22 Formas bilineales complejas

Sea V un espacio vectorial sobre el cuerpo de los complejos $\mathbb{C}$. Una función $A : V \times V \to \mathbb{C}$ se dice que es una forma bilineal *hermítica* si

$$A(v_1 + v_2, w) = A(v_1, w) + A(v_2, w), \quad A(v, w_1 + w_2) = A(v, w_1) + A(v, w_2)$$

$$A(v, \alpha w) = \alpha A(v, w), \quad A(\alpha v, w) = \alpha^* A(v, w)$$

$\forall\ v, v_1, v_2, w, w_1, w_2 \in V$ y $\alpha \in \mathbb{C}$. Nótese que α sale como conjugado cuando está en el primer miembro.
Si V es de dimensión finita n y $e = (e_1, \ldots, e_n)$ es una base de V, escribiendo $v = \sum_{1=1}^{n} \alpha_i e_i$, $w = \sum_{j=1}^{n} \beta_j e_j$, con $[v]_e = (\alpha_1, \ldots, \alpha_n)^t$, $[w]_e = (\beta_1, \ldots, \beta_n)^t$, se obtiene

$$A(v, w) = \sum_{i,j=1}^{n} \alpha_i^* \beta_j A(e_i, e_j) = [v]_e^\dagger [A]_e [w]_e$$

donde el símbolo $\dagger$ denota traspuesto conjugado ($[v]_e^\dagger \equiv ([v]_e^t)^*$) y $[A]_e$ es la matriz de $n \times n$ de elementos

$$([A]_e)_{ij} = A(e_i, e_j)$$

Ejemplo: La siguiente es una forma bilineal de $\mathbb{C} \times \mathbb{C} \to \mathbb{C}$:

$$\begin{aligned} A(v, w) &= \alpha_1^* \beta_1 + (1+i)\alpha_1^* \beta_2 + (1-i)\alpha_2^* \beta_1 + 2\alpha_2^* \beta_2 \\ &= (\alpha_1^*, \alpha_2^*) \begin{pmatrix} 1 & 1+i \\ 1-i & 2 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \end{aligned}$$

donde hemos escrito en la base canónica $v = (\alpha_1, \alpha_2) = \alpha_1 e_1 + \alpha_2 e_2$, $w = (\beta_1, \beta_2) = \beta_1 e_1 + \beta_2 e_2$. A queda entonces representada en esta base por la matriz

$$[A]_e = \begin{pmatrix} 1 & 1+i \\ 1-i & 2 \end{pmatrix}$$

con $A(e_1, e_1) = 1$, $A(e_1, e_2) = 1 + i$, $A(e_2, e_1) = 1 - i$, $A(e_2, e_2) = 2$.

Si $A(v, w) = A(w, v)^*\ \forall v, w \Rightarrow$ la forma bilineal se dice que es hermíticamente simétrica y si $A(v, w) = -A(w, v)^*$, hermíticamente antisimétrica En el primer caso, la matriz que la representa es *hermítica*: $[A]_e^\dagger = [A]_e$, ya que $A(e_i, e_j) = A(e_j, e_i)^*$, y en el segundo caso *antihermítica*: $[A]_e^\dagger = -[A]_e$. Análogamente, si $[A]_e^\dagger = \pm[A]_e$, A es herm. simétrica (+) o antisimétrica (−). El ejemplo anterior corresponde a una forma bilineal herm. simétrica.
Notemos que una forma bilineal compleja que satisface las 4 condiciones no puede ser simplemente simétrica o antisimétrica a no ser que sea nula: Si $A(v, w) = \pm A(w, v)\ \forall\ v, w, \Rightarrow A(\alpha v, w) = \alpha^* A(v, w) = \pm A(w, \alpha v) = \pm\alpha A(w, v) = \alpha A(v, w)\ \forall\ \alpha,\ v, w$, lo que implica $(\alpha - \alpha^*)A(v, w) = 0$, es decir $A(v, w) = 0\ \forall\ v, w$.
Notemos también que toda forma bilineal puede expresarse como suma de una forma bilineal herm. simétrica y una forma bilineal herm. antisimétrica:

$$A(v, w) = \frac{1}{2}[A(v, w) + A(w, v)^*] + \frac{1}{2}[A(v, w) - A(w, v)^*]$$

22.1 Formas cuadráticas complejas

En forma análoga al caso real, la función $Q : V \to \mathbb{C}$ definida por

$$Q(v) = A(v, v)$$

se denomina *forma cuadrática* y satisface

$$Q(\alpha v) = A(\alpha v, \alpha v) = \alpha^* \alpha A(v, v) = |\alpha|^2 Q(v)$$

Una diferencia importante con las formas bilineales reales es que ahora la forma cuadrática determina completamente la forma bilineal (y no solamente la parte simétrica, como en el caso real). En efecto, podemos expandir $Q(v + w) = A(v + w, v + w)$ y $Q(v + iw) = A(v + iw, v + iw)$ como

$$Q(v + w) = Q(v) + Q(w) + A(v, w) + A(w, v),$$

$$Q(v+iw)=Q(v)+Q(w)+i[A(v,w)-A(w,v)]$$

de donde

$$\begin{array}{rcl} A(v,w) & = & Q(v+w)-iQ(v+iw)-(1-i)(Q(v)+Q(w)) \\ A(w,v) & = & Q(v+w)+iQ(v+iw)-(1+i)(Q(v)+QA(w)) \end{array}$$

De aquí se deduce también una propiedad fundamental:
$Q(v)$ es *real* $\forall\ v$ si y sólo si $A(v,w)=[A(w,v)]^*\ \forall\ v,w$, es decir, sii la forma bilineal asociada es hermíticamente simétrica.
En efecto, de las expresiones anteriores se ve que si $Q(v)\in\mathbb{R}\ \forall\ v\in V \Rightarrow A(w,v)=[A(v,w)]^*\ \forall\ v,w\in V$. Y si $A(w,v)=[A(v,w)]^*\ \forall\ v,w \Rightarrow Q(v)=A(v,v)=[A(v,v)]^*$ es real. Formas cuadráticas *reales* determinan pues *formas bilineales hermíticamente simétricas y viceversa.*
Por otro lado, si $A(v,w)$ es hermíticamente simétrica, $A'(v,w)=iA(v,w)$ es hermíticamente antisimétrica. La forma cuadrática asociada $Q'(v)=A'(v,v)$ es obviamente imaginaria.
Ejemplo: La forma bilineal del ej. anterior origina la forma cuadrática

$$\begin{array}{rcl} Q(v)=A(v,v)=(\alpha_1^*,\alpha_2^*)\begin{pmatrix} 1 & 1+i \\ 1-i & 2 \end{pmatrix}\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} & = & \alpha_1^*\alpha_1+2\alpha_2^*\alpha_2+(1+i)\alpha_1^*\alpha_2+(1-i)\alpha_2^*\alpha_1 \\ & = & |\alpha_1|^2+2|\alpha_2|^2+2\mathrm{Re}[(1+i)\alpha_1^*\alpha_2] \end{array}$$

que es obviamente *real*.

22.2 Cambio de base

Si efectuamos un cambio de base $e'_i=\sum_{j=1}^n S_{ji}e_j$, con $|S|\neq 0$,

$$A(e'_i,e'_j)=A(\sum_{k=1}^n S_{ki}e_k,\sum_{l=1}^n S_{lj}e_l)=\sum_{k,l}S_{ki}^*S_{lj}A(e_k,e_l)$$

por lo que

$$[A]_{e'}=S^\dagger[A]_eS$$

donde $\dagger$ denota por su puesto la operación de traspuesto+conjugado.
Notemos que $\mathrm{Det}([A]_{e'})=|\mathrm{Det}(S)|^2\mathrm{Det}([A]_e)$, por lo que la fase del determinante es la misma en cualquier base. Obtenemos entonces

$$A(v,w)=[v]_e^\dagger[A]_e[w]_e=[v]_{e'}^\dagger[A]_{e'}[w]_{e'}$$

donde $[w]_{e'}=R[w]_e$, $[v]_{e'}^\dagger=[v]_eR^\dagger$ y $R=S^{-1}$.

22.3 Base canónica: Si A es herm. simétrica, *existe una base* e' (base canónica) donde $[A]'_e$ es diagonal:

$$[A]_{e'}=S^\dagger[A]_eS=\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ & & \dots & \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

En esta base, si $v=\sum_{i=1}^n\alpha'_ie'_i$, $w=\sum_{i=1}^n\beta'_ie'_i$, tenemos

$$A(v,w)=\sum_{i=1}^n\lambda_i{\alpha'_i}^*\beta'_i$$

La demostración de la existencia de esta base puede efectuarse en forma similar al caso real, completando ahora módulos cuadrados, y se deja comos ejercicio. Sug.: Llamando $a_{ij}=([A]_e)_{ij}$ (con $a_{ji}=a_{ij}^*$) y asumiendo $a_{nn}\neq 0$, escribir la parte que contiene α_n y α_n^* en $A(v,v)$ como

$$a_{nn}\alpha_n^*\alpha_n+\sum_{j=1}^{n-1}(a_{nj}\alpha_n^*\alpha_j+a_{nj}^*\alpha_j^*\alpha_n)=a_{nn}{\alpha'_n}^*\alpha'_n-(\sum_{j=1}^{n-1}a_{nj}^*\alpha_j^*)(\sum_{j=1}^{n-1}a_{nj}\alpha_j)/a_{nn}$$

con $\alpha'_n=\alpha_n+\frac{1}{a_{nn}}\sum_{j=1}^{n-1}a_{nj}\alpha_j$, y proceder luego por inducción. Si $a_{nn}=0$ se comienza con una variable α_i tal que $a_{ii}\neq 0$, y si $a_{ii}=0\ \forall i$ se efectúa un cambio de variables simple para que a_{ii} sea no nulo para algún i (por ej., si $a_{ij}=a^*_{ji}\neq 0$, $a_{ij}\alpha^*_i\alpha_j+a_{ji}\alpha^*_j\alpha_i=2|a_{ij}|^2(|\alpha'_i|^2-|\alpha'_j|^2)$, con $\alpha_i=a_{ij}(\alpha'_i+\alpha_j)$, $\alpha'_j=\alpha_i-\alpha_j$. El cambio $\alpha'_i=\sum_{j=1}^n R_{ij}\alpha_j$ define una base $e'_i=\sum_{j=1}^n S_{ji}e_j$, con $S=R^{-1}$, en la que $[A]_{e'}=S^\dagger[A]S$ es diagonal.

Otra forma de demostrar la existencia es directamente diagonalizando la matriz $[A]_e$, que es en este caso hermítica y por lo tanto diagonalizable en una base ortonormal, tal que $S^{-1}=S^\dagger$ y $S^\dagger[A]_eS$ es diagonal. No obstante esto supone haber demostrado antes que tales matrices son diagonalizables, lo que nosotros realizaremos luego.

La base canónica no es única. Una base canónica puede obtenerse, al igual que en el caso real, completando módulos cuadrados o bien diagonalizando la matriz $[A]_e$.

Ejemplo: Hallar una base canónica para el ejemplo previo. Completando módulos cuadrados, obtenemos

$$Q(v)=(\alpha^*_1+(1-i)\alpha^*_2)(\alpha_1+(1+i)\alpha_2)+|\alpha_2|^2[2-(1+i)(1-i)]=|\alpha'_1|^2+0|\alpha'_2|^2$$

donde $\alpha'_1=\alpha_1+(1+i)\alpha_2$, $\alpha'_2=\alpha_2$, o sea $\binom{\alpha'_1}{\alpha'_2}=\left(\begin{smallmatrix}1&1+i\\0&1\end{smallmatrix}\right)\binom{\alpha_1}{\alpha_2}$. La matriz de cambio de base es entonces

$$S=\begin{pmatrix}1&1+i\\0&1\end{pmatrix}^{-1}=\begin{pmatrix}1&-1-i\\0&1\end{pmatrix}$$

y se verifica

$$[A]_{e'}=S^\dagger[A]_eS=\begin{pmatrix}1&0\\0&0\end{pmatrix}$$

Alternativamente, diagonalizando la matriz $[A]_e$ se obtienen los autovalores y autovectores

$$\lambda_1=1,\ v'_1=(1+i,2),\qquad \lambda_2=0,\ v_2=(-1-i,1)$$

Normalizando los autovectores, la matriz de cambio de base es entonces

$$S=\begin{pmatrix}(1+i)/\sqrt{6}&-(1+i)/\sqrt{3}\\2/\sqrt{6}&1/\sqrt{3}\end{pmatrix}$$

con $S^{-1}=S^\dagger$ (pues los autovectores en S están normalizados). Se obtiene así la reprentación diagonal

$$[A]_{e'}=S^\dagger[A]_eS=\begin{pmatrix}3&0\\0&0\end{pmatrix}$$

Vemos que el número de coeficientes diagonales positivos y nulos en las dos formas diagonales obtenidas es el mismo. Esta propiedad es general y constituye el

22.4 Teorema de Inercia para formas cuadráticas hermíticas: Si QA es una forma cuadrática herm. simétrica, el número de términos diagonales positivos, negativos, y nulos en una representación diagonal arbitraria es siempre el mismo. Se demuestra igual que en el caso real (Demostrar como ejercicio).
Es importante notar que el teorema de inercia no vale para formas cuadráticas comunes extendidas a los complejos: Si $Q(v)=\alpha_1^2+\alpha_2^2$, la transformación $\alpha'_1=i\alpha_1$, $\alpha'_2=\alpha_2$ la lleva a $-\alpha'^2_1+\alpha'^2_2$. Tal forma cuadrática no proviene de una forma bilineal hermítica, ya que no cumple $Q(\alpha v)=|\alpha|^2Q(v)$.

22.5 Formas cuadráticas positivas:
Una forma cuadrática se denomina *definida positiva* (o estríctamente positiva) si $Q(v)>0\ \forall\ v\neq 0$, y semipositiva (o no negativa) si $Q(v)\geq 0\ \forall\ v$ (obviamente, en cualquier caso, $Q(0)=0$). Por ser reales, estas formas cuadráticas están necesariamente asociadas a formas bilineales hermíticamente simétricas.
La forma cuadrática es pues definida positiva si y sólo si los coeficientes diagonales en una base canónica satisfacen $a_{ii}>0\ \forall\ i$ y semipositiva sii $a_{ii}\geq 0\ \forall\ i$. El teorema de inercia asegura que el número de coeficientes diagonales positivos y nulos para estas formas cuadráticas (pero no su valor particular) es siempre el mismo en cualquier base canónica.
En form análoga, una matriz $A\in\mathbb{C}^{n\times n}$ se dice definida positiva si $X^\dagger AX>0\ \forall\ X\neq 0$, $X\in\mathbb{C}^n$, en cuyo

caso podemos considerarla como la representación en la base canónica de $V = \mathbb{C}^n$ de una forma cuadrática definida positiva. Notemos que necesariamente A debe ser *hermítica* ($A^\dagger = A$), para que $X^\dagger AX$ sea real. **Una matriz hermítica A es pues definida positiva si y sólo si todos sus autovalores son positivos.**

Notemos que si $Q_A(v)$ es una forma cuadrática definida positiva $\Rightarrow$ *existe una base canónica e'' donde* $A(e''_i, e''_j) = \delta_{ij}$, es decir,

$$[A]_{e''} = I_n$$

(matriz identidad). En efecto, existirá una base canónica, obtenida completando módulos cuadrados o diagonalizando, en la que $A(e'_i, e'_j) = ([A]_{e'})_{ij} = \lambda_i \delta_{ij}$, con $\lambda_i > 0 \ \forall\ i$. En la nueva base definida por $e''_i = e'_i/\sqrt{\lambda_i}$ tendremos $A(e''_i, e''_j) = A(e'_i, e'_j)/\sqrt{\lambda_i\lambda_j} = \lambda_i\delta_{ij}/\sqrt{\lambda_i\lambda_j} = \delta_{ij}$ para $i, j = 1, \ldots, n$.
Esto implica que existe una matriz $S = [I]_e^{e''}$ no singular tal que

$$[A]_{e''} = S^\dagger [A]_e S = I_n$$

con I_n la matriz identidad.
Esto implica a su vez que toda matriz $A \equiv [A]_e$ definida positiva puede escribirse como

$$A = R^\dagger R$$

con $R = S^{-1}$ no singular. La recíproca es también válida: La matriz $R^\dagger R$ es definida positiva $\forall\ R$ no singular (probar como ejercicio).

Para saber si una forma cuadrática es definida positiva o semipositiva se completan módulos cuadrados o se obtienen los autovalores de la matriz que la representa en alguna base, y se observan lo signos de los coeficientes diagonales resultantes. El determinante de una matriz asociada a una forma cuadrática definida positiva es obviamente positivo en cualquier base ($\mathrm{Det}[A] = \mathrm{Det}[R^\dagger R] = |\mathrm{Det}[R]|^2$), aunque esta condición no garantiza que A sea definida positiva.

Es válido no obstante el siguiente teorema: **Una matriz hermítica A de $n \times n$ es definida positiva si y sólo si todos sus determinantes principales son positivos ($\mathrm{Det}(A_m) > 0$, $m = 1, \ldots, n$).**
Dem.: Si A es definida positiva la forma cuadrática asociada debe ser positiva en cualquier subespacio de V y por lo tanto cualquier submatriz de A (obtenida quitando un conjunto de filas y las resp. columnas) es definida positiva. En particular, todas las submatrices principales ($A_m = [a_{ij}], i, j = 1, \ldots, m$, $m = 1, \ldots, n$) son definidas positivas y sus determinantes por ende positivos.
Recíprocamente, si todos los determinantes principales son positivos, procedemos por inducción sobre n. Para $n = 1$ se cumple trivialmente. Asumiendo la submatriz principal de $(n-1) \times (n-1)$ definida positiva, existirá una base $(e'_1, \ldots, e'_{n-1})$ del subespacio correspondiente en la que $A(e'_i, e'_j) = \delta_{ij}$. Definimos ahora $e'_n = e_n - \sum_{i=1}^{n-1} \alpha_i e'_i$, con $\alpha_i = A(e'_i, e_n)$, tal que $A(e'_i, e'_n) = A(e'_i, e_n) - \alpha_i = 0$ para $i = 1, \ldots, n-1$. En tal caso, $[A]_{e'}$ será diagonal y con todos sus elementos diagonales positivos, pues $A(e'_i, e'_i) = 1$ para $i < n$ y $A(e'_n, e'_n) = \mathrm{Det}([A]_{e'}) > 0$ por hipótesis (el signo del determinante no cambia al cambiar la base).

Recordemos también aqui los **círculos de Gershgorin:** Si $A \in \mathbb{C}^{n\times n}$ es una matriz cuadrada de elementos a_{ij}, entonces sus autovalores se encuentran en la unión de los círculos (en el plano complejo)

$$|\lambda - a_{ii}| \leq |\sum_{j\neq i} |a_{ij}|$$

En efecto, sea $v = (x_1, \ldots, x_n)^t \in \mathbb{C}^{n\times 1}$, $v \neq 0$, un autovector de A asociado al autovalor λ, tal que $Av = \lambda v$, es decir, $\sum_j a_{ij}x_j = \lambda x_i$. Si $|x_i| = \mathrm{Max}[|x_1|, \ldots, |x_n|] \neq 0$ es el módulo de la coordenada de v de módulo máximo, tenemos, dado que $(\lambda - a_{ii})x_i = \sum_{j\neq i} a_{ij}x_j$,

$$|\lambda - a_{ii}| = |\sum_{j\neq i} a_{ij}x_j/x_i| \leq \sum_{j\neq i} |a_{ij}|x_j/x_i| \leq \sum_{j\neq i} |a_{ij}|$$

Una desigualdad simular es válida para sumas sobre columnas, ya que los autovalores de A son idénticos a los de A^t.

En el caso de matrices hermíticas, tanto los elementos diagonales como los autovalores son todos reales. La cota anterior implica entonces la siguiente condición suficiente (aunque no necesaria) de positividad de una matriz hermítica A: Si $a_{ii} > 0\ \forall\ i$ y $\sum_{j\neq i} |a_{ij}| < a_{ii}\ \forall\ i$, los autovalores serán todos positivos y por ende A será definida positiva.

23 Espacios Unitarios (Espacios de Hilbert)

Un espacio vectorial V sobre $\mathbb{C}$ se denomina unitario o espacio de Hilbert si está equipado con una operación $V\times V\to\mathbb{C}$, denominada *producto interno* o producto escalar, y denotada por (v,w), que satisface

$$(v,w)=(w,v)^*,\quad (v,\alpha w)=\alpha(v,w),\quad (v,w_1+w_2)=(v,w_1)+(v,w_2)$$

$$(v,v)>0\;\forall\; v\neq 0$$

Es decir, el producto interno no es otra cosa que una forma bilineal hermíticamente simétrica y *definida positiva*. En el caso de dimensión infinita, un espacio de Hilbert debe ser además completo: Si $\{u_n\}$ es una sucesión de vectores tal que $\sum_{n=0}^{\infty}||u_n||$ es convergente entonces $\lim_{n\to\infty}u_n$ debe pertenecer al espacio.

En en el caso de dimensión finita, en una base arbitraria e tendremos, denotando con $[A]_e$ la matriz de elementos $a_{ij}=(e_i,e_j)=a_{ji}^*$,

$$(v,w)=[v]_e^\dagger[A]_e[w]_e=\sum_{i,j=1}^{n}\alpha_i^*a_{ij}\beta_j$$

donde $v=\sum_{i=1}^n\alpha_ie_i$, $w=\sum_{i=1}^n\beta_ie_i$
Y si e denota ahora la base canónica en la que $(e_i,e_j)=\delta_{ij}$, obtenemos la forma corriente

$$(v,w)=[v]_e^\dagger[w]_e=\sum_{i=1}^{n}\alpha_i^*\beta_i$$

Esta base es una *base ortonormal* para el producto escalar ($(e_i,e_j)=\delta_{ij}$). En esta base,

$$(v,v)=[v]_e^\dagger[v]_e=\sum_{i=1}^{n}|\alpha_i|^2$$

Ejemplo: En $\mathbb{C}^n$, el producto interno usual en la base canónica está dado por

$$(v,w)=\sum_{i=1}^{n}x_i^*y_i,\quad (w,v)=\sum_{i=1}^{n}y_i^*x_i=(v,w)^*$$

para $v=(x_1,\ldots,x_n)$, $w=(y_1,\ldots,w_n)$, lo que implica $(v,v)=\sum_{i=1}^n|x_i|^2$.

Ejemplo: En el espacio de funciones complejas de parte real e imaginaria continua, $C_{[a,b]}=\{f:\mathbb{R}\to\mathbb{C},\;\; f=f^r+if^i,\;f^r,f^i\in R_{[a,b]}\}$, con $a<b$, el producto interno usual está dado por

$$(f,g)=\int_a^b f^*(x)g(x)dx=(g,f)^*$$

con $(f,f)=\int_a^b f^*(x)f(x)dx=\int_a^b|f(x)|^2dx>0$ si $f\neq 0$.
Ejemplo: En el caso de matrices complejas de $m\times n$, $V=\mathbb{C}^{m\times n}$, podemos definir el producto escalar

$$(A,B)=\mathrm{Tr}\,[A^\dagger\,B]=\sum_{i=1}^{m}\sum_{j=1}^{n}A_{ij}^*B_{ij}$$

con $(A,A)=\sum_{i,j}|A_{ij}|^2>0\;\forall\;A\neq 0$.

En los espacios unitarios son válidas propiedades similares a las de espacios euclídeos. En particular:

La **norma** de un vector se define por

$$||v||=\sqrt{(v,v)}\geq 0$$

con $||v||=0$ sii $v=0$ y $||\alpha v||=|\alpha|\,||v||$. La distancia entre dos vectores es $d(v,w)=||v-w||$.

La desigualdad de **Cauchy-Schwarz** también se verifica:

$$|(v,w)|\leq ||v||\,||w||$$

donde la igualdad vale si y sólo si $\{v, w\}$ son LD.
Demostración: Si $v = 0$ o $w = 0$ la igualdad se cumple trivialmente: $0 = (v, w) = ||v|| \, ||w||$.
Idem si v y w son LD: En tal caso $w = \alpha v$ (o $v = \alpha w$) y por lo tanto $|(v, w)| = |\alpha(v, v)| = |\alpha| \, ||v||^2 = ||v|| \, ||w||$.
Si $v \neq 0$ y $w \neq 0$, denotemos con $v_n = v/||v||$, $w_n = w/||w_n||$ los vectores normalizados ($||v_n|| = ||w_n|| = 1$), tal que $(v, w) = (v_n, w_n)||v|| \, ||w||$. Se obtiene, para s un número complejo arbitrario de módulo 1 ($|s| = 1$),

$$0 \leq (v_n - sw_n, v_n - sw_n) = ||v_n||^2 + |s|^2 ||w_n||^2 - s(v_n, w_n) - s^*(w_n, v_n) = 2 - 2\mathrm{Re}[s(v_n, w_n)] = 2(1 - \mathrm{Re}[s(v_n, w_n)])$$

Recordemos ahora que todo número complejo z puede escribirse como $z = |z|e^{i\phi}$, con $|z| = \sqrt{zz^*}$ (módulo) y ϕ reales. Por lo tanto, si $z = (v_n, w_n) = |(v_n, w_n)|e^{i\phi}$, eligiendo $s = e^{-i\phi}$ se obtiene

$$0 \leq 1 - |(v_n, w_n)|$$

de donde $|(v_n, w_n)| \leq 1$. Por lo tanto, $|(w, v)| = |(v, w)| \leq ||v|| \, ||w||$, q.e.d.
Además, si $|(w, v)| = 1 \Rightarrow |(w_n, v_n)| = 1$ y $(v_n - sw_n, v_n - sw_n) = 0$, por lo que $v_n - sw_n = 0$, es decir, $v = sw||v_n||/||w_n||$, lo que implica que v, w son L.D.
Las desigualdades triangulares *permanecen válidas* en espacios unitarios, por la vigencia de la desigualdad anterior: $|||v|| - ||w||| \leq ||v + w|| \leq ||v|| + ||w||$.
No obstante, no se pueden definir ahora ángulos entre vectores pues $(v, w)/(||v|| \, ||w||)$, si bien tiene módulo menor que 1, es en general complejo.
Ejemplo: Dados $v = (1 + i, i), w = (i, 1 + i) \in \mathbb{C}^2$, tenemos

$$(v, w) = (1 - i)i + (-i)(1 + i) = 2 \leq ||v|| \, ||w|| = \sqrt{|1 + i|^2 + 1}\sqrt{1 + |1 + i|^2} = \sqrt{3}\sqrt{3} = 3$$

Notación de Mecánica Cuántica:
La notación empleada en mecánica cuántica para los vectores de estado de un sistema (que pertenecen a un espacio de Hilbert) es $|v\rangle$, y para el producto interno $\langle w|v\rangle$. Es decir, $v \to |v\rangle$, $(w, v) \to \langle w|v\rangle$, con $\langle v|w\rangle = \langle w|v\rangle^*$.

23.1 Ortogonalidad y Método de ortogonalización Gram-Schmidt

Las propiedades de ortogonalidad son análogas al caso euclídeo. Dos vectores v, w de un espacio unitario son **ortogonales** si $(v, w) = 0$.
Al igual que en el caso euclídeo, dado un conjunto de m vectores v_i L.I., es posible construir con el método de Gram-Schmidt un conjunto ortogonal de vectores que genera el mismo espacio que los v_i, dados por:

$$w_1 = v_1\,, \quad w_i = v_i - \sum_{j=1}^{i-1} P_{w_j}(v_i)\,, \quad i = 2, \ldots, m$$

donde

$$P_{w_j}(v_i) = \frac{(w_j, v_i)}{||w_j||^2} w_j$$

es la *proyección ortogonal* de v_i sobre w_j. Notemos que en el caso complejo es necesario ser cuidadoso con el orden en el producto escalar, ya que $(w_j, v_i) \neq (v_i, w_j) = (w_j, v_i)^*$. Es fácil verificar que de esta forma, $(w_i, w_j) = 0$ si $i \neq j$, siendo los w_i no nulos si los vectores originales son L.I.
Dada una base arbitraria de V, es pues siempre posible por este método construir una base ortogonal de V, que puede convertirse en *ortonormal* normalizando los vectores resultantes.
Notemos que el cuadrado de la norma de los w_i está dado, para $i > 1$, por

$$||w_i||^2 = (w_i, w_i) = (v_i, w_i) = ||v_i||^2 - \sum_{j=1}^{i-1} \frac{|(w_j, v_i)|^2}{||w_j||^2} \leq ||v_i||^2$$

Notemos también que la matriz que representa al proyector sobre w_i en la base canónica es

$$[P_{w_i}]_e = \frac{[w_i]_e [w_i]_e^\dagger}{||w_i||^2}$$

Ejemplo 1 : Consideremos los vectores $v_1 = (1+i, i, 0)$, $v_2 = (i, 1+i, 1)$. Tenemos

$$w_1 = v_1 = (1+i, i, 0), \quad w_2 = v_2 - \frac{(w_1, v_2)}{||w_1||^2} w_1 = (i, 1+i, 1) - \frac{2}{3}(1+i, i, 0) = (-2+i, 3+i, 3)/3$$

que verifican $(w_1, w_2) = 0$.

Ejemplo 2: Las funciones $f_k(x) = e^{ikx}$, con k *entero*, son ortogonales con el producto interno $(f, g) = \int_{-\pi}^{\pi} f^*(x) g(x) dx$:

$$(f_{k'}, f_k) = \int_{-\pi}^{\pi} e^{-ik'x} e^{ikx} dx = \int_{-\pi}^{\pi} e^{ix(k-k')} dx = \begin{cases} 2\pi & k = k' \\ \frac{e^{ix(k-k')}}{i(k-k')}|_{-\pi}^{\pi} = 0 & k \neq k' \end{cases}$$

Ejemplo 3 (Transformada de Fourier discreta): Sea $V = \mathbb{C}^n$ y sea $e = (e_1, \ldots, e_n)$ una base canónica ($(e_i, e_j) = \delta_{ij}$). Los n vectores

$$\tilde{e}_k = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} e^{i2\pi kj/n} e_j$$

forman también una base ortonormal: $(\tilde{e}_k, \tilde{e}_l) = \delta_{kl}$.
En efecto, utilizando que $(e_i, e_j) = \delta_{ij}$ obtenemos, para $k, l = 1, \ldots, n$,

$$(\tilde{e}_k, \tilde{e}_l) = \frac{1}{n} \sum_{j=1}^{n} e^{i2\pi j(l-k)/n} = \begin{cases} 1 & k = l \\ \frac{1}{\sqrt{n}} \frac{1 - e^{i2\pi(l-k)}}{1 - e^{i2\pi(l-k)/n}} = 0 & k \neq l \end{cases}$$

Ejemplo 4: Obtener una base ortonormal de $\mathbb{C}^{2\times 2}$ (con escalares complejos) para el producto escalar $(A, B) = \mathrm{Tr}\, A^\dagger B$, partiendo de $v_1 = I_2 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$.
Consideremos las matrices $v_1 = I_2$, $v_2 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$, $v_3 = \frac{1}{2}\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, $v_4 = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, que forman una base no ortogonal de $\mathbb{C}^{2\times 2}$. Obtenemos, $w_1 = v_1 = I_2$,

$$w_2 = v_2 - \tfrac{1}{2}(w_1, v_2) w_1 = v_2 - \tfrac{1}{2} w_1 = \tfrac{1}{2}\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right), \quad w_3 = v_3 - \tfrac{1}{2}(w_1, v_3) w_1 - 2(w_2, v_3) w_2 = v_3 = \tfrac{1}{2}\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right),$$

$$w_4 = v_4 - \tfrac{1}{2}(w_1, v_4) w_1 - 2(w_2, v_3) w_2 - 2(w_3, v_4) w_3 = v_4 - w_3 = \tfrac{1}{2}\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$$

Las matrices de Pauli se definen precisamente como

$$\sigma_0 = I_2 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right), \quad \sigma_x = 2w_3 = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right), \quad \sigma_y = -2iw_4 = \left(\begin{smallmatrix} 0 & -i \\ i & 0 \end{smallmatrix}\right), \quad \sigma_z = 2w_2 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$$

y forman una base ortogonal y hermítica de $\mathbb{C}^{2\times 2}$: $\sigma_\mu^\dagger = \sigma_\mu$, $(\sigma_\mu, \sigma_\nu) = \mathrm{Tr}\, \sigma_\mu \sigma_\nu = 2\delta_{\mu\nu}$.

Considerando ahora $\mathbb{C}^{2\times 2}$ sobre escalares reales, estas 4 matrices forman también una base del subespacio de matrices hermíticas de 2×2. Las matrices que representan las componentes del espín $\mathbf{s} = (s_x, s_y, s_z)$ en la base estándar de autoestados de s_z son precisamente

$$s_\mu = \frac{1}{2} \hbar \sigma_\mu, \quad \mu = x, y, z$$

23.2 Expansión en una base ortonormal

Si $e = (e_1, \ldots, e_n)$ es una base ortonormal de V ($(e_i, e_j) = \delta_{ij}$) y

$$v = \sum_{i=1}^{n} x_i e_i$$

entonces $(e_i, v) = \sum_{j=1}^{n} x_j (e_i, e_j) = x_j$. Por lo tanto,

$$x_i = (e_i, v)$$

Se cumple entonces

$$v = \sum_{i=1}^{n} P_{e_i}(v)$$

Se verifica también, por la ortogonalidad de los e_i, la generalización del teorema de Pitágoras,

$$||v|| = (v, v) = \sum_{i=1}^{n} ||x_i e_i||^2 = \sum_{i=1}^{n} |x_i|^2$$

Notemos que en la notación de mecánica cuántica, $e_i \to |i\rangle$ y $|v\rangle = \sum_i \alpha_i |i\rangle$, con $\alpha_i = \langle i|v\rangle$.

23.3 Proyectores ortogonales y matriz de Gram

Dado un subespacio $S \subset V$, es posible construir el complemento ortogonal $S_\perp = \{v \in V | (w, v) = 0 \; \forall \; w \in S\}$, cumpliéndose que $V = S \oplus S_\perp$ y por lo tanto, dim $S+$ dim $S_\perp = n$
Si $v \in V$, podemos escribir

$$v = v_s + (v - v_s)$$

con $v_s \in S$ y $v - v_s \in S_\perp$. Si $(w_1, \ldots, w_m)$ es una base *ortogonal* de S, escribiendo $v_s = \sum_{i=1}^m \alpha_i w_i$, la condición $(w_i, v - v_s) = 0$ para $i = 1, \ldots, m$ implica $\alpha_i = (w_i, v)/||w_i||^2$ y por lo tanto

$$v_s = \sum_{i=1}^m P_{w_i}(v) = P_S(v)\,, \qquad P_S = \sum_{i=1}^m P_{w_i}$$

El vector v_s es el vector de S con distancia mínima a v: Si $w_s \in S$,

$$||v - w_s||^2 = ||(v - v_s) + (v_s - w_s)||^2 = ||v - v_s||^2 + ||w_s - v_s||^2 \geq ||v - v_s||^2$$

En general, para una base arbitraria $(w_i, \ldots, w_m)$ de S no necesariamente ortogonal,

$$[P_S]_e = R(R^\dagger R)^{-1} R^\dagger$$

donde $R = ([w_1]_e, \ldots, [w_m]_e)$ es la matriz de $n \times m$ donde cada columna son las coordenadas de los m vectores w_i de la base de S en una base canónica de V $((e_i, e_j) = \delta_{ij})$. Las fórmulas del caso euclídeo se generalizan pues directamente al caso unitario reemplazando t (traspuesta) por $\dagger$ (traspuesto conjugado).
Recordemos que

$$P_S + P_{S_\perp} = I$$

Notemos también que la matriz de Gram

$$G = R^\dagger R$$

de $m \times m$, con $G_{ij} = (w_i, w_j) = G^*_{ji}$, es ahora una matriz *hermítica*, que posee las mismas propiedades anteriores: $|G| \neq 0$ sii los m vectores w_i son LI, los autovalores λ_i de G son reales y no negativos, los autovectores X_i $(GX_i = \lambda_i X_i)$ correspondientes a autovalores no nulos determinan vectores u_i de componentes $[u_i]_e = RX_i$, que son ortogonales $((u_i, u_j) = 0$ si $\lambda_i \neq \lambda_j)$, y aquellos correspondientes a autovalores nulos dan las combinaciones lineales nulas de los w_i (Demostraciones totalmente similares al caso euclídeo).

Ejemplo: Proyectar el vector $v = (1, i, 1+i) \in \mathbb{C}^3$ sobre el espacio generado por los vectores $v_1 = (1+i, i, 0)$, $v_2 = (i, 1+i, 1)$. Aplicando la representación general, tenemos

$$R = \begin{pmatrix} 1+i & i \\ i & 1+i \\ 0 & 1 \end{pmatrix}$$

con $R^\dagger R = \left(\begin{smallmatrix} 3 & 2 \\ 2 & 4 \end{smallmatrix}\right)$, $(R^\dagger R)^{-1} = \left(\begin{smallmatrix} 4 & -2 \\ -2 & 3 \end{smallmatrix}\right)/8$ y

$$[P_S]_e = R(R^\dagger R)^{-1} R^\dagger = \begin{pmatrix} 7 & 1-i & -2+i \\ 1+i & 6 & 3+i \\ -2-i & 3-i & 3 \end{pmatrix} /8$$

Podemos arribar a este mismo resultado considerando también la base ortogonal de S obtenida previamente al ortogonalizar v_1 y v_2 por Gram-Schmidt, dada por $w_1 = v_1$, $w_2 = (-2+i, 3+i, 3)/3$:

$$[P_S]_e = [P_{w_1}]_e + [P_{w_2}]_e = \frac{[w_1]_e [w_1]_e^\dagger}{||w_1||^2} + \frac{[w_2]_e [w_2]_e^\dagger}{||w_2||^2}$$

$$= \frac{1}{3} \begin{pmatrix} 1+i \\ i \\ 0 \end{pmatrix} (1-i, -i, 0) + \frac{1}{24} \begin{pmatrix} -2+i \\ 3+i \\ 3 \end{pmatrix} (-2-i, 3-i, 3) = \begin{pmatrix} 7 & 1-i & -2+i \\ 1+i & 6 & 3+i \\ -2-i & 3-i & 3 \end{pmatrix} /8$$

Se obtiene finalmente

$$[P_S(v)]_e = [P_S]_e [v]_e = \begin{pmatrix} 5 \\ 3+11i \\ 2+5i \end{pmatrix} /8$$

La distancia mínima al plano es $||v - v_s|| = 3/\sqrt{8}$.

23.4 Operadores adjuntos y autoadjuntos en espacios unitarios

Sea $F : V \to V$ un operador lineal. El operador adjunto $F^\dagger$ se define por la relación

$$(v, F(w)) = (F^\dagger(v), w)$$

$\forall\ v, w \in V$. Considerando una base canónica e de V ($(e_i, e_j) = \delta_{ij}$), y teniendo en cuenta que $[F(v)]_e = [F]_e[v]_e$, y $(v, w) = [v]_e^\dagger[w]_e$, se obtiene $(v, F(w)) = [v]_e^\dagger[F]_e[w]_e$, $(F^\dagger(v), w) = [v]_e^\dagger[F^\dagger]_e^\dagger[w]_e$ y por lo tanto

$$[F^\dagger]_e = [F]_e^\dagger$$

La matriz que representa al operador adjunto de F en una base canónica es pues la traspuesta conjugada de la que representa a F en dicha base. Notemos que:
1) si $G = \alpha F$, con $\alpha \in \mathbb{C} \Rightarrow G^\dagger = \alpha^* F^\dagger$ (pues $(\alpha^* F^\dagger(v), w) = \alpha(F^\dagger(v), w) = \alpha(v, F(w)) = (v, \alpha F(w))$)
2) $(F^\dagger)^\dagger = F$ (pues $((F^\dagger)^\dagger(v), w) = (v, F^\dagger(w)) = (F(v), w)\ \forall\ v, w$)
3) $(FG)^\dagger = G^\dagger F^\dagger$ (pues $(v, FG(w)) = (F^\dagger(v), G(w)) = (G^\dagger F^\dagger(v), w)$).

Un operador F es *autoadjunto* si $F^\dagger = F$. En tal caso la matriz que lo representa en una base canónica es *hermítica*:

$$[F]_e^\dagger = [F]_e$$

Una propiedad importante de operadores adjuntos es que **si S es un subespacio invariante por $F \Rightarrow$ $S_\perp$ es invariante por $F^\dagger$.**
Demostración: si $F(v) \in S\ \forall\ v \in S$, y $w \in S_\perp \Rightarrow (w, F(v)) = 0\ \forall\ w \in S_\perp$ y $v \in S$. Por lo tanto,

$$(F^\dagger(w), v) = (w, F(v)) = 0$$

$\forall\ w \in S_\perp$ y $v \in S$, de modo que $F^\dagger(w) \in S_\perp$
En particular, si F es *autoadjunto* y S es invariante por $F \Rightarrow S_\perp$ *es también invariante por* F.

Comentemos finalmente que en una base B general, donde $(v, w) = [v]_B^\dagger A[w]_B$ con A es una matriz hermítica definida positiva ($A^\dagger = A$, $X^\dagger AX > 0\ \forall\ X \neq 0$, $X \in \mathbb{C}^{n\times 1}$) la condición $(v, F(w)) = (F^\dagger(v), w)$ $\forall$ $v, w \in V$ implica $[F^\dagger]_B^\dagger A = A[F]_B$, y por lo tanto,

$$[F^\dagger]_B = A^{-1}[F]_B^\dagger A$$

La matriz que representa el operador adjunto $F^\dagger$ en una base arbitraria es pues semejante (pero no necesariamente igual) a $[F]_B^\dagger$.

23.5 Operadores Unitarios

Un operador lineal $U : V \to V$ que conserva el producto interno en un espacio unitario se denomina unitario:

$$(U(v), U(w)) = (v, w)$$

$\forall\ v, w \in V$. Como $(U(v), U(w)) = (U^\dagger U(v), w) \Rightarrow U^\dagger U = I$ (identidad), por lo que en una base canónica tenemos

$$[U]_e^\dagger[U]_e = I_n$$

y por lo tanto $[U]_e[U]_e^\dagger = I_n$. Las matrices que representan a un operador unitario en una base canónica se denominan unitarias y satisfacen $[U]_e^{-1} = [U]_e^\dagger$, lo que implica filas y columnas ortonormales:

$$\sum_{j=1}^{n} S_{ji}^* S_{jk} = \delta_{ik} \quad \sum_{j=1}^{n} S_{ij} S_{kj}^* = \delta_{ik}$$

donde aquí $S_{ij} = ([U]_e)_{ij}$. El determinante de un operador unitario tiene módulo 1:

$$1 = \mathrm{Det}[U^\dagger U] = \mathrm{Det}[U]^*\mathrm{Det}[U] = |\mathrm{Det}[U]|^2$$

por lo que

$$|\mathrm{Det}[U]| = 1$$

Podemos entonces escribir $\mathrm{Det}[U]=e^{i\phi}$, con ϕ real.
Debe remarcarse que los operadores unitarios transforman bases ortonormales en bases ortonormales: si $e'_i=U(e_i)=\sum_{j=1}^m S_{ji}e_j$, $i=1,\ldots,n \Rightarrow$

$$(e'_i,e'_j)=(U(e_i),U(e_j))=(e_i,e_j)=\delta_{ij}$$

Análogamente, cualquier par de bases ortonormales e,e' de V están relacionadas por una transformación unitaria, es decir, por una matriz de cambio de base S que satisface $S^\dagger S=SS^\dagger=I_n$, como es fácil verificar: Si $e'_i=\sum_j S_{ji}e_j$ y $(e'_i,e'_j)=(e_i,e_j)=\delta_{ij}$ entonces

$$(e'_i,e'_j)=\sum_{k,l}(S_{ki}e_i,S_{lj}e_l)=\sum_{k,l}S^*_{ki}S_{lj}(e_k,el)=\sum_k S^\dagger_{ik}S_{kj}=(S^\dagger S)_{ij}=\delta_{ij}$$

Remarquemos también que el *producto* (pero no la suma) de operadores unitarios es *unitario*: Si U,W son unitarios $\Rightarrow (UW)^{-1}=W^{-1}U^{-1}=W^\dagger U^\dagger=(UW)^\dagger$, por lo que UW es unitario. Está propiedad es también obvia a partir de la definición.

24. Autovalores y Autovectores de operadores autoadjuntos

1) Si $F:V\to V$ es un operador lineal autoadjunto $\Rightarrow$ sus autovalores son todos reales y los autovectores correspondientes a autovalores distintos son ortogonales.
Demostración: Si $F(v)=\lambda v$,

$$(v,F(v))=(v,\lambda v)=\lambda(v,v)$$

pero por ser F autoadjunto,

$$(v,F(v))=(F(v),v)=(\lambda v,v)=\lambda^*(v,v)$$

por lo que

$$(\lambda-\lambda^*)(v,v)=0$$

lo que implica, si $v\neq 0$, $\lambda-\lambda^*=0$, es decir, λ *real*. Todos los autovalores de F serán pues *reales*.
Además, si $F(v)=\lambda v$ y $F(v')=\lambda' v'$, entonces

$$(v',F(v))=\lambda(v',v)=(F(v'),v)=\lambda'(v',v)$$

por lo que

$$(v',v)(\lambda-\lambda')=0$$

lo que implica

$$(v',v)=0 \text{ si } \lambda\neq\lambda'$$

2) Si $F:V\to V$ es un operador lineal autoadjunto en un espacio V de dimensión finita, existe siempre una base *ortonormal* de V formada por autovectores de F: $\exists\ e'=(e'_1,\ldots,e'_n)$, tal que

$$F(e'_i)=\lambda_i e'_i,\ \ i=1,\ldots,n,\ \ (e'_i,e'_j)=\delta_{ij}$$

Es decir, F *es siempre diagonalizable* y además lo es en una base ortonormal, la cual estará relacionada con la base canónica original por una transformación unitaria U:

$$[F]_{e'}=S^\dagger[F]_eS=\begin{pmatrix}\lambda_1 & 0 & \ldots & 0\\ 0 & \lambda_2 & \ldots & 0\\ & & \ldots & \\ 0 & 0 & \ldots & \lambda_n\end{pmatrix},\quad S^\dagger S=SS^\dagger=I$$

con $S=[U]_e$ y $e'_i=U(e_i)$.
Demostración: Por inducción sobre n. Para $n=1$ todo F es trivialmente diagonal en cualquier base. Considerando ahora $n>1$, si los n autovalores de F son todos distintos, entonces esta propiedad es inmediata, ya que por 1) existirán n autovectores ortogonales entre si, que puede ser convertidos en ortonormales luego de normalización ($e'_i\to e'_i/||e'_i||$).

En general, supongamos que e'_1 es un autovector normalizado de F ($F(e'_1) = \lambda_1 e'_1$, $((e'_1, e'_1) = 1)$ y sea S_1 el subespacio de V generado por e'_1. En tal caso S_1 es invariante por F y por lo tanto, el complemento ortogonal $S_{1\perp}$, de dimensión $n-1$, **será también invariante por** $F^\dagger = F$. F restringido a $S_{1\perp}$ es obviamente también autoadjunto. Por lo tanto, por hipótesis inductiva, existe una base ortonormal de $S_{1\perp}$ en la que F es diagonal. F resulta así diagonal en la base ortonormal de V formada por e'_1 y la base anterior de $S_{1\perp}$. F será entonces diagonalizable $\forall\ n$ en una base ortonormal.

3) Si F y G son dos operadores autoadjuntos y $[F, G] = 0$ (o sea, $FG = GF$) $\Rightarrow$ *existe una base ortonormal común* e' en la que ambos operadores son *simultáneamente diagonales:*

$$F(e'_i) = \lambda_i^F e'_i\,, \quad G(e'_i) = \lambda_i^G e'_i, \quad i = 1, \ldots, n$$

Demostración: Como F es autoadjunto, existe una base ortonormal donde F es diagonal. Como $[G, F] = 0$ $\Rightarrow$ si $F(e'_i) = \lambda_i^F e'_i$, $FG(e'_i) = GF(e'_i) = \lambda_i^F G(e'_i)$, por lo que $G(e'_i) \in V_F(\lambda_i^F)$ (espacio propio). $V_F(\lambda_i^F)$ es pues también invariante por G. Pero G restringido a $V_F(\lambda_i^F)$ es asimismo autoadjunto, por lo que es siempre posible elegir una base ortonormal de $V_F(\lambda_i^F)$ en la que G será también diagonal, con autovalores λ_i^G. Los elementos de dicha base serán, por pertenecer a $V_F(\lambda_i^F)$, también autovectores de F. Repitiendo esto para todos los autovalores, vemos que existirá una base ortonormal de V en la que tanto F y G serán diagonales.

24.1 Operadores normales

Un operador lineal $A : V \to V$ se dice *normal* si $A^\dagger A = AA^\dagger$, es decir, si

$$[A, A^\dagger] = 0$$

donde $[A, B] = AB - BA$ denota el conmutador.
Así, los operadores autoadjuntos ($F^\dagger = F$) son obviamente normales, y también son normales los unitarios ($U^\dagger = U^{-1}$, con $UU^\dagger = U^\dagger U = I$). Otro caso de operador normal son las antiautadjuntos ($F^\dagger = -F$).

Teorema de diagonalización para operadores normales:
Si $A : V \to V$ es un operador normal, entonces existe una base *ortonormal* e' en la cual $[A]_{e'}$ es *diagonal*:

$$[A]_{e'} = S^\dagger [A]_e S = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ & & \ldots & \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}, \quad S^\dagger S = SS^\dagger = I_n$$

Además si $A : V \to V$ es diagonal en una base ortonormal $\Rightarrow$ es normal.
Demostración: Hemos ya demostrado que para todo operador autoadjunto existe una base ortonormal donde es diagonal. La extensión para todo operador normal se basa en la descomposición

$$A = A^r + iA^i, \quad A^r = \frac{A + A^\dagger}{2}, \quad A^i = \frac{A - A^\dagger}{2i}$$

válida para *cualquier* operador A, donde A_r y A_i son claramente operadores *autoadjuntos*: $(A^r)^\dagger = A^r$, $(A^i)^\dagger = A^i$. Esta descomposición del operador es similar a la de un número complejo $z = x + iy$ en parte real x e imaginaria iy (caso particular $n = 1$).
Si A es *normal* $\Rightarrow$

$$[A^r, A^i] = \frac{1}{4i}[A + A^\dagger, A - A^\dagger] = 0$$

y por lo tanto, *existe una base ortonormal común e' donde A^r y A^i son simultáneamente diagonales*. Los autovalores de A serán entonces de la forma

$$\lambda_j = \lambda_j^r + i\lambda_j^i$$

con λ_j^r y λ_j^i *reales* y autovalores de A^r y A^i respect., por lo que λ_j será en general complejo.
Si A es autoadjunto ($A^\dagger = A$) $\Rightarrow$ $A_i = 0$ y por lo tanto $\lambda_j^i = 0$. Los autovalores de A son entonces todos reales, como ya habíamos demostrado.
Si A es antiautoadjunto ($A^\dagger = -A$) $\Rightarrow$ $A_r = 0$ y por lo tanto $\lambda_j^r = 0$. Los autovalores de A son entonces

todos *imaginarios puros.*
Finalmente, si A es unitario, $[A]^\dagger_{e'}[A]_{e'} = I_n$, lo que implica $\lambda_j\lambda_j^* = |\lambda_j|^2 = 1$, es decir $|\lambda_j| = 1$. Esto implica

$$\lambda_j = e^{i\phi_j} = \cos\phi_j + i\sin\phi_j$$

con $\lambda_j^r = \cos\phi_j$, $\lambda_j^i = \sin\phi_j$.

Por otro lado, si A es diagonal en una base e' ortonormal $\Rightarrow A^\dagger$ es también diagonal en dicha base, con

$$[A^\dagger]_{e'} = [A]^\dagger_{e'} = \begin{pmatrix} \lambda_1^* & 0 & \dots & 0 \\ 0 & \lambda_2^* & \dots & 0 \\ & & \dots & \\ 0 & 0 & \dots & \lambda_n^* \end{pmatrix}$$

Por lo tanto

$$[AA^\dagger - A^\dagger A]_{e'} = [A]_{e'}[A^\dagger]_{e'} - [A^\dagger]_{e'}[A]_{e'} = 0$$

lo que implica $AA^\dagger - A^\dagger A = 0$. A es entonces normal.

En resumen, el teorema implica que en un espacio unitario, *un operador tiene representación diagonal en una base ortonormal si y sólo si es un operador normal.* En términos matriciales, si A es una matriz de $n\times n$, entonces *existe una matriz unitaria S tal que $A' = S^\dagger AS$ es diagonal si y sólo si A es normal ($[A^\dagger, A] = 0$).* Esto comprende en particular las matrices hermíticas ($A^\dagger = A$), antihermíticas ($A^\dagger = -A$) y unitarias ($A^\dagger = A^{-1}$). Destaquemos también que todo $v \in V$ puede expandirse en la base e' de autovectores de un operador normal A,

$$v = \sum_{i=1}^n \alpha_i e'_i = \sum_{i=1}^n P_{e'_i}(v)$$

donde $\alpha_i = (e'_i, v)$ y $P_{e'_i}(v) = (e'_i, v)e'_i = \alpha'_i e'_i$. Por lo tanto

$$A(v) = \sum_{i=1}^n A(\alpha_i e'_i) = \sum_{i=1}^n \alpha_i\lambda_i e'_i = \sum_{i=1}^n \lambda_i P_{e'_i}(v)$$

Como v es arbitrario, esto implica

$$A = \sum_{i=1}^n \lambda_i P_{e'_i}$$

Un operador normal puede pues expresarse como *combinación lineal de proyectores ortogonales sobre sus espacios propios.*

De lo anterior se desprende además que todo operador unitario U puede escribirse en la forma

$$U = \exp[iF]$$

con F autoadjunto: Como los autovalores de U son de la forma $e^{i\phi_j}$, podemos definir F como el operador autoadjunto que es también diagonal en la base ortonormal e' en que U es diagonal y que tiene autovalores reales ϕ_j. En tal caso, $[U]_{e'} = \exp[i[F]_{e'}] = [\exp[iF]]_{e'}$, lo que implica $[U]_e = [\exp(iF)]_e$ en cualquier base. Esto conduce a $U = \exp[iF]$.

Ejercicio: Utilizando la representación diagonal, mostrar que si $F : V \to V$ es autoadjunto, entonces $\forall$ $v \in V$, con $v \neq 0$, se tiene

$$\lambda_m \leq \frac{(v, F(v))}{(v, v)} \leq \lambda_M$$

donde λ_m y λ_M denotan resp. el menor y mayor autovalor de F.

24.2 Isometrías en espacios euclideos

Hemos visto que los autovalores de un operador unitario U son necesariamente de la forma $\lambda = e^{i\phi} = \cos(\phi)+i\sin(\phi)$, con ϕ real. Mediante el "embedding" de un espacio euclídeo en un espacio unitario discutido en clase, esto permite demostrar que las isometrías U en espacios euclídeos sólo pueden ser rotaciones (Det$U = 1$) o rotaciones seguidas o precedidas de una reflexión (Det$U = -1$).

En efecto, si $S \equiv [U]_e$ es una matriz real que representa una isometría U en una base ortonormal de un espacio euclideo ($S^t = S^{-1}$), considerada en un espacio complejo representa una transformación unitaria ($S^\dagger = S^{-1}$). Dado que S es real, los autovalores vendrán de a pares conjugados con autovectores conjugados:

$$SX = \lambda X, \quad SX^* = \lambda^* X^*$$

Escribiendo $\lambda = \lambda_r + i\lambda_i$, $X = X_r + iX_i$, con $\lambda_r = \cos\phi$, $\lambda_i = \sin\phi$ y X_r, X_i reales, esto implica

$$SX_r = \lambda_r X_r - \lambda_i X_i\,, \;\; SX_i = \lambda_r X_i + \lambda_i X_r$$

Si λ no es real ($\lambda_i \neq 0$) $\Rightarrow$ $X_i \neq 0$ (pues S es real) y la ortogonalidad de los autovectores para autovalores distintos (válido para cualquier matriz normal $[S^\dagger, S] = 0$) implica $(X^*)^\dagger X = 0$, o sea,

$$(X_r + iX_i)^t (X_r + iX_i) = X_r^t X_r - X_i^t X_i + 2iX_i^t X_r = 0$$

de donde $X_r^t X_r = X_i^t X_i$ y $X_i^t X_r = 0$. Por lo tanto, vemos que en el subespacio generado por X^*, X, existe una base *real* y *ortonormal con el producto escalar euclideo*, formada por (X_i, X_r), en la que el bloque correspondiente de $[U]_{e'}$ tiene la forma

$$S'_\phi = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$$

que representa una rotación de ángulo ϕ ($\mathrm{Det}[S'_\phi] = 1$). Y en el espacio euclideo completo, vemos entonces que existe una base ortonormal e' donde $S' \equiv [U]_{e'}$ tiene la forma

$$S' = \begin{pmatrix} {S'}_{\phi_1} & 0 & \ldots & 0 & 0 \\ 0 & {S'}_{\phi_2} & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & \pm 1 & 0 \\ 0 & 0 & \ldots & 0 & \pm 1 \end{pmatrix}$$

donde S'_{ϕ_i} son bloques de la forma anterior que representan rotaciones en subespacios de dimensión 2, y los elementos ± 1 representan los posibles autovalores reales. U representa pues rotaciones ($\mathrm{Det}[S'] = 1$) o rotaciones compuestas con reflexiones ($\mathrm{Det}[S'] = -1$). Por ej., en $\mathbb{R}^3$, las posibilidades son un bloque A_ϕ seguido de $+1$ (rotación) o -1 (rotación compuesta con reflexión).

24.3 Elementos de matriz de un operador lineal en una base ortonormal

Recordemos que si $F: V \to V$ es un operador lineal $\Rightarrow$ la matriz $T = [F]_e$ ($\equiv [F]_e^e$) que lo representa en una base e de V queda definida por

$$F(e_i) = \sum_{j=1}^{n} T_{ji} e_j$$

En un espacio unitario y en una base ortonormal e, los elementos de matriz $T_{ji} = ([F]_e)_{ji}$ pueden entonces obtenerse, por ortonormalidad de los e_i, como

$$T_{ji} = (e_j, F(e_i))$$

De esta forma,

$$v = \sum_{i=1}^{n} \alpha_i e_i, \quad \alpha_i = (e_i, v)$$

y

$$F = \sum_{i,j=1}^{n} T_{ji} E_{ji}, \quad T_{ji} = (e_j, F(e_i))$$

con E_{ji} el operador lineal definido por

$$E_{ji}(v) = (e_i, v) e_j$$

ya que $F(e_i) = \sum_{j=1}^n T_{ji} e_j = \sum_{j,k=1}^n T_{jk}(e_k, e_i) e_j = (\sum_{j,k=1}^n T_{jk} E_{jk})(e_i)$.
Notemos que $([E_{ji}]_e)_{kl} = \delta_{kj}\delta_{il}$.

Notación de Mecánica cuántica:

$$T_{ji} = (e_j, F(e_i)) \to \langle j|F|i\rangle, \quad E_{ji} \to |j\rangle\langle i|$$

donde $|i\rangle \equiv e_i$ y $\langle i| \equiv f^i$ (vector asociado del espacio dual). Por lo tanto

$$F = \sum_{i,j} F_{ij}|i\rangle\langle j|, \quad F_{ij} = \langle i|F|j\rangle$$

Por ej., el proyector ortogonal sobre e_i se escribe como $P_{e_i} = |i\rangle\langle i|$ (ya que $(e_j, P_{e_i}(e_i)) = (e_j, e_i) = \delta_{ji}$) mientras que el operador identidad es $I = \sum_i |i\rangle\langle i|$. En general, para todo operador normal $F : V \to V$ existe entonces una base ortonormal $\{|i\rangle\}$ de V formada de autovectores de F en la que $\langle i|F|j\rangle = \delta_{ij}\lambda_i$ y

$$F = \sum_i \lambda_i |i\rangle\langle i|$$

25 Descomposición en valores singulares (DVS)

Sea $F : V \to W$ una transformación lineal arbitraria entre espacios unitarios V y W de dimensiónes n y m respectivamente, y e_V, e_W bases ortonormales de V y W. Entonces existen bases ortonormales e'_V, e'_W en las que F queda representado por una matriz *diagonal* de elementos no negativos. En otras palabras, dada una matriz A de $m \times n$, existen matrices unitarias U de $m \times m$ y V de $n \times n$ tales que

$$A = UA'V^\dagger$$

con $U^\dagger U = I_m$, $V^\dagger V = I_n$ y A' de $m \times n$ *diagonal* de elementos $A'_{kj} = \sigma_j \delta_{kj}$, con $\sigma_j \geq 0$. Aquí $A = [F]_{e_W}^{e_V}$, $A' = [F]_{e'_W}^{e'_V}$, $U = [I]_{e_W}^{e'_W}$ y $V = [I]_{e_V}^{e'_V}$, con $V^\dagger = [I]_{e'_V}^{e_V}$. Los σ_j no nulos se denominan *valores singulares* y son las *raíces* de los autovalores no nulos de la matriz hermítica $A^\dagger A$, de $n \times n$ (que posee autovalores no negativos). V es la correspondiente matriz de autovectores normalizados (tal que $V^\dagger(A^\dagger A)V$ es diagonal).
La demostración es similar al caso de matrices reales (espacios euclideos) y se deja como ejercicio.
Recordemos que si k es el número de autovalores no nulos de $A^\dagger A$, las primeras k columnas de U son los vectores $u_i = Av_i/\sigma_i$, $i = 1, \ldots, k$, $\sigma_i \neq 0$, obteniéndose las restantes $m - k$ columnas ortonormales de U por el método de Gram-Schmidt complejo.

Ejercicios: Para A de $m \times n$ compleja general, demostrar (en forma similar al caso euclideo) que:
0) Las matrices $A^\dagger A$ y $AA^\dagger$ son ambas hermíticas.
1) Los autovalores de $A^\dagger A$ son todos no negativos.
2) El número de autovalores no nulos de $A^\dagger A$ es igual al rango de A.
3) Los autovalores no nulos de las matrices $A^\dagger A$ y $AA^\dagger$ son iguales.
4) $||A||_2 = \sigma_M$, siendo σ_M el máximo valor singular y $||A||_2 \equiv \mathrm{Max}_{v\neq 0}||Av||/||v||$, con $||v|| = \sqrt{v^\dagger v}$.
5) Si $m = n$ y λ es autovalor de $A \Rightarrow |\lambda| \leq \sigma_M$.
6) Si $m = n$ y A es invertible $\Rightarrow n_c(A) = \sigma_M/\sigma_m$, donde σ_m es el mínimo valor singular de A y $n_c(A)$ es el número de condición.

25.1 Forma polar de un operador lineal

En el caso de $V = W$, la DVS permite obtener en forma inmediata la denominada *forma polar* de un operador: Si $F : V \to V$ es un operador lineal en un espacio unitario V entonces F puede escribirse como

$$F = WM = \tilde{M}W$$

donde W es un operador *unitario* y M, $\tilde{M}$ operadores autoadjuntos *positivos*.
Dem.: Utilizando la DVS para la representación $A = [F]_e$ de F en una base ortonormal e de V, se tiene

$$\begin{aligned} A &= UA'V^\dagger \\ &= (UV^\dagger)(VA'V^\dagger) = WM, \quad W = UV^\dagger, \quad M = VA'V^\dagger = \sqrt{A^\dagger A} \\ &= (UA'U^\dagger)(UV^\dagger) = \tilde{M}W, \quad \tilde{M} = UA'U^\dagger = \sqrt{AA^\dagger} \end{aligned}$$

donde W es unitario ($W^\dagger = VU^\dagger = W^{-1}$) y $A^\dagger A = VA'^2V^\dagger$, $AA^\dagger = UA'^2U^\dagger$.
Ejercicio: Discutir la DVS y la descomposición polar de una matriz hermítica.

26 Desigualdad de Cauchy Schwarz y relaciones de incerteza

Consideremos dos operadores autoadjuntos F, G. El valor medio de un operador F en un estado normalizado $|\psi\rangle$ ($\langle\psi|\psi\rangle = 1$) es

$$\langle F\rangle_\psi = \langle\psi|F|\psi\rangle$$

(o sea $\langle F\rangle_\psi = (\psi, F(\psi))$ en notación de A.L.). Si F es autoadjunto, $\langle F\rangle_\psi$ es *real*, ya que $\langle\psi|F|\psi\rangle = \langle\psi|F^\dagger|\psi\rangle^* = \langle\psi|F|\psi\rangle^*$ (o sea, $(\psi, F(\psi)) = (F(\psi), \psi)^* = (\psi, F^\dagger(\psi))^* = (\psi, F(\psi))^*$).

La varianza de un operador F en el estado $|\psi\rangle$ se define como el valor medio del cuadrado de la diferencia entre F y $\langle F\rangle_\psi$ y es una medida de la dispersión alrededor de la media:

$$\Delta^2 F = \langle(F - \langle F\rangle_\psi I)^2\rangle_\psi = \langle\psi|(F - \langle F\rangle_\psi I)^2|\psi\rangle = \langle\psi|F^2|\psi\rangle - \langle\psi|F|\psi\rangle^2$$

La desviación estándar es la raíz de la varianza: $\Delta F = \sqrt{\Delta^2 F} = \sqrt{\langle(F - \langle F\rangle_\psi I)^2\rangle_\psi}$.

Definamos ahora

$$\tilde{F} = F - \langle F\rangle_\psi I, \quad \tilde{G} = G - \langle G\rangle_\psi I$$

tal que $\Delta F = \sqrt{\langle\psi|\tilde{F}^2|\psi\rangle}$, $\Delta G = \sqrt{\langle\psi|\tilde{G}^2|\psi\rangle}$, y consideremos el producto escalar $(\tilde{F}(\psi), \tilde{G}(\psi)) = (\psi, \tilde{F}\tilde{G}(\psi))$, es decir $\langle\psi|\tilde{F}\tilde{G}|\psi\rangle$ en notación cuántica. La desigualdad de Cauchy-Schwarz implica $|(\tilde{F}(\psi), \tilde{G}(\psi))| \leq ||\tilde{F}(\psi)||\,||\tilde{G}(\psi)||$, con $||\tilde{F}(\psi)||^2 = (\tilde{F}(\psi), \tilde{F}(\psi)) = (\psi, \tilde{F}^2(\psi))$, o sea,

$$|\langle\psi|\tilde{F}\tilde{G}|\psi\rangle| \leq \sqrt{\langle\psi|\tilde{F}^2|\psi\rangle}\sqrt{\langle\psi|\tilde{G}^2|\psi\rangle} = (\Delta F)(\Delta G)$$

Por otro lado, si $[F, G] = FG - GF$ denota el conmutador de F y G, entonces

$$\langle\psi|[F, G]|\psi\rangle = \langle\psi|[\tilde{F}, \tilde{G}]|\psi\rangle = \langle\psi|\tilde{F}\tilde{G}|\psi\rangle - \langle\psi|\tilde{G}\tilde{F}|\psi\rangle = 2i\,\mathrm{Im}[\langle|\psi|\tilde{F}\tilde{G}|\psi\rangle]$$

donde Im denota la parte imaginaria, ya que $\langle\psi|\tilde{G}\tilde{F}|\psi\rangle = \langle\psi|(\tilde{G}\tilde{F})^\dagger|\psi\rangle^* = \langle\psi|\tilde{F}\tilde{G}|\psi\rangle^*$. Por lo tanto

$$\tfrac{1}{2}|\langle\psi|[F, G]|\psi\rangle| \leq |\langle\psi|\tilde{F}\tilde{G}|\psi\rangle| \leq (\Delta F)(\Delta G)$$

es decir,

$$(\Delta F)(\Delta G) \geq \tfrac{1}{2}|\langle[F, G]\rangle_\psi|$$

Esta es la denominada relación de incerteza entre dos operadores: Si el conmutador es no nulo entonces el producto de sus "incertezas" $(\Delta F)(\Delta G)$ en un estado $|\psi\rangle$ no puede ser menor que el módulo del valor medio del conmutador en dicho estado.

Como ejemplo fundamental, consideremos el espacio L^2 de funciones $\psi(x)$ de $\mathbb{R} \to \mathbb{C}$ de norma finita ($||\psi||^2 = \int_{-\infty}^{\infty} |\psi(x)|^2 dx < \infty$) y que tienden a 0 para $x \to \pm\infty$, tal que el producto escalar

$$(\psi, \phi) = \int_{-\infty}^{\infty} \psi^*(x)\phi(x)dx$$

esté bien definido. Los operatores X y $P = -i\hbar\partial_x = -i\hbar\frac{\partial}{\partial x}$, donde $\hbar = h/(2\pi)$, con $\hbar$ la constante de Planck, son autoadjuntos en este espacio: $(\psi, X\phi) = \int_{-\infty}^{\infty} \psi^*(x)x\phi(x)dx = (X\psi, \phi)$, y

$$(\psi, P\phi) = -i\hbar\int_{-\infty}^{\infty}\psi^*(x)\phi'(x)dx = -i\hbar[\,\psi^*(x)\phi(x)\,|_{-\infty}^{\infty} - \int_{-\infty}^{\infty}\psi^{*\prime}(x)\phi(x)dx] = \int_{-\infty}^{\infty}[-i\hbar\psi'(x)]*\phi(x)dx = (P(\psi), \phi)$$

Dado que $[X, P]\psi(x) = -i\hbar(x\psi'(x) - (x\psi(x))') = i\hbar\psi(x)\ \forall\ \psi$, es decir, $[X, P] = i\hbar I$, obtenemos $|\langle[X, P]\rangle_\psi| = \hbar\ \forall\ \psi$ y el resultado anterior implica entonces

$$(\Delta P)(\Delta X) \geq \frac{\hbar}{2}$$

El operador P representa en Mecánica Cuántica el operador impulso de una partícula (en una dimensión). Por lo tanto, en cualquier estado cuántico el producto de las desviaciones estándar de X y P es no nulo y mayor que $\hbar/2$.