

Formalismo en dimensión finita.

En este capítulo presentaremos la descripción de los elementos matemáticos que nos permitirán representar los estados y observables cuánticos. Como dijimos, los estados estarán representados por vectores que viven en espacios vectoriales complejos cuya dimensión es igual al número máximo de resultados diferentes en una medición exhaustiva. Veremos en el siguiente capítulo que la dimensión del espacio vectorial que describe a un estado cuántico también puede ser infinita (no numerable), como lo es para los casos en donde intervienen grados de libertad continuos como el de traslación.

2.1. Espacios vectoriales.

Un *espacio vectorial lineal* es un conjunto de elementos, denominados “vectores”, que es cerrado frente a una operación que llamamos “suma” y frente al producto con elementos de otro conjunto (que debe ser un “cuerpo”) a los que llamamos “escalares”. Un espacio vectorial consta de cuatro ingredientes $\{\mathcal{V}, +, \mathcal{K}, \times\}$, donde \mathcal{V} es un conjunto de vectores, \mathcal{K} es un cuerpo (que típicamente puede ser el de los números reales o los números complejos), y las operaciones $+$ y \times son la suma de vectores y el producto por un escalar.

Denotaremos a los vectores de \mathcal{V} de la siguiente manera: $|v_1\rangle, |v_2\rangle, \dots |v_n\rangle$. Entonces, si $|v_1\rangle$ y $|v_2\rangle$ son elementos de \mathcal{V} , y λ_1 y λ_2 son elementos de \mathcal{K} , tenemos que $|w\rangle = \lambda_1 \times |v_1\rangle + \lambda_2 \times |v_2\rangle$ es también un elemento de \mathcal{V} . En general omitiremos el símbolo \times para indicar el producto de un vector por un escalar y escribiremos simplemente $|w\rangle = \lambda_1 |v_1\rangle + \lambda_2 |v_2\rangle$.

Diremos que un conjunto de vectores $\{|v_n\rangle\}$ es *linealmente independiente* si la única combinación lineal entre ellos que resulta cero es la trivial. Es decir se cumple que $\sum_n c_n |v_n\rangle = 0$ sólo si $c_n = 0$ para todo n .

La *dimensión del espacio vectorial* $\dim(\mathcal{V})$ es el número máximo de vectores linealmente independientes que es posible encontrar en \mathcal{V} .

Producto interno hermitiano

Trabajaremos con espacios vectoriales sobre los que es posible definir un producto interno. Dados dos vectores $|v\rangle$ y $|w\rangle$ el producto interno entre ambos se denominará $(|v\rangle, |w\rangle)$ y es tal que

- $(|v\rangle, |w\rangle) \in \mathcal{C}$, donde \mathcal{C} es el conjunto de los números complejos.
- $(|v\rangle, |w\rangle) = (|w\rangle, |v\rangle)^*$ donde el superíndice $*$ se usará para denotar el complejo conjugado.
- $(|w\rangle, \lambda_1|v_1\rangle + \lambda_2|v_2\rangle) = \lambda_1(|w\rangle, |v_1\rangle) + \lambda_2(|w\rangle, |v_2\rangle)$. Y por otro lado, $(\lambda_1|v_1\rangle + \lambda_2|v_2\rangle, |w\rangle) = \lambda_1^*(|v_1\rangle, |w\rangle) + \lambda_2^*(|v_2\rangle, |w\rangle)$. El producto interno es lineal en la segunda entrada y antilineal en la primera.
- $(|v\rangle, |v\rangle) \geq 0$ y $(|v\rangle, |v\rangle) = 0$ si y sólo si $|v\rangle = 0$ (donde 0 denota aquí el vector que es el elemento neutro de la operación suma de vectores).

El producto escalar nos permite definir también una noción de ortogonalidad entre vectores. Diremos que $|u\rangle$ es *ortogonal* a $|v\rangle$ si y solo si $(|u\rangle, |v\rangle) = 0$.

Norma

Con un producto escalar hermitiano podemos definir la norma de un vector así como también una noción de distancia entre vectores.

- La norma de un vector se define como $\| |v\rangle \| = \sqrt{(|v\rangle, |v\rangle)}$.
- La distancia entre dos vectores se define como la norma del vector diferencia entre ambos: $\text{dist}(|u\rangle, |v\rangle) = \| |u\rangle - |v\rangle \|$.

Desigualdad de Schwartz y Triangular.

Todo producto interno con las propiedades mencionadas más arriba satisface la *desigualdad de Schwartz* que establece que:

$$(|v\rangle, |u\rangle)|^2 \leq (|v\rangle, |v\rangle) \times (|u\rangle, |u\rangle). \quad (2.1)$$

Esta propiedad es fácil de probar: Consideremos dos vectores $|u\rangle$ y $|v\rangle$. Definamos un tercer vector $|z\rangle$ como la parte de $|u\rangle$ que es ortogonal a $|v\rangle$. O sea: $|z\rangle = |u\rangle - \frac{(|v\rangle, |u\rangle)}{(|v\rangle, |v\rangle)}|v\rangle$ (es trivial probar que $(|v\rangle, |z\rangle) = 0$). De aquí vemos que el vector $|u\rangle$ puede escribirse como suma de dos vectores ortogonales simplemente invirtiendo la expresión anterior: $|u\rangle = |z\rangle + \frac{(|v\rangle, |u\rangle)}{(|v\rangle, |v\rangle)}|v\rangle$. Calculando ahora la norma de $|u\rangle$ obtenemos $\| |u\rangle \|^2 = \| |z\rangle \|^2 + \| |v\rangle \|^2 \times \frac{(|v\rangle, |u\rangle)^2}{\| |v\rangle \|^4} \geq \frac{(|v\rangle, |u\rangle)^2}{\| |v\rangle \|^2}$. Multiplicando ambos lados de la igualdad por $\| |v\rangle \|^2$ obtenemos la desigualdad de Schwartz.

Por otro lado, como todo espacio vectorial normado también satisface la conocida *desigualdad triangular* que establece:

$$\| |u\rangle + |v\rangle \| \leq \| |u\rangle \| + \| |v\rangle \|. \quad (2.2)$$

Esta desigualdad es sencilla de demostrar expandiendo el producto interno de la suma de vectores y luego utilizando la desigualdad de Schwartz.

Bases ortonormales

Un conjunto de D vectores $B = \{ |u_i\rangle, i = 1, \dots, D \}$ en un espacio vectorial de dimensión D es una base ortonormal si y solo si:

1. Los D vectores son linealmente independientes
2. Los vectores satisfacen que $(|u_i\rangle, |u_j\rangle) = \delta_{i,j}$.

Todo vector $|v\rangle$ puede escribirse como combinación lineal de los elementos de una base ortonormal. En efecto, $|v\rangle = \sum_j v_j |u_j\rangle$ donde los coeficientes v_j pueden calcularse tomando el producto interno con los estados $|u_k\rangle$. De este modo obtenemos $v_k = (|u_k\rangle, |v\rangle)$. O sea que la descomposición de cualquier vector es $|v\rangle = \sum_k (|v_k\rangle, |v\rangle) |v_k\rangle$.

Funcionales lineales y notación de Dirac

Diremos que f es una funcional lineal si f es una aplicación $f : \mathcal{V} \rightarrow \mathcal{C}$ (a todo vector le asigna un número complejo) tal que satisface linealidad $f(\lambda_1 |v_1\rangle + \lambda_2 |v_2\rangle) = \lambda_1 f(|v_1\rangle) + \lambda_2 f(|v_2\rangle)$.

Podemos notar que una vez definido el producto escalar $(|u\rangle, |v\rangle)$ queda definida una asociación entre vectores y funcionales. Esto es, para todo vector $|v\rangle$ podemos definir una funcional, que denotaremos $F_{|v\rangle}$, de modo tal que su acción sobre cualquier otro vector $|w\rangle$ es tal que $F_{|v\rangle}(|w\rangle) = (|v\rangle, |w\rangle)$. Es decir, la funcional asociada al vector $|v\rangle$ le asigna un número a todo otro vector $|w\rangle$ que es igual a la proyección de $|w\rangle$ sobre $|v\rangle$. Mas aún, es posible mostrar (teorema de Riez) que existe una correspondencia uno a uno entre vectores y funcionales.

De esta manera, a partir de ahora usaremos la llamada "Notación de Dirac" para las funcionales. A la funcional $F_{|v\rangle}$ la denotaremos como $\langle v|$ (o sea: $F_{|v\rangle} = \langle v|$) y al producto interno entre dos vectores lo denotaremos como

$$(|v\rangle, |w\rangle) \equiv \langle v|w\rangle. \quad (2.3)$$

De esta manera, a los vectores $|v\rangle$ se los suele llamar 'kets' y a las funcionales $\langle v|$ que viven en el espacio dual a \mathcal{V} 'bras'.

2.2. Operadores lineales

Un operador lineal \hat{A} es una aplicación $\hat{A} : \mathcal{V} \rightarrow \mathcal{V}$ tal que:

$$\hat{A}(\lambda_1|v_1\rangle + \lambda_2|v_2\rangle) = \lambda_1\hat{A}(|v_1\rangle) + \lambda_2\hat{A}(|v_2\rangle) = |v_3\rangle, \quad (2.4)$$

donde los vectores $|v_i\rangle$ pertenecen todos a \mathcal{V} .

Para caracterizar un operador, es suficiente con indicar su acción en una base dada, ya que cualquier vector puede obtenerse como combinación lineal de los elementos de la base. Sobre cada base $B = \{|u_j\rangle, j = 1, \dots, D\}$ un operador lineal tiene asociada una matriz con elementos A_{jk} . En efecto, supongamos que aplicamos el operador \hat{A} al vector $|w\rangle = \sum_k w_k|u_k\rangle$. Por linealidad tenemos que $\hat{A}|w\rangle = |w'\rangle = \sum_k w_k \hat{A}|u_k\rangle$. Pero el vector $|w'\rangle$ también puede desarrollarse en la misma base: $|w'\rangle = \sum_j w'_j |u_j\rangle$. De manera que:

$$|w'\rangle = \sum_j w'_j |u_j\rangle = \sum_k w_k \hat{A}|u_k\rangle. \quad (2.5)$$

De aquí tomando el producto interno con $|u_j\rangle$ podemos despejar $w'_j = \sum_k w_k A_{jk}$, donde $A_{jk} = \langle u_j|\hat{A}|u_k\rangle$ (o sea, $A_{jk} = \langle u_j|\hat{A}|u_k\rangle$). La matriz del operador \hat{A} en la base B nos permite obtener al vector transformado por \hat{A} como combinación lineal de los elementos de la base B .

Naturalmente, los elementos de matriz del operador dependen de la base. Para relacionar la matriz del operador en dos bases diferentes $B = \{|u_j\rangle, j = 1, \dots, D\}$ y $B' = \{|u'_j\rangle, j = 1, \dots, D\}$ podemos proceder de la siguiente forma: si vinculamos las dos bases escribiendo $|u'_j\rangle = \sum_k \langle u_k|u'_j\rangle|u_k\rangle$ y reemplazamos esta expresión en la fórmula para $A'_{jk} = \langle u'_j|\hat{A}|u'_k\rangle$ obtenemos:

$$\begin{aligned} A'_{jk} &= \langle u'_j|\hat{A}|u'_k\rangle = \sum_{lm} \langle u'_j|u_l\rangle \langle u_l|\hat{A}|u_m\rangle \langle u_m|u'_k\rangle \\ A'_{jk} &= \sum_{lm} U_{jl} A_{lm} U_{km}^* \end{aligned} \quad (2.6)$$

donde definimos la matriz de cambio de base $U_{jl} = \langle u'_j|u_l\rangle$.

2.3. Proyectores y algunas virtudes de la notación de Dirac.

A partir de dos vectores $|v\rangle$ y $|w\rangle$ podemos construir dos operadores lineales distintos:

$$P_{uv} = |u\rangle\langle v|, \quad P_{vu} = |v\rangle\langle u| \quad (2.7)$$

donde la interpretación de cada uno es simplemente que P_{uv} aplicado a cualquier vector $|w\rangle$ siempre apunta en la dirección de $|u\rangle$ y tiene un módulo proporcional a

la proyección de $|w\rangle$ sobre $|v\rangle$. Asimismo, es fácil notar que hay un operador al que podemos denominar el “proyector sobre $|v\rangle$ ” que simplemente resulta ser

$$P_{|v\rangle} = |v\rangle\langle v| \quad (2.8)$$

Es evidente que este operador cumple que $P_{|v\rangle}^2 = P_{|v\rangle}$ (que es la regla básica que define a un proyector). El proyector $P_{|v\rangle}$ proyecta sobre un único vector y por eso se dice que es un proyector de rango 1 (su matriz tiene un único autovalor no nulo e igual a 1). Podemos definir proyectores de mayor rango simplemente sumando los proyectores sobre dos vectores ortogonales $|v_1\rangle$ y $|v_2\rangle$. En efecto, el operador $P_{1,2} = P_{|v_1\rangle} + P_{|v_2\rangle}$ también es un proyector pero su rango es 2. Esto puede generalizarse a proyectores de rango más alto, tal como veremos en lo que sigue.

Consideremos una base ortonormal $B = \{|v_j\rangle, j = 1, \dots, D\}$ y los proyectores $P_j = |v_j\rangle\langle v_j|$. Entonces, podemos demostrar que el operador identidad $\mathbb{1}$ puede escribirse como

$$\mathbb{1} = \sum_j |v_j\rangle\langle v_j| = \sum_j P_j \quad (2.9)$$

Esto es válido para cualquier base ortonormal. Es decir, la identidad se puede escribir como la suma de los proyectores sobre cualquier base ortonormal. Esta descomposición de la identidad es muy útil ya que nos permite obtener fácilmente muchos resultados. Por ejemplo, a partir de ella podemos obtener de manera inmediata la relación que existe entre la matriz del operador en dos bases diferentes. En efecto:

$$\begin{aligned} A'_{jk} &= \langle v'_j | \hat{A} | v'_k \rangle = \langle v'_j | \mathbb{1} \hat{A} \mathbb{1} | v'_k \rangle \\ &= \langle v'_j | \left(\sum_l |v_l\rangle\langle v_l| \right) \hat{A} \left(\sum_m |v_m\rangle\langle v_m| \right) | v'_k \rangle \\ &= \sum_{lm} \langle v'_j | v_l \rangle \langle v_l | \hat{A} | v_m \rangle \langle v_m | v'_k \rangle \\ &= \sum_{lm} U_{jl} A_{lm} U_{km}^* \end{aligned} \quad (2.10)$$

También es evidente que la matriz U satisface que $U \times U^{T*} = \mathbb{1}$ ya que $(U \times U^{T*})_{jk} = \sum_l U_{jl} U_{kl}^* = \sum_l \langle v'_j | v_l \rangle \langle v_l | v'_k \rangle = \langle v'_j | (\sum_l |v_l\rangle\langle v_l|) | v'_k \rangle = \langle v'_j | \mathbb{1} | v'_k \rangle = \delta_{jk}$.

Operador Adjunto. Operadores hermíticos y unitarios

Dado un operador \hat{A} definimos el operador adjunto hermitiano, que se denota como \hat{A}^\dagger , como aquel operador tal que para todo par de vectores $|u\rangle$ y $|v\rangle$ vale que

$$(\vec{u}, \hat{A}|v\rangle) = (\hat{A}^\dagger|u\rangle, |v\rangle). \quad (2.11)$$

Es fácil encontrar una relación simple entre los elementos de matriz de \hat{A} y de \hat{A}^\dagger . En efecto:

$$\begin{aligned}\hat{A}_{jk}^\dagger &= (\langle v_j |, \hat{A}^\dagger | v_k \rangle) \\ &= (\hat{A}^\dagger | v_k \rangle, | v_j \rangle)^* = (\langle v_k |, \hat{A} | v_j \rangle)^* \\ &= \hat{A}_{kj}^*.\end{aligned}\tag{2.12}$$

O sea, la matriz de \hat{A}^\dagger es la transpuesta y conjugada de la matriz de \hat{A} .

Hay propiedades importantes de los operadores adjuntos. En particular se cumple que $(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger$. Esta identidad se demuestra trivialmente usando que la matriz adjunta es la transpuesta y conjugada: $((\hat{A}\hat{B})^\dagger)_{jk} = (\hat{A}\hat{B})_{kj}^* = \sum_l A_{kl}^* B_{lj} = \sum_l (\hat{B}^\dagger)_{jl} (\hat{A}^\dagger)_{lk} = (\hat{B}^\dagger\hat{A}^\dagger)_{jk}$.

Definimos los siguientes tipos de operadores:

- **Hermíticos.** \hat{A} es un operador hermítico si y solo si satisface que $\hat{A} = \hat{A}^\dagger$. Es evidente que para estos operadores todos los elementos diagonales deben ser reales.
- **Unitarios.** \hat{U} es un operador unitario si y solo si satisface que $\hat{U}^\dagger\hat{U} = \hat{U}\hat{U}^\dagger = \mathbb{1}$. Es decir, que para estos operadores la matriz adjunta es la inversa. Un ejemplo importante de este tipo de operadores es el operador de cambio de base, que transforma vectores ortogonales en vectores ortogonales. Dadas dos bases $B = \{|v_j\rangle, j = 1, \dots, D\}$ y $B' = \{|v'_j\rangle, j = 1, \dots, D\}$ la matriz definida por los productos escalares entre los elementos de ambas bases es unitaria. Es decir, si definimos $(\hat{U})_{jk} = \langle v'_j | v_k \rangle$ entonces se cumple que $\hat{U}\hat{U}^\dagger = \hat{U}^\dagger\hat{U} = \mathbb{1}$. Esto puede demostrarse apelando a la versatilidad de la notación de Dirac de manera muy sencilla:

$$\begin{aligned}(\hat{U}^\dagger\hat{U})_{jk} &= \sum_l \langle v_j | \hat{U}^\dagger | v_l \rangle \langle v_l | \hat{U} | v_k \rangle \\ &= \sum_l \langle v_l | \hat{U} | v_j \rangle^* \langle v_l | \hat{U} | v_k \rangle \\ &= \sum_l \langle v'_l | v_j \rangle^* \langle v'_l | v_k \rangle = \sum_l \langle v_j | v'_l \rangle \langle v'_l | v_k \rangle \\ &= \langle v_j | v_k \rangle = \delta_{jk}\end{aligned}$$

Dejamos la demostración simétrica de que $\hat{U}\hat{U}^\dagger = \mathbb{1}$ como ejercicio.

- **Proyectores.** Como ya definimos mas arriba, los proyectores son operadores que cumplen que $\hat{P}^2 = \hat{P}$.
- **Normales.** Un operador \hat{A} es “normal” si y solo si se cumple que $\hat{A}^\dagger\hat{A} = \hat{A}\hat{A}^\dagger$ (o sea, si \hat{A} “conmuta” con su adjunto). Evidentemente, los operadores hermíticos y los unitarios son normales.

Traza de un operador

Dado un operador \hat{A} se define la traza de \hat{A} como el funcional lineal que cumple $\text{Tr}(\hat{A}) = \sum_j \langle v_j | \hat{A} | v_j \rangle$. O sea, la traza es la suma de todos los elementos diagonales de \hat{A} en una base. Es simple ver que la traza es la misma cualquiera sea la base en la que la calculemos. En efecto, si usamos la base B' en lugar de la base B para escribir la traza, podemos ver que

$$\begin{aligned}
 \sum_j \langle v_j | \hat{A} | v_j \rangle &= \sum_{jkm} \langle v_j | v'_k \rangle \langle v'_k | \hat{A} | v'_m \rangle \langle v'_m | v_j \rangle \\
 &= \sum_{km} \left(\sum_j \langle v'_m | v_j \rangle \langle v_j | v'_k \rangle \right) \langle v'_k | \hat{A} | v'_m \rangle \\
 &= \sum_{km} \langle v'_m | v'_k \rangle \langle v'_k | \hat{A} | v'_m \rangle \\
 &= \sum_m \langle v'_m | \hat{A} | v'_m \rangle
 \end{aligned} \tag{2.13}$$

Una propiedad muy importante de la traza, que se deduce de las expresiones anteriores, es que $\text{Tr}(\hat{A}\hat{B}) = \text{Tr}(\hat{B}\hat{A})$. De hecho, puede demostrarse que la única funcional lineal con esta propiedad es la traza, definida más arriba.

Operadores diagonalizables

A lo largo del apunte prestaremos especial atención los operadores hermíticos (asociados a los observables) y unitarios (asociados a las evoluciones temporales). Estos operadores son normales por definición y por lo tanto se puede mostrar además que son diagonalizables. Esto implica que, dado un operador normal \hat{A} que actúa en un espacio vectorial finito, existe una base ortonormal en el que la matriz que representa al operador \hat{A} es diagonal. Si \hat{A} es diagonal en la base B entonces se cumple que $\hat{A}|v_j\rangle = a_j|v_j\rangle$. Cuando esto sucede, se dice que los vectores $|v_j\rangle$ son los autovectores de \hat{A} y a_j son los correspondientes autovalores. Es decir, en la base B el operador \hat{A} se representa por una matriz diagonal cuyos elementos son los autovalores a_j . Asimismo, es sencillo notar que las matrices hermíticas poseen autovalores reales, y las unitarias autovalores complejos de la forma $e^{i\phi}$.

En general, es importante saber cuándo un operador \hat{A} es diagonalizable. Para esto previamente definimos el polinomio característico de \hat{A} , y lo denotamos como $p(x) = \det(A - x\mathbb{1})$, donde $\det()$ es el funcional “determinante”. Dado el polinomio característico $p(x)$ podemos encontrar sus raíces. Si el grado del polinomio es D a lo sumo hay D raíces distintas que llamamos x_1, x_2, \dots, x_D . O sea $p(x) = \prod_j (x - x_j)$ (a menos de un factor multiplicativo). Si \hat{A} es diagonalizable entonces las raíces de $p(x)$ son los autovalores de \hat{A} . Por último, se define el polinomio “minimal” de \hat{A} , denotado $m(x)$, como el polinomio de menor grado tal que $m(\hat{A}) = 0$. Es posible demostrar, que $m(x)$ divide a $p(x)$ y que $m(x)$ tiene las mismas raíces que $p(x)$. Dados estos elementos podemos formular la condición necesaria y suficiente para que \hat{A}

sea diagonalizable: \hat{A} es diagonalizable si y solo si el polinomio minimal $m(x)$ no tiene raíces múltiples.

El método de diagonalización de operadores es bien conocido: una vez conocido el polinomio característico y sus raíces, para cada una de ellas es necesario resolver un sistema de ecuaciones para encontrar los autovectores asociados a cada autovalor. El sistema es de la forma $\hat{A}|w_j\rangle = x_j|w_j\rangle$.

Para finalizar, como ejemplo de matriz no diagonalizable podemos mencionar a la matriz $\sigma_- = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ cuyo polinomio característico es $p(x) = x^2$ y cuyo polinomio minimal es $m(x) = p(x)$. La matriz (que es nilpotente) no es diagonalizable ya que $m(x)$ tiene una raíz doble.

Descomposición espectral de un operador.

Supongamos que \hat{A} es diagonalizable en la base $B = \{|v_j\rangle, j = 1, \dots, D\}$ y que los autovalores asociados a cada autovector son a_j . Entonces el operador puede escribirse como

$$\hat{A} = \sum_j a_j |v_j\rangle\langle v_j|. \quad (2.14)$$

En un caso más general, puede haber muchos autovectores que tengan el mismo autovalor. La degeneración del autovalor a_j será igual al número máximo de autovectores ortogonales que tienen autovalor a_j . La denotaremos como g_j y no es otra cosa que la dimensión del subespacio asociado al autovalor a_j . En ese caso, la base de autovectores de \hat{A} puede escribirse como $B = \{|v_{j,\mu_j}\rangle, j = 1, \dots, K, \mu_j = 1, \dots, g_j\}$ donde K es el número de autovalores diferentes (en este caso, la dimensión del espacio de estados es $D = \sum_{j=1}^K g_j$). Entonces, la descomposición espectral del operador es

$$\hat{A} = \sum_{j=1}^K a_j \sum_{\mu=1}^{g_j} |v_{j,\mu_j}\rangle\langle v_{j,\mu_j}| \quad (2.15)$$

En cualquier caso, esta descomposición es del tipo $\hat{A} = \sum_j a_j \Pi_j$ donde $\Pi_j = \sum_{\mu=1}^{g_j} |v_{j,\mu_j}\rangle\langle v_{j,\mu_j}|$ es el proyector sobre el subespacio asociado al autovalor a_j (cuya dimensión es g_j).

Operadores compatibles. Teorema: compatibles si y solo si conmutan.

Sean dos operadores diagonalizables \hat{A} y \hat{B} . Se dice que estos operadores son compatibles cuando son diagonalizables en la misma base (o sea, tienen una base común de autovectores). Podemos demostrar que dos operadores \hat{A} y \hat{B} son compatibles si y sólo si conmutan. O sea, si y sólo si

$$[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A} = 0. \quad (2.16)$$

Este teorema es muy sencillo de probar en una dirección. En efecto, si los operadores son compatibles entonces es evidente que conmutan ya que dos matrices que son diagonales en la misma base conmutan (su producto, en cualquier orden, es diagonal en la misma base). En la otra dirección, el teorema es menos trivial. En efecto, supongamos que \hat{A} y \hat{B} conmutan y supongamos que \hat{A} es diagonal en la base de autovectores $B = \{|v_{j,\mu_j}\rangle, j = 1, \dots, K, \mu_j = 1, \dots, g_j\}$ (o sea, estamos suponiendo que el operador \hat{A} puede ser degenerado y que la degeneración del autovalor a_j es, como antes, igual a g_j). Como estos vectores son autovectores de \hat{A} se cumple que $\hat{A}|v_{j,\mu_j}\rangle = a_j|v_{j,\mu_j}\rangle$. Veamos ahora que propiedad tienen los vectores $\hat{B}|v_{j,\mu_j}\rangle$. Se puede probar que estos vectores siguen siendo autovectores de \hat{A} con el mismo autovalor a_j . Esto se ve usando la conmutatividad de los dos operadores $\hat{A}\hat{B}|v_{j,\mu_j}\rangle = \hat{B}\hat{A}|v_{j,\mu_j}\rangle = a_j\hat{B}|v_{j,\mu_j}\rangle$. En consecuencia, el operador \hat{B} deja invariantes a los subespacios de dimensión g_j que están asociados a los distintos autovalores a_j . Por lo tanto, en esta base las matrices de los operadores \hat{A} y \hat{B} son

$$\hat{A} = \begin{bmatrix} a_1 \mathbb{1}_{g_1 \times g_1} & 0 & 0 & \dots & 0 \\ 0 & a_2 \mathbb{1}_{g_2 \times g_2} & 0 & \dots & 0 \\ 0 & 0 & a_3 \mathbb{1}_{g_3 \times g_3} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_K \mathbb{1}_{g_K \times g_K} \end{bmatrix},$$

$$\hat{B} = \begin{bmatrix} B_{g_1 \times g_1}^{(1)} & 0 & 0 & \dots & 0 \\ 0 & B_{g_2 \times g_2}^{(2)} & 0 & \dots & 0 \\ 0 & 0 & B_{g_3 \times g_3}^{(3)} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & B_{g_K \times g_K}^{(K)} \end{bmatrix}. \quad (2.17)$$

Los bloques son de $g_j \times g_j$ y en esos bloques la matriz de \hat{A} son proporcionales a la identidad. O sea, en esta base la matriz de \hat{B} es diagonal por bloques. Si hacemos cualquier cambio de base en esos subespacios no modificaremos la matriz de \hat{A} (ya que es proporcional a la identidad). Entonces, como \hat{B} es diagonalizable, cada una de las submatrices $B^{(j)}$ es diagonalizable, por lo tanto podemos encontrar una base en ese subespacio en la cual $B^{(j)}$ es diagonal. Haciendo eso en cada uno de los subespacios asociados a los autovalores a_j encontramos una base en la cual \hat{B} será diagonal y en la cual \hat{A} es también diagonal (y tiene la misma forma que antes). Por lo tanto, existe una base en la que ambos operadores son diagonales.

Función de un operador

Veremos dos formas de definir la función de un operador. Por un lado, dado un operador \hat{A} que posee una base completa de autovectores $\{|v\rangle_i\}$ con autovalores a_i

y una función $f(x) : \mathcal{C} \rightarrow \mathcal{C}$, definimos:

$$f(\hat{A}) = \sum_{n=1}^N f(a_i) |v_i\rangle \langle v_i|. \quad (2.18)$$

Por otro lado, sea una función $f(x) : \mathcal{C} \rightarrow \mathcal{C}$ que admite un desarrollo de Taylor de la forma $f(x) = \sum_{n \geq 0} f^{(n)}(x)|_{x=0} x^n / n!$. Entonces, dado un operador \hat{A} cualquiera, podemos definir al operador lineal $f(\hat{A})$ como

$$f(\hat{A}) = \sum_{n \geq 0} \frac{\hat{A}^n}{n!} f^{(n)}(x)|_{x=0}. \quad (2.19)$$

Por ejemplo, el operador $\exp(\hat{A}) = \sum_{n \geq 0} \hat{A}^n / n!$. Resulta inmediato comprobar que ambas definiciones coinciden para operadores que poseen una base ortonormal completa de autovectores. Sin embargo, para operadores que no posean una base ortonormal completa de autovectores sólo tendrá sentido la última definición.

Descomposición en valores singulares de una matriz

El siguiente es un resultado sumamente útil (y no demasiado conocido). Sea A una matriz de $N \times M$ entonces, existen matrices unitarias U (de $N \times N$) y V (de $M \times M$) y existe una matriz semidiagonal y positiva D (de $N \times M$) tal que $A = UDV$. La matriz D , si $N \leq M$ (lo cual se puede suponer sin pérdida de generalidad) es tal que

$$D = \begin{pmatrix} d_1 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & d_2 & 0 & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & d_N & \cdot & 0 \end{pmatrix}_{N \times M} \quad (2.20)$$

Esto puede demostrarse de la siguiente forma: Dada la matriz A , podemos construir dos operadores hermíticos de la siguiente manera $E = A^\dagger A$ y $F = AA^\dagger$. El carácter hermítico de ambos operadores surge inmediatamente de su definición. Además ambos operadores son positivos. Como ambos operadores son hermíticos entonces son diagonalizables. En particular, podemos escribir $F = UKU^\dagger$, donde K es una matriz de diagonal de $N \times N$ con elementos positivos en la diagonal. Ahora bien, cualquier matriz K de ese tipo puede escribirse como $K = DD^\dagger$, donde las matrices "diagonales" D son las definidas más arriba en la ecuación (2.20). En efecto, siempre podemos escribir $K = DD^\dagger$ donde los elementos no nulos de D son las raíces cuadradas de los elementos diagonales de K . De este modo podemos escribir $F = UDD^\dagger U^\dagger$. En esta expresión podemos introducir la matriz unitaria V , de $M \times M$ y su adjunta de modo tal que $F = UDVV^\dagger D^\dagger U^\dagger$. En consecuencia, podemos identificar $A = UDV$ donde U es la matriz que diagonaliza $F = AA^\dagger$. Razonando en forma análoga para el operador $E = A^\dagger A$, podemos ver que V es la matriz que diagonaliza E .

2.4. Spin 1/2

Para concluir, revisamos los conceptos introducidos aplicándolos al caso de los sistemas de spin 1/2. Como vimos anteriormente, podemos describir los estados de spin por medio de vectores de norma unidad que notaremos como:

$$|\Psi\rangle. \quad (2.21)$$

Combinaciones lineales de cualquier base de dos vectores también describirán algún otro estado. En particular para un spin 1/2 podremos escribir:

$$|\Psi\rangle = \alpha|0_z\rangle + \beta|1_z\rangle. \quad (2.22)$$

donde α y β son dos números complejos.

Los observables estarán representados por matrices de 2×2 que siempre pueden escribirse como combinaciones lineales de la identidad y las matrices de Pauli σ_x , σ_y y σ_z . De esta manera, cualquier observable para sistemas de spin 1/2 podrá escribirse como $\hat{A} = a_0 I + \vec{b} \cdot \vec{\sigma}$ y por lo tanto está definido por cuatro parámetros reales. Aunque en algunos casos resultará conveniente reescribir a \hat{A} como $\hat{A} = a_0 \left(I + \frac{\vec{b}}{a_0} \cdot \vec{\sigma} \right)$. Si definimos al versor $\vec{n} = \vec{b}/|\vec{b}|$, entonces cualquier operador puede escribirse como

$$\hat{A} = a_0 \left(I + \frac{|\vec{b}|}{a_0} \vec{n} \cdot \vec{\sigma} \right). \quad (2.23)$$

Este operador obviamente conmuta con $\vec{n} \cdot \vec{\sigma}$ y por lo tanto puede ser diagonalizado en la misma base que $\vec{n} \cdot \vec{\sigma}$. Esto implica que cualquier operador compatible con \hat{A} es, en definitiva, un múltiplo de $\vec{n} \cdot \vec{\sigma}$ mas un factor proporcional a la identidad. Por lo tanto, los únicos observables no triviales para un espín son las componentes del vector $\vec{\sigma}$ en alguna dirección.

Asimismo, resulta sencillo obtener los autovalores y autovectores de $\vec{n} \cdot \vec{\sigma}$ (tarea que haremos en detalle más adelante). En efecto, usando que $(\vec{n} \cdot \vec{\sigma})^2 = I$ es evidente que el proyector sobre los autoestados de autovalor ± 1 de $\vec{n} \cdot \vec{\sigma}$ son

$$\hat{P}_{\vec{n}, \pm} = \frac{1}{2} (I \pm \vec{n} \cdot \vec{\sigma}). \quad (2.24)$$