrar en [10] y [6], algunos con ligeras variaciones, debido a que la definición de familia entrelazada que hemos dado en el capítulo 1 es más general que la definición que aparece en estas fuentes.

Comenzaremos exponiendo el concepto de estabilidad real en polinomios, el cual es una generalización de la propiedad de polinomios en una variable de tener todas las raíces reales a polinomios multivariantes, y varios resultados relacionados con este, que emplearemos más adelante para deducir que la esperanza de $\chi\left(\sum_{i=1}^m v_i v_i^*\right)$ tiene raíces reales.

Definición 3.1 *Decimos que un polinomio en m variables con coeficientes complejos, $f \in \mathbb{C}[z_1,\ldots,z_m]$, es estable si no tiene raíces con las partes imaginarias de todas sus coordenadas mayores estrictamente mayor que 0, es decir, si:*

$$f(z_1,...,z_m) \neq 0 \text{ siempre que para todo } i = 1,...,m \text{ se tenga } Im(z_i) > 0.$$

En particular, decimos que f es real estable si es estable y tiene todos sus coeficientes reales.

Si un polinomio f tiene todos sus coeficientes reales, z es raíz de p si y sólo si lo es su complejo conjugado $\overline{z}$. Por tanto, la estabilidad real de un polinomio garantiza que tampoco tiene raíces con las partes imaginarias de todas sus coordenadas menores estrictamente que 0.

Por lo tanto, en el caso partícular de un polinomio p(z) en una variable, la condición de estabilidad real implicará que no tiene raíces con Im(z)>0 ni con Im(z)<0, luego esta condición es equivalente a tener todas sus raíces reales. Esto justifica haber introducido la condición de estabilidad real como una generalización de esta última a polinomios multivariantes.

En algunos casos resultará de utilidad la siguiente caracterización de la estabilidad (y de la estabilidad real).

Lema 3.2 *Sea $f \in \mathbb{C}[z_1,...,z_m]$. Entonces f es estable (respectivamente, real estable) si y sólo si para todo $a \in \mathbb{R}^m$ y para todo $v \in (0,\infty)^m$ se tiene que el polinomio de una variable compleja $f(a + vt)$ es estable (respectivamente, real estable).*

Demostraremos que la esperanza tiene raíces reales obteniendo una expresión para ella como el resultado de una aplicación de transformaciones que preservan la estabilidad real a un polinomio real estable, de tal manera que el polinomio resultante de estas aplicaciones sea de una variable, y por tanto de raíces reales por ser real estable.

Para ello veremos a continuación un tipo de polinomio real estable y dos transformaciones que preservan la estabilidad real.

Proposición 3.3 *Sean A_1, ..., A_m matrices semidefinidas positivas y sea B una matriz hermítica, todas ellas de tamaño d x d. Entonces el polinomio:*

$$f(z_1, ..., z_m) = det\left(\sum_{i=1}^{m} z_i A_i + B\right)$$

es, o bien real estable, o bien idénticamente nulo.

Demostración: Comenzamos tratando el caso en el que A_1, ..., A_m son matrices definidas positivas

Utilizamos la caracterización del Lema 3.2. Tomamos z(t) = a + vt, con a = $(a_1,...a_m) \in \mathbb{R}^m$ y v = $(v_1,...,v_m) \in (0,\infty)^m$ cualquiera, y $t \in \mathbb{C}$; y debemos demostrar que el polinomio:

$$f(z(t)) = det\left(\sum_{i=1}^{m} a_i A_i + \sum_{i=1}^{m} t v_i A_i + B\right)$$

tiene todas sus raíces reales.

Denotamos $Q \equiv \sum_{i=1}^{m} t v_i A_i$, la cual es una matriz definida positiva por ser todos los $v_i > 0$. Al ser definida positiva es invertible y posee una raíz cuadrada $Q^{(1/2)}$. Denotamos $H \equiv \sum_{i=1}^{m} a_i A_i + B$, la cual es una matriz hermítica.

Entonces, multiplicando por $Q^{(1/2)}$ y su inversa a izquierda y derecha de H:

$$f(z(t)) = det\,(tQ + H) = det\,(Q)\,det\left(tI + Q^{(-1/2)} H Q^{(-1/2)}\right)$$

Obtenemos que f(z(t)) es proporcional, con constante det(Q) >0 (ya que Q es definida positiva), al polinomio característico de la matriz hermítica $Q^{(-1/2)}HQ^{(-1/2)}$, por lo que tiene todas sus raíces y coeficientes reales.

El caso general, en el que A_1, ..., A_m son matrices semidefinidas positivas, se demuestra tomando el límite de matrices definidas positivas y aplicando que en cada restricción univariante el límite debe tener todas sus raíces reales o ser el polinomio cero. □

Para comprobar que las transformaciones que nos interesan preservan la estabilidad, necesitaremos trabajar con sucesiones convergentes de polinomios en variable compleja y obtener información de las raíces de su límite a partir de información sobre las raíces de los polinomios de la sucesión. Para ello utilizaremos el Teorema de Hurwitz, un resultado central en el análisis complejo.

Teorema 3.4 (de Hurwitz) *Sea $\{f_k\}_{k=1}^{\infty}$ una sucesión de funciones analíticas en un dominio (un conjunto abierto y conexo) $D \subset \mathbb{C}^m$. Si las funciones de la sucesión no tienen raíces en D y convergen uniformemente a una función f en subconjuntos compactos de D, entonces, o bien f no tiene raíces en D, o bien f es idénticamente nula.*

Ahora presentamos una transformación que preserva la estabilidad real que es de especial utilidad por eliminar una variable de un polinomio.

Lema 3.5 *La transformación consistente en evaluar una variable, z_i, de un polinomio estable $f(z_1,\ldots,z_m)$ en un valor c con Im(c) >0 preserva la estabilidad. Es decir, $f(z_1,\ldots,z_{i-1},c,z_{i+1},\ldots,z_m)$ es estable.*

Por otra parte, la transformación consistente en evaluar una variable z_i, de un polinomio real estable $f(z_1,\ldots,z_m)$ en un valor c real, preserva la estabilidad real. Es decir, $f(z_1,\ldots,z_{i-1},c,z_{i+1},\ldots,z_m)$ es real estable.

Demostración: La primera parte del teorema es inmediata razonando por reducción al absurdo. Dado un polinomio estable $f(z_1,\ldots,z_m)$, si el polinomio en m-1 variables $f(z_1,\ldots,z_{i-1},c,z_{i+1},\ldots,z_m)$ tuviera una raíz en un punto con todas las coordenadas con parte imaginaria mayor que 0, el polinomio original f se anularía en el punto cuyas m coordenadas son las m-1 coordenadas de este más la coordenada c en el lugar de la variable i-ésima, en contra de la hipótesis de que f es estable.

Veamos ahora la segunda parte del teorema. Sea ahora $f(z_1,\ldots,z_m)$ un polinomio real estable. Claramente esta transformación de evaluación en un número real, conserva el hecho de que todos los coeficientes sean reales.

Resta por tanto demostrar que se preserva la estabilidad en este caso. Para ello definimos para todo k natural, definimos los polinomios f_k en m-1 variables, resultado de evaluar la variable i-ésima en $c+\frac{1}{k}$, $f_k(z_1,...,z_{i-1},z_{i+1},...,z_m) = f(z_1,...,z_{i-1},c+\frac{1}{k},z_{i+1},...,z_m)$.

Por tener $c+\frac{1}{k}$ parte imaginaria estrictamente positiva, los polinomios f_k son estables, luego, aplicando el Lema 3.2, para todo $a \in \mathbb{R}^{m-1}$ y para todo $v \in (0,\infty)^{m-1}$ se tiene que los polinomios $f_k(a + vt)$ son estables. Dado que estos son univariantes, no tienen raíces en el semiplano complejo superior $D = \{t \in \mathbb{C}$ tal que $\mathrm{Im}(t) > 0\}$, que es un dominio.

Además la sucesión $\{f_k(a + vt)\}_{k=1}^{\infty}$ converge a $f(a + vt)$ en los compactos de D. Por el Teorema de Hurwitz, $f(a + vt)$ no tiene raices en D, luego es estable. Como $f(a + vt)$ es estable para a y v cualquiera, aplicando el Lema 3.2 (la implicación recíproca a la que usamos en el párrafo anterior), f es estable. □

El último resultado de estabilidad que necesitaremos es otra transformación que preserva la estabilidad real.

Lema 3.6 *Si $f(z_1,...,z_m)$ es un polinomio estable (respectivamente, real estable), entonces para todo $i = 1, ..., m$:*

$$(1 - \partial_i)f(z_1, ..., z_m)$$

es un polinomio estable (respectivamente, real estable).

Por ∂_i denotamos a la derivada parcial i-ésima.

Demostración: Esta transformación conserva el hecho de que los coeficientes sean reales, luego si conserva la estabilidad, también conservará la estabilidad real. Por tanto, basta con probar que conserva la estabilidad.

Sea $f(z_1,...,z_m)$ un polinomio estable. Fijamos $z_1,...,z_{i-1},z_{i+1},...,z_m$ cualquiera con parte imaginaria estrictamente mayor que 0. El polinomio de una variable $g(z_i) = f(z_1,...,z_m)$ es estable por la aplicación recurrente del Lema 3.5 al fijar las variables $z_1,...,z_{i-1},z_{i+1},...,z_m$. Además, $(g - g')(z_i) = (1 - \partial_i)\ f(z_1,...,z_m)$, luego el problema queda resuelto si g - g' no tiene raíces con parte imaginaria positiva.

Llamamos $r_1,...,r_d$ a las raíces de g. Si g y g - g' tuvieran alguna raíz común r_j, tendríamos $g(r_j) = g'(r_j) = 0$ y por tanto, g tendría un cero de orden $k \geq 2$. Dividiendo q por $(z - r_j)^k$, tenemos un polinomio h sin raíces comunes con h - h' y con las mismas raíces (sin contar el orden de estas) que g. Por tanto sin pérdida de generalidad, suponemos que g y g - g' no tienen raíces comunes. Con esta suposición, si z es raíz de g - g':

$$1 = \frac{q'(z)}{q(z)} = \sum_{j=1}^{d} \frac{1}{z - r_j} = \sum_{j=1}^{d} \frac{\overline{z - r_j}}{|z - r_j|^2}$$

Llamamos $\zeta = \sum_{j=1}^{d} |z-r_j|^{-2} > 0$. Por un lado, por definición, $1 = \dfrac{1}{\zeta}\sum_{j=1}^{d} |z-r_j|^{-2}$ y por otro lado, tomando el conjugado complejo en la igualdad anterior:

$$1 = \sum_{j=1}^{d} \frac{z-r_j}{|z-r_j|^2} = z\zeta - \sum_{j=1}^{d} r_j|z-r_j|^2$$

De las dos últimas ecuaciones, despejando z:

$$z = \frac{1}{\zeta}\sum_{j=1}^{d} |z-r_j|^{-2}\left(r_i + \frac{1}{\zeta}\right)$$

vemos que z es una combinación lineal convexa de las raíces de g desplazadas $\dfrac{1}{\zeta}$ en la dirección del eje real. Como g es estable, sus raíces están contenidas en el conjunto $\{$t $\in \mathbb{C}$ tal que Im(t) $\leq 0\}$, que es invariante ante traslaciones en la dirección del eje real y es convexo, luego z pertenece a este conjunto, con lo que queda probado el resultado. $\square$

A continuación buscamos llevar la esperanza del polinomio característico $\chi\left(\sum_{i=1}^{m} v_i v_i^*\right)$ a una expresión que nos permita aplicar estos resultados de estabilidad real. Para ello exponemos a continuación unos hechos de álgebra lineal que relacionan la esperanza con transformaciones que hemos visto que preservan la estabilidad real.

Lema 3.7 *Si A es una matriz compleja invertible de tamaño d x d, y u y v* $\in \mathbb{C}^d$ *son vectores, se tiene:*

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

Demostración: Expresando I + uv* como producto de tres matrices y utilizando el hecho de que el determinante de un producto es el producto de determinantes, tenemos:

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

$$= det\begin{pmatrix} I & u \\ 0 & 1 + v^*u \end{pmatrix} = 1 + v^*u$$

Y aplicando lo anterior, con vectores A$^{-1}u, v^*$, se sigue el resultado:

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

$\square$

El siguiente teorema involucra las derivadas parciales de determinantes.

Teorema 3.8 (de Jacobi) *Si A es una matriz compleja invertible y B es una matriz compleja del mismo tamaño que A, se tiene:*

$$\partial_t det(A + tB)|_{t=0} = det(A)Tr(A^{-1}B)$$

Hemos denotado por $|_{t=0}$ la evaluación de la función en 0.

Notas sobre la demostración del Teorema 3.8: Sin entrar en detalles, la demostración se obtiene a partir del Lema 3.7 anterior y dos propiedades de la traza:

- 1. Tr(AB) $\geq$ 0 para cualquier matriz A y B semidefinidas positivas del mismo tamaño.
- 2. Tr(AB) = Tr(BA) para cualquier matriz A de tamaño k x n y B de tamaño n x k. En general se tiene que la traza del producto de n matrices (de dimensiones adecuadas para que este producto esté bien definido) es invariante ante permutaciones cíclicas de estas.

El último lema de álgebra lineal que necesitamos se deriva de los resultados anteriores.

Lema 3.9 *Sea A una matriz compleja de tamaño d x d, y v un vector aleatorio en $\mathbb{C}^d$, entonces:*

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

Demostración: Primeramente, demostramos el lema en el caso en que A es invertible. Partiendo del lado izquierdo de la ecuación y aplicando el Lema 3.7, llegamos a:

$$\mathbf{E}det(A - vv^*) = \mathbf{E}det(A)(1 - v^*A^{-1}v) = \mathbf{E}det(A)(1 - Tr(A^{-1}vv^*)) =$$

$$= det(A) - det(A)\mathbf{E}Tr(A^{-1}vv^*)) = det(A) - det(A)Tr(A^{-1}\mathbf{E}vv^*))$$

En la segunda igualdad hemos usado, que por ser v*A^{-1}v una matriz de tamaño 1 x 1, es igual a su propia traza; y la propiedad de la traza de ser invariante ante permutaciones cíclicas. La última igualdad se obtiene de la linealidad de la esperanza.

Ahora, partiendo del lado derecho de la ecuación, y aplicando el Teorema de Jacobi (3.8), llegamos la misma expresión:

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

Para probar el lema en el caso general, si A no es invertible, tomamos una sucesión de matrices invertibles que tienda a A. Para todas las matrices de la sucesión se cumple la igualdad del teorema. Como la expresión en ambos lados de esta es un polinomio en los elementos de matriz, por continuidad concluimos que el teorema también se cumple para A. □

Ahora estamos en condiciones de dar una expresión para la esperanza de del polinomio característico $\chi\left(\sum_{i=1}^m v_i v_i^*\right)$ adecuada para probar que todas sus raíces son reales.

Teorema 3.10 (del polinomio característico mixto) *Sean v_1, ..., v_m vectores aleatorios independientes en $\mathbb{C}^d$ con soporte finito. Denotamos $A_i \equiv \mathbf{E} v_i v_i^*$ para todo i. Entonces:*

$$\mathbf{E}\chi\left(\sum_{i=1}^m v_i v_i^*\right) = \left(\prod_{i=1}^m (1 - \partial_{z_i})\right) det\left(xI + \sum_{i=1}^m z_i A_i\right)\Bigg|_{z_1=...=z_m=0}$$

Vemos que la esperanza es igual a un polinomio de una variable que es función de las matrices de rango 1 A_1, ..., A_m. Llamamos a este polinomio el polinomio característico mixto de A_1, ..., A_m, $\mu\left[A_1, ..., A_m\right]$

Demostración: Demostraremos que si v_1, ..., v_m vectores aleatorios independientes en $\mathbb{C}^d$ con soporte finito se tiene la igualdad:

$$\mathbf{E}det\left(M - \sum_{i=1}^m v_i v_i^*\right) = \left(\prod_{i=1}^m (1 - \partial_{z_i})\right) det\left(M + \sum_{i=1}^m z_i A_i\right)\Bigg|_{z_1=...=z_m=0}$$

para toda matriz M de tamaño d x d. La afirmación del teorema es que se tiene esta igualdad en el caso particular en que M = xI.

Probaremos la igualdad por inducción sobre el número de vectores aleatorios k. El caso k=0 es trivial. Suponemos a continuación que la igualdad se cumple para k-1 vectores y veamos que se cumple para k:

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

La primera igualdad se da por la hipótesis de independencia de los vectores aleatorios v_1,..,v_m. Al ser la expresión de la esperanza de $\chi\left(\sum_{i=1}^m v_i v_i^*\right)$ que proporciona este teorema necesaria para demostrar que tiene todas sus raíces reales, esta hipótesis de independencia es imprescindible en el Teorema 2.10.

La segunda igualdad se obtiene aplicando el Lema 3.9 a la esperanza en una variable z_k. La tercera igualdad se obtiene por linealidad. Por último, la siguiente igualdad se obtiene aplicando la hipótesis de inducción a k-1 vectores. $\square$

Ahora es sencillo ver que el polinomio característico mixto tiene todas sus raíces reales.

Teorema 3.11 (de las raíces reales) *Sean v_1, ..., v_m vectores aleatorios independientes en $\mathbb{C}^d$ con soporte finito. Entonces el polinomio característico mixto de A_1, ..., A_m tiene todas sus raíces reales.*

Demostración: Para todo i = 1,..., m; la matriz $A_i \equiv \mathbf{E} v_i v_i^*$ es semidefinida positiva, dado que la matriz $v_i {v_i}^*$ es semidefinida positiva para cualquier valor posible de la variable v_i, como mostramos en la introducción del capítulo; y la esperanza es una combinacion lineal convexa estas matrices semidefinidas positivas. Además, xI es una matriz hermítica. Entonces, aplicando la Proposición 3.3, el polinomio multivariante:

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

es real estable.

Ahora, el polinomio característico mixto se obtiene de la aplicación a esta matriz del operador $(1 - \partial_{z_i})$ y la evaluación de la variable i-ésima en $z_i = 0$ para todo i, sucesivamente. Por el Lema 3.6 y el Lema 3.5, respectivamente, sabemos que estas transformaciones preservan la estabilidad real, luego el el polinomio característico mixto es real estable, y al ser un polinomio en una variable, tiene todas sus raíces reales. $\square$

Ahora que sabemos que el polinomio característico mixto tiene todas sus raíces reales, sólo necesitamos, para demostrar el Teorema 2.10, acotar sus raíces y demostrar que los posibles valor del polinomio $\chi\left(\sum_{i=1}^m v_i v_i^*\right)$ forman una familia entrelazada.

En el artículo de la prueba original de la Conjetura de Kadison-Singer [6], se encuentra la siguiente cota para las raíces del polinomio característico mixto, utilizando una generalización del argumento de la función de barrera establecido en [2] por Batson, Spielman, and Srivastava, que exponemos sin demostración.

Teorema 3.12 *Sea ϵ >0. Sean A_1, ..., A_m matrices hermíticas semidefinidas positivas tal que $Tr(A_i) \leq \epsilon$ para todo i y tal que $\sum_{i=1}^m A_i = I$. Entonces la raíz más grande* *del polinomio característico mixto de A_1,...,A_m es a lo sumo $(1+\sqrt{\epsilon})^2$, es decir:*

$$\lambda_1 \left(\mu \left[A_1, ..., A_m\right](x)\right) \leq (1+\sqrt{\epsilon})^2$$

Para probar que los posibles valor del polinomio $\chi\left(\sum_{i=1}^{m} v_i v_i^*\right)$ forman una familia entrelazada utilizaremos el hecho de que su esperanza tiene todas sus raíces reales.

Teorema 3.13 (principal de entrelazado) *Sean v_1, ..., v_m vectores aleatorios independientes en $\mathbb{C}^d$ con soporte finito. Entonces, los posibles valor del polinomio $\chi\left(\sum_{i=1}^{m} v_i v_i^*\right)$ forman una familia entrelazada.*

Demostración: Para todo i = 1, ..., m; sea v_i un vector aleatorio en $\mathbb{C}^{l_i}$ que toma los valores $w_{i,1}$, ..., w_{i,l_i} con probabilidad $p_{i,1}$, ..., p_{i,l_i}.

Llamaremos $f_{j_1,...,j_m}$ al polinomio $\chi\left(\sum_{i=1}^{m} v_i v_i^*\right)$ que resulta de la asignación de los vectores aleatorios $v_1 = w_{1,j_1}, ..., v_m = w_{m,j_m}$. Estos son los polinomios que identificamos con las hojas del árbol.

Comenzamos fijando los vértices que identificamos con padres de las hojas como como la esperanza respecto de la variable v_m de los polinomios que tienen la misma asignación de los m-1 primeros vectores, los cuales son sus hijos. En cada paso para construir los vértices hasta la raíz tomamos la esperanza respecto de una variable más, siguiendo el orden: v_m, v_{m-1},...v_1; de las hojas que tienen la misma asignación en todas las variables anteriores. Es decir, para todo k = 1,...,m-1; las hojas son:

$$S_{j_1,...,j_k} = \mathbf{E}_{v_{k+1},..v_m}\left(\chi\left(\sum_{i=1}^{k} w_{i,j_i} w_{i,j_i}^* + \sum_{i=k+1}^{m} v_i v_i^*\right)\right).$$

Y la raíz del árbol es el polinomio característico mixto: $S_\emptyset = \mu\left[A_1, ..., A_m\right]$.

Al ser la esperanza una combinación convexa, cada vértice está bien definido, como combinación convexa de sus hijos.

Ahora, para conseguir las condiciones de la definición de familia entrelazada (1.7), nos falta comprobar que para cada k, el conjunto de vértices $\{S_{j_1,...,j_k,t}\}_{t=1}^{l_{k+1}}$ que son hijos de un vértice correspondiente a una asignación parcial $v_1 = w_{1,j_1}, ..., v_k = w_{k,j_k}$ cualquiera, tiene un entrelazado común.

Por el Lema 1.6, la condición para que los polinomios en $\{S_{j_1,...,j_k,t}\}_{t=1}^{l_{k+1}}$ tengan un entrelazado común es equivalente a que cualquier combinación convexa de ellos tenga todas sus raíces reales. Sea $\sum_{t=1}^{l_{k+1}} \rho_t \, S_{j_1,...,j_k,t}$, con $\rho_t \geq 0$ y con $\sum_{t=1}^{l_{k+1}} \rho_t = 1$, una combinación convexa cualquiera.

Ahora, si tomamos un vector aleatorio u_{k+1} que tome cada valor $w_{k+1,t}$ con probabilidad λ_t, entonces:

$$\sum_{t=1}^{l_{k+1}} \rho_t S_{j_1,...,j_k,t} = \mathbf{E}_{u_{k+1},v_{k+2},..v_m}\left(\chi\left(\sum_{i=1}^{k} w_{i,j_i} w_{i,j_i}^* + u_i u_i^* + \sum_{i=k+2}^{m} v_i v_i^*\right)\right).$$

Ahora, el Teorema del polinomio característico mixto nos dice que, para toda matriz M:

$$\mathbf{E}_{u_{k+1},v_{k+2},..v_m} det\left(M - u_i u_i^* - \sum_{i=k+2}^{m} v_i v_i^*\right) =$$

$$= \left(\prod_{i=k+1}^{m}(1-\partial_{z_i})\right) det\left(M + z_{k+1}\mathbf{E}u_i u_i^* + \sum_{i=k+2}^{m} z_i A_i\right)\Bigg|_{z_{k+1}=...=z_m=0}$$

Tomando M $= xI + \sum_{i=1}^{k} w_{i,j_i} w_{i,j_i}^*$, que es una matriz hermítica, y razonando por estabilidad con la expresión de la derecha, como en la demostración del Teorema de las raíces reales, obtenemos el resultado. □

Resaltamos la sencillez con la que hemos probado el Teorema principal de entrelazado, a pesar de que la relativa complejidad de la definición de familia entrelazada dada en 1.7 puede llevar a creer erróneamente que la comprobación de que un conjunto cumple todas sus propiedades será muy difícil en casos de interés.

Ahora estamos en condiciones de probar el Teorema 2.10, con el que completamos la demostración de la Conjetura de Kadison-Singer.

Demostración: del Teorema 2.10.

Sea $\epsilon > 0$. Sean v$_1$, ..., v$_m$ vectores aleatorios independientes en $\mathbb{C}^d$ con soporte finito tal que $\mathbf{E}||$v$_i|| \leq \epsilon$ para todo i y tal que

$$\sum_{i=1}^{m} \mathbf{E} v_i v_i^* = I_d$$

Las matrices A$_i$ son semidefinidas positivas, cumplen que $\sum_{i=1}^{m} A_i = \sum_{i=1}^{m} \mathbf{E} v_i v_i^* =$ I$_d$ y también se tiene que, para todo i, Tr(A$_i$) $= \mathbf{E}||$v$_i|| \leq \epsilon$. Entonces, por el Teorema 3.12,

$$\lambda_1\left(\mu\left[A_1, ..., A_m\right](x)\right) \leq (1+\sqrt{\epsilon})^2$$

Ahora, por el Teorema principal de entrelazado, los posibles valores que puede tomar $\chi\left(\sum_{i=1}^{m} v_i v_i^*\right)$ forman una familia de polinomios entrelazados. Por tanto, existe una asignación de las variables v$_1$ = w$_1$, ..., v$_m$ = w$_m$ tal que

$$\left\|\sum_{i=1}^{m} w_i w_i^*\right\| = \lambda_1\left(\chi\left(\sum_{i=1}^{m} w_i w_i^*\right)\right) \leq \lambda_1\left(\mu\left[A_1, ..., A_m\right]\right) \leq (1+\sqrt{\epsilon})^2$$

En caso de que el polinomio $\chi\left(\sum_{i=1}^{m} w_i w_i^*\right)$ tenga probabilidad 0, contribuirá con coeficiente 0 en el cálculo del polinomio característico mixto. Por tanto, podemos eliminar esta asignación v$_1$ = w$_1$, ..., v$_m$ = w$_m$ sin que varíe el polinomio característico mixto, lo que nos permite asegurar que existe otra asignación de las variables con la misma propiedad. Repetimos el proceso hasta encontrar una asignación con esta propiedad y probabilidad positiva. Entonces

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

y queda probado el resultado. □

Referencias

[1] Akemann, C. A., and Anderson, J. Lyapunov theorems for operator algebras. *Mem. Amer. Math. Soc. 94*, 458 (1991), iv+88.

[2] Batson, J., Spielman, D. A., and Srivastava, N. Twice-Ramanujan sparsifiers. *SIAM Rev. 56*, 2 (2014), 315–334.

[3] Kadison, R. V., and Ringrose, J. R. *Fundamentals of the theory of operator algebras. Vol. I*, vol. 15 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1997. Elementary theory, Reprint of the 1983 original.

[4] Kadison, R. V., and Singer, I. M. Extensions of pure states. *Amer. J. Math. 81* (1959), 383–400.

[5] Marcus, A. W., Spielman, D. A., and Srivastava, N. Interlacing families I: Bipartite Ramanujan graphs of all degrees. *Ann. of Math. (2) 182*, 1 (2015), 307–325.

[6] Marcus, A. W., Spielman, D. A., and Srivastava, N. Interlacing families II: Mixed characteristic polynomials and the Kadison-Singer problem. *Ann. of Math. (2) 182*, 1 (2015), 327–350.

[7] Perales Anaya, D. *Familias de polinomios que se entrelazan*. CIMAT, 2018.

[8] Samperio Valdivieso, A. *Formulacion algebraica de la mecánica cuántica. La Conjetura de Kadison-Singer*. Universidad de Valladolid, 2019.

[9] Sanz Torres, A. *Funciones Zeta en los grafos*. Universidad de Valladolid, 2016.

[10] Stevens, M. *The Kadison-Singer property*, vol. 14 of *SpringerBriefs in Mathematical Physics*. Springer, Cham, 2016. With a foreword by Klaas Landsman.

[11] Weaver, N. The Kadison-Singer problem in discrepancy theory. *Discrete Math. 278*, 1-3 (2004), 227–239.