



Universidad de Valladolid

Facultad de Ciencias

TRABAJO FIN DE GRADO

Grado en Física

Formulación algebraica de la mecánica cuántica.

La Conjetura de Kadison-Singer.

Autor: Álvaro Samperio Valdivieso

Tutor/es: Fernando Gómez Cubillo

Índice

Introducción.	1
1. Observables y estados.	3
2. Formulación usual de la mecánica cuántica.	14
3. Formulación algebraica de la mecánica cuántica.	22
4. La Conjetura de Kadison-Singer	28
Referencias.	45

Introducción

El TFG estudia los formalismos matemáticos de la mecánica cuántica y detalla el planteamiento y la prueba de la conjetura de Kadison-Singer, de interés en este contexto. El significado físico de los elementos fundamentales de los formalismos de la teoría cuántica se enmarca en una descripción probabilística, válida también para la teoría clásica. Los ingredientes básicos en esta descripción son “estados” y “observables”, los cuales se expresan matemáticamente mediante espacios de Hilbert y operadores lineales, en base a los postulados propuestos por Dirac y Von Neumann definidos en ellos, y la correspondiente teoría de C^* -álgebras, introducida por Von Neumann y Murray y desarrollada posteriormente por autores como Gelfand, Naimark, Segal...entre otros. La conjetura de Kadison-Singer, formulada en 1959, afirma (sucintamente) que cada estado puro definido en la C^* -álgebra de operadores diagonales sobre un espacio de Hilbert separable tiene una única extensión a la C^* -álgebra de todos los operadores de dicho espacio. Los especialistas siempre han creído que la conjetura no era cierta. Resultados previos de von Neumann permiten tratar el problema en tres casos diferentes denominados “continuo”, “mixto” y “discreto”. Kadison y Singer probaron que la respuesta es negativa en los dos primeros. El caso discreto ha sido resuelto positivamente por Marcus, Spielman y Srivastava en 2013 mediante técnicas algebraicas que involucran polinomios entrelazados. La reducción de la conjetura de Kadison-Singer a enunciados equivalentes en espacios de Hilbert de dimensión finita, como la Conjetura de Pavimentación de Anderson (1991) o la conjetura KS_r de Weaver (2004), juega un papel fundamental en la demostración. El TFG centra su atención en este caso.

El primer capítulo está dedicado al estudio la descripción probabilística de los formalismos de la mecánica cuántica, en particular a la introducción de los conceptos de estado y observable y sus propiedades. El segundo capítulo presenta la formulación usual de la mecánica cuántica en términos de espacios de Hilbert y operadores lineales definidos en ellos. El tercer capítulo introduce la formulación algebraica de la mecánica cuántica en términos de C^* -álgebras y discute su conveniencia sobre la formulación anterior. Por último, el cuarto capítulo está dedicado al planteamiento de la Conjetura de Kadison-Singer y, en el caso discreto, se presenta su equivalencia con la Conjetura de Pavimentación de Anderson y la conjetura KS_r de Weaver y las técnicas algebraicas utilizadas por Marcus, Spelman y Srivastava para probar que la conjetura es cierta en este caso.

1. Observables y estados.

En este primer capítulo detallaremos una descripción de la mecánica cuántica, para lo cual empezaremos estableciendo propiedades de los estados y observables, conceptos físicos elementales que serán la base de las dos descripciones de la mecánica cuántica, que desarrollaremos en los capítulos siguientes.

Siguiendo el criterio de Araki en [2], asumimos que en cualquier proceso de medida físico, están involucrados los siguientes cuatro elementos:

- 1. El sistema físico, la porción del universo, sobre el que estamos midiendo una determinada propiedad.
- 2. El instrumento de medida de la propiedad que queremos determinar.
- 3. El observador que realiza el experimento.
- 4. El entorno, la porción de universo que no está incluida en ninguno de los elementos anteriores.

En nuestra descripción del sistema total nos restringiremos a un caso ideal en el que no tendremos en cuenta el efecto de los elementos 3 y 4.

Denotamos los sistemas físicos con las etiquetas $\alpha_1, \alpha_2, \dots$, donde cada etiqueta distingue un sistema preparado de una manera específica.

Denotamos los instrumentos de medida con letras mayúsculas Q, Q', Q_1, Q_2, \dots , donde cada etiqueta distingue un instrumento de medida con una preparación y una manera de medir específicas.

Admitimos que podemos asignar un carácter numérico al resultado de cualquier medida. El resultado de una medida será un elemento de \mathbb{R}^n . Lo denotamos con una letra minúscula: q, q', q_1, q_2, \dots

En mecánica cuántica, una medida con un instrumento concreto Q en un sistema concreto α puede dar lugar a diferentes resultados. En cada caso, el conjunto de todos los resultados posibles Ω_α^Q es un conjunto finito o infinito, con una probabilidad asociada discreta o continua de ocurrencia w_α^Q .

En el caso de que el conjunto de resultados Ω_α^Q sea discreto, la probabilidad w_α^Q se determina con las probabilidades puntuales $w_\alpha^Q(q)$, donde w_α^Q es una función que asigna a cada resultado q un número real no negativo de tal manera que la suma de probabilidades de todos los resultados sea 1:

$$\sum_q w_\alpha^Q(q) = 1$$

La probabilidad de obtener al medir un resultado q de un subconjunto de resultados $S \subset \Omega_\alpha^Q$ es la suma de las probabilidades puntuales de los resultados que pertenecen a S . Debido a que la suma de todas las probabilidades puntuales es 1, la probabilidad del conjunto total de resultados es 1 (obtener algún resultado en la medida es un suceso seguro).

En el caso de que el conjunto de resultados Ω_α^Q sea continuo, la probabilidad asociada no se determina a partir de las probabilidades puntuales de cada resultado q . En su lugar, la probabilidad es una función definida sobre un conjunto determinado de subconjuntos de Ω_α^Q . La probabilidad de un subconjunto de resultados S es

$$w_\alpha^Q(S) = \int_S f(q) dq$$

donde f es la función densidad de la probabilidad continua.

Al igual que en el caso discreto, la probabilidad de cualquier subconjunto debe ser no negativa, la probabilidad del conjunto total debe ser 1, y la probabilidad de una unión disjunta de subconjuntos debe ser igual a la suma de las probabilidades de cada subconjunto.

En lo sucesivo, únicamente escribiremos los resultados y definiciones en el caso discreto. Los resultados y definiciones para el caso continuo son análogos y se pueden obtener reemplazando las sumas por integrales, y las probabilidades puntuales por probabilidades de subconjuntos.

El caso de una medida en un sistema clásico concreto, en el que cada medida da como resultado un único valor posible q' , se describe como caso un caso particular de una medida en un sistema cuántico en el que todas las probabilidades $w_\alpha^Q(q)$ son nulas, salvo en el caso $q=q'$, que se tiene $w_\alpha^Q(q')=1$.

La asunción de la naturaleza probabilista de la medida tiene justificación experimental. Si realizamos N medidas en condiciones idénticas y obtenemos n_q veces el valor q , vemos que el cociente entre n_q y N tiende a un valor fijo cuando el número de experimentos tiende a infinito, el cual, por la Ley de los Grandes Números es igual a la probabilidad de obtener el valor q :

$$\lim_{N \rightarrow \infty} \frac{n_q}{N} = w_\alpha^Q(q)$$

Ahora, fijando un conjunto de sistemas y un conjunto de instrumentos de medida, el concepto de "estado" surge al identificar entre sí los sistemas que son indistinguibles utilizando esos instrumentos.

De manera formal, definimos en el conjunto de los sistemas físicos una relación de la siguiente manera: α_i está relacionado con α_j ($\alpha_i R \alpha_j$) si para cada instrumento de medida Q , el conjunto de resultados en ambos sistemas y su probabilidad asociada

son idénticos, es decir, si:

$$\begin{cases} \Omega_{\alpha_i}^Q(q) = \Omega_{\alpha_j}^Q(q) \\ w_{\alpha_i}^Q(q) = w_{\alpha_j}^Q(q) \end{cases}$$

para todo observable Q y resultado q .

Es trivial ver que esta relación es una relación de equivalencia, es decir, cumple las siguientes propiedades:

- 1. Reflexiva: $\alpha_i R \alpha_i$ para todo i . Es decir, todo sistema está relacionado consigo mismo.
- 2. Simétrica: si $\alpha_i R \alpha_j$, entonces $\alpha_j R \alpha_i$. Es decir, si un sistema está relacionado con otro, este último también está relacionado con el primero.
- 3. Transitiva: si $\alpha_i R \alpha_j$ y $\alpha_j R \alpha_k$, entonces $\alpha_i R \alpha_k$. Es decir, si un primer sistema está relacionado con un segundo sistema, y este segundo está relacionado con otro tercer sistema, el primer sistema está relacionado con el tercero.

Esta relación, al ser de equivalencia, induce una partición del conjunto de sistemas en clases de equivalencia, que llamamos estados. Con esta definición, dos sistemas están en el mismo estado si pertenecen a la misma clase de equivalencia, es decir, si la medida de todos los instrumentos da el mismo conjunto de resultados con la misma probabilidad.

Como cada estado α queda determinado unívocamente por las probabilidades puntuales $w_{\alpha}^Q(q)$ para todo Q y q , éstas pueden ser vistas como unas coordenadas del estado.

Análogamente, definimos una relación de equivalencia en el conjunto de resultados que identifique los instrumentos Q_i, Q_j , que al medir dan los mismos resultados con la misma probabilidad para todos los estados ($w_{\alpha_i}^Q(q) = w_{\alpha_j}^Q(q)$ para todo estado α y resultado q).

Esta relación también es de equivalencia, y estas clases de equivalencia en el conjunto de los instrumentos son llamadas observables.

Por otra parte, reflexionemos sobre los resultados de una medida. Hemos admitido que el resultado de una medida siempre se puede expresar como un número real o un elemento de \mathbb{R}^n y que al fijar un observable, fijamos la manera de etiquetar los resultados. Sin embargo, esto se puede hacer de diferentes maneras, a veces equivalentes, en el sentido de que hay un isomorfismo entre dos conjuntos de resultados y la probabilidad definida en uno de ellos es igual a la probabilidad definida en el otro compuesta con ese isomorfismo.

Un ejemplo hipotético de lo anterior es un termómetro de mercurio que mide la temperatura de un sistema en un estado que puede tener un valor de $0\text{ }^{\circ}\text{C}$ con una probabilidad de 0.3 y un valor de $100\text{ }^{\circ}\text{C}$ una probabilidad de 0.7 (que diferenciamos por la altura del mercurio) podría ser descrito equivalentemente como un que puede tener un valor de 273.15 K con una probabilidad de 0.3 y un valor de 373.15 K con una probabilidad de 0.7, con un isomorfismo entre las etiquetas de los resultados $\{0, 100\}$ y $\{273.15, 373.15\}$.

Esto motiva la definición de función de un observable. Sea α un estado y sea Q un observable, cuyo conjunto de resultados posibles es $\Omega_{\alpha}^Q \subset \mathbb{R}^n$ con una probabilidad asociada w_{α}^Q . Dada cualquier función medible f definida en un subconjunto de \mathbb{R}^n que contenga a Ω_{α}^Q y con imagen en \mathbb{R}^m , definimos la función del observable Q , $f(Q)$, como el observable que tiene por resultados el conjunto imagen de Ω_{α}^Q , $f(\Omega_{\alpha}^Q)$, con una probabilidad $w_{\alpha}^{f(Q)} = w_{\alpha}^Q \circ f^{-1}$.

Por función medible (con respecto a la medida dada por la probabilidad) entendemos una función tal que para cualquier subconjunto $S \subset \mathbb{R}^m$, su contraimagen por f , $f^{-1}(S)$ tenga una probabilidad asociada. En la definición anterior, exigimos la medibilidad de f con la intención de excluir las funciones f que no hagan posible que la probabilidad en la imagen de f , $w_{\alpha}^{f(Q)}$, se pueda definir correctamente. Esta condición es superflua en el caso discreto, pues cualquier subconjunto de resultados tiene una probabilidad asociada.

En el caso particular que discutíamos en el ejemplo anterior de que f sea un isomorfismo, es decir, una correspondencia uno a uno entre los resultados, $f(Q)$ resulta ser el mismo observable con una re-etiqueta de los resultados.

Conociendo la función f , las coordenadas de $f(Q)$ quedan determinadas a partir de las de Q , de la siguiente manera:

$$w_{\alpha}^{f(Q)}(q') = \sum_{q: f(q)=q'} w_{\alpha}^Q(q)$$

Un postulado de la mecánica cuántica es la transformación de cualquier estado en otro estado concreto al realizar una medida (es decir, con la actuación de un observable sobre un estado), lo cual lleva a la imposibilidad de medir simultáneamente ciertas parejas de observables, que llamamos incompatibles. Llamamos compatibles a los observables que sí se pueden medir simultáneamente.

Un conjunto finito de observables Q_1, Q_2, \dots, Q_r es compatible si todos son función de un único operador Q , es decir, si existen funciones f_1, f_2, \dots, f_r tal que $Q_i = f_i(Q)$ para todo i .

El hecho de que sean funciones de un mismo observable Q garantiza que podemos medir los resultados y probabilidades de Q , y a partir de esto, podemos determinar

a la vez los resultados y probabilidades de Q_1, Q_2, \dots, Q_r si conocemos las funciones f_1, f_2, \dots, f_r .

Estamos interesados en conocer, dado un subconjunto finito de observables compatibles del conjunto de observables que consideramos en un proceso de medida, cuál es el subconjunto maximal de observables compatibles que podemos obtener que contenga al subconjunto, es decir, hasta dónde podemos llegar añadiendo a este subconjunto observables compatibles con todos los del subconjunto. Esto motiva la siguiente definición:

Decimos que un conjunto (no necesariamente finito) de observables es un Conjunto Completo de Observables Compatibles (CCOC) si todos sus subconjuntos finitos son de observables compatibles y es maximal respecto a la inclusión de conjuntos (es decir, no existe un conjunto con esta propiedad que lo contenga salvo él mismo).

Un conjunto C es un CCOC si y solo si cumple las tres propiedades siguientes:

- 1. Para cualquier número finito de observables Q_1, Q_2, \dots, Q_r de C, existe un observable Q en C tal que Q_1, Q_2, \dots, Q_r son función de Q.
- 2. Cualquier función de un observable de C pertenece a C.
- 3. Si un número finito de observables Q_1, Q_2, \dots, Q_r de C son función de dos observables Q y Q' pertenecientes a C, entonces la distribución de probabilidad conjunta de los resultados q_i de Q_i obtenida a través de la medida de Q es igual a la obtenida a través de la medida de Q'. Escribiendo $Q_i = f_i(Q) = f'_i(Q')$, esta propiedad se traduce en:

$$w_\alpha^{(f_1(Q), f_2(Q), \dots, f_r(Q))}(q_1, q_2, \dots, q_r) = w_\alpha^{(f'_1(Q'), f'_2(Q'), \dots, f'_r(Q'))}(q_1, q_2, \dots, q_r)$$

para todo estado α .

Un caso particular de CCOC es el conjunto de todas las funciones medibles de un observable Q, el conjunto $\{f(Q) \text{ tal que } f \text{ es medible}\}$.

Ahora veamos cómo podemos aprovechar estos conceptos de función de un observable para definir unas coordenadas alternativas que caractericen a un estado α de manera equivalente a las coordenadas consistentes en las probabilidades puntuales de cada resultado q de cada observable Q, $w_\alpha^Q(q)$.

Definimos la esperanza $\alpha(Q)$ del observable Q con conjunto de resultados Ω_α^Q en el estado α como

$$\alpha(Q) = \sum_{q \in \Omega_\alpha^Q} qw_\alpha^Q(q)$$

De la definición, se sigue trivialmente que la esperanza es lineal, es decir, dados Q_1, Q_2, \dots, Q_r , la esperanza de cualquier función de observable $f(Q_1, Q_2, \dots, Q_r) = c_1 Q_1 +$

$c_2Q_2 + \dots + c_rQ_r$ proveniente de una función f real lineal (con c_1, c_2, \dots, c_r constantes reales) es:

$$\alpha(c_1Q_1 + c_2Q_2 + \dots + c_rQ_r) = c_1\alpha(Q_1) + c_2\alpha(Q_2) + \dots + c_r\alpha(Q_r)$$

Además, la probabilidad un subconjunto $S \subset \Omega_\alpha^Q$ se obtiene como la esperanza del operador $\chi_S(Q)$, donde

$$\chi_S(q) = \begin{cases} 1 & \text{si } q \in S \\ 0 & \text{si } q \notin S \end{cases}$$

Trivialmente, χ_S es una función medible. Por tanto, si en el conjunto de observables incluimos el CCOC de funciones de Q para cada observable Q , fijando la esperanza de todos los observables podemos determinar en particular todas las coordenadas de Q del estado α :

$$w_\alpha^Q(q) = \alpha(\chi_{\{q\}}(Q))$$

Por tanto, para determinar un estado α , las coordenadas $\alpha(Q)$ para todo observable Q tienen la misma información que las coordenadas $w_\alpha^Q(q)$. En particular, si para todo observable Q se tiene $\alpha_1(Q) = \alpha_2(Q)$, entonces $\alpha_1 = \alpha_2$.

Esta equivalencia entre las dos coordenadas nos permitirá elegir en lo sucesivo cualquiera de ellas para etiquetar un estado. El uso de las coordenadas de la esperanza matemática es más conveniente para realizar una formulación detallada de la mecánica cuántica que el de las coordenadas dadas por las probabilidades puntuales, como veremos en las secciones restantes del capítulo.

Hasta ahora hemos descrito un sistema físico a partir de un conjunto de observables y un conjunto de estados, pero sin dotar de una estructura concreta a éste último. En esta última parte del capítulo daremos al conjunto de estados una estructura de subconjunto de espacio vectorial topológico, lo que nos permitirá definir con precisión cuáles son en nuestro sistema los estados puros y cuáles son mezclas estadísticas (introducimos estos conceptos a continuación).

Fijando un conjunto de observables etiquetados por Q , con resultados de medida q , decimos que un estado α , de coordenadas $w_\alpha^Q(q)$ es una mezcla estadística de los n estados $\alpha_1, \alpha_2, \dots, \alpha_n$ de coordenadas $w_{\alpha_1}^Q(q), w_{\alpha_2}^Q(q), \dots, w_{\alpha_n}^Q(q)$ si las coordenadas de α son una combinación lineal convexa de las coordenadas de $\alpha_1, \alpha_2, \dots, \alpha_n$. Es decir, si existen $\lambda_1, \lambda_2, \dots, \lambda_n \geq 0$ con $\lambda_1 + \lambda_2 + \dots + \lambda_n = 1$ tal que

$$w_\alpha^Q(q) = \lambda_1 w_{\alpha_1}^Q(q) + \lambda_2 w_{\alpha_2}^Q(q) + \dots + \lambda_n w_{\alpha_n}^Q(q)$$

para todo Q y todo q .

Esta definición es equivalente en términos de las coordenadas de esperanza matemática, sustituyendo la ecuación anterior por:

$$\alpha(Q) = \lambda_1\alpha_1(Q) + \lambda_2\alpha_2(Q) + \dots + \lambda_n\alpha_n(Q)$$

para todo Q .

La correspondencia uno a uno entre estados y coordenadas nos permite ver a un estado α que es mezcla estadística de los n estados $\alpha_1, \alpha_2, \dots, \alpha_n$ como una combinación lineal convexa de estos en el espacio de estados del sistema. Escribimos:

$$\alpha = \lambda_1\alpha_1 + \lambda_2\alpha_2 + \dots + \lambda_n\alpha_n$$

con $\lambda_1, \lambda_2, \dots, \lambda_n \geq 0$ y $\lambda_1 + \lambda_2 + \dots + \lambda_n = 1$.

Con esta definición, una mezcla estadística es un estado que tiene una probabilidad λ_i de estar en el estado α_i para todo i (la propiedad $\lambda_1 + \lambda_2 + \dots + \lambda_n = 1$ garantiza que las probabilidades λ_i estén bien definidas).

Si podemos preparar los estados $\alpha_1, \alpha_2, \dots, \alpha_n$, la mezcla estadística $\alpha = \lambda_1\alpha_1 + \lambda_2\alpha_2 + \dots + \lambda_n\alpha_n$ se puede preparar de manera que la medida en α consiste en elegir un número i con probabilidad λ_i en $\{1, 2, \dots, n\}$ y a continuación realizar una medida en el estado α_i .

Por tanto podemos exigir que el conjunto de estados, que llamaremos Σ , contenga a todos los estados mezcla de estados de Σ . Es decir, Σ es un conjunto convexo (es cerrado para combinaciones convexas de elementos de Σ).

El concepto de mezcla estadística nos permite ver a un estado $\alpha = \lambda_1\alpha_1 + \lambda_2\alpha_2 + \dots + \lambda_n\alpha_n$ como un estado que admite una descomposición en estados más puros que él $\alpha_1, \alpha_2, \dots, \alpha_n$. Nos preguntamos si siempre es posible asegurar la existencia de estados en nuestro conjunto Σ que sea imposible descomponer en estados más puros. Llamamos a estos estados estados puros.

Cualquier mezcla estadística de n estados puede expresarse como mezcla de dos estados. Por ejemplo, la mezcla de tres estados $\alpha = \lambda_1\alpha_1 + \lambda_2\alpha_2 + \lambda_3\alpha_3$ se puede expresar como:

$$\alpha = \lambda_1\alpha_1 + \lambda_2\alpha_2 + \lambda_3\alpha_3 = \alpha = \lambda_1\alpha_1 + (1 - \lambda)\beta$$

donde el estado $\beta \equiv \lambda'_2\alpha_2 + (1 - \lambda'_2)\alpha_3$, con $\lambda'_2 \equiv (1 - \lambda_1)^{-1}\lambda_2$, es un estado del sistema por ser Σ convexo. Además se cumple la condición $(1 - \lambda_1)(1 - \lambda'_2) = \lambda_3$ dado que $\lambda_1 + \lambda_2 + \lambda_3 = 1$.

Aplicando recurrentemente esta identificación de mezcla estadística de tres estados con una mezcla estadística de dos estados, llegamos a que cualquier mezcla estadística de n estados puede expresarse como mezcla de dos estados. Esto nos permite caracterizar de manera sencilla los estados puros.

Un estado $\alpha \in \Sigma$ es puro si es extremal en el conjunto de estados Σ , es decir, si no admite una expresión $\alpha = \lambda\alpha_1 + (1 - \lambda)\alpha_2$ como mezcla estadística de dos estados $\alpha_1, \alpha_2 \in \Sigma$ excepto en los casos triviales:

- 1. $\lambda = 1$, donde necesariamente $\alpha = \alpha_1$
- 2. $\lambda = 0$, donde necesariamente $\alpha = \alpha_2$
- 3. $\alpha_1 = \alpha_2$, donde λ puede tomar cualquier valor

Dado que cualquier mezcla estadística no trivial de un número n de estados puede expresarse como una combinación no trivial de dos estados, un estado puro es aquel que no admite ninguna expresión como mezcla estadística no trivial de estados de Σ .

Un estado mezcla es aquel que no es estado puro.

La cuestión acerca de la existencia o no de estados puros en nuestro sistema es equivalente a la cuestión de si nuestro conjunto de estados Σ tiene o no puntos maximales. La respuesta es que únicamente con la asunción de que Σ es un conjunto convexo, no es suficiente para garantizar su existencia.

El siguiente teorema, que exponemos sin demostración, garantiza la existencia de puntos extremales en un conjunto convexo con hipótesis adicionales.

Teorema de Krein–Milman: un subconjunto convexo y compacto de un espacio vectorial localmente convexo es la envolvente convexa cerrada de sus puntos extremales.

La envolvente convexa de un conjunto S se define como el conjunto formado por todas las combinaciones lineales convexas de S (los elementos de S también están en la envolvente convexa de S , ya que son combinaciones convexas triviales de sí mismos).

La idea fundamental de este teorema es que, bajo sus hipótesis, en un conjunto S siempre podemos encontrar un subconjunto E de puntos extremales de S a partir del cuál podemos "recuperar" el conjunto S , es decir, podemos expresar cualquier elemento de S como límite de combinaciones lineales convexas de elementos de E .

Si admitimos que nuestro conjunto de estados cumple las condiciones del teorema de Krein–Milman, quedaría garantizada la existencia de estados puros en los que se puede descomponer cualquier estado del sistema, ya que si no existieran, el subconjunto de puntos extremales del conjunto de estados sería vacío, y su envoltura convexa sería vacía, en contra del resultado del teorema.

La búsqueda de la compacidad del conjunto de estados Σ nos lleva a definir en él una topología.

El conjunto de pares (Q, q) , donde Q es un observable y q es uno de los posibles resultados de ese observable Q , puede ser finito, infinito numerable o incluso infinito no numerable, correspondiendo cada par a una coordenada $P_\alpha^Q(q)$ de los estados α ,

sin embargo, al realizar un proceso de medida sobre un estado sólo podemos medir un número finito de pares (Q, q) , es decir, sólo podemos determinar un número finito n de coordenadas de un estado α . Llamaremos (Q_i, q_i) con $i = 1, \dots, n$ a los pares (Q, q) que elegimos medir en cada caso

Además, la medida de cada observable sólo se puede realizar una cantidad finita de veces N , tras lo cual obtenemos n_{q_i} veces el valor n_{q_i} . La información que obtenemos de la medida es un valor $w_i = \frac{n_{q_i}}{N}$ que aproxima a la coordenada $w_\alpha^{Q_i}(q_i)$ con un error $\epsilon_i > 0$ que atribuimos a que el número de medidas N es finito y a aspectos del proceso de medida en sí.

Es decir, del proceso de medida obtenemos que nuestro sistema está en un estado cuyas coordenadas correspondientes a los pares (Q, q) son desconocidas si $(Q, q) \neq (Q_i, q_i)$ para todo $i = 1, \dots, n$, y son iguales a w_i en caso contrario, con:

$$|w_i - w_\alpha^{Q_i}(q_i)| < \epsilon_i$$

para todo $i = 1, \dots, n$.

Resulta natural definir la base de entornos abiertos de la topología en Σ como el conjunto de los estados que cumplen la condición anterior, con $N < \infty$ tomando Q_i, q_i, w_i y $\epsilon_i > 0$ todos los posibles valores del sistema.

De esta manera hemos definido en Σ una topología, que se conoce como topología física. Al definirla de esta manera, un entorno de un estado que tiene un número finito de coordenadas iguales a las w_i que hemos obtenido a través de la medida, está formado por los estados α cuyo valor real de las coordenadas $w_\alpha^{Q_i}(q_i)$ es cercano a los valores medidos w_i .

Ahora bien, identificando cada estado con sus coordenadas, podemos ver el conjunto de estados Σ como un subconjunto del espacio vectorial V producto de los números reales por sí mismo tantas veces como número de coordenadas tenga cada estado, o tantas veces como sea el cardinal del conjunto de parejas (Q, q) en el caso infinito.

En general escribimos los elementos de este espacio V como $x = \{x_j\}_{j \in J}$, donde J es un cierto conjunto de etiquetas de cada coordenada, y puede ser no numerable.

Es claro que, con esta identificación, la topología física en Σ coincide con la topología producto en este espacio, ya que una base de entornos abiertos de un punto x para la topología producto es el conjunto de los vectores $y \in V$ para los que existe un número finito n de coordenadas $j_1, j_2, \dots, j_n \in J$ tal que:

$$|x_{j_i} - y_{j_i}| < \epsilon_i$$

para todo $i = 1, 2, \dots, n$, con $\epsilon_1, \epsilon_2, \dots, \epsilon_n \in > 0$.

Además, dado que las coordenadas son probabilidades puntuales, toman valores en el intervalo $[0, 1]$. La identificación de los estados con sus coordenadas permite ver a Σ como subconjunto de:

$$\{\{x_j\}_{j \in J} \in V \text{ tal que } 0 \leq x_j \leq 1 \text{ para todo } j \in J\}$$

Este subconjunto de V es el producto del intervalo $[0, 1]$ una determinada cantidad de veces, y por el Teorema de Tychonoff, es compacto en la topología producto.

Por tanto, podemos ver al espacio de estados Σ como un subconjunto de un conjunto compacto, pero para poder aplicar el Teorema de Krein–Milman a Σ necesitamos que Σ sea compacto.

Al ser Σ subconjunto de un espacio compacto, es compacto si y sólo si es cerrado en la topología física, es decir, si Σ es igual a su adherencia, $\bar{\Sigma}$.

Los elementos de $\bar{\Sigma}$ tienen las propiedades que permiten definirlos correctamente como estados. Por estar en la adherencia de Σ , existirá una red de estados de Σ , con un número finito de coordenadas que convergerán a las coordenadas del estado límite. Las coordenadas del estado límite serán no negativas y sumarán 1, por ser una propiedad de todos los elementos de la red.

Incluir los elementos de $\bar{\Sigma}$ en nuestro conjunto de estados significa aceptar que si tenemos un estado cuyas coordenadas pretendemos determinar cada vez con mayor precisión (ya sea aumentando el número de medidas N , minimizando cualquier contaminación del sistema...), es posible llegar a medir exactamente el valor límite. Esta es una situación ideal, pero es la inclusión de estos estados ideales la que garantiza que al conjunto de estados le podemos aplicar el Teorema de Krein–Milman, asegurando la existencia de estados puros en el sistema y asegurando que cualquier estado mezcla se puede expresar como límite de mezclas estadísticas de estos.

Resulta natural que la existencia de estados puros se introduzca en esta situación ideal, es decir, los estados puros son ideales, ya que debido a factores pertenecientes a la influencia del entorno y el observador, que no es posible eliminar completamente, un estado puro es, en realidad imposible de preparar exactamente.

Ilustremos estas consideraciones sobre topología con un ejemplo sencillo. Supongamos un sistema físico con un único observable Q que puede tomar tres valores distintos q_1, q_2 y q_3 , con probabilidad $w_\alpha^Q(q_1), w_\alpha^Q(q_2)$ y $w_\alpha^Q(q_3)$ en un estado α . Las coordenadas de probabilidades puntuales en un estado α entonces son $(w_\alpha^Q(q_1), w_\alpha^Q(q_2), w_\alpha^Q(q_3))$, lo que nos permite verlo como un elemento de \mathbb{R}^3 .

La condición de que las probabilidades deben sumar 1, nos restringe al plano $w_\alpha^Q(q_1) + w_\alpha^Q(q_2) + w_\alpha^Q(q_3) = 1$, y dado que las probabilidades deben estar en $[0, 1]$, nuestro espacio de estados debe ser un subconjunto acotado de un plano.

Ahora la existencia o no de estados puros depende de si el espacio de estados σ es

cerrado. Si, por ejemplo, σ es el conjunto de estados que forman un triángulo cerrado, sus estados puros son los vértices del triángulo, son sus tres puntos extremales. Además, cualquier punto del interior del triángulo o de sus bordes, puede ser escrito como combinación lineal convexa de esos tres puntos. El espacio de estados se divide de esta manera en tres estados puros y sus estados mezcla, que son las mezclas estadísticas de los tres.

Ahora bien, la condición de ser σ un conjunto cerrado es fundamental. Si por ejemplo eliminamos los estados correspondientes al borde del triángulo, resulta que σ es un triángulo abierto y no tiene puntos extremales, luego el sistema físico no tendría estados puros en esta descripción.

Nótese que esta última descripción del estado es más realista que la anterior, ya que los estados del borde del triángulo son aquellos cuyo porcentaje de alguno de los estados puros de la descripción anterior es exactamente nulo, y corresponden a casos ideales en los que se ha eliminado del todo la contaminación de uno de los estados. Sin embargo es más conveniente la inclusión de estos estados ideales en σ como en la descripción anterior, dado que permite expresar los estados no ideales como combinación de estados puros, que no existen en esta última descripción.

2. Formulación usual de la mecánica cuántica.

En esta sección realizaremos una descripción de un sistema cuántico cualquiera, partiendo de los postulados de la mecánica cuántica de Paul Adrien Maurice Dirac y John Von Neumann que podemos encontrar en [3] y teniendo en cuenta las consideraciones con base experimental que hemos obtenido en el capítulo anterior. Utilizamos varios resultados que se pueden encontrar en [4] o [2].

Como establecimos en el apartado anterior, un sistema físico queda caracterizado por un conjunto de estados y un conjunto de observables que tienen un conjunto de posibles valores medidos por cada uno de ellos. Sólo tratamos el caso discreto, es decir, en el que el conjunto de medidas posibles valores medidos por cada operador es discreto.

Consideramos una descripción en un instante de tiempo fijo.

Según los postulados de Dirac y Von Neumann, cada estado α_Ψ tiene asociado un vector Ψ de un espacio de Hilbert H que lo determina unívocamente. Además este vector tiene norma 1, $\|\Psi\| = 1$.

Por otra parte, un observable Q es un operador en H , $Q: H \rightarrow H$ autoadjunto. El conjunto de sus posibles resultados de medida es el espectro de Q .

Por último, el valor esperado de un observable Q en un estado α_Ψ viene dado por el producto escalar: $\langle Q(\Psi), \Psi \rangle$.

Presentamos a continuación conceptos como norma y producto escalar para llegar a la definición de espacio de Hilbert, operadores autoadjuntos y espectro.

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

Además, si V es completo con la métrica definida por la norma, como $d(x, y) = \|x - y\|$, decimos que V es un espacio de Banach.

Un producto escalar definido en un espacio vectorial V sobre un cuerpo K es una aplicación $\langle \cdot, \cdot \rangle: V \times V \rightarrow \mathbb{R}$ que cumple:

- 1. $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$ para todo $x, y, z \in V$.
- 2. $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle$ para todo $x, y \in V$, para todo $\lambda \in K$.

- 3. $\langle x, y \rangle = \overline{\langle y, x \rangle}$ para todo $x, y \in V$ (donde la línea horizontal superior es la conjugación compleja).
- 4. $\langle x, x \rangle \geq 0$ para todo $x \in V$; y además $\langle x, x \rangle = 0$ si y solo si $x = 0$.

Además un producto escalar cumple la desigualdad de Cauchy-Schwarz: $|\langle x, y \rangle|^2 \leq \langle x, x \rangle \langle y, y \rangle$ para todo $x, y \in V$.

Decimos que un espacio H es de Hilbert si es un espacio con producto escalar que es espacio de Banach con la norma definida por $\|x\| = \langle x, x \rangle^{1/2}$.

Si H es un espacio de Hilbert y $T: H \rightarrow H$ es un operador lineal y continuo, llamamos adjunto de T al único operador T^* tal que:

$$\langle T(x), y \rangle = \langle x, T^*(y) \rangle$$

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

Decimos que $a \in \mathbb{C}$ es un autovalor de un operador Q si existe un elemento $x \in H$ tal que $T(x) = ax$. Llamamos espectro de T al conjunto de todos sus autovalores.

El postulado que afirma que un observable debe ser un operador autoadjunto se basa en que un operador autoadjunto tiene sólo autovalores reales, y el resultado de una medida siempre se identifica con un valor real. En efecto si a es autovalor de un operador autoadjunto T ,

$$a \langle x, x \rangle = \langle T(x), x \rangle = \langle x, T(x) \rangle = \langle x, x \rangle \bar{a}$$

para todo $x \in H$. Luego a es real.

En la descripción establecemos que cualquier función $f(Q)$ de un observable Q , (concepto definido en el capítulo anterior), es un observable del sistema.

En particular, para todo observable Q , la función indicadora de un subconjunto de resultados de medida $S \subset \mathbb{R}$, $\chi_S(q)$, que definimos en el capítulo anterior como

$$\chi_S(q) = \begin{cases} 1 & \text{si } q \in S \\ 0 & \text{si } q \notin S \end{cases}$$

da lugar al observable $\chi_S(Q)$.

Además, dado que $\chi_S(q)^2 = \chi_S(q)$, se tiene $\chi_S(Q)^2 = \chi_S(Q)$. Por tanto, si a es autovalor de $\chi_S(Q)$ para algún vector $x \in H$, $a^2 = a$, ya que $\chi_S(Q)(x) = \chi_S(Q)^2(x) = \chi_S(Q) \chi_S(Q)(x) = \chi_S(Q)(ax) = a^2(x)$.

Por otra parte, en cualquier estado α_Ψ correspondiente a un vector Ψ del espacio de Hilbert H , la esperanza de estos observables $\chi_S(Q)$ es:

$$\alpha_\Psi(\chi_S(Q)) = \langle \chi_S(Q)(\Psi), \Psi \rangle = w_{\alpha_\Psi}^Q(S) \geq 0$$

De estas dos últimas afirmaciones, deducimos que los autovalores de $\chi_S(q)$ son 0 o 1. Decimos que este observable es una "pregunta", puesto que para todo estado correspondiente a un vector del espacio de Hilbert, la medida será 1 ó 0 (que identifiquemos con una respuesta positiva o negativa, respectivamente). El conjunto de todas las preguntas es el mismo que el de todas las proyecciones ortogonales (operadores P tal que $P = P^2 = P^*$). Llamamos a este conjunto $P(H)$.

La suma de dos elementos $\chi_{S_1}(Q)$ y $\chi_{S_2}(Q)$ de $P(H)$ es un elemento de $P(H)$ si y solo si la intersección de los conjuntos de estos resultados de Q , S_1 y S_2 , tiene intersección vacía. En este caso, el observable que resulta de la suma es la proyección sobre la unión de la imagen de las proyecciones $\chi_{S_1}(Q)$ y $\chi_{S_2}(Q)$, que es igual a $\chi_{S_1 \cup S_2}(Q)$. Esto puede verse a partir de la suma de las funciones a partir de las que se definen estos observables, $\chi_{S_1}(q) + \chi_{S_2}(q) = \chi_{S_1 \cup S_2}(q)$.

Esta condición de que dos elementos $\chi_{S_1}(Q)$ y $\chi_{S_2}(Q)$ de $P(H)$ sean proyecciones sobre conjuntos disjuntos también es equivalente a que sean ortogonales entre sí. Es decir, $\chi_{S_1}(Q) \chi_{S_2}(Q) = \chi_{S_2}(Q) \chi_{S_1}(Q) = 0$. El resultado se puede ver teniendo en cuenta que, por ser la suma una proyección, $(\chi_{S_1}(Q) + \chi_{S_2}(Q))^2 = \chi_{S_1}(Q) + \chi_{S_2}(Q)$.

Ahora retomamos el punto de vista del capítulo 1 en el que un estado sobre un conjunto de observables queda determinado por las coordenadas consistentes en la esperanza matemática de estos en el estado.

Dado que la esperanza de un observable $\chi_S(Q)$ de $P(H)$ en un estado α_Ψ es la probabilidad $\alpha_\Psi(\chi_S(Q)) = w_{\alpha_\Psi}^Q(S)$, la definición de estado en $P(H)$ debe ser compatible con las propiedades de esta probabilidad sobre el conjunto de resultados de Q .

Establecemos una primera definición de estado α en $P(H)$ que se corresponde con una probabilidad con aditividad finita para conjuntos disjuntos S_1, \dots, S_n de resultados de Q en α , $w_\alpha^Q(S_1 \cup \dots \cup S_n) = \sum_i w_\alpha^Q(S_i)$, sin exigir por el momento la aditividad para una suma infinita, que asumimos en la definición de w_α^Q en el capítulo 1.

Definición 2.1 *Un estado con aditividad finita en $P(H)$ es una función $\alpha: P(H) \rightarrow [0,1]$ tal que $\alpha(I) = 1$ y tal que si para cualquier número finito de proyecciones $P_1, \dots, P_n \in P(H)$ que sean ortogonales dos a dos (es decir, si $P_i P_j = 0$ para todo $i, j = 1, \dots, n$), se tiene:*

$$\alpha\left(\sum_i^n P_i\right) = \sum_i^n \alpha(P_i)$$

A continuación nos preguntamos cómo podemos extender este concepto de estado a un conjunto de observables que contenga a $P(H)$.

La imagen por un estado de un observable debe ser la esperanza de éste en el estado. Por tanto, un estado debe ser un funcional lineal.

Dado un espacio vectorial V sobre el cuerpo K , decimos que una función $f: V \rightarrow$

K es un funcional lineal en V si $f(\lambda_1x + \lambda_2y) = \lambda_1f(x) + \lambda_2f(y)$ para todo $x, y \in V$, y para todo $\lambda \in K$.

Por otra parte, sólo extenderemos la definición de estado al conjunto $B(H)$ de los operadores acotados en el espacio de Hilbert H .

La norma que utilizamos para un operador lineal y continuo T entre los espacios normados V y W es $\|T\| = \sup\{\|T(x)\|_{(W)} \mid \|x\|_{(V)} \leq 1\} = \sup\{\|T(x)\|_{(W)} \mid \|x\|_{(V)} = 1\}$, 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 el caso en el que V y/o W sea de dimensión finita, consideramos que la norma en ese espacio es la norma euclídea sin indicarlo explícitamente.

La definición de un estado en $B(H)$ es la siguiente.

Definición 2.2 *Dado un espacio de Hilbert H , un estado α en $B(H)$ es un funcional lineal complejo en $B(H)$ que cumple:*

- 1. $\alpha(A) \geq 0$ para todo $A \in B(H)$ positivo.
- 2. $\alpha(I) = 1$.

La primera condición se llama positividad de α . Decimos que un operador $A \in B(H)$ es positivo si su espectro está contenido en $[0, \infty)$, o equivalentemente, si existe un operador $C \in B(H)$ tal que $A = C^*C$. Abusando de la notación, para $A, B \in B(H)$, decimos que $A \geq B$ si $A - B$ es positivo.

La segunda condición da la normalización del estado, $\|\alpha\| = 1$.

A partir de esta definición es trivial ver que la restricción de cualquier estado en $B(H)$ a $P(H) \subset B(H)$ es un estado con aditividad finita.

Ahora, decimos que un estado α' en $B(H)$ es una extensión de un estado con aditividad finita α en $P(H)$ si $\alpha = \alpha'$ en $P(H)$.

El siguiente teorema es un resultado fundamental sobre la unicidad de este tipo de extensiones.

Teorema 2.3 (de Gleason) [4] *Si H es un espacio de Hilbert de dimensión mayor o igual que 3, para cada estado con aditividad finita α en $P(H)$ existe un único estado en $B(H)$ que es una extensión de α .*

A partir de este teorema llegamos a la conclusión de que, en un espacio de dimensión mayor o igual que 3, las definiciones 2.1 y 2.2 son equivalentes para determinar un estado, dado que hay una correspondencia uno a uno entre estados en $P(H)$ y $B(H)$.

Por otro lado, el estudio del caso particular de estados en $P(H)$ que se corresponden con una probabilidad con aditividad numerable, nos lleva a conclusiones muy interesantes.

Definición 2.4 *Un estado α completamente aditivo en $P(H)$ es un estado con aditividad finita en $P(H)$ tal que si para cualquier conjunto de proyecciones $\{P_i\}_{i \in I} \in P(H)$ que sean ortogonales dos a dos, se tiene:*

$$\alpha\left(\sum_i P_i\right) = \sum_i \alpha(P_i)$$

Si un espacio de Hilbert H es separable, cualquier base ortonormal en H es numerable, por lo tanto, cualquier conjunto de proyecciones ortogonales dos a dos también es numerable.

Una base ortonormal en un espacio de Hilbert es un conjunto de vectores ortonormales entre sí que cumple que al añadirle cualquier vector de H , no es posible que el conjunto resultante sea ortonormal.

Entonces los estados en $P(H)$ que se corresponden con una probabilidad con aditividad numerable son exactamente los estados completamente aditivos en $P(H)$ si H es separable. La extensión de uno de estos estados a $B(H)$ da lugar a un estado en $B(H)$ con la siguiente propiedad:

Definición 2.5 *Decimos que un estado α en $B(H)$ es un estado normal si para cualquier sucesión $\{A_n\}_{n=1}^{\infty}$ creciente y acotada de elementos positivos de $B(H)$, se tiene:*

$$\alpha(\sup A_n) = \sup \alpha(A_n)$$

La correspondencia entre estados completamente aditivos y estados normales hace que estas dos propiedades sean equivalentes a la hora de determinar un estado, en analogía con el teorema de Gleason para los estados de las definiciones 2.1 y 2.2.

Teorema 2.6 *Sea H un espacio de Hilbert separable. Si un estado es completamente aditivo en $P(H)$, su extensión a un estado en $B(H)$ es un estado normal. Si un estado en $B(H)$ es normal, su restricción a $P(H)$ es un estado completamente aditivo.*

Los estados normales en un álgebra $B(H)$, con H un espacio de Hilbert separable, admiten una caracterización muy sencilla en términos de una "matriz densidad". Entendemos por matriz densidad a un operador en H positivo y de traza igual a 1.

Se define la traza de un operador en H como la siguiente generalización de la traza de una matriz (de dimensión finita):

$$Tr(A) = \sum_{n=1}^{\infty} \langle Ae_n, e_n \rangle$$

donde $\{e_n\}_{n=1}^{\infty}$ es cualquier base ortonormal de H . La suma que define la traza es convergente y no depende de la elección de la base ortonormal de H .

Teorema 2.7 *Sea H un espacio de Hilbert separable. Entonces, si α es un estado normal en $B(H)$, existe una única matriz densidad ρ_α de manera que, para todo operador $A \in B(H)$,*

$$\alpha(A) = \text{Tr}(\rho_\alpha A)$$

Asimismo, para cualquier matriz densidad ρ , el funcional definido por $\varphi(A) = \text{Tr}(\rho A)$, para todo operador $A \in B(H)$, es un estado normal en $B(H)$.

Veamos ahora cómo encajan los estados que corresponden a un vector del espacio de Hilbert en esta descripción de los estados como funcionales en $B(H)$.

Por los postulados de la mecánica cuántica, si Ψ es un vector de H con norma 1, existe un estado α_Ψ tal que la esperanza matemática de cualquier observable Q en este estado es: $\alpha_\Psi(Q) = \langle Q\Psi, \Psi \rangle$.

Ahora, la ecuación

$$\alpha_\Psi(A) = \langle A\Psi, \Psi \rangle$$

para todo operador $A \in B(H)$, define un estado en $B(H)$, luego la definición de estado establecida en 2.2 incluye en particular a los estados que corresponden a un vector del espacio de Hilbert.

La condición de normalización del estado α_Ψ es consecuencia de que el vector esté normalizado Ψ , $\|\Psi\| = 1$, y la positividad del estado es trivial de la definición.

Además, si definimos la siguiente proyección:

$$P_\Psi(x) = \langle x, \Psi \rangle \Psi$$

para todo $x \in H$, vemos que es la matriz densidad del estado α_Ψ

Claramente, P_Ψ es una proyección de rango 1, es decir, cuya imagen tiene dimensión 1. Además, si tomamos una base ortonormal $\{e_n\}_{n=1}^{\infty}$ que contenga al vector Ψ , la traza de P_Ψ es igual a $\langle \Psi, \Psi \rangle = \|\Psi\|^2 = 1$. Calculando en esta base la traza de $P_\Psi A$, para todo operador $A \in B(H)$:

$$\text{Tr}(P_\Psi A) = \sum_{n=1}^{\infty} \langle P_\Psi e_n, e_n \rangle = \langle A\Psi, \Psi \rangle$$

vemos que, efectivamente, P_Ψ es la matriz densidad de α_Ψ . Luego los estados correspondientes a un vector de H son normales.

Aún podemos decir más de este tipo de estados. Si tomamos un estado normal cualquiera, le corresponde una matriz densidad ρ . Por ser esta un operador positivo,

existe en H una base ortonormal de autovectores de H , correspondientes a autovalores positivos, $\rho e_n = \lambda_n e_n$, con $\lambda_n \geq 0$.

Además, por ser una matriz densidad, su traza es 1, y ésta es igual a la suma de sus autovalores. Ahora calculemos el valor del estado en un operador A :

$$\text{Tr}(\rho A) = \sum_{n=1}^{\infty} \langle \rho A e_n, e_n \rangle = \sum_{n=1}^{\infty} \lambda_n \langle A e_n, e_n \rangle \quad (1)$$

Dado que, para todo n , $\langle A e_n, e_n \rangle$ es un estado correspondiente a un vector unitario del espacio de Hilbert H , y los λ_n son positivos y suman 1, deducimos que cualquier estado normal es límite de combinaciones convexas de estados correspondiente a un vector del espacio de Hilbert H .

En particular, cualquier estado normal puro, admite una expresión como la de la ecuación (1). Por ser un elemento extremal, todos los coeficientes λ_n son 0 menos uno, que es igual a 1, luego, cualquier estado normal puro es igual a un estado correspondiente a un vector unitario del espacio de Hilbert H . De hecho, es sencillo probar la implicación inversa.

Teorema 2.8 *Sea H un espacio de Hilbert separable. Los estados normales puros en $B(H)$ son exactamente los que corresponden a un vector unitario del espacio de Hilbert H .*

Terminamos la sección con una reflexión sobre la definición 2.2 de estado que hemos establecido. A primera vista podría parecer más conveniente considerar como definición de estado la definición 2.5 (de estado normal), dado que estos estados provienen de una probabilidad en el conjunto de los resultados que cumple los axiomas de Kolmogorov (específicamente, que es σ -aditiva), tienen una caracterización interesante en términos de matrices densidad, y una relación general con los estados correspondientes a un vector unitario del espacio de Hilbert H .

Sin embargo, es posible dar una justificación para la elección de la definición 2.2 como definición de estado con un argumento topológico.

Decimos que una sucesión de estados $\{\alpha_n\}_{n=1}^{\infty}$, converge débilmente a un estado α si para cada operador $A \in B(H)$, se tiene

$$\varphi(A) = \lim_{n \rightarrow \infty} \varphi_n(A) \quad (2)$$

La topología débil [5] (que es caracterizada por esta convergencia) es equivalente a la topología física que definimos en el capítulo 1 en el conjunto de estados.

Además, el conjunto de todos los estados que cumplen la definición 2.2 es compacto en esta topología. (En particular, es cerrado, luego, si $\{\alpha_n\}_{n=1}^{\infty}$ es una sucesión de

estados, y el límite de la ecuación (2) existe para todo operador $A \in B(H)$, este límite es un estado en $B(H)$).

Por otra parte, con la topología débil, el conjunto de todos los estados normales es denso en el conjunto de todos los estados. Por tanto, los estados que no son normales pueden ser vistos como un caso límite de estados normales.

Estamos en la situación que discutimos al final del capítulo 1. Aunque los estados normales son aquellos cuya identificación con el estado de un sistema físico surge de una manera más natural, la inclusión en la definición de estado de los estados que no son normales, que pueden ser vistos como un caso ideal (debido que hay sucesiones de estados normales que convergen a ellos), hace que el conjunto de estados sea compacto y que podamos aplicar el Teorema de Krein–Milman.

Siguiendo los argumentos expuestos al final del capítulo 1, es conveniente incluir estos estados en nuestra descripción, lo cuál justifica la definición de estado dada en 2.2.

3. Formulación algebraica de la mecánica cuántica.

En esta sección realizaremos una formulación de la mecánica cuántica alternativa a la anterior, que puede verse como una generalización de ésta, a la que llamamos formulación algebraica. Utilizamos varios resultados que se pueden encontrar en [5] y algún resultado de [2].

La formulación algebraica, a diferencia de la formulación usual, no parte de los postulados de la mecánica cuántica de Dirac y Von Neumann que establecen que los observables de un sistema son operadores autoadjuntos en un espacio de Hilbert, cuyos elementos se corresponden con estados.

En su lugar, partiremos de que los observables en la descripción de un sistema son elementos autoadjuntos (en un sentido que precisaremos más adelante) de una C^* -álgebra.

La estructura de C^* -álgebra permite realizar las operaciones habituales que involucran a los operadores de un conjunto $B(H)$, como el producto, la suma, toma de límites, una involución...

Ahora pasamos a definir varios conceptos relacionados con álgebras.

Definición 3.1 *Un álgebra normada U es un álgebra sobre un cuerpo 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.*

Definición 3.2 *Una C^* -álgebra U es un álgebra compleja de Banach con una involución denotada por $*$, es decir, un automorfismo de álgebras de Banach en U cuyo cuadrado es la identidad:*

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

Con esta definición, vemos que, dado un espacio de Hilbert H , el conjunto $B(H)$, con la involución consistente en tomar el operador adjunto, es un caso particular C^* -álgebra.

Debido a que las C^* -álgebras son una generalización de los conjuntos $B(H)$, llamamos adjunto de T al elemento T^* que es la involución de T , para cualquier elemento T de una C^* -álgebra. Asimismo, decimos que T es autoadjunto si $T = T^*$.

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

Sea U un álgebra de Banach. También, generalizando conceptos definidos en espacios $B(H)$, para $A \in U$, definimos el espectro de A en U , $\text{sp}_U(A)$ como $\text{sp}_U(A) = \{\lambda \in \mathbb{C} \text{ tal que } A - \lambda 1 \text{ no tiene inverso en } U \text{ (por los dos lados)}\}$.

Decimos que un operador A es positivo si $\text{sp}_U(A) \subset [0, \infty]$. 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 definición de estado en una C^* -álgebra es la misma que en la formulación usual (Definición 2.2), sustituyendo el conjunto $B(H)$ por el álgebra de operadores en general.

Definición 3.3 *Dada una C^* -álgebra U , un estado α en U es un funcional lineal complejo en U que cumple:*

- 1. $\alpha(A) \geq 0$ para todo $A \in U$ positivo.
- 2. $\alpha(I) = 1$.

El hecho de haber definido las C^* -álgebras como generalizaciones de los conjuntos de operadores acotados en un espacio de Hilbert $B(H)$ permite definir homomorfismos de una C^* -álgebra en espacios de la forma $B(H)$, que llamaremos representaciones.

El nombre de representación se debe a que, al ser un homomorfismo una aplicación que preserva las operaciones definidas en ambos conjuntos, podemos ver la imagen de una C^* -álgebra como una C^* -álgebra de $B(H)$.

Definición 3.4 *Sea U una C^* -álgebra y sea H un espacio de Hilbert. Decimos que $\pi: U \rightarrow B(H)$ es una representación si es un homomorfismo, es decir, si es una aplicación que cumple:*

- 1. $\pi(cA + dB) = c\pi(A) + d\pi(B)$ para todo A positivo.
- 2. $\pi(AB) = \pi(A)\pi(B)$
- 3. $\pi(A^*) = \pi(A)^*$

para todo $A, B \in U$ y para todo $b, d \in \mathbb{C}$.

Si $\pi(A) = 0$ implica $A = 0$, entonces se dice que esta representación es fiel.

Si π_1 es una representación de U en H_1 y π_2 es una representación de U en H_2 , se dice que π_1 y π_2 son unitariamente equivalentes si existe un operador unitario W de H_1 en H_2 tal que $W \pi_1(A) = \pi_2(A) W$ para todo $A \in U$.

El nombre de representación fiel en el caso en que $\pi(A) = 0$ implica $A = 0$, se debe a que en este caso, la representación es un isomorfismo entre la C^* -álgebra U y su imagen, lo que permite ver a la C^* -álgebra U como C^* -subálgebra de $B(H)$.

Para cualquier representación π , $\|\pi(A)\| \leq \|A\|$, por lo que $\|\pi\| \leq 1$ y las representaciones son funciones continuas. En particular, si π es fiel, $\|\pi(A)\| = \|A\|$, independientemente de π .

La principal motivación para realizar esta formulación algebraica de la mecánica cuántica es que la C^* -álgebra que es el conjunto de operadores de nuestro sistema, puede ser representada de diferentes maneras en función de la situación física.

La construcción de Gelfand–Naimark–Segal (GNS), que será el resultado más importante de esta sección, nos permite asegurar que, dado cualquier estado en una C^* -álgebra, existe una representación en la que ese estado corresponde a un vector unitario de un espacio de Hilbert. Utilizando conjuntamente varias de estas representaciones, obtendremos el Teorema de Gelfand–Naimark, que asegura que cualquier C^* -álgebra tiene una representación fiel, lo cual nos permitirá trabajar con cualquier C^* -álgebra como C^* -subálgebra de un conjunto de operadores acotados en un espacio de Hilbert en el siguiente capítulo.

Teorema 3.5 (Construcción GNS) *Sea U una C^* -álgebra. Para cada estado α en U , existen un espacio de Hilbert H_α , una representación π_α de U en H_α y un vector unitario $x_\alpha \in H_\alpha$ que cumplen:*

- 1. Para todo $A \in U$,

$$\alpha(A) = \langle \pi_\alpha(A)x_\alpha, x_\alpha \rangle$$

- 2. x_α es un vector cíclico de π_α , es decir, el conjunto

$$\pi_\alpha(U)x_\alpha \equiv \{\pi_\alpha(A)x_\alpha \text{ tal que } A \in U\}$$

es denso en H_α .

Además, la elección de los tres elementos $(H_\alpha, x_\alpha, \pi_\alpha)$ es única salvo equivalencias unitarias. Es decir, si existe otro $(H'_\alpha, x'_\alpha, \pi'_\alpha)$ cumpliendo estas propiedades, entonces, existe un operador unitario W de H_α en H'_α tal que $W \pi_\alpha(A) = \pi'_\alpha(A) W$ para todo $A \in U$ y $W x_\alpha = x'_\alpha$.

A continuación damos alguna pauta sobre cómo se construye la construcción GNS de un estado, la cual se basa en el siguiente resultado.

Teorema 3.6 Sea U una C^* -álgebra. Si α es un estado en U , el conjunto

$$L_\alpha = \{A \in U \text{ tal que } \alpha(A^*A) = 0\}$$

es un ideal por la izquierda cerrado en U , (es decir, si $A \in L_\alpha$ y $B \in U$, entonces $AB \in L_\alpha$), y $\alpha(B^*A) = 0$ para cualquier $A \in L_\alpha$ y cualquier $B \in U$.

La ecuación

$$\langle A + L_\alpha, B + L_\alpha \rangle = \alpha(B^*A),$$

con $A, B \in U$, define un producto escalar en el espacio cociente U/L_α .

Este teorema permite definir sin ambigüedad el operador $\pi(A)$ en el espacio cociente U/L_α como: $\pi(A)(B + L_\alpha) = AB + L_\alpha$, el cual podemos extender por continuidad al espacio H_α , que es el completado de U/L_α , y por tanto es un espacio de Hilbert. Llamaremos $\pi_\alpha(A)$ al operador extendido.

La representación de la construcción GNS es la que asigna a cada operador $A \in U$, el operador $\pi_\alpha(A) \in B(H_\alpha)$.

Por último, el vector x_α de la representación GNS del estado α es $x_\alpha = I + L_\alpha \in U/L_\alpha \subset B(H_\alpha)$.

Ahora, para el Teorema de Gelfand-Naimark, necesitamos demostrar un lema previo, que es consecuencia del Teorema de Hahn Banach, el cual es un resultado central del análisis funcional, que garantiza la existencia de extensiones de funcionales con ciertas propiedades.

Teorema 3.7 (de Hahn Banach) Sea V un espacio vectorial complejo, sea p una seminorma en V , esto es, una aplicación de $p: V \rightarrow \mathbb{R}$ tal que

- 1. $p(x) \geq 0$ para todo $x \in V$.
- 2. $p(\lambda x) = |\lambda| p(x)$ 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 $\Psi = \varphi$ en M .

Una de las múltiples consecuencias de este teorema es que el conjunto de funcionales acotados de un espacio normado V (el espacio dual topológico de V) separa puntos. Esto es, si $x \in V$ y $f(x) = 0$ para todo funcional acotado f , entonces $x = 0$.

Es sencillo ver que esta propiedad de separar puntos se tiene para el conjunto total de estados en una C^* -álgebra (que son un tipo específico de funcionales), y, dado que cualquier estado es límite en la topología débil de combinaciones convexas de estados

puros por el Teorema de Krein-Milman, también el conjunto de estados puros de una C^* -álgebra separa puntos. El siguiente lema es inmediato a partir de este resultado.

Lema 3.8 *Sea U una C^* -álgebra. Sea $A \in U$. Si para todo estado puro α en U , la imagen de A por la representación dada por la construcción GNS de α es nula (es decir, $\pi_\alpha(A) = 0$), entonces $A = 0$.*

Demostración: Sea α un estado puro en una C^* -álgebra U , y sea $A \in U$. Por la propiedad 1. de la construcción GNS, la imagen de A por α es

$$\alpha(A) = \langle \pi_\alpha(A)x_\alpha, x_\alpha \rangle$$

Como por hipótesis $\pi_\alpha(A) = 0$, $\alpha(A) = 0$ para todo estado puro, luego $A = 0$. \square

Para demostrar el Teorema de Gelfand-Naimark, necesitamos definir qué es una suma directa de espacios de Hilbert y qué es una suma directa de representaciones.

Definición 3.9 *Sea $\{H_\nu\}_{\nu \in I}$ una familia de espacios de Hilbert, donde I puede ser numerable o no numerable. Decimos que el espacio de Hilbert H es suma directa de esta familia si cumple lo siguiente:*

- 1. Los vectores de H son las familias de vectores $\{x_\nu\}_{\nu \in I}$, con $x_\nu \in H_\nu$ que cumplen:

$$\sum_{\nu} \|x_\nu\|^2 < \infty$$

- 2. Las combinaciones lineales de vectores $x, y \in H$ son:

$$cx + dy = \{cx_\nu + dy_\nu\}_{\nu \in I}$$

- 2. El producto escalar de vectores $x, y \in H$ es:

$$\langle x, y \rangle = \sum_{\nu} \langle x_\nu, y_\nu \rangle$$

Nótese que la condición 1. asegura que la norma dada por el producto escalar en 3. está bien definida, pues $\langle x, x \rangle$ siempre es una suma convergente.

Representamos la suma directa de vectores y de espacios de Hilbert por: $x = \bigoplus_{\nu} x_\nu$ y $H = \bigoplus_{\nu} H_\nu$, respectivamente.

Sea U una C^* -álgebra. Sea $\{H_\nu\}_{\nu \in I}$ una familia de espacios de Hilbert, de tal manera que existe una representación π_ν de U en cada espacio H_ν . Si H es la suma directa de esta familia de estados, existe una representación π de U en H , que llamamos la suma directa de las representaciones π_ν dada por:

$$\pi(A)(\bigoplus_{\nu} x_\nu) = \bigoplus_{\nu} \pi_\nu(A)x_\nu.$$

Representamos la suma directa de vectores y de espacios de Hilbert por: $\pi = \bigoplus_{\nu} \pi_{\nu}$.

Ahora estamos en condiciones de probar el Teorema de Gelfand-Naimark.

Teorema 3.10 (de Gelfand-Naimark) *Toda C^* -álgebra tiene una representación fiel.*

Demostración: Sea U una C^* -álgebra. Sea F cualquier conjunto de estados de U que contenga a todos los estados puros. Sea π la suma directa de $\{\pi_{\alpha}\}_{\alpha \in F}$, donde π_{α} es la representación del estado α dada por la construcción GNS.

Si $A \in U$ y $\pi(A) = 0$, se tiene $\pi(A) = \bigoplus_{\alpha} \pi_{\alpha}(A) = 0$, y, por tanto $\pi_{\alpha}(A) = 0$ para todo $\alpha \in F$. Como F contiene a todos los estados puros, por el lema 3.8, $A = 0$. Luego π es una representación fiel. \square

Como hemos comentado anteriormente, este resultado proporciona una representación que permite ver cualquier C^* -álgebra como una C^* -subálgebra de un álgebra de operadores acotados en un espacio de Hilbert $B(H)$, lo cuál refuerza la conexión de esta formulación algebraica con la formulación usual expuesta en el capítulo 2. Dado que la formulación algebraica es más general, la tomaremos como marco de trabajo para el problema que planteamos en el capítulo siguiente.

4. La Conjetura de Kadison-Singer

En esta sección trataremos un problema acerca de la extensión de estados en C^* -subálgebras de observables, desde el marco de trabajo que nos proporciona la formulación algebraica de la mecánica cuántica.

Comenzaremos por formular el problema y exponer algunos problemas equivalentes.

Dada una C^* -álgebra, los observables que conmutan se denominan compatibles puesto que pueden ser medidos simultáneamente.

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).

Desde un punto de vista físico, las C^* -subálgebras abelianas maximales están en correspondencia con los Conjuntos Completos de Observables Compatibles introducidos en el capítulo 1.

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 $n \times n$, y D es equivalente a la C^* -subálgebra de ésta consistente en el conjunto de matrices diagonales complejas de tamaño $n \times n$.

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. $l^\infty(\mathbb{N}) \subset B(l^2(\mathbb{N}))$ (caso discreto)
- 2. $L^\infty(0,1) \subset B(L^2(0,1))$ (caso continuo),
- 3. $L^\infty(0,1) \oplus l^\infty(\{1, \dots, n\}) \subset B(L^2(0,1) \oplus l^2(\{1, \dots, n\}))$ (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 $a, b \in l^2$ $a = (a(1), a(2), \dots)$ y $b = (b(1), b(2), \dots)$ es $\langle a, b \rangle = a(1)\overline{b(1)} + a(2)\overline{b(2)} + \dots$

En el caso continuo, $L^\infty(0,1)$ 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 $\{1, \dots, n\}$, y $l^2(\mathbb{N})$ es el espacio de Hilbert de las sucesiones de números en $\{1, \dots, n\}$ de cuadrado integrable.

En el caso discreto, que es el de interés en este texto, la identificación de un elemento $a = (a(1), a(2), \dots) \in \ell^\infty$ con un operador $\phi_a \in B(\ell^2)$ es clara: el operador actúa multiplicando la componente n -ésima de un elemento $x \in \ell^2$ por $a(n)$, es decir $\phi_a x = a(1)x(1) + a(2)x(2) + \dots$. 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 [6] plantearon el siguiente problema:

Conjetura 4.1 (de Kadison-Singer) *¿Cada estado puro definido en una C^* -subálgebra abeliana maximal de la C^* -álgebra $B(\ell^2)$ de operadores acotados en ℓ^2 admite una única extensión a un estado puro en $B(\ell^2)$?*

Este problema plantea la ambigüedad (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 [7] 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 [10] 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 4.2 (KS_r , $r \geq 2$) *Existen constantes universales $\eta \geq 2$ y $\theta > 0$ de manera que se tiene lo siguiente: Sea $w_1, \dots, w_m \in \mathbb{C}^d$ 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 $u \in \mathbb{C}^d$. Entonces existe una partición S_1, S_2, \dots, S_r de $\{1, \dots, m\}$ tal que

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

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

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 [7].

Expondremos las principales ideas de la demostración de la conjetura KS_2 al final de este capítulo. 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 [10] 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 4.3 (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 $n \times n$, existen proyecciones diagonales P_1, \dots, P_r cuya suma es la matriz identidad tal que*

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

para todo $i = 1, 2, \dots, r$.

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 $\|T\| = \sup\{\|T(x)\|_{(W)} \mid \|x\|_{(V)} \leq 1\} = \sup\{\|T(x)\|_{(W)} \mid \|x\|_{(V)} = 1\}$, 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 euclídea 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 $n \times n$ 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 4.4 *Existen constantes universales $0 \leq \epsilon \leq 1$ y $\delta > 0$ de manera que se tiene lo siguiente: Para toda proyección $P = (p_{ij})$ de tamaño $n \times n$ cumpliendo $\max_i p_{ii}$*

$\leq \delta$, existen proyecciones diagonales Q_1, \dots, Q_r cuya suma es la matriz identidad tal que

$$\|Q_i P Q_i\| \leq 1 - \epsilon$$

para todo $i = 1, 2, \dots, r$.

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 $\epsilon = \frac{\theta}{\eta}$, $\delta = \frac{1}{\eta}$ 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 $\max_i p_{ii} \leq \frac{1}{\eta}$. Queremos demostrar que existen proyecciones diagonales Q_1, \dots, Q_r cuya suma es la matriz identidad tal que

$$\|Q_i P Q_i\| \leq 1 - \frac{\theta}{\eta}$$

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 $i = 1, \dots, n$ 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 $i = 1, \dots, n$; cumplen las hipótesis de la conjetura KS_r

Por un lado, para todo i ,

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

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, \dots, S_r de $\{1, \dots, n\}$ tal que

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

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

Para todo $j = 1, \dots, r$, definimos la proyección diagonal Q_j como:

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

para todo $i = 1, \dots, n$.

Ahora buscamos una cota de $\|Q_j P Q_j\|$ para todo j . Primero vemos que

$$\|Q_j P Q_j\| = \|Q_j P P Q_j\| = \|Q_j P (Q_j P)^*\| = \|Q_j P\|^2$$

Por otra parte,

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

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

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

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

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

Por tanto, para todo j , $\|Q_j P Q_j\| \leq 1 - \frac{\theta}{\eta}$ 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 $n \times n$ 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 encontrar en [9].

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

Teorema 4.5 Para todo $\epsilon > 0$ existe un $l \in \mathbb{N}$ tal que para toda matriz T compleja, con todos sus elementos diagonales nulos y de tamaño $n \times n$, existen proyecciones diagonales P_1, \dots, P_r cuya suma es la matriz identidad tal que

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

para todo $i = 1, 2, \dots, l$.

Demostración: Sea $\epsilon > 0$. 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 $n \times n$, existen proyecciones diagonales P_1, \dots, P_r cuya suma es la matriz identidad tal que

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

para todo $i = 1, 2, \dots, r$.

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 $n \times n$. 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 $\|B\| \leq \|T\|$.

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

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

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

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

para todo $i = 1, 2, \dots, r$.

Ahora definimos para cada uno de los $l = m^2$ pares $(i, j) \in \{1, \dots, r\} \times \{1, \dots, r\}$, 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 $n \times n$ con elementos diagonales iguales a 0 ó 1 (resultado del producto 0×0 , 0×1 ó 1×1 , 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:

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

□

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

Las matrices $n \times n$ (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 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 $\text{diag}: B(l_2) \rightarrow l^\infty$, definida como: $\text{diag}(T)(i) = \langle T(e_i), e_i \rangle$, 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 4.6 Para todo $\epsilon > 0$ existe un $l \in \mathbb{N}$ tal que para todo $T \in B(l_2)$ con $\text{diag}(T) = 0$, existen proyecciones diagonales P_1, \dots, P_l con $\sum_{i=1}^l P_i = 1$ tal que

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

para todo $i = 1, 2, \dots, l$.

Demostración: Sea $\epsilon > 0$. 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 $R_1, \dots, R_l \in M_n(\mathbb{C})$ con $\sum_{i=1}^l R_i = 1$ tal que

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

para todo $i = 1, 2, \dots, l$.

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

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

Trivialmente, vemos que, para todo n , $\text{diag}(\phi_n(T)) = 0$ y $\|\phi_n(T)\| = 1$. Por lo tanto, existen proyecciones diagonales $R_{n,1}, \dots, R_{n,l} \in \mathbb{N}$ 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 $i = 1, 2, \dots, l$.

Ahora veamos que existe una función estrictamente creciente $f: \mathbb{N} \rightarrow \mathbb{N}$ 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:

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

para cada $i = 1, \dots, l$.

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 $y_1 = (1, 1, 1, \dots)$ y para la función identidad $f = \text{Id}$, 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 $g: \mathbb{N} \rightarrow \mathbb{N}$ 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:

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

para cada $i = 1, \dots, l-1$.

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

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

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 $h: \mathbb{N} \rightarrow \mathbb{N}$ 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 $i = 1, \dots, l-1$; 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:

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

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

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

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 $\sum_{i=1}^l P_i = 1$. Veámoslo 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:

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

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 [9].

Sea H un espacio de Hilbert con una base ortonormal $\{v_i\}_{i \in I}$. Sea $A: H \rightarrow H$ 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 $\{i \in I \text{ tal que } \langle x, v_i \rangle \neq 0\}$ y $\{i \in I \text{ tal que } \langle y, v_i \rangle \neq 0\}$ 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 $c_N: l^2 \rightarrow \mathbb{N}$ 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 $i = 1, 2, \dots, l$; 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:

$$\begin{aligned} \langle P_i T P_i(u), v \rangle &= \langle T P_i(u), P_i(v) \rangle = \\ &= \langle \phi_n(T) R_{N,i}(c_N(u)), R_{N,i}(c_N(v)) \rangle = \langle R_{N,i} \phi_n(T) R_{N,i}(c_N(u)), (c_N(v)) \rangle \end{aligned}$$

Ahora, tomando el módulo del producto escalar:

$$\begin{aligned} |\langle P_i T P_i(u), v \rangle| &= |\langle R_{N,i} \phi_n(T) R_{N,i}(c_N(u)), (c_N(v)) \rangle| \leq \\ &\leq \|R_{N,i} \phi_n(T) R_{N,i}(c_N(u))\| \|c_N(v)\| \leq \end{aligned}$$

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

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 $i = 1, 2, \dots, l$; $\|P_i T P_i\| \leq \epsilon \|T\|$, y el teorema queda probado. \square

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 4.7 *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 $A \in U$. Es posible escribir $A = B + iC$, con B y $C \in U$ autoadjuntos. En efecto, la ecuación se cumple para $B = \frac{A + A^*}{2}$ y $C = \frac{A - A^*}{2i}$, que son autoadjuntos.

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

Una demostración precisa de este hecho se encuentra en la Proposición 4.2.3 de[5]. 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 U . La descomposición de Q (cuyo espectro es real) en dos elementos positivos (con espectro real no negativo) es análoga a 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. \square

Lema 4.8 *Sea U una C^* -álgebra con elemento unidad. Sea φ un estado puro en U . Entonces, para todo funcional lineal $\Psi: U \rightarrow \mathbb{N}$ con $0 \leq \Psi \leq \varphi$, existe un t con $0 \leq t \leq 1$ tal que se tiene $\Psi = t\varphi$.*

Demostración: Sea φ un estado puro en U y $\Psi: U \longrightarrow \mathbb{N}$ 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, $0 \leq \frac{A}{\|A\|} \leq 1$, 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 $\Psi(A) = 0$. 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 idénticamente nulo en U , $\Psi = 0$.

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

Por último, estudiamos el caso en que $0 < \Psi(1) < 1$. Podemos definir los funcionales $\Psi_1 = \frac{1}{1 - \Psi(1)} (\varphi - \Psi)$ y $\Psi_2 = \frac{1}{\Psi(1)} \Psi$, que son positivos por serlo Ψ y $\varphi - \Psi$, 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 $t = \Psi(1)$. \square

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

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

Primeramente, tratamos el caso $0 \leq B \leq 1$. Sea $C \in U$ positivo. Dado que $B, 1 - B$ y C son positivos, podemos escribir $B = D^*D$, $(1 - B) = U^*U$ 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 $\Psi: U \rightarrow \mathbb{C}$ como $\Psi(Z) = \varphi(ZB)$ para todo $Z \in U$. Como φ es positivo por ser un estado, y $ZB \geq 0$ para todo $Z \geq 0$, Ψ es positivo.

Además, si $Z \geq 0$, también se tiene $ZB \geq 0$, y

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

con lo que $\varphi - \Psi$ también es positivo.

Combinando los dos resultados, tenemos $0 \leq \Psi \leq \varphi$. Aplicando el Lema 2.8, existe t con $0 \leq t \leq 1$ tal que $\Psi = t\varphi$. 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 $0 \leq B \leq 1$, utilizamos la descomposición de B en elementos positivos cuya existencia garantiza el Lema 2.7. Escribimos $B = \sum_{k=0}^3 i^k \|B_k\| \frac{B_k}{\|B_k\|}$, con $B_k \geq 0$. Ahora, por la linealidad de φ y el hecho de que $0 \leq \frac{B_k}{\|B_k\|} \leq 1$:

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

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

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 Hahn 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 $P_1, \dots, P_n \in l^\infty$ con $\sum_{i=1}^n P_i = 1$ tal que

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

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

Sea $T \in B(l^2)$ con $\text{diag}(T) = 0$, y sean P_1, \dots, 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 $i = 1, 2, \dots, n$. Ahora bien, la 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 $\Psi: B(l^2)^2 \rightarrow \mathbb{C}$ definida como $\Psi(X, Y) = \Psi(X^*Y)$ es un producto escalar en $B(l^2)$ (cumple las propiedades de la definición 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,

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

para todo $i, j = 1, 2, \dots, n$. 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:

$$\begin{aligned} |\Psi(T)| &= \left| \Psi \left(\left(\sum_i P_i \right) T \left(\sum_j P_j \right) \right) \right| = \left| \sum_{i,j} \Psi(P_i T P_j) \right| = \\ &= |\Psi(P_{i_0} T P_{i_0})| \leq \|P_{i_0} T P_{i_0}\| \leq \epsilon \|T\| \end{aligned}$$

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 $\text{diag}(T) = 0$, 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 $\text{diag}(T) = 0$, $\|\Psi\| \leq \epsilon$.

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 $\Psi = \varphi \circ \text{diag}$ en $B(l^2)$. En efecto, si $A \in B(l^2)$, $\text{diag}(A - \text{diag}(A)) = \text{diag}(A) - \text{diag}(A) = 0$, luego aplicando el resultado anterior, $\Psi(A - \text{diag}(A)) = 0$, por lo que:

$$\Psi(A) = \Psi(\text{diag}(A)) = \varphi(\text{diag}(A))$$

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

Hemos demostrado que si Ψ es una extensión de φ , entonces $\Psi = \varphi \circ \text{diag}$, por lo que si existen dos funciones Ψ_1 y Ψ_2 que sean extensiones de φ , debe ser $\Psi_1 = \varphi \circ \text{diag} = \Psi_2$.

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 $r \geq 2$, la conjetura de Kadison-Singer tiene resultado positivo.

Dedicamos buena parte del TFG del Grado en Matemáticas [8] a detallar una prueba positiva de la conjetura KS_2 , lo que, por lo dicho anteriormente, es suficiente para que la Conjetura de Kadison-Singer quede demostrada.

Además de dar un resultado positivo a la Conjetura de Kadison-Singer, esta prueba es un ejemplo de aplicación de la técnica de familias de polinomios entrelazados, una técnica de álgebra elemental que es útil en otros problemas matemáticos.

Exponemos a continuación las ideas principales de esta demostración.

Comenzamos por reducir la prueba del teorema KS_2 a la demostración del siguiente teorema.

Teorema 4.10 *Sea $\epsilon > 0$. Sean v_1, \dots, 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

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

Definimos varios conceptos que aparecen en el teorema.

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

Sin entrar en detalles a la hora de definir los conceptos probabilísticos que aparecen en la formulación de este teorema, un vector aleatorio v en \mathbb{C}^d con soporte finito tiene una ley de probabilidad asociada en el conjunto finito de valores $z_1, \dots, z_n \in \mathbb{C}^d$ que puede tomar. Denotamos por $\mathbf{P}(v = z_i)$ a la probabilidad de que v tome el valor z_i , según esa ley.

Por otra parte, $\|\sum_{i=1}^m v_i v_i^*\|$ es la norma de una matriz de tamaño $d \times d$. En este caso, la norma que definimos anteriormente para operadores entre espacios normados es igual al mayor autovalor de la matriz $\sum_{i=1}^m v_i v_i^*$.

Los autovalores de esta matriz son las raíces de su polinomio característico, $\chi(\sum_{i=1}^m v_i v_i^*)$.

Dado que cada vector aleatorio v_i puede tomar un número finito de valores, el polinomio

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

también puede tomar un número finito de valores (ser igual a un número finito de polinomios), lo que nos permite verlo como una variable aleatoria con soporte finito en un conjunto de polinomios.

En particular, definimos la ley de probabilidad de esta variable aleatoria y su esperanza, que es una combinación convexa de los posibles valores (polinomios) que puede tomar.

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, \dots, v_m con probabilidad no nula, tal que $\|\sum_{i=1}^m v_i v_i^*\| \leq (1 + \sqrt{\epsilon})^2$.

Antes de continuar con la prueba de este teorema, veamos que, efectivamente es una reducción del Teorema KS₂.

Una elección adecuada de la distribución de los vectores aleatorios en este teorema, lleva sin dificultad al siguiente corolario, en términos de proyecciones, del que se deduce trivialmente el Teorema KS₂:

Corolario 4.11 *Sea $\delta > 0$ y r un entero positivo. Sean $u_1, \dots, u_m \in C^d$ 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, \dots, S_r de $\{1, \dots, m\}$ 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 $j = 1, \dots, r$.

Tomando $r = 2$ y $\delta = 1/18$ en este corolario, la prueba del Teorema KS₂ es inmediata para $d = 18$ y $\epsilon = 2$.

Por tanto, basta con probar el Teorema 4.10 para probar el Teorema KS₂ y por tanto completar la prueba de la Conjetura de Kadison-Singer.

Por lo que hemos comentado acerca del significado del teorema, éste afirma 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, \dots, v_m , a partir de los cuales se construye el polinomio, cumpla las hipótesis del teorema).

Para obtener información acerca de de las raíces de los polinomios que están involucrados en el teorema, resulta clave la técnica de familias de polinomios entrelazados.

Comenzaremos definiendo qué es un entrelazado.

Definición 4.12 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

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

Decimos que los polinomios f_1, \dots, 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 4.13 Sean f_1, \dots, f_m polinomios de grado $n \geq 2$ de coeficientes y raíces reales con coeficiente principal positivo y entrelazado común.

Llamamos:

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

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

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

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

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

La demostración de este teorema es muy sencilla, a partir del Teorema de Bolzano, sin embargo, es un poco larga. En su lugar discutimos el teorema en el caso de dos parábolas (polinomios de grado 2), el cual, a pesar de ser el caso más sencillo, es análogo al caso general.

Sean $f(x) = a(x - \lambda_1(f))(x - \lambda_2(f))$ y $g(x) = b(x - \lambda_1(g))(x - \lambda_2(g))$ de coeficientes y raíces reales con coeficiente principal positivo y entrelazado común. Sea S una combinación lineal de los polinomios con coeficientes reales no negativos

Tratamos el caso en que no tienen raíces comunes ni dobles, que es fácil, pero habría que estudiar aparte

Sea $\lambda_1(f) > \lambda_1(g)$ (el caso contrario es análogo). Como ambos polinomios tienen coeficiente principal positivo, deben ser positivos para $x > \lambda_1(f)$, ya que tienden a $+\infty$ para x tendiendo a $+\infty$. Por tanto $S(x) > 0$ para $x > \lambda_1(f)$.

Como f y g tienen un entrelazado común, la siguiente raíz por orden de mayor a menor debe ser $\lambda_1(g)$. En cada raíz, los polinomios cambian de signo, luego para $x < \lambda_1(g)$, $S(x) < 0$. Por el Teorema de Bolzano, S tiene que tener una raíz en $(\lambda_1(f), \lambda_1(g))$. Con la otra raíz se razona análogamente.

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 4.14 *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.*

En la práctica, para la comprobación de la propiedad de entrelazado se utiliza una caracterización más sencilla que el orden de las raíces. Además el concepto de familia entrelazada permite aplicar este último teorema recursivamente únicamente exigiendo tan solo el entrelazado común de varios subconjuntos y de conjuntos de combinaciones de estos.

La idea principal de esta demostración del Teorema 2.10 es sortear la dificultad de tener que buscar una cota para las raíces de alguno de los valores (polinomios) que puede tomar la variable $\chi(\sum_{i=1}^m v_i v_i^*)$, de los cuales puede haber una cantidad finita, pero arbitrariamente grande. Para ello se reduce 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 $\chi(\sum_{i=1}^m v_i v_i^*)$, al ser esta de soporte finito, es una combinación lineal convexa de los posibles valores que puede tomar la variable y, por tanto, después de encontrar que, bajo las hipótesis adecuadas, está acotada por $(1+\sqrt{\epsilon})^2$, podemos aplicar el Teorema 4.13 si el conjunto soporte de posibles valores del polinomio es una familia entrelazada. Así la raíz más grande de la esperanza está en $[m_1, M_1]$. Por tanto, debe existir un polinomio del conjunto cuya mayor raíz sea menor o igual que ésta y, por tanto, que sea menor o igual que $(1+\sqrt{\epsilon})^2$, lo que prueba el resultado.

Un desarrollo completo de estas ideas puede encontrarse en el TFG del Grado en Matemáticas [8].

Referencias

- [1] AKEMANN, C. A., AND ANDERSON, J. Lyapunov theorems for operator algebras. *Mem. Amer. Math. Soc.* 94, 458 (1991), iv+88.
- [2] ARAKI, H. *Mathematical theory of quantum fields*, vol. 101 of *International Series of Monographs on Physics*. Oxford University Press, Oxford, 2009. Translated from the 1993 Japanese original by Ursula Carow-Watamura, Reprint of the 1999 edition [MR1799198].
- [3] COHEN-TANNOUJDI C., DIU B., L. F. *Quantum mechanics*. 2005.
- [4] GLEASON, A. M. Measures on the closed subspaces of a Hilbert space. *J. Math. Mech.* 6 (1957), 885–893.
- [5] 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.
- [6] KADISON, R. V., AND SINGER, I. M. Extensions of pure states. *Amer. J. Math.* 81 (1959), 383–400.
- [7] 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.
- [8] SAMPERIO VALDIVIESO, A. *Familias de polinomios entrelazados. La Conjetura de Kadison-Singer*. Universidad de Valladolid, 2019.
- [9] STEVENS, M. *The Kadison-Singer property*, vol. 14 of *SpringerBriefs in Mathematical Physics*. Springer, Cham, 2016. With a foreword by Klaas Landsman.
- [10] WEAVER, N. The Kadison-Singer problem in discrepancy theory. *Discrete Math.* 278, 1-3 (2004), 227–239.