Formas Hermitianas y Producto escalar complejo

Pedro Montero

Departamento de Matemática, Universidad Técnica Federico Santa María

Estas notas son material adicional al curso de Álgebra Lineal para estudiantes de segundo año en Matemáticas, y tienen como objetivo recordar y recopilar resultados sobre formas hermitianas, y que permiten generalizar al caso complejo las nociones de producto escalar, normas y ortogonalidad. Una gran parte de la teoría es muy similar a la teoría de formas bilineales simétricas reales, por lo que el objetivo de estas notas será sobre todo explicar las diferencias.

Valparaíso, Julio 2020

Índice

1. Recuerdo sobre los números complejos

Sea $\mathbb{C}$ el cuerpo de los números complejos. Denotamos por $i \in \mathbb{C}$ una raíz cuadrada de -1, la cual fijaremos durante todo el texto. Así, todo $z \in \mathbb{C}$ se escribe de manera *única* como

$$z = a + ib,$$

donde $a, b \in \mathbb{R}$. En particular, $\mathbb{R}^2 \to \mathbb{C},\ (a,b) \mapsto a + ib$ es un isomorfismo de espacios vectoriales reales. Orientaremos al espacio euclideano $\mathbb{C} \cong \mathbb{R}^2$ declarando que la imagen de la base canónica de $\mathbb{R}^2$ por el isomorfismo anterior es una base directa, i.e., la base ordenada $(1, i)$ es una base directa[1]. Por convención, identificamos $\mathbb{R}$ y $\{a + ib \in \mathbb{C} \mid b = 0\}$, y diremos que todo complejo perteneciente a este último conjunto es **real**, y escribiremos $z \in \mathbb{R}$. Del mismo modo, los elementos de $i\mathbb{R} := \{a + ib \in \mathbb{C} \mid a = 0\}$ serán llamados **imaginarios puros**.

Definición 1.1. Sea $z \in \mathbb{C}$ un número complejo.

1. Si escribimos $z = a + ib$ con $a, b \in \mathbb{R}$. Decimos que a (resp. b) es la **parte real** (resp. **parte imaginaria**) de z, y la denotaremos $\mathrm{Re}(z)$ (resp. $\mathrm{Im}(z)$). Luego, $z = \mathrm{Re}(z) + i\,\mathrm{Im}(z)$.
2. La **conjugación compleja** es la aplicación que a todo $z = a + ib$ asocia $\overline{z} := a - ib$.
3. Si escribimos $z = a + ib$ con $a, b \in \mathbb{R}$, entonces $z\overline{z} = a^2 + b^2$, y definimos el **módulo** de z mediante $|z| := \sqrt{z\overline{z}}$.

[1] Informalmente, esto quiere decir que declaramos como ángulos *positivos* aquellos que van en sentido anti-horario.

4. Si $z \neq 0$, entonces $z/|z|$ tiene módulo 1, y luego es de la forma $e^{i\theta} = \cos(\theta) + i\sin(\theta)$. Así, todo $z \in \mathbb{C} \setminus \{0\}$ posee una **forma polar**

$$z = \rho e^{i\theta}, \quad \text{donde } \rho = |z| \in \mathbb{R}^{\geq 0} \text{ y } \theta \in \mathbb{R} \mod 2\pi\mathbb{Z}.$$

En particular, si suponemos $\theta \in]-\pi, \pi]$ entonces θ es único, y lo llamamos el **argumento** de z.

Ejercicio 1.2. Sean $z, z' \in \mathbb{C}$ dos números complejos.

a) Probar que $\mathrm{Re}(-iz) = \mathrm{Im}(z)$ y que $\mathrm{Im}(iz) = \mathrm{Re}(z)$.

b) Probar que $z + \overline{z} = \mathrm{Re}(z)$, $z - \overline{z} = 2i\,\mathrm{Im}(z)$, y que $z \in \mathbb{R}$ si y sólo si $z = \overline{z}$.

c) Probar que $\overline{z + z'} = \overline{z} + \overline{z'}$, que $\overline{zz'} = \overline{z}\overline{z'}$ y $|zz'| = |z||z'|$..

d) Probar que si $z \in \mathbb{C}$ cumple que $z \in \mathbb{R}$ y que $iz \in \mathbb{R}$, entonces necesariamente $z = 0$.

2. Formas hermitianas

Antes de dar la definición formal de una forma hermitiana, veamos el ejemplo fundamental a considerar.

Ejemplo 2.1. En $\mathbb{R}^2$, la norma euclideana $\|(x, y)\| = \sqrt{x^2 + y^2}$ está definida a partir (de la raíz cuadrada) del producto escalar canónico $\langle (x_1, y_1), (x_2, y_2) \rangle = x_1x_2 + y_1y_2$. Si consideramos ahora los números complejos $z_1 = x_1 + iy_1$ y $z_2 = x_2 + iy_2$, nos gustaría construir una función $h : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ tal que para $z_1 = z_2$ obtengamos $h(z_1, z_1) = x_1^2 + y_1^2$.

Notar que si considerásemos símplemente el producto $z_1z_2 = (x_1 + iy_1)(x_2 + iy_2) = (x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1)$ entonces para $z_1 = z_2$ obtendríamos $z_1^2 = (x_1^2 - y_1^2) + i2x_1y_1$, el cual **no** es necesariamente real y que no podemos usar para calcular el módulo de z_1.

Por otra parte, dado que $|z|^2 = z\overline{z}$, si consideramos cualquiera de las dos funciones

$$h_1(z_1, z_2) := z_1\overline{z_2} \quad \text{o} \quad h_2(z_1, z_2) := \overline{h_1(z_1, z_2)} = \overline{z_1}z_2,$$

entonces obtenemos para $z_1 = z_2$ el resultado esperado.

Es muy importante notar que h_1 y h_2 **no** son formas bilineales complejas, puesto que h_1 (resp. h_2) es $\mathbb{C}$–lineal en la primera variable (resp. segunda variable), pero **no** es lineal en la segunda (resp. en la primera). Por ejemplo, para todos $z_1, z_2, z_3 \in \mathbb{C}$ (pensados como vectores) y todo $\lambda \in \mathbb{C}$ (pensado como escalar) se tiene:

$$h_1(z_1 + \lambda z_2, z_3) = h(z_1, z_3) + \lambda h(z_2, z_3) \text{ y } h_1(z_3, z_1 + \lambda z_2) = h_1(z_3, z_1) + \overline{\lambda} h_1(z_3, z_2).$$

Luego, h_1 *anti-lineal* (o *lineal conjugada*) respecto a la segunda variable: diremos que h_1 es **sesquilineal**.

En lo que sigue supondremos que los espacios vectoriales son de dimensión *finita*, a pesar que de la mayoría de las definiciones tienen sentido en dimensión arbitraria.

Definición 2.2 (aplicación anti-lineal)**.** Sean V, W dos $\mathbb{C}$-espacios vectoriales. Decimos que una función $\varphi : V \to W$ es una **aplicación anti-lineal** si para todos $v_1, v_2 \in V$ y todo $\lambda \in \mathbb{C}$ se tiene que

$$\varphi(v_1 + v_2) = \varphi(v_1) + \varphi(v_2), \quad \varphi(\lambda v_1) = \overline{\lambda}\varphi(v).$$

Observación 2.3. *Es muy importante observar que si V es un espacio vectorial sobre $\mathbb{C}$, entonces podemos definir el* **espacio vectorial conjugado** $\overline{V}$ *de la manera siguiente: como* <u>grupo abeliano</u> *definimos $\overline{V} := V$ y, si $v \in V$ es un vector y $\lambda \in \mathbb{C}$ un escalar, entonces* <u>definimos</u> *la multiplicación por escalares en $\overline{V}$ mediante*

$$\lambda \cdot v := \overline{\lambda} v,$$

donde el lado derecho lo definimos a partir de la estructura de espacio vectorial complejo de V. Luego, una función $\varphi : V \to W$ es anti-lineal si y sólo si la función $\varphi : V \to \overline{W}$ es una aplicación lineal (en el sentido usual).

Definición 2.4 (forma sesquilineal)**.** Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo. Decimos que una función $h : V \times V \to \mathbb{C}$, $(x, y) \mapsto h(x, y)$ es una **forma sesquilineal** si:

1. h es $\mathbb{C}$-lineal en la *primera* variable, i.e., para todos $x, y, z \in V$ y $\lambda \in \mathbb{C}$ se tiene que
$$h(\lambda x, y) = \lambda h(x, y) \quad \text{y} \quad h(x + y, z) = h(x, z) + h(y, z).$$

2. h es anti-lineal en la *segunda* variable, i.e., para todos $x, y, z \in V$ y $\lambda \in \mathbb{C}$ se tiene que
$$h(x, \lambda y) = \overline{\lambda} h(x, y) \quad \text{y} \quad h(x, y + z) = h(x, y) + h(x, z).$$

Denotamos por $\mathrm{Sesq}(V)$ al espacio vectorial *complejo* de formas sesquilineales en V. Finalmente, dada una forma sesquilineal $h : V \times V \to \mathbb{C}$, definimos la **forma sesquilineal adjunta** $h^* : V \times V \to \mathbb{C}$ mediante la fórmula $h^*(x, y) := \overline{h(y, x)}$.

Ejercicio 2.5. Probar que la forma sesquilineal adjunta $h^* : V \times V \to \mathbb{C}$ es una forma sesquilineal.

Observación 2.6. *En vista de la observación anterior, una forma sesquilineal $h : V \times V \to \mathbb{C}$ es lo mismo que una forma $\mathbb{C}$-bilineal $h : V \times \overline{V} \to \mathbb{C}$.*

⚠ Es importante señalar que, tal como lo vimos en el Ejemplo 2.1, podemos considerar también formas anti-linealeales en la *primera* variable y $\mathbb{C}$-lineales en la *segunda* variable. Esta última convención es comunmente usada en Física, puesto que es coherente con la notación *bra-ket* introducida por Dirac.

Definición 2.7 (forma (anti-)hermitiana)**.** Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo. Decimos que una forma sesquilineal $h : V \times V \to \mathbb{C}$ es:

1. una **forma hermitiana** si $h^* = h$, i.e., $h(x, y) = \overline{h(y, x)}$ para todos $x, y \in V$.
2. una **forma anti-hermitiana** si $h^* = -h$, i.e., $h(x, y) = -\overline{h(y, x)}$ para todos $x, y \in V$.

Denotaremos por $\mathrm{Herm}(V)$ (resp. $\mathrm{AntiHerm}(V)$) al espacio vectorial <u>real</u> de formas hermitianas (resp. anti-hermitianas) de V.

Proposición 2.8. *Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo. Una forma sesquilineal h sobre V es hermitiana si y sólo si $h(x, x) \in \mathbb{R}$ es real para todo $x \in V$.*

Demostración. Si h es hermitiana entonces $h(x, y) = \overline{h(y, x)}$ implica (tomando $x = y$) que $h(x, x) = \overline{h(x, x)}$, i.e., $h(x, x) \in \mathbb{R}$. Por otra parte, si la forma sesquilineal h verifica $h(x, x) \in \mathbb{R}$ para todo $x \in V$ entonces para todos $x, y \in V$ tenemos que
$$h(x, y) + h(y, x) = h(x + y, x + y) - h(x - y, x - y) \in \mathbb{R},$$
por lo que $h(x, y) - \overline{h(y, x)} = \underbrace{(h(x, y) + h(y, x))}_{\in \mathbb{R}} + \underbrace{(h(y, x) + \overline{h(y, x)})}_{= 2\,\mathrm{Re}(h(y,x)) \in \mathbb{R}} \in \mathbb{R}$. Si reemplazamos x por ix obtenemos
$$i(h(x, y) - \overline{h(y, x)}) = h(ix, y) + \overline{ih(y, x)} = h(ix, y) + \overline{\overline{(-i)}h(y, x)} = h(ix, y) - \overline{h(y, ix)} \in \mathbb{R}.$$
Por otra parte, si $a \in \mathbb{C}$ verifica $a \in \mathbb{R}$ y $ia \in \mathbb{R}$ entonces necesariamente $a = 0$, de donde concluimos que $h(x, y) - \overline{h(y, x)} = 0$ y luego h es hermitiana. □

Observación 2.9. *Sea $h : V \times V \to \mathbb{C}$ forma sesquilineal en $V \cong \mathbb{C}^n$. Dado que $(ih)^* = -ih^*$, se tiene*
$$h \text{ hermitiana} \Leftrightarrow ih \text{ anti-hermitiana}, \quad h \text{ anti-hermitiana} \Leftrightarrow ih \text{ hermitiana}.$$
En particular, h es anti-hermitiana si y sólo si $h(v, v) \in i\mathbb{R}$ (imaginario puro) para todo $v \in V$.

Notemos que toda forma sesquilineal h se descompone de manera única en una suma
$$h = \sigma + \alpha,$$
donde σ es hermitiana y α anti-hermitiana, donde
$$\sigma = \frac{1}{2}(h + h^*) \quad \text{y} \quad \alpha = \frac{1}{2}(h - h^*).$$
En particular, tenemos el siguiente resultado.

Proposición 2.10. *Sea $V \cong \mathbb{C}^n$ espacio vectorial complejo. Entonces, tenemos una descomposición en suma directa de espacios vectoriales* <u>*reales*</u>

$$\mathrm{Sesq}(V) = \mathrm{Herm}(V) \oplus \mathrm{AntiHerm}(V),$$

donde

$$\dim_{\mathbb{R}} \mathrm{Herm}(V) = \dim_{\mathbb{R}} \mathrm{AntiHerm}(V) = \frac{1}{2} \dim_{\mathbb{R}} \mathrm{Sesq}(V) = \dim_{\mathbb{C}} \mathrm{Sesq}(V).$$

Demostración. La discusión anterior implica directamente que $\mathrm{Sesq}(V) = \mathrm{Herm}(V) \oplus \mathrm{AntiHerm}(V)$ como $\mathbb{R}$-espacios vectoriales. Más aún, tenemos un isomorfismo $\mathbb{R}$-lineal

$$\mathrm{Herm}(V) \xrightarrow{\sim} \mathrm{AntiHerm}(V),\ h \mapsto ih,$$

de donde se deduce la relación entre las dimensiones. □

Tal como en el caso de formas bilineales reales, podemos describir formas sesquilineales en términos matriciales. Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo y sea $\mathscr{B} = (e_1, \ldots, e_n)$ una base de V. Dados vectores

$$x = \sum_{j=1}^{n} x_j e_j \in V,\ y = \sum_{j=1}^{n} y_j e_j \in V,$$

la hipótesis de sesquilinealidad de $h : V \times V \to \mathbb{C}$ implica que

$$h(x,y) = h\left(\sum_{j=1}^{n} x_j e_j, y\right) = \sum_{j=1}^{n} x_j h(e_j, y) = \sum_{j=1}^{n} x_j h\left(e_j, \sum_{k=1}^{n} y_k e_k\right) = \sum_{j=1}^{n}\sum_{k=1}^{n} x_j \overline{y_k}\, h(e_j, e_k).$$

Luego, si definimos $h_{jk} := h(e_j, e_k) \in \mathbb{C}$ entonces

$$h(x,y) = \sum_{1 \le j,k \le n} h_{jk} x_j \overline{y_k}.$$

Recíprocamente, toda expresión de esta forma con coeficientes $h_{jk} \in \mathbb{C}$ arbitrarios define una forma sesquilineal $h : V \times V \to \mathbb{C}$. En particular, $\mathrm{Sesq}(V) \cong M_n(\mathbb{C})$ y luego $\dim_{\mathbb{C}} \mathrm{Sesq}(V) = n^2$.

Definición 2.11. Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo y sea $\mathscr{B} = (e_1, \ldots, e_n)$ una base de V. Dada una forma sesquilineal $h : V \times V \to \mathbb{C}$, definimos la **matriz de h en la base $\mathscr{B}$** mediante

$$H = \mathrm{Mat}_{\mathscr{B}}(h) := (h_{jk})_{1 \le j,k \le n} = (h(e_j, e_k))_{1 \le j,k \le n}.$$

Más aún, si $A = (a_{jk}) \in M_{n\times p}(\mathbb{C})$ es una matriz compleja de n filas y p columnas, llamamos la **matriz adjunta** a la matriz $A^* \in M_{p\times n}(\mathbb{C})$ dada por la transpuesta conjugada

$$A^* = {}^{\mathrm{t}}\overline{A} = (\overline{a_{kj}}).$$

En particular, $\mathrm{Mat}_{\mathscr{B}}(h^*) = H^*$.

Ejercicio 2.12. Sean $A \in M_{n\times p}(\mathbb{C})$ y $B \in M_{p\times r}(\mathbb{C})$ matrices complejas.

a) Probar que $(A^*)^* = A$ y que la aplicación $A \mapsto A^*$ es anti-lineal.

b) Probar que $(AB)^* = B^* A^*$.

c) Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo y sea $\mathscr{B} = (e_1, \ldots, e_n)$ una base de V. Si $h : V \times V \to \mathbb{C}$ es una forma sesquilineal de matriz $H = (h_{jk})$ respecto a $\mathscr{B}$, probar que para $x = \sum_{j=1}^{n} x_j e_j$ e $y = \sum_{k=1}^{n} y_k e_k$ se cumple que

$$h(x,y) = {}^{\mathrm{t}}XH\overline{Y},$$

donde

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{C}^n \quad \text{y} \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{C}^n$$

son los vectores coordenadas de x e y respecto a $\mathscr{B}$. Deducir que si $\mathscr{B}'$ es otra base de V y $P = \mathrm{Mat}_{\mathscr{B}}(\mathscr{B}')$ es la matriz de cambio de base, entonces

$$H' = \mathrm{Mat}_{\mathscr{B}'}(h) = {}^{\mathrm{t}}PH\overline{P}.$$

⚠ Usualmente en Física se utiliza la notación A^* para referirse a la matriz conjugada de A (que nosotros denotamos $\overline{A}$), y se usa $A^\dagger$ para referirse a la matriz adjunta de A (que nosotros denotamos A^*).

Una consecuencia inmediata de la definición de matriz adjunta es la siguiente caracterización matricial de las formas hermitianas y anti-hermitianas.

Proposición 2.13. *Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo, y sea $h : V \times V \to \mathbb{C}$ una forma sesquilineal. Entonces, para toda base $\mathscr{B}$ de V se tiene que:*

1. *La forma h es* **hermitiana** *si y sólo $H^* = H$ (i.e., ${}^{\mathrm{t}}H = \overline{H}$).*
2. *La forma h es* **anti-hermitiana** *si y sólo si $H^* = -H$ (i.e., ${}^{\mathrm{t}}H = -\overline{H}$).*

donde $H = \mathrm{Mat}_{\mathscr{B}}(h)$ es la matriz de h en la base $\mathscr{B}$.

Demostración. Sea $\mathscr{B} = (e_1, \ldots, e_n)$ una base arbitraria de V. Gracias al Ejercicio 2.12 (c), si escribimos $x = \sum_{j=1}^n x_j e_j$ e $y = \sum_{k=1}^n y_k e_k$ entonces se cumple que

$$h(x, y) = {}^{\mathrm{t}}XH\overline{Y},$$

donde

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{C}^n \quad \text{y} \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{C}^n$$

son los vectores coordenadas de x e y respecto a $\mathscr{B}$. En particular, $h^*(x, y) = {}^{\mathrm{t}}XH^*\overline{Y}$ y luego $h^* = h$ (resp. $h^* = -h$) si y sólo si $H^* = H$ (resp. $H^* = -H$). □

Definición 2.14 (matriz (anti-)hermitiana)**.** Sea $A \in M_n(\mathbb{C})$ una matriz compleja. Decimos que A es una **matriz hermitiana** (resp. **matriz anti-hermitiana**) si $A^* = A$ (resp. $A^* = -A$). Denotaremos por $\mathrm{Herm}_n(\mathbb{C})$ (resp. $\mathrm{AntiHerm}_n(\mathbb{C})$) al espacio vectorial real de matrices hermitianas (resp. anti-hermitianas).

Observación 2.15.

1. *Si $A = (a_{jk}) \in M_n(\mathbb{C})$ es una matriz hermitiana, los coeficientes diagonales a_{jj} verifican $a_{jj} = \overline{a_{jj}}$, y luego $a_{jj} \in \mathbb{R}$. De manera análoga, si A es una matriz anti-hermitiana entonces sus coeficientes diagonales son imaginarios puros.*
2. *Notemos que $\dim_{\mathbb{C}} \mathrm{Sesq}(V) = n^2$ y la relación*

$$\dim_{\mathbb{R}} \mathrm{Herm}(V) = \dim_{\mathbb{R}} \mathrm{AntiHerm}(V) = \frac{1}{2} \dim_{\mathbb{R}} \mathrm{Sesq}(V) = \dim_{\mathbb{C}} \mathrm{Sesq}(V)$$

implican que $\dim_{\mathbb{R}} \mathrm{Herm}_n(\mathbb{C}) = \dim_{\mathbb{R}} \mathrm{AntiHerm}_n(\mathbb{C}) = n^2$.

Definición 2.16 (forma cuadrática hermitiana)**.** Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo y sea $h : V \times V \to \mathbb{C}$ una forma *hermitiana*. La **forma cuadrática hermitiana** asociada a h es por definición la función $Q : V \to \mathbb{R}$ dada por

$$Q(x) = h(x, x) \quad \text{para todo } x \in V.$$

En particular, para todo $x \in V$ y todo $\lambda \in \mathbb{C}$ se cumple

$$Q(\lambda x) = h(\lambda x, \lambda x) = \lambda\overline{\lambda}h(x, x) = |\lambda|^2 Q(x).$$

Observación 2.17. *Dado que h es una forma hermitiana, tenemos que $Q(x) = h(x, x) \in \mathbb{R}$ para todo $x \in V$ (c.f. Proposición 2.8). Luego, Q está asociada también a la forma* bilineal real *$B : V \times V \to \mathbb{R}$ dada por $(x, y) \mapsto B(x, y) := \mathrm{Re}\, h(x, y)$, la cual es simétrica pues*

$$B(y, x) = \mathrm{Re}\, h(y, x) = \mathrm{Re}\, \overline{h(x, y)} = \mathrm{Re}\, h(x, y) = B(x, y).$$

En particular, Q es una forma cuadrática real *si pensamos $V \cong \mathbb{R}^{2n}$ como un espacio vectorial real. Más aún, la forma cuadrática real Q verifica la propiedad adicional*

$$Q(\lambda x) = |\lambda|^2 Q(x) \quad \forall \lambda \in \mathbb{C},\ \forall x \in V.$$

Veremos que esta última propiedad permite distinguir a las formas cuadráticas hermitianas entre las formas cuadráticas reales (ver Proposición 2.19). Finalmente, es importante destacar que una forma cuadrática hermitiana **no es** *una forma cuadrática compleja asociada a una forma bilineal compleja sobre $V \cong \mathbb{C}^n$, puesto que en el último caso deberíamos tener $Q(\lambda x) = \lambda^2 Q(x)$ para todo $\lambda \in \mathbb{C}$ (lo cual impide que Q tenga valores reales, a menos que $Q = 0$).*

Tal como en el caso real, la **fórmula de polarización** nos permitirá deducir de manera única la forma hermitiana asociada a una forma cuadrática hermitiana.

Lema 2.18 (polarización hermitiana)**.** *Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo y sea $h : V \times V \to \mathbb{C}$ una forma hermitiana. Si denotamos por $Q : V \to \mathbb{R}$ la forma cuadrática hermitiana $Q(x) = h(x, x)$ asociada a h, entonces para todos $x, y \in V$ se tiene:*

1. $\operatorname{Re} h(x, y) = \frac{1}{2}(Q(x + y) - Q(x) - Q(y)) = \frac{1}{4}(Q(x + y) - Q(x - y))$.
2. $\operatorname{Im} h(x, y) = \frac{1}{2}(Q(x + iy) - Q(x) - Q(y)) = \frac{1}{4}(Q(x + iy) - Q(x - iy))$.

Luego,

$$h(x, y) = \frac{1}{4}(Q(x + y) - Q(x - y) + iQ(x + iy) - iQ(x - iy)).$$

Demostración. Para todos $x, y \in V$ tenemos que

$$\begin{aligned} Q(x + y) &= h(x + y, x + y) = Q(x) + Q(y) + h(x, y) + h(y, x) \\ Q(x - y) &= h(x - y, x - y) = Q(x) + Q(y) - h(x, y) - h(y, x) \end{aligned}$$

y, dado que $h(x, y) + h(y, x) = h(x, y) + \overline{h(x, y)} = 2 \operatorname{Re} h(x, y)$, obtenemos[2] (1). Dado que para todo $z \in \mathbb{C}$ se tiene que $\operatorname{Im}(z) = \operatorname{Re}(-iz)$, obtenemos

$$\operatorname{Im} h(x, y) = \operatorname{Re}(-ih(x, y)) = \operatorname{Re} h(x, iy),$$

por lo que (2) se obtiene de (1) reemplazando y por iy, y usando que $Q(iy) = |i|^2 Q(y) = Q(y)$. Finalmente, dado que $h(x, y) = \operatorname{Re} h(x, y) + i \operatorname{Im} h(x, y)$, la fórmula de polarización hermitiana

$$h(x, y) = \frac{1}{4}(Q(x + y) - Q(x - y) + iQ(x + iy) - iQ(x - iy)).$$

se obtiene de (1) y (2). □

La siguiente proposición prueba que las formas cuadráticas reales sobre $V \cong \mathbb{C}^n \cong \mathbb{R}^{2n}$ que verifican la propiedad adicional $Q(\lambda x) = |\lambda|^2 Q(x)$ para todo $\lambda \in \mathbb{C}$ corresponden exactamente a formas cuadráticas hermitianas. Mejor aún, veremos que basta considerar $\lambda = i$.

Proposición 2.19. *Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo. Si $Q : V \to \mathbb{R}$ es una forma cuadrática real en $V \cong \mathbb{R}^{2n}$ verificando adicionalmente que $Q(ix) = Q(x)$ para todo $x \in V$, entonces la fórmula de polarización hermitiana define una forma sesquilineal hermitiana h. En particular, Q es una forma cuadrática hermitiana.*

Demostración. El hecho que $Q : V \to \mathbb{R}$ sea una forma cuadrática real implica que

$$B(x, y) = \frac{1}{4}(Q(x + y) - Q(x - y))$$

es una forma bilineal simétrica real. Deducimos que

$$h(x, y) := \frac{1}{4}(Q(x + y) - Q(x - y) + iQ(x + iy) - iQ(x - iy)) = B(x, y) - iB(x, iy)$$

[2]Notemos que (1) es exactamente la fórmula de polarización *real* aplicada a la forma bilineal simétrica $(x, y) \mapsto \operatorname{Re} h(x, y)$.

es una forma $\mathbb{R}$-bilineal con valores complejos. Más aún, por una parte tenemos que

$$h(x,iy)=\frac{1}{4}(Q(x+iy)-Q(x-iy)+iQ(x-y)-iQ(x+y))=-ih(x,y).$$

Por otra parte, la hipótesis $Q(ix)=Q(x)$ implica que $Q(x)=Q(-x)=Q(-ix)$, de donde deducimos

$$\begin{aligned}h(ix,y)&=\frac{1}{4}(Q(ix+y)-Q(ix-y)+iQ(ix+iy)-iQ(ix-iy))\\&=\frac{1}{4}(Q(x-iy)-Q(x+iy)+iQ(x+y)-iQ(x-y))=ih(x,y).\end{aligned}$$

De lo anterior y de la $\mathbb{R}$-bilinealidad de h se obtiene que $h:V\times V\to\mathbb{C}$ es una forma hermitiana. $\square$

Terminemos esta sección definiendo el rango de una forma hermitiana o, equivalentemente, de una forma cuadrática hermitiana.

Definición 2.20 (rango)**.** Sea $V\cong\mathbb{C}^n$ un espacio vectorial complejo y sea $h:V\times V\to\mathbb{C}$ una forma hermitiana. Dada una base $\mathscr{B}=(e_1,\ldots,e_n)$ de V, definimos el **rango** de h (o de la forma cuadrática hermitiana $Q:V\to\mathbb{R}$ asociada a h) como $\mathrm{rg}(h)=\mathrm{rg}(Q):=\mathrm{rg}(H)$, donde $H=\mathrm{Mat}_{\mathscr{B}}(h)$ es la matriz de h respecto a la base[3] $\mathscr{B}$. Diremos que la forma hermitiana h (o que Q) es **no-degenerada** si $\mathrm{rg}(h)=\dim_{\mathbb{C}}(V)$, i.e., si la matriz H es invertible.

3. Ortogonalidad hermitiana y Teorema de Sylvester

En esta sección estudiaremos el análogo hermitiano de la noción de ortogonalidad respecto a una forma bilineal simétrica (o, equivalentemente, respecto a una forma cuadrática).

Definición 3.1. Sea $V\cong\mathbb{C}^n$ un espacio vectorial complejo y sea $h:V\times V\to\mathbb{C}$ una forma hermitiana. Diremos que:

1. Dos vectores $x,y\in V$ son **ortogonales** respecto a h si $h(x,y)=0$ (o equivalemente, si $h(y,x)=0$).
2. Si $U\subseteq V$ es un sub-conjunto no-vacío, el conjunto[4]

$$U^{\perp}=\{y\in V\mid \forall x\in U,\ h(x,y)=0\}$$

es el **ortogonal de** U.
3. El **kernel** de h es

$$\ker(h):=V^{\perp}=\{x\in V\mid \forall y\in V,\ h(x,y)=0\}.$$

4. El conjunto de **vectores isótropos** de la forma cuadrática hermitiana $Q(x)=h(x,x)$ asociada es $\{x\in V\mid Q(x)=0\}$.

Observación 3.2. *Cabe destacar que a pesar que $U\subseteq V$ pueda ser símplemente un sub-conjunto (y no necesariamente un* sub-espacio vectorial*) de V, el ortogonal $U^{\perp}\subseteq V$ es* siempre *un sub-espacio vectorial*[5] *de V. Por otro lado, el conjunto de vectores isótropos de una forma cuadrática hermitiana Q (asociada a la forma hermitiana h) es un* **cono complejo***, i.e., si $x\in V$ es un vector isótropo entonces λx es un vector isótropo para todo $\lambda\in\mathbb{C}$. En general, se tiene la inclusión* $\ker(h)\subseteq\{x\in V\mid Q(x)=0\}$ *pero dicha inclusión puede ser estricta, tal como lo muestra el siguiente ejemplo.*

[3]Notamos que dicho rango **no depende** de la base escogida, pues si $\mathscr{B}'$ es otra base y $P=\mathrm{Mat}_{\mathscr{B}}(\mathscr{B}')$ es la matriz de cambio de base, entonces $H'=\mathrm{Mat}_{\mathscr{B}'}(h)={}^{\mathrm{t}}PH\overline{P}$ y, dado que ${}^{\mathrm{t}}P,\overline{P}\in\mathrm{GL}_n(\mathbb{C})$, se tiene que $\mathrm{rg}(H')=\mathrm{rg}(H)$.

[4]En ocasiones, para insistir en el hecho que $U^{\perp}$ es el ortogonal respecto a h (o equivalentemente, respecto a la forma cuadrática hermitiana Q asociada a h), es que se denota también $U^{\perp_h}$ o $U^{\perp_Q}$.

[5]Si $y_1,y_2\in U^{\perp}$ y $\lambda\in\mathbb{C}$ entonces para todo $x\in U$ se cumple $h(x,\lambda y_1+y_2)=\overline{\lambda}h(x,y_1)+h(x,y_2)=0$, i.e., $\lambda y_1+y_2\in U^{\perp}$.

Ejemplo 3.3. Sea $V = \mathbb{C}^2$ y consideremos la forma cuadrática hermitiana

$$Q(z_1, z_2) = z_1\overline{z_1} - z_2\overline{z_2} = |z_1|^2 - |z_2|^2.$$

La fórmula de polarización hermitiana implica que $h((z_1, z_2), (w_1, w_2)) = z_1\overline{w_1} - z_2\overline{w_2}$ es la forma hermitiana asociada. La matríz H de la forma h respecto a la base canónica $\mathscr{B}$ de $\mathbb{C}^2$ está dada por

$$H = \mathrm{Mat}_{\mathscr{B}}(h) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

la cual es una matriz invertible y luego h es no-degenerada. En particular,

$$\ker(h) = \{0\} \neq \{(z_1, z_2) \in \mathbb{C}^2 \mid Q(z_1, z_2) = 0\} = \{(z_1, z_2) \in \mathbb{C}^2 \mid |z_1| = |z_2|\}.$$

Ejercicio 3.4. Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo y sea $h : V \times V \to \mathbb{C}$ una forma hermitiana. Sean $U, W \subseteq V$ sub-conjuntos no-vacíos de V.

a) Probar que si $U \subseteq V$ entonces $V^\perp \subseteq U^\perp$.

b) Probar que $U^\perp = \mathrm{Vect}_{\mathbb{C}}(U)^\perp$.

c) Supongamos que $U \subseteq V$ es un sub-espacio vectorial y que $(e_1, \ldots, e_r)$ es un conjunto generador de U (no necesariamente linealmente independiente), probar que

$$U^\perp = \{e_1, \ldots, e_r\}^\perp = \{x \in V \mid h(x, e_j) = 0 \ \forall j = 1, \ldots, r\}.$$

Teorema 3.5. *Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo y sea $h : V \times V \to \mathbb{C}$ una forma hermitiana. Entonces, para todo sub-espacio vectorial $U \subseteq V$ se tiene que:*

1. $U \subseteq (U^\perp)^\perp$ y $\dim_{\mathbb{C}}(U^\perp) \geq \dim_{\mathbb{C}}(V) - \dim_{\mathbb{C}}(U)$.

2. $\ker(h) = \{0\}$ si y sólo si h es no-degenerada.

3. Si h es no-degenerada, entonces $U = (U^\perp)^\perp$ y $\dim_{\mathbb{C}}(U^\perp) = \dim_{\mathbb{C}}(V) - \dim_{\mathbb{C}}(U)$.

4. Si $U \cap U^\perp = \{0\}$ entonces $V = U \oplus U^\perp$.

Demostración. Sea $x \in U$, entonces para todo $y \in U^\perp$ se tiene $h(x, y) = \overline{h(y, x)} = 0$ por lo que $x \in (U^\perp)^\perp$, de donde se prueba la primera parte de (1). Para probar la segunda parte de (1) consideramos una base $(e_1, \ldots, e_r)$ una base de U, donde $r = \dim_{\mathbb{C}}(U)$, la cual completamos en una base $\mathscr{B} = (e_1, \ldots, e_n)$ de V. Sea $H = (h_{jk})$ la matriz de h respecto a la base $\mathscr{B}$ (i.e., $h_{jk} = h(e_j, e_k)$ para todo $j, k = 1, \ldots, n$). Sea $x = x_1e_1 + \ldots + x_ne_n \in V$, entonces $x \in U^\perp$ si y sólo si $h(x, e_j) = 0$ para $j = 1, \ldots, r$. Dado que

$$h(x, e_j) = h(x_1e_1 + \ldots + x_ne_n, e_j) = \sum_{k=1}^{n} x_k h(e_k, e_j) = \sum_{k=1}^{n} h_{kj}x_k,$$

la condición $x \in U^\perp$ equivale a decir que el vector de coordenadas $X = (x_1, \ldots, x_n) \in \mathbb{C}^n$ es solución del sistema lineal

$$\begin{cases} a_{11}x_1 & + \cdots + & a_{n1}x_n = & 0 \\ \vdots & & \vdots & \\ a_{1r}x_1 & + \cdots + & a_{nr}x_n = & 0 \end{cases}$$

cuya matriz asociada $A \in M_{n\times r}(\mathbb{C})$ es la matriz obtenida al considerar las r primeras filas de ${}^{\mathrm{t}}H$. Dado que el espacio de soluciones del sistema es de dimensión $n - \mathrm{rg}(A)$, obtenemos

$$\dim_{\mathbb{C}}(U^\perp) = n - \mathrm{rg}(A) \geq n - r,$$

de donde obtenemos la segunda parte de (1). Más aún, en el caso particular donde $U = V$, tenemos que $A = {}^{\mathrm{t}}H$ y luego $\dim_{\mathbb{C}}(U^\perp) = n - \mathrm{rg}({}^{\mathrm{t}}H) = n - \mathrm{rg}(H)$. Luego, $\ker(h) = V^\perp$ es nulo si y sólo si $\mathrm{rg}(H) = n$, de donde obtenemos (2).

Para probar (3) supongamos que h es no-degenerada, i.e. las columnas de la matriz H son linealmente independientes. En particular, las primeras r columnas de H son linealmente independientes y luego $\mathrm{rg}(A) = r$, de donde deducimos que $\dim_{\mathbb{C}}(U^{\perp}) = n - r$. Reemplazando U por $U^{\perp}$ obtenemos la igualdad $\dim_{\mathbb{C}}(U^{\perp})^{\perp} = n - (n - r) = r$ y luego la inclusión $U \subseteq (U^{\perp})^{\perp}$ es una igualdad, de donde obtenemos (3).

Finalmente, si suponemos que $U \cap U^{\perp} = \{0\}$ (sin suponer necesariamente que h es no-degenerada), entonces U y $U^{\perp}$ están en suma directa y la dimensión del sub-espacio $U \oplus U^{\perp}$ de V es $d = r + \dim_{\mathbb{C}}(U^{\perp})$. Gracias a (1), tenemos que $d \geq n$ y luego $V = U \oplus U^{\perp}$ (en particular, $\dim_{\mathbb{C}}(U^{\perp}) = n - r$), lo que prueba (4). □

Un caso particular importante es cuando los vectores de una base son ortogonales respecto a una forma hermitiana.

Definición 3.6 (base ortogonal)**.** Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo, sea $h : V \times V \to \mathbb{C}$ una forma hermitiana y sea $Q : V \to \mathbb{R}$ la forma cuadrática hermitiana asociada. Decimos que una base $\mathscr{B} = (e_1, \ldots, e_n)$ de V es una **base ortogonal** respecto a h (o respecto a Q) si

$$h(e_j, e_k) = 0 \text{ para todos } j \neq k.$$

Equivalentemente, la matriz $H = \mathrm{Mat}_{\mathscr{B}}(h)$ es una matriz *diagonal*

$$H = \begin{pmatrix} \lambda_1 & 0 & \cdots & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 & \lambda_n \end{pmatrix}$$

donde los $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ son números *reales* (pues H es hermitiana), siendo esto último a su vez equivalente a que si $(z_1, \ldots, z_n)$ son coordenadas respecto a la base $\mathscr{B}$ entonces

$$Q(z_1, \ldots, z_n) = \lambda_1 |z_1|^2 + \ldots + \lambda_n |z_n|^2.$$

El siguiente resultado afirma que siempre podemos encontrar al menos una base ortogonal respecto a una forma hermitiana dada.

Lema 3.7. *Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo y sea $h : V \times V \to \mathbb{C}$ una forma hermitiana. Entonces, existe una base $\mathscr{B}$ de V que es ortogonal respecto a h.*

Demostración. Demostraremos la existencia de una base ortogonal por inducción en la dimensión $n = \dim_{\mathbb{C}}(V)$. Dado que el resultado es cierto si $n = 0$ o si $h = 0$, podemos suponer que el resultado es cierto para $n - 1$, y que además $h \neq 0$. En particular, la forma cuadrática hermitiana $Q : V \to \mathbb{R}$ asociada a h es no-nula y luego existe un vector $e_1 \in V$ tal que $Q(e_1) \neq 0$.
Sea $U = \mathrm{Vect}_{\mathbb{C}}(e_1)$ la recta generada por e_1. Dado que $Q(e_1) = h(e_1, e_1) \neq 0$, tenemos que $U \cap U^{\perp} = \{0\}$ y luego $V = U \oplus U^{\perp}$. La hipótesis de inducción implica que existe una base $(e_2, \ldots, e_n)$ de $U^{\perp}$ tal que $h(e_j, e_k) = 0$ para todos $j \neq k$ en $\{2, \ldots, n\}$. Luego, $\mathscr{B} = (e_1, e_2, \ldots, e_n)$ es una base de V ortogonal respecto a h. □

Concluímos esta sección con el análogo hermitiano del teorema de Sylvester de clasificación de formas cuadráticas reales.

Teorema 3.8 (de Sylvester, caso hermitiano)**.** *Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo, sea $h : V \times V \to \mathbb{C}$ una forma hermitiana, y sea $Q : V \to \mathbb{R}$ la forma cuadrática hermitiana asociada. Sea $\mathscr{B} = (e_1, \ldots, e_n)$ una base de V ortogonal respecto a h y sea $H = \mathrm{Mat}_{\mathscr{B}}(h)$ la matriz diagonal real de coeficientes $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. Entonces,*

1. *Módulo permutación de los elementos de la base $\mathscr{B}$, podemos suponer que $\lambda_1, \ldots, \lambda_r \neq 0$ y que $\lambda_j = 0$ si $j > r$. En particular, si $(z_1, \ldots, z_n)$ son las coordenadas respecto a la base $\mathscr{B}$, entonces*

$$Q(z_1, \ldots, z_n) = \lambda_1 |z_1|^2 + \ldots + \lambda_r |z_r|^2.$$

2. *Sea $p \in \mathbb{N}$ (resp. $q \in \mathbb{N}$) el número de índices $j \in \{1, \ldots, n\}$ tales que $Q(e_j) > 0$ (resp. $Q(e_j) < 0$). Entonces, p y q* no dependen *de la base ortogonal escogida. El par (p, q) se llama la* **signatura**[6] *de la forma hermitiana h (o de Q), y se cumple que $p + q = r = \mathrm{rg}(h)$.*

3. *El kernel* $\ker(h)$ *es el sub-espacio* $\mathrm{Vect}_{\mathbb{C}}(e_{r+1}, \ldots, e_n)$ *dado por las ecuaciones $z_1 = \ldots = z_r = 0$.*

Más aún, podemos escoger $\mathscr{B}$ de tal suerte que

$$H = \left(\begin{array}{c|c|c} \mathrm{I}_p & 0 & 0 \\ \hline 0 & -\mathrm{I}_q & 0 \\ \hline 0 & 0 & \mathbf{0}_{n-r} \end{array}\right).$$

Demostración. El hecho que $\mathscr{B}$ sea una base ortogonal respecto a h implica directamente (1), así como la igualdad $p + q = r = \mathrm{rg}(h)$ en (2).

Comencemos por probar (3). La fórmula de polarización hermitiana implica que, respecto a la base $\mathscr{B}$, la forma hermitiana h está dada por

$$h(x, y) = \lambda_1 x_1 \overline{y_1} + \ldots + \lambda_r x_r \overline{y_r},$$

donde $x = \sum_{j=1}^n x_j e_j$ e $y = \sum_{j=1}^n y_j e_j$. Supongamos que $x = \sum_{j=1}^n x_j e_j$ pertenece a $\ker(h)$ y consideremos $y = e_j$ para $j = 1, \ldots, r$ (i.e., $y_j = 1$ e $y_k = 0$ para $k \neq j$), de donde obtenemos que $x_j = 0$ y en particular $x \in \mathrm{Vect}_{\mathbb{C}}(e_{r+1}, \ldots, e_n)$. Recíprocamente, la fórmula de h en la base $\mathscr{B}$ implica que todo vector $x = \sum_{j=1}^n x_j e_j$ en $\mathrm{Vect}_{\mathbb{C}}(e_{r+1}, \ldots, e_n)$ (i.e., tal que $x_1 = \ldots = x_r = 0$) pertenece a $\ker(h)$, de donde se obtiene (3).

Veamos ahora (2), i.e., que el par (p, q) es independiente de la base ortogonal escogida. Recordemos que $r = \mathrm{rg}(h)$. Sean $\mathscr{B} = (e_1, \ldots, e_n)$ y $\mathscr{B}' = (e'_1, \ldots, e'_n)$ dos bases de V ortogonales respecto a h. Denotemos por p (resp. p') al número de índices $j \in \{1, \ldots, n\}$ tales que $Q(e_j) > 0$ (resp. $Q(e'_j) > 0$) y por q (resp. q') al número de índices $j \in \{1, \ldots, n\}$ tales que $Q(e_j) < 0$ (resp. $Q(e'_j) < 0$). Entonces,

$$p + q = r = p' + q'.$$

En particular, basta probar que $q = q'$ para obtener que $(p, q) = (p', q')$. Reordenando los elementos de $\mathscr{B}$ y $\mathscr{B}'$ si fuese necesario, podemos suponer que

$$(\star) \qquad \begin{cases} Q(e_j) > 0 & \text{si } j = 1, \ldots, p \\ Q(e_j) < 0 & \text{si } j = p+1, \ldots, p+q = r \\ Q(e_j) = 0 & \text{si } j > p+q = r \end{cases} \qquad \begin{cases} Q(e'_j) > 0 & \text{si } j = 1, \ldots, p' \\ Q(e'_j) < 0 & \text{si } j = p'+1, \ldots, p'+q' = r \\ Q(e'_j) = 0 & \text{si } j > p'+q' = r \end{cases}$$

Sea P_+ el sub-espacio vectorial de V generado por los $n - q$ vectores e_j tales que $Q(e_j) \geq 0$. En particular, $\dim_{\mathbb{C}} P_+ = n - q$. Sea $x \in P_+$ y escribamos $x = \sum_{j \in I} x_j e_j$ donde $J = \{1, \ldots, p\} \cup \{r+1, \ldots, n\}$. Luego, las relaciones $(\star)$ implican que

$$Q(x) = \sum_{j=1}^{p} |x_j|^2 Q(e_j) \geq 0.$$

Por otra parte, si P'_- es el sub-espacio vectorial de V generado por los q' vectores e'_j tales que $Q(e'_j) < 0$ entonces $\dim_{\mathbb{C}} P'_- = q'$. Sea $y \in P'_-$ un vector no-nulo y escribamos $y = \sum_{j=p'+1}^r y_j e'_j$, donde al menos uno de los $y_j \neq 0$ (pues $y \neq 0$). Entonces, las relaciones $(\star)$ implican que

$$Q(y) = \sum_{j=p'+1}^{r} |y_j|^2 Q(e'_j) < 0.$$

En consecuencia, tenemos que $P_+ \cap P'_- = \{0\}$ y por lo tanto

$$n = \dim_{\mathbb{C}}(V) \geq \dim_{\mathbb{C}}(P_+) + \dim_{\mathbb{C}}(P'_-) = n - q + q',$$

de donde deducimos que $q \geq q'$. Intercambiando los roles de $\mathscr{B}$ y $\mathscr{B}'$ en la discusión anterior, obtenemos del mismo modo que $q' \geq q$. Así, $q = q'$ y por ende $p = p'$, lo cual prueba (2).

[6] En ocasiones, se define también la signatura de h como el trío $(\sigma_+, \sigma_-, \sigma_0)$, donde $\sigma_+ = p$, $\sigma_- = q$ y $\sigma_0 = \dim_{\mathbb{C}} \ker(h)$.

Finalmente, si ordenamos los elementos de una base ortogonal $\mathscr{B} = (e_1, \dots, e_n)$ como en $(\star)$ y consideramos el valor absoluto *real* $|Q(e_j)| > 0$ para $j = 1, \dots, p+q = r$, entonces al reemplazar e_j por el vector $e_j/\sqrt{|Q(e_j)|}$ para cada $j = 1, \dots, r$ obtenemos una base ortogonal $\mathscr{B}'$ tal que $\mathrm{Mat}_{\mathscr{B}'}(h)$ es de la forma

$$\left(\begin{array}{c:c:c} \mathrm{I}_p & 0 & 0 \\ \hdashline 0 & -\mathrm{I}_q & 0 \\ \hdashline 0 & 0 & \mathbf{0}_{n-r} \end{array}\right),$$

de donde se obtiene el resultado. □

4. Grupo unitario $\mathbf{U}(p,q)$ y grupo especial unitario $\mathbf{SU}(p,q)$

Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo y sea $Q : V \to \mathbb{R}$ una forma cuadrática hermitiana. Denotamos por $h : V \times V \to \mathbb{C}$ la única forma hermitiana asociada a Q mediante la fórmula de polarización hermitiana.

Ejercicio 4.1. Sea $u : V \to V$ un endomorfismo. Probar (usando la fórmula de polarización hermitiana) que las propiedades siguientes son equivalentes:

1. u preserva la forma hermitiana h, i.e., $h(u(x), u(y)) = h(x, y)$ para todos $x, y \in V$.
2. u preserva la forma cuadrática hermitiana Q, i.e., $Q(u(x)) = Q(x)$ para todo $x \in V$.

Definición 4.2 (automorfismo unitario)**.** Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo y sea $Q : V \to \mathbb{R}$ una forma cuadrática hermitiana. Decimos que un automorfismo $u : V \xrightarrow{\sim} V$ (i.e., $u \in \mathrm{GL}(V) \cong \mathrm{GL}_n(\mathbb{C})$) es **unitario** respecto a Q si u preserva Q (o equivalentemente, preserva la forma hermitiana asociada h). Denotamos por

$$\mathbf{U}(Q) = \{u \in \mathrm{GL}(V) \mid u \text{ unitario respecto a } Q\} \subseteq \mathrm{GL}(V)$$

al conjunto de automorfismos unitarios respecto a Q.

Proposición 4.3 (grupo unitario)**.** *Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo y sea $Q : V \to \mathbb{R}$ una forma cuadrática hermitiana. El conjunto $\mathbf{U}(Q)$ es un* sub-grupo *de* $\mathrm{GL}(V)$*, llamado el* **grupo unitario de** Q.

Demostración. La identidad Id_V preserva la forma cuadrática hermitiana Q y luego $\mathrm{Id}_V \in \mathbf{U}(Q)$. Por otra parte, si $u, v \in \mathbf{U}(Q)$ preservan Q entonces para todo $x \in V$ se tiene

$$Q((u \circ v)(x)) = Q(u(v(x)) = Q(v(x)) = Q(x),$$

por lo que la composición $u \circ v \in \mathbf{U}(Q)$ preserva Q. Finalmente, si $u \in \mathbf{U}(Q)$ preserva Q entonces para todo $x \in V$ escribimos $y = u^{-1}(x)$ y calculamos

$$Q(u^{-1}(x)) = Q(y) = Q(u(y)) = Q(x),$$

por lo que $u^{-1} \in \mathbf{U}(Q)$ preserva Q. Así, concluímos que $\mathbf{U}(Q)$ es un sub-grupo de $\mathrm{GL}(V)$. □

⚠ Si la forma cuadrática hermitiana $Q : V \to \mathbb{R}$ es *no-degenerada*, entonces todo endomorfismo $u : V \to V$ que preserva Q es automáticamente biyectivo (i.e., $u \in \mathrm{GL}(V)$). En efecto, si denotamos por $h : V \times V \to \mathbb{C}$ la forma hermitiana asociada a Q y si $x \in V$ verifica $u(x) = 0$, entonces

$$h(x, y) = h(u(x), u(y)) = 0 \quad \text{para todo } y \in V.$$

En particular, $x \in V^\perp$ y por ende $x = 0$ (pues $V^\perp = \{0\}$ si Q es no-degenerada). Luego, u es inyectivo y por ende biyectivo (pues $V \cong \mathbb{C}^n$ es de dimensión finita).

Convención: Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo y $Q : V \to \mathbb{R}$ una forma cuadrática hermitiana. Si $\mathscr{B}$ es una base de V, diremos que la **matriz de** Q respecto a la base $\mathscr{B}$ es la matríz de la única forma hermitiana $h : V \times V \to \mathbb{C}$ asociada a Q mediante la fórmula de polarización hermitiana. En otras palabras,

$$\mathrm{Mat}_{\mathscr{B}}(Q) := \mathrm{Mat}_{\mathscr{B}}(h) = H \in M_n(\mathbb{C}).$$

Proposición 4.4 (grupo especial unitario)**.** *Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo y $Q : V \to \mathbb{R}$ una forma cuadrática hermitiana. Sea $\mathscr{B}$ una base de V y denotemos por H la matriz de Q respecto a la base $\mathscr{B}$. Dado un automorfismo $u : V \xrightarrow{\sim} V$ de V de matriz $A = \mathrm{Mat}_{\mathscr{B}}(u)$ respecto a $\mathscr{B}$, tenemos que:*

1. *$u \in \mathbf{U}(Q)$ (i.e., u es unitario) si y sólo si ${}^{\mathrm{t}}AH\overline{A} = H$.*

2. *Si Q es* no-degenerada *y $u \in \mathbf{U}(Q)$ entonces $|\det(u)| = 1$. Más aún, el conjunto*

$$\mathbf{SU}(Q) = \{u \in \mathbf{U}(Q) \mid \det(u) = 1\}$$

es un sub-grupo de $\mathbf{U}(Q)$, llamado el **grupo especial unitario de** *Q.*

Demostración. Sean $x, y \in V$. Si representamos x e y por los vectores columna $X, Y \in \mathbb{C}^n$ dados por sus coordenadas respecto a la base $\mathscr{B}$, entonces sabemos que

$$h(x, y) = {}^{\mathrm{t}}XH\overline{Y},$$

por lo que $h(u(x), u(y)) = {}^{\mathrm{t}}(AX)H\overline{(AY)} = {}^{\mathrm{t}}X({}^{\mathrm{t}}AH\overline{A})\overline{Y}$. Luego, la matriz de la forma hermitiana $(x, y) \mapsto h(u(x), u(y))$ respecto a la base $\mathscr{B}$ es ${}^{\mathrm{t}}AH\overline{A}$. Así, u es unitario respecto a Q si y sólo si las formas hermitianas $h(\cdot,\cdot)$ y $h(u(\cdot), u(\cdot))$ coínciden, lo cual equivale a la igualdad $H = {}^{\mathrm{t}}AH\overline{A}$ en (1).

Si Q es no-degenerada entonces $H \in \mathrm{GL}_n(\mathbb{C})$ y luego $\det(H) \neq 0$. En particular, si $u \in \mathbf{U}(Q)$ entonces, gracias a (1), tenemos que $H = {}^{\mathrm{t}}AH\overline{A}$ y por ende $\det(H) = \det(A)\det(H)\det(\overline{A})$. Esto último implica que $\det(A)\det(\overline{A}) = \det(A)\overline{\det(A)} = |\det(A)|^2 = 1$, de donde deducimos que en este caso $|\det(u)| = 1$. Finalmente, dado que $\det(\mathrm{Id}_V) = 1$ y que $\det(u \circ v) = \det(u)\det(v)$, concluímos directamente que $\mathrm{Id}_V \in \mathbf{SO}(U)$, que $u, v \in \mathbf{SU}(Q)$ entonces $u \circ v \in \mathbf{SU}(Q)$, y que si $u \in \mathbf{SU}(Q)$ entonces $u^{-1} \in \mathbf{SU}(Q)$. En otras palabras, $\mathbf{SU}(Q)$ es un sub-grupo del grupo unitario $\mathbf{U}(Q)$, probando así (2). ☐

Definición 4.5 (equivalencia de formas cuadráticas hermitianas)**.** Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo, y sean $Q : V \to \mathbb{R}$ y $Q' : V \to \mathbb{R}$ dos formas cuadráticas hermitianas. Decimos que las formas Q y Q' son **equivalentes** si existe un automorfismo $v \in \mathrm{GL}(V)$ tal que $Q(x) = Q'(v(x))$ para todo $x \in V$, y escribiremos $Q \sim Q'$.

Proposición 4.6. *Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo, y sean $Q : V \to \mathbb{R}$ y $Q' : V \to \mathbb{R}$ dos formas cuadráticas hermitianas. Si las formas $Q \sim Q'$ son equivalentes, entonces los grupos unitarios asociados $\mathbf{U}(Q) \cong \mathbf{U}(Q')$ son isomorfos.*

Demostración. Sea $v \in \mathrm{GL}(V)$ automorfismo tal que $Q(x) = Q'(v(x))$ para todo $x \in V$. Dado $u : V \to V$ endomorfismo unitario respecto a Q (i.e., $u \in \mathbf{U}(Q)$) y dado $x \in V$, tenemos que:

$$Q'(x) = Q'((v \circ v^{-1})(x)) = Q(v^{-1}(x)) = Q((u \circ v^{-1}(x)) = Q'((v \circ u \circ v^{-1})(x)).$$

En particular, $v \circ u \circ v^{-1} \in \mathbf{U}(Q')$. Finalmente, la función

$$\varphi : \mathbf{U}(Q) \longrightarrow \mathbf{U}(Q'),\ u \longmapsto \varphi(u) = v \circ u \circ v^{-1}$$

es biyectiva (pues su inversa está dada explícitamente por $\varphi^{-1}(u') = v^{-1} \circ u' \circ v$ para todo $u' \in \mathbf{U}(Q')$) y preserva las estructuras de grupos, i.e., $\varphi(u_1 \circ u_2) = \varphi(u_1) \circ \varphi(u_2)$. Concluímos así que $\mathbf{U}(Q) \cong \mathbf{U}(Q')$. ☐

Caso particular importante: Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo. Gracias al Teorema de Sylvester y a la proposición anterior, el estudio de los grupos unitarios de formas cuadráticas hermitianas **no-degeneradas** de signatura (p, q), con $p + q = n$ en este caso, se reduce a estudiar

$$Q(z_1, \ldots, z_n) = |z_1|^2 + \ldots + |z_p|^2 - |z_{p+1}|^2 - \ldots - |z_n|^2.$$

El grupo unitario correspondiente es denotado $\mathbf{U}(p, q)$ y el grupo especial unitario asociado es denotado $\mathbf{SU}(p, q)$. Explícitamente, si consideramos

$$H_{p,q} := \left(\begin{array}{c|c} \mathrm{I}_p & 0 \\ \hline 0 & -\mathrm{I}_q \end{array}\right) \in \mathrm{GL}_n(\mathbb{C}),$$

entonces $\mathbf{U}(p, q) \cong \{A \in \mathrm{GL}_n(\mathbb{C}) \mid {}^{\mathrm{t}}AH_{p,q}\overline{A} = H_{p,q}\}$ y $\mathbf{SU}(p, q) \cong \{A \in \mathbf{U}(p, q) \mid \det(A) = 1\}$.

Si $(p, q) = (n, 0)$ (i.e., Q es una forma cuadrática hermitiana **definida positiva**) escribimos $\mathbf{U}(n)$ y $\mathbf{SU}(n)$ en lugar de $\mathbf{U}(n, 0)$ y $\mathbf{SU}(n, 0)$. Explícitamente,

$$\mathbf{U}(n) \cong \{A \in \mathrm{GL}_n(\mathbb{C}) \mid {}^{\mathrm{t}}A\overline{A} = \mathrm{I}_n\} = \{A \in \mathrm{GL}_n(\mathbb{C}) \mid A^*A = \mathrm{I}_n\}.$$

5. Producto escalar complejo y espacios hermitianos

Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo. Recordemos que una forma hermitiana $Q : V \to \mathbb{R}$ es **definida positiva** si su signatura es $(n, 0)$. En otras palabras, existe una base $\mathscr{B}$ de V tal que, si $(z_1, \ldots, z_n)$ son las coordenadas respecto a dicha base, entonces se tiene

$$Q(z_1, \ldots, z_n) = |z_1|^2 + \ldots + |z_n|^2 = z_1\overline{z_1} + \ldots + z_n\overline{z_n}.$$

En particular, toda forma cuadrática hermitiana definida positiva es *no-degenerada*, i.e., si $x \in V$ entonces $Q(x) = 0$ implica que $x = 0$. Más aún, $Q(x) > 0$ para todo $x \in V \setminus \{0\}$ no-nulo.

Definición 5.1 (producto escalar y espacio hermitiano)**.** Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo. Un **producto escalar complejo** es una forma hermitiana $h : V \times V \to \mathbb{C}$ tal que la forma cuadrática hermitiana asociada es *definida positiva*. Un **espacio hermitiano**[7] es un espacio vectorial complejo dotado de un producto escalar complejo.

Notación: El producto escalar entre dos vectores $x, y \in V \cong \mathbb{C}^n$ se denota usualmente como $\langle x, y\rangle$ en lugar de $h(x, y)$. Así, la forma hermitiana

$$\langle\cdot,\cdot\rangle : V \times V \longrightarrow \mathbb{C},\ (x, y) \longmapsto \langle x, y\rangle$$

es:

1. lineal en la *primera* variable, i.e., para todos $x, y, z \in V$ y $\lambda \in \mathbb{C}$ se tiene que
$$\langle\lambda x, y\rangle = \lambda\langle x, y\rangle \quad \text{y} \quad \langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle.$$
2. anti-lineal en la *segunda* variable, i.e., para todos $x, y, z \in V$ y $\lambda \in \mathbb{C}$ se tiene que
$$\langle x, \lambda y\rangle = \overline{\lambda}\langle x, y\rangle \quad \text{y} \quad \langle x, y + z\rangle = \langle x, y\rangle + \langle x, z\rangle.$$
3. hermitiana, i.e., para todos $x, y \in V$ se tiene que
$$\langle x, y\rangle = \overline{\langle y, x\rangle}.$$
4. definida positiva, i.e., $\langle x, x\rangle \in \mathbb{R}^{>0}$ para todo $x \in V \setminus \{0\}$ no-nulo. En otras palabras, el único vector isótropo es el vector nulo.

Ejemplo 5.2. Sea $V = \mathbb{C}^n$. Para todos $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in \mathbb{C}^n$, la relación

$$\langle x, y\rangle := x_1\overline{y_1} + x_2\overline{y_2} + \ldots + x_n\overline{y_n}$$

define un producto escalar complejo en $\mathbb{C}^n$, llamado el *producto escalar canónico*. Así, $\mathbb{C}^n$ es un espacio hermitiano.

⚠ **Terminología:** La generalización de los conceptos de espacio euclideano y de espacio hermitiano a *dimensión infinita* es lo que conocemos como un **espacio de Hilbert**[8]. Así, un espacio euclideano (resp. espacio hermitiano) no es nada más que un espacio de Hilbert real (resp. complejo) de dimensión finita.

Definición 5.3 (base ortonormal)**.** Sea $V \cong \mathbb{C}^n$ un espacio hermitiano. Diremos que

1. una familia de vectores $(e_j)_{j\in J}$, donde J es un conjunto de índices arbitrario, es **ortonormal** si $\langle e_j, e_k\rangle = \delta_{jk}$ para todos $j, k \in J$ (i.e., $\langle e_j, e_j\rangle = 1$ y $\langle e_j, e_k\rangle = 0$ si $j \neq k$).
2. una base $\mathscr{B} = (e_1, \ldots, e_n)$ es una **base ortonormal** de V si $\mathscr{B}$ es una familia ortonormal.

Lema 5.4. *Sea $V \cong \mathbb{C}^n$ un espacio hermitiano. Entonces, toda familia ortonormal de vectores es linealmente independiente. En particular, toda familia ortonormal $(e_1, \ldots, e_n)$ de cardinal $n = \dim_{\mathbb{C}}(V)$ es una base ortonormal de V.*

[7]En honor al matemático francés Charles Hermite (1822–1901).
[8]En honor al matemático alemán David Hilbert (1862–1943).

Demostración. Sea $(e_j)_{j\in J}$ una familia ortonormal y supongamos que existe una relación lineal

$$\lambda_1 e_{j_1} + \ldots + \lambda_r e_{j_r} = 0$$

donde $j_1, \ldots, j_r \in J$ son índices distintos y $\lambda_1, \ldots, \lambda_r \in \mathbb{C}$. Al considerar el producto escalar de la igualdad anterior con el vector e_{j_k}, donde $k \in \{1, \ldots, r\}$, obtenemos

$$\langle e_{j_k}, \lambda_1 e_{j_1} + \ldots + \lambda_r e_{j_r}\rangle = \overline{\lambda_1}\langle e_{j_k}, e_{j_1}\rangle + \ldots + \overline{\lambda_r}\langle e_{j_k}, e_{j_r}\rangle.$$

Como $\langle e_{j_k}, e_{j_\ell}\rangle = 0$ si $k \neq \ell$, se tiene que $0 = \overline{\lambda_k}\langle e_{j_k}, e_{j_k}\rangle = \overline{\lambda_k}$, y luego $\lambda_k = 0$ para todo $k \in \{1, \ldots, r\}$. En otras palabras, $(e_j)_{j\in J}$ es una familia linealmente independiente. □

Teorema 5.5 (existencia de bases ortonormales)**.** *Sea $V \cong \mathbb{C}^n$ un espacio hermitiano. Entonces V admite una base ortonormal.*

Demostración. Esto es consecuencia directa del Teorema de Sylvester en el caso hermitiano, y del hecho que la signatura de la forma cuadrática hermitiana asociada al producto escalar es de signatura $(n, 0)$. □

Observación 5.6 (sub-espacios)**.** *Si $V \cong \mathbb{C}^n$ es un espacio hermitiano y $U \subseteq V$ es un sub-espacio vectorial. Entonces la restricción a U del producto escalar $\langle\cdot,\cdot\rangle|_U : U \times U \to \mathbb{C}$ es un producto escalar en U. Luego, todo sub-espacio de un espacio hermitiano es también un espacio hermitiano.*

Teorema 5.7 (proyección ortogonal)**.** *Sea $V \cong \mathbb{C}^n$ un espacio hermitiano y sea $U \subseteq V$ un sub-espacio vectorial. Si denotamos por $U^\perp$ el ortogonal de U respecto a la forma hermitiana dada por el producto escalar de V, entonces:*

1. *Se tiene que $V = U \oplus U^\perp$. Más aún, la aplicación $p_U : V \to V$ definida por esta descomposición es lineal, donde* $\mathrm{Im}(p_U) = U$ *y* $\ker(p_U) = U^\perp$*. Decimos que $p_U : V \to V$ es la* **proyección ortogonal** *sobre U.*

2. *Sea $(e_1, \ldots, e_r)$ una base* ortonormal *de U, donde $r = \dim_{\mathbb{C}}(U)$. Entonces, para todo $x \in V$ se tiene que*
$$p_U(x) = \langle x, e_1\rangle e_1 + \ldots + \langle x, e_r\rangle e_r,$$
donde los coeficientes $\langle x, e_j\rangle \in \mathbb{C}$ son llamados los **coeficientes de Fourier** *de $x \in V$ respecto a la base $(e_1, \ldots, e_r)$.*

3. *Se tiene que $(U^\perp)^\perp$. En particular, la proyección ortonogonal $p_{U^\perp} : V \to V$ sobre $U^\perp$ coincide con* $\mathrm{Id}_V - p_U$*, i.e.,* $\mathrm{Id}_V = p_U + p_{U^\perp}$.

Demostración. Gracias al Teorema 3.5, y dado que el producto escalar complejo es una forma hermitiana no-degenerada, tenemos que $U = (U^\perp)^\perp$. Además, dado que no hay vectores isótropos no-nulos pues la forma cuadrática hermitiana asociada es definida positiva, tenemos que $V = U \oplus U^\perp$. Así, todo vector $x \in V$ se escribe de manera única como $x = a + b$, donde $a \in U$ y $b \in U^\perp$. Luego, si definimos $p_U(x) = a$ entonces $p_U : V \to V$ es lineal y verifica por construcción $\mathrm{Im}(p_U) = U$ y $\ker(p_U) = U^\perp$. Más aún, dado que $(U^\perp)^\perp = U$, tenemos que $p_{U^\perp}$ está definido por $p_{U^\perp}(x) = b$ y luego $\mathrm{Id}_V = p_U + p_{U^\perp}$. Con esto hemos probado (1) y (3).

Para probar (2) escribamos provisoriamente $a := \langle x, e_1\rangle e_1 + \ldots + \langle x, e_r\rangle e_r \in U$. Entonces, para todo $j \in \{1, \ldots, r\}$ tenemos que

$$\langle x - a, e_j\rangle = \langle x, e_j\rangle - \left\langle \sum_{k=1}^{r} \langle x, e_k\rangle e_k, e_j \right\rangle = \langle x, e_j\rangle - \sum_{k=1}^{r} \langle x, e_k\rangle \underbrace{\langle e_k, e_j\rangle}_{\delta_{jk}} = \langle x, e_j\rangle - \langle x, e_j\rangle = 0,$$

por lo que $x = a + (x - a)$, donde $a \in U$ y $(x - a) \in U^\perp$. Dado que $V = U \oplus U^\perp$, la escritura anterior es única y luego $p_U(x) = a = \langle x, e_1\rangle e_1 + \ldots + \langle x, e_r\rangle e_r$. □

Al igual que en el caso de espacios euclideanos, el producto escalar complejo nos permite definir una noción de *distancia* entre vectores en un espacio hermitiano. Mejor aún, nos permite definir una *norma*.

Definición 5.8 (norma)**.** Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo. Una **norma** $\|\cdot\|$ en V es una función $\|\cdot\| : V \to \mathbb{R}^{\geq 0}$, $x \mapsto \|x\|$ verificando las propiedades siguientes:

1. $\|x\| = 0$ si y sólo si $x = 0$.
2. $\|\lambda x\| = |\lambda|\|x\|$ para todos $x \in V$ y $\lambda \in \mathbb{C}$.
3. Para todos $x, y \in V$ se tiene que $\|x + y\| \leq \|x\| + \|y\|$ («desigualdad triangular»).

Un **espacio vectorial normado** es un espacio vectorial dotado de una norma.

Teorema 5.9 (desigualdad de Cauchy-Schwarz y norma hermitiana)**.** *Sea $V \cong \mathbb{C}^n$ un espacio hermitiano, y sea $Q(x) = \langle x, x\rangle$ la forma cuadrática hermitiana asociada al producto escalar complejo. Entonces, para todos $x, y \in V$ se tiene la* desigualdad de Cauchy-Schwarz

$$|\langle x, y\rangle|^2 \leq Q(x)Q(y),$$

con igualdad si y sólo si x e y son colineales (i.e., son linealmente dependientes). En consecuencia, la aplicación $x \mapsto \|x\| := \sqrt{Q(x)} = \sqrt{\langle x, x\rangle}$ define una norma en V, llamada la **norma hermitiana** *asociada al producto escalar $\langle\cdot, \cdot\rangle$. Luego, la desigualdad de Cauchy-Schwarz se reescribe como*

$$|\langle x, y\rangle| \leq \|x\|\|y\|$$

para todos $x, y \in V$.

Demostración. Ya observamos anteriormente que la forma cuadrática hermitiana $Q(x) = \langle x, x\rangle$ es también la forma cuadrática *real* asociada a la forma $\mathbb{R}$-bilineal simétrica $B(x, y) = \operatorname{Re}\langle x, y\rangle$ en $V \cong \mathbb{R}^{2n}$. Para $x, y \in V \cong \mathbb{R}^{2n}$ (visto como espacio vectorial real), la desigualdad de Cauchy-Schwarz real implica que

$$|\operatorname{Re}\langle x, y\rangle| \leq \sqrt{Q(x)}\sqrt{Q(y)}.$$

Consideremos la escritura del número complejo $z = \langle x, y\rangle \in \mathbb{C}$ en su forma polar:

$$\langle x, y\rangle = re^{i\theta},\ r = |\langle x, y\rangle|.$$

Entonces, dado que $e^{-i\theta} = \overline{e^{i\theta}}$ y $r \in \mathbb{R}$, se tiene que

$$r = e^{-i\theta}\langle x, y\rangle = \langle x, e^{i\theta}y\rangle = \operatorname{Re}\langle x, e^{i\theta}y\rangle,$$

por lo que

$$|\langle x, y\rangle| = |\operatorname{Re}\langle x, e^{i\theta}y\rangle| \leq \sqrt{Q(x)}\sqrt{Q(e^{i\theta}y)} = \sqrt{Q(x)}\sqrt{|e^{i\theta}|^2Q(y)} = \sqrt{Q(x)}\sqrt{Q(y)}.$$

Finalmente, por el cálculo anterior y por el caso real de la desigualdad de Cauchy-Schwarz, tenemos la igualdad ocurre si y sólo si x e $e^{i\theta}y$ son $\mathbb{R}$-linealmente dependientes. Esto implíca en particular que x e y son $\mathbb{C}$-linealmente dependientes. Recíprocamente, si x e y son $\mathbb{C}$-linealmente dependientes y $x \neq 0$, entonces $y = \lambda x$ para cierto $\lambda \in \mathbb{C}$ y por ende

$$\langle x, y\rangle = \overline{\lambda}\langle x, x\rangle = \overline{\lambda}Q(x), \quad \sqrt{Q(x)}\sqrt{Q(y)} = |\lambda|\sqrt{Q(x)}\sqrt{Q(x)} = |\lambda|Q(x),$$

de donde $|\langle x, y\rangle = \sqrt{Q(x)}\sqrt{Q(y)}$ en este caso. El caso $x = 0$ es similar, pero más simple.

Finalmente, dado que para todo número complejo $z = a + ib$ se tiene $\operatorname{Re}(z) = a \leq \sqrt{a^2 + b^2} = |z|$, para todos $x, y \in V$ tenemos por la fórmula de polarización y por la desigualdad de Cauchy-Schwarz que:

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\operatorname{Re}\langle x, y\rangle \leq \|x\|^2 + \|y\|^2 + 2|\langle x, y\rangle| \leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| = (\|x\|^2 + \|y\|^2),$$

de donde se deduce la desigualdad triangular. $\square$

Ejemplo 5.10. Sea $V = \mathbb{C}^n$. La norma hermitiana asociada al producto escalar canónico de $\mathbb{C}^n$ está dada por

$$\|(z_1, \ldots, z_n)\| = \sqrt{|z_1|^2 + \ldots + |z_n|^2},$$

donde $(z_1, \ldots, z_n) \in \mathbb{C}^n$. En particular, si $(w_1, \ldots, w_n) \in \mathbb{C}^n$ es otro vector, entonces la desigualdad de Cauchy-Schwarz se escribe como

$$|z_1\overline{w_1} + \ldots + z_n\overline{w_n}| \leq \sqrt{\sum_{j=1}^{n} |z_j|^2} \cdot \sqrt{\sum_{j=1}^{n} |w_j|^2}.$$

Observación 5.11. *Vale la pena reescribir con la nueva notación las fórmulas de polarización discutidas anteriormente en Lema 2.18. Explícitamente, si $V \cong \mathbb{C}^n$ es un espacio hermitiano y $x, y \in V$, entonces:*

1. $\mathrm{Re}\langle x, y\rangle = \frac{1}{2}(\|x + y\|^2 - \|x\|^2 - \|y\|^2) = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2)$.
2. $\mathrm{Im}\langle x, y\rangle = \frac{1}{2}(\|x + iy\|^2 - \|x\|^2 - \|y\|^2) = \frac{1}{4}(\|x + iy\|^2 - \|x - iy\|^2)$.

Luego,

$$\langle x, y\rangle = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2 + i\|x + iy\|^2 - i\|x - iy\|^2).$$

Ejercicio 5.12. Sea $V \cong \mathbb{C}^n$ un espacio hermitiano. Probar que si $x_1, \ldots, x_m \in V$ son ortogonales, entonces se verifica el «Teorema de Pitágoras»:

$$\|x_1 + \ldots + x_m\|^2 = \|x_1\|^2 + \ldots + \|x_m\|^2.$$

Tal como en el caso euclideano, una clase importante de aplicaciones lineales son aquellas que preservan la norma hermitiana (o equivalentemente, el producto escalar complejo).

Proposición 5.13 (isometrías)**.** *Sean $V \cong \mathbb{C}^n$ un espacio hermitiano, y sea $u : V \to V$ un endomorfismo. Las siguientes propiedades son equivalentes:*

1. *u preserva la norma, i.e., para todo $x \in V$ se tiene que*

$$\|x\| = \|u(x)\|.$$

2. *u preserva el producto escalar, i.e., para todos $x, y \in V$ se tiene que*

$$\langle x, y\rangle = \langle u(x), u(y)\rangle.$$

3. Para toda *base ortonormal $\mathscr{B}$ de V, la imagen $u(\mathscr{B})$ es una base ortonormal de V.*
4. Existe *una base ortonormal $\mathscr{B}$ de V tal que la imagen $u(\mathscr{B})$ es una base ortonormal de V.*
5. *La matriz $A = \mathrm{Mat}_{\mathscr{B}}(u)$ respecto a una base* ortonormal *$\mathscr{B}$ de V es una* **matriz unitaria**, *i.e., $A^*A = \mathrm{I}_n$ (o equivalentemente, ${}^{\mathrm{t}}A\overline{A} = \mathrm{I}_n$).*

Un endomorfismo $u : V \to V$ que cumple estas propiedades es necesariamente biyectivo (i.e., $u \in \mathrm{GL}(V)$), y es llamado una **isometría** *de V. En particular, el conjunto de isometrías (lineales) de un espacio hermitiano $V \cong \mathbb{C}^n$ se identifica naturalmente al grupo unitario $\mathbf{U}(n)$.*

Demostración. Vimos que (1) y (2) son equivalentes gracias a la fórmula de polarización hermitiana. Por otra parte, claramente (2) implica (3), y (3) implica (4). Veamos que (4) implica (1):

Sea $\mathscr{B} = (e_1, \ldots, e_n)$ una base ortonormal tal que $u(\mathscr{B}) = (u(e_1), \ldots, u(e_n))$ es una base ortonormal de V. Sea $x \in V$ y escribamos $x = \sum_{j=1}^{n} x_j e_j$. Entonces, $u(x) = \sum_{j=1}^{n} x_j u(e_j)$ y, dado que tanto $(e_1, \ldots, e_n)$ como $(u(e_1), \ldots, u(e_n))$ son bases ortonormales de V, tenemos que los coeficientes de Fourier respecto a dichas bases verifican

$$x_j = \langle x, e_j\rangle \quad \text{y} \quad x_j = \langle u(x), u(e_j)\rangle,$$

por lo que $\langle u(x), u(e_j)\rangle = \langle x, e_j\rangle$ para todo $j \in \{1, \ldots, n\}$. En particular, el Teorema de Pitágoras implica

$$\|x\|^2 = \sum_{j=1} \langle x, e_j\rangle^2 = \sum_{j=1} \langle u(x), u(e_j)\rangle^2 = \|u(x)\|^2,$$

de donde obtenemos $\|u(x)\| = \|x\|$ y por ende (1) se verifica.

Finalmente, sabemos por la Proposición 4.4 que un endomorfismo $u : V \to V$ preserva la forma hermitiana *no-degenerada* $h = \langle\cdot,\cdot\rangle$ (y luego, u es automáticamente biyectivo) si y sólo si ${}^{\mathrm{t}}AH\overline{A} = H$, donde $H = (h_{jk})$ con $h_{jk} = \langle e_j, e_k\rangle = \delta_{jk}$ si $\mathscr{B} = (e_1, \ldots, e_n)$ es una base ortonormal. En otras palabras, $H = \mathrm{I}_n$, y luego (1) es equivalente a (5). $\square$

Como consecuencia inmediata, obtenemos las siguientes caracterizaciones diferentes del grupo unitario.

Corolario 5.14. *Sea $V \cong \mathbb{C}^n$ un espacio hermitiano y sea $\mathscr{B}$ una base ortonormal de V. Entonces, el grupo unitario de isometrías de V es isomorfo a*

$$\begin{aligned}\mathbf{U}(n) &= \{A \in M_n(\mathbb{C}) \mid A^*A = I_n\} = \{A \in M_n(\mathbb{C}) \mid {}^{\mathrm{t}}A\overline{A} = I_n\} = \{A \in \mathrm{GL}_n(\mathbb{C}) \mid A^{-1} = A^*\} \\ &= \{A \in M_n(\mathbb{C}) \mid \langle AX, AY\rangle = \langle X, Y\rangle \text{ para todos } X, Y \in \mathbb{C}^n\} \\ &= \{A \in M_n(\mathbb{C}) \mid \|AX\| = \|X\| \text{ para todo } X \in \mathbb{C}^n\} \\ &= \{A \in M_n(\mathbb{C}) \mid (Av_1, \ldots, Av_n) \text{ es base ortonormal para toda base ortonormal } (v_1, \ldots, v_n) \text{ de } \mathbb{C}^n\} \\ &= \{A \in M_n(\mathbb{C}) \mid (Ae_1, \ldots, Ae_n) \text{ es base ortonormal, donde } (e_1, \ldots, e_n) \text{ es la base canónica de } \mathbb{C}^n\} \\ &= \{A \in M_n(\mathbb{C}) \mid \text{las columnas de } A \text{ son ortonormales}\},\end{aligned}$$

donde $\langle\cdot,\cdot\rangle$ y $\|\cdot\|$ denotan el producto escalar y la norma hermitiana canónicos en $\mathbb{C}^n$. $\square$

6. Endomorfismo adjunto y endomorfismos normales

Comencemos por introducir la noción de *adjunto* de un endomorfismo.

Teorema 6.1 (adjunto)**.** *Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo y sea $h : V \times V \to \mathbb{C}$ una forma hermitiana* no-degenerada*. Entonces, para todo endomorfismo $u : V \to V$ existe un único endomorfismo $u^* : V \to V$ que verifica*

$$h(u(x), y) = h(x, u^*(y)) \text{ para todos } x, y \in V,$$

llamado el **adjunto** *de u (respecto a h). Más aún, para toda base $\mathscr{B}$ de V se tiene que*

$$\mathrm{Mat}_{\mathscr{B}}(u^*) = \overline{H^{-1}}\,\mathrm{Mat}_{\mathscr{B}}(u)^*\overline{H},$$

donde $H = \mathrm{Mat}_{\mathscr{B}}(h)$ es la matriz de h respecto a $\mathscr{B}$. En particular, $(u^*)^* = u$.

Demostración. Supongamos que existe un endomorfismo $u^* : V \to V$ verificando $h(u(x), y) = h(x, u^*(y))$ para todos $x, y \in V$. Sea $\mathscr{B}$ una base de V y denotemos $H = \mathrm{Mat}_{\mathscr{B}}(h)$, $A = \mathrm{Mat}_{\mathscr{B}}(u)$ y $B = \mathrm{Mat}_{\mathscr{B}}(u^*)$. Sean $x, y \in V$ arbitrarios y denotemos por $X, Y \in \mathbb{C}^n$ los vectores columnas asociadas (i.e., los vectores coordenadas respecto a la base $\mathscr{B}$). Entonces,

$$h(u(x), y) = {}^{\mathrm{t}}(AX)H\overline{Y} = {}^{\mathrm{t}}X\,{}^{\mathrm{t}}AH\overline{Y} = h(x, u^*(y)) = {}^{\mathrm{t}}XH\overline{BY},$$

por lo que ${}^{\mathrm{t}}AH = H\overline{B}$. Así, dado que H es invertible (puesto que la forma hermitiana h es no-degenerada), tenemos que $B = \overline{H^{-1}\,{}^{\mathrm{t}}AH} = \overline{H^{-1}}A^*\overline{H}$, donde por definición $A^* = {}^{\mathrm{t}}\overline{A}$ es la matriz adjunta de A. Esto último muestra que u^*, en caso de existir, verifica la relación

$$\mathrm{Mat}_{\mathscr{B}}(u^*) = \overline{H^{-1}}\,\mathrm{Mat}_{\mathscr{B}}(u)^*\overline{H}$$

y por ende está únicamente determinado por u y por h. Recíprocamente, y con la notación anterior, si denotamos por $u^* : V \to V$ el endomorfismo de V cuya matriz respecto a la base $\mathscr{B}$ está dada por $\overline{H^{-1}}A^*\overline{H}$ entonces para todos $x, y \in V$ se tiene que

$$h(x, u^*(y)) = {}^{\mathrm{t}}XH\overline{BY} = {}^{\mathrm{t}}XH(\overline{\overline{H^{-1}}A^*\overline{H}})\overline{Y} = {}^{\mathrm{t}}XHH^{-1}\overline{A^*}H\overline{Y} = {}^{\mathrm{t}}X\,{}^{\mathrm{t}}AH\overline{Y} = {}^{\mathrm{t}}(AX)H\overline{Y} = h(u(x), y),$$

por lo que $u^* : V \to V$ verifica la relación pedida. Esto demuestra la existencia y unicidad del endomorfismo adjunto. Finalmente, notamos que para todos $x, y \in V$ se tiene que

$$h(u^*(x), y) = \overline{h(y, u^*(x))} = \overline{h(u(y), x)} = h(x, u(y))$$

y por ende u es el adjunto de u^*, i.e., $(u^*)^* = u$. $\square$

Ejercicio 6.2. Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo. Probar que la aplicación

$$\mathrm{End}_{\mathbb{C}}(V) \longrightarrow \mathrm{End}_{\mathbb{C}}(V),\ u \longmapsto u^*$$

es anti-lineal.

En el caso de espacios hermitianos el teorema anterior se reescribe de la manera siguiente.

Teorema 6.3 (adjunto en espacio hermitiano)**.** *Sea $V \cong \mathbb{C}^n$ un espacio hermitiano. Entonces, para todo endomorfismo $u : V \to V$ existe un único endomorfismo $u^* : V \to V$ que verifica*

$$\langle u(x), y\rangle = \langle x, u^*(y)\rangle \text{ para todos } x, y \in V,$$

llamado el **adjunto** *de u. Más aún, para toda* base ortonormal *$\mathscr{B}$ de V se tiene que*

$$\mathrm{Mat}_{\mathscr{B}}(u^*) = \mathrm{Mat}_{\mathscr{B}}(u)^* = {}^{t}\overline{\mathrm{Mat}_{\mathscr{B}}(u)}.$$

Demostración. Si $\mathscr{B}$ es una base ortonormal de V, entonces la matriz H de la forma hermitiana $h = \langle\cdot,\cdot\rangle$ es la identidad I_n, de donde obtenemos el resultado. $\square$

El siguiente lema sobre estabilidad de sub-espacios vectoriales será útil en lo que sigue. Recordemos que si $u : V \to V$ es un endomorfismo de un espacio vectorial V, entonces un sub-espacio vectorial $U \subseteq V$ es **estable** por u si $u(U) \subseteq U$, i.e., si para todo $x \in U$ se tiene que $u(x) \in U$.

Lema 6.4 (estabilidad)**.** *Sea $V \cong \mathbb{C}^n$ un espacio hermitiano y sea $u : V \to V$ un endomorfismo. Entonces, para todo sub-espacio vectorial $U \subseteq V$ se tiene que*

$$u(U) \subseteq U \Leftrightarrow u^*(U^\perp) \subseteq U^\perp.$$

En otras palabras, U es estable por u si y sólo si $U^\perp$ es estable por u^*.

Demostración. Supongamos que $u(U) \subseteq U$ y sea $y \in U^\perp$. Entonces, para todo $x \in U$ se tiene

$$\langle x, u^*(y)\rangle = \langle u(x), y\rangle = 0,$$

donde la última igualdad se obtiene gracias a que $u(x) \in U$ e $y \in U^\perp$, por lo que $u^*(y) \in U^\perp$. Así, $u^*(U^\perp) \subseteq U^\perp$. Recíprocamente, supongamos que $u^*(U^\perp) \subseteq U^\perp$ y, por el mismo argumento anterior pero reemplazando u por u^* y U por $U^\perp$, tenemos en este caso que $(u^*)^*((U^\perp)^\perp) \subseteq (U^\perp)^\perp$. Dado que $(U^\perp)^\perp = U$ y $(u^*)^*$=u, la inclusión $(u^*)^*((U^\perp)^\perp) \subseteq (U^\perp)^\perp$ es equivalente a $u(U) \subseteq U$. $\square$

Las siguientes clases de endomorfismos de un espacio hermitiano son especialmente importantes.

Definición 6.5 (endomorfismo normal)**.** Sea $V \cong \mathbb{C}^n$ un espacio hermitiano, y sea $u : V \to V$ un endomorfismo. Diremos que u es:

1. **auto-adjunto** (o **hermitiano**) si $u^* = u$.

2. **anti-hermitiano** si $u^* = -u$.

3. **unitario** si $u \circ u^* = \mathrm{Id}_V$ (i.e., $u \in \mathrm{GL}(V)$ y $u^{-1} = u^*$).

De manera más general, diremos que u es un endomorfismo **normal** si conmuta con su adjunto u^*, i.e., $u \circ u^* = u^* \circ u$. En particular, esta noción engloba los tres casos precedentes.

En dimensión finita, el *espectro* de un endomorfismo es por definición el conjunto de sus valores propios. La *teoría espectral* es el estudio de valores propios y vectores propios de endomorfismos (eventualmente generalizando resultados de dimensión finita a dimensión infinita, para poder tomar en cuenta *operadores*, tales como el operador de Schrödinger en mecánica cuántica). En este texto introductorio, nos contentaremos con estudiar la teoría espectral de endomorfismos normales en dimensión finita (que cubre en particular el caso de endomorfismos simétricos y anti-simétricos reales, hermitianos y anti-hermitianos complejos, unitarios y ortogonales reales). El siguiente resultado es conocido como **teorema espectral**.

Teorema 6.6. *Sea $V \cong \mathbb{C}^n$ un espacio hermitiano y sea $u : V \to V$ un* endomorfismo normal. *Entonces, u es diagonalizable y los espacios propios asociados a valores propios distintos son ortogonales. En particular, existe una base ortonormal $\mathscr{B}$ de V formada por vectores propios de u. Más aún:*

1. *Si u es* hermitiano*, entonces los valores propios de u son* reales $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ *y luego*
$$\mathrm{Mat}_{\mathscr{B}}(u) = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} \in M_n(\mathbb{R}).$$

2. *Si u es* anti-hermitiano*, entonces los valores propios de u son* imaginarios puros $i\lambda_1, \dots, i\lambda_n \in i\mathbb{R}$ *y luego*
$$\mathrm{Mat}_{\mathscr{B}}(u) = \begin{pmatrix} i\lambda_1 & 0 & \dots & 0 \\ 0 & i\lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & i\lambda_n \end{pmatrix} \in M_n(\mathbb{C}).$$

3. *Si u es* unitario*, entonces los valores propios de u son* de módulo 1*, i.e., $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ con $|\lambda_j| = 1$, y luego*
$$\mathrm{Mat}_{\mathscr{B}}(u) = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} \in \mathrm{GL}_n(\mathbb{C}).$$

Demostración. La demostración es por inducción en la dimensión $n = \dim_{\mathbb{C}}(V)$. Notamos que el resultado es cierto si $n = 1$, por lo que podemos suponer que $n \geq 2$ y que el resultado se verifica para $n - 1$. Dado que $\mathbb{C}$ es algebraicamente cerrado, el polinomio característico $P_u(X)$ posee al menos una raíz λ en $\mathbb{C}$, y λ es un valor propio de u.

Sea V_λ el espacio propio asociado, el cual es *estable* por u^* (i.e., $u^*(V_\lambda) \subseteq V_\lambda$). En efecto, dado que u y u^* conmutan, tenemos que para todo $x \in V_\lambda$ se cumple que
$$u(u^*(x)) = u^*(u(x)) = u^*(\lambda x) = \lambda u^*(x),$$
por lo que $u^*(x) \in V_\lambda$. Luego, gracias al lema de estabilidad (ver Lema 6.4) tenemos que el ortogonal $V_\lambda^\perp$ es estable tanto por u como por u^*. Por otro lado, dado que $V = V_\lambda \oplus V_\lambda^\perp$ y $\dim_{\mathbb{C}}(V_\lambda) > 0$ (por definición de espacio propio), tenemos que $\dim_{\mathbb{C}}(V_\lambda^\perp) = \dim_{\mathbb{C}}(V) - \dim_{\mathbb{C}}(V_\lambda) < \dim_{\mathbb{C}}(V) = n$.

Denotemos por $v := u|_{V_\lambda^\perp} : V_\lambda^\perp \to V_\lambda^\perp$ la restricción de u a $V_\lambda^\perp$, y por v^* la restricción de u^* a $V_\lambda^\perp$. Entonces, para todos $x, y \in V_\lambda^\perp$ tenemos que $v(x) = u(x)$ y $v^*(y) = u^*(y)$, por lo que
$$\langle v(x), y \rangle = \langle u(x), y \rangle = \langle x, u^*(y) \rangle = \langle x, v^*(y) \rangle,$$
lo que muestra que el adjunto de v es v^*, vistos como endomorfismos de $V_\lambda^\perp$. Dado que u y u^* conmutan, las restricciones v y v^* también conmutan, i.e., v es un endomorfismo normal de $V_\lambda^\perp$. Luego, la hipótesis de inducción implica que v es diagonalizable y que espacios propios asociados a valores propios distintos son ortogonales. Dado que $V = V_\lambda \oplus V_\lambda^\perp$, lo anterior también se cumple para u. En particular, considerando bases ortonormales de cada espacio propio de u, obtenemos que existe una base ortonormal $\mathscr{B}$ de V formada por vectores propios de u.

Finalmente, si consideramos la matriz diagonal $D = \mathrm{Mat}_{\mathscr{B}}(u)$ con coeficientes diagonales $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$, entonces se tiene que si u es además

1. hermitiano, entonces $D^* = \overline{D} = D$, i.e., $\lambda_j \in \mathbb{R}$ para todo $j \in \{1, \ldots, n\}$.
2. anti-hermitiano, entonces $D^* = \overline{D} = -D$, i.e., $\lambda_j \in i\mathbb{R}$ para todo $j \in \{1, \ldots, n\}$.
3. unitario, entonces $D^*D = \overline{D}D = \mathrm{I}_n$, i.e., $|\lambda_j| = 1$ para todo $j \in \{1, \ldots, n\}$.

de donde se prueba el resultado. □

Observación 6.7. *Tal como mencionamos anteriormente, el teorema espectral anterior cubre también el caso de matrices* reales *simétricas, anti-simétricas y ortogonales. En efecto, si $A \in M_n(\mathbb{R})$ es una matriz real y la pensamos como elemento de $M_n(\mathbb{C})$, entonces $A^* = {}^{\mathrm{t}}\overline{A} = {}^{\mathrm{t}}A$. Por lo que A es hermitiana (resp. anti-hermitiana, resp. unitaria) si y sólo si es simétrica (resp. anti-simétrica, resp. ortogonal).*

7. Apéndice A: Método de reducción de Gauss (caso hermitiano)

El presente apéndice tiene por objetivo probar e ilustrar con ejemplos el método de reducción de Gauss en el caso de formas cuadráticas hermitianas. Los cálculos son un poco más complicados que en el caso real a causa de la presencia de complejos conjugados que tenemos que tener en cuenta.

Ejemplo 7.1. Consideremos la forma cuadrática hermitiana $Q : \mathbb{C}^2 \to \mathbb{R}$ en $V = \mathbb{C}^2$ dada por

$$Q(z, w) = 3z\overline{z} - 2iz\overline{w} + 2iw\overline{z} - 5w\overline{w}.$$

Primero que todo, es importante destacar que los coeficientes de $z\overline{w}$ y $w\overline{z}$ son necesariamente conjugados, puesto que en caso contrario Q no sería real ni tampoco sería hermitiana.

La fórmula de polarización hermitiana implica que la forma hermitiana asociada a Q está dada por

$$h((z_1, w_1), (z_2, w_2)) = 3z_1\overline{z_2} + 2iw_1\overline{z_2} - 2iz_1\overline{w_2} - 5w_1\overline{w_2}.$$

Notar que sólamente los conjugados $\overline{z_2}$ y $\overline{w}_2$ pueden aparecer, puesto que h es lineal en la primera variable y anti-lineal en la segunda variable. Además, sus coeficientes h_{jk} deben verificar $\overline{h_{kj}} = h_{jk}$. La matriz de h respecto a la base canónica $\mathscr{B} = (e_1, e_2)$ de $\mathbb{C}^2$ es

$$H = \mathrm{Mat}_{\mathscr{B}}(h) = \begin{pmatrix} 3 & -2i \\ 2i & -5 \end{pmatrix}.$$

El método de reducción de Gauss busca descomponer $Q(z, w)$ como suma de cuadrados $|\ell_j(z, w)|^2$ de formas lineales ℓ_j. La idea es que cada vez que tenemos un *término cuadrado* del tipo $z\overline{z} = |z|^2$ en Q, reagrupamos todos los términos conteniendo la variable z. En este ejemplo:

$$3z\overline{z} - 2iz\overline{w} + 2iw\overline{z} = 3\left(z + \frac{2i}{3}w\right)\overline{\left(z + \frac{2i}{3}w\right)} - \frac{4}{3}w\overline{w}.$$

De donde obtenemos

$$Q(z, w) = 3\left(z + \frac{2i}{3}w\right)\overline{\left(z + \frac{2i}{3}w\right)} - \frac{19}{3}w\overline{w} = 3|\ell_1(z, w)|^2 - \frac{19}{3}|\ell_2(z, w)|^2,$$

donde $\ell_1(z, w) = z + \frac{2i}{3}w$ y $\ell_2(z, w) = w$ son formas lineales independientes.

Para encontrar una base ortogonal, consideramos el cambio de variable

$$\begin{cases} u = z + \frac{2i}{3}w \\ v = w \end{cases} \Leftrightarrow \begin{cases} z = u - \frac{2i}{3}v \\ w = v \end{cases}$$

o equivalentemente $X' = LX \Leftrightarrow X = PX'$, donde $L \in \mathrm{GL}_2(\mathbb{C})$ es la matriz asociada a las formas lineales (ℓ_1, ℓ_2) y $P = L^{-1}$. Así, encontramos una base $\mathscr{B}' = (e_1', e_2')$ dada por las columnas de la matriz de cambio de base

$$\begin{pmatrix} 1 & -\frac{2i}{3} \\ 0 & 1 \end{pmatrix}.$$

En otras palabras, $e_1' = (1,0)$ y $e_2' = (-\frac{2i}{3}, 1)$. Así, las formas Q y h se escriben en las nuevas coordenadas (u,v) como

$$Q(z,w) = Q(u,v) = 3|u|^2 - \frac{19}{3}|v|^2,$$

por lo que $h((z_1,w_1),(z_2,w_2)) = h((u_1,v_1),(u_2,v_2)) = 3u_1\overline{u_2} - \frac{19}{3}v_1\overline{v_2}$, y en particular

$$\mathrm{Mat}_{\mathscr{B}'}(h) = \begin{pmatrix} 3 & 0 \\ 0 & -\frac{19}{3} \end{pmatrix}.$$

Obtenemos así una base ortogonal respecto a Q, y deducimos de la discusión anterior que la signatura de Q es (1,1) y su rango es 2 (i.e., es no-degenerada).

Teorema 7.2 (reducción de Gauss). *Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo y sea $\mathscr{B} = (e_1, \ldots, e_n)$ una base de V. Toda forma cuadrática hermitiana $Q : V \to \mathbb{R}$ en V admite una descomposición en suma de cuadrados*

$$Q = \sum_{j=1}^{r} a_j |\ell_j|^2$$

donde $\ell_1, \ldots, \ell_r \in V^*$ son formas lineales independientes, $a_j \in \mathbb{R} \setminus \{0\}$, y $r \in \{0, 1, \ldots, n\}$. La forma hermitiana h asociada a la forma cuadrática hermitiana Q está dada por

$$h(x,y) = \sum_{j=1}^{r} a_j \ell_j(x)\overline{\ell_j(y)}.$$

Si completamos[9] *los ℓ_j en una base $(\ell_1, \ldots, \ell_n)$ de V^*, entonces obtenemos nuevas coordenadas $x_j' = \ell_j(x)$ (i.e., matricialmente: $X' = LX \Leftrightarrow X = PX'$, donde $P = L^{-1}$) en las cuales la matriz de Q es diagonal real. Explícitamente, las columnas $(e_1', \ldots, e_n')$ de P definen una base $\mathscr{B}'$ (ortogonal respecto a Q) de V tal que*

$$\mathrm{Mat}_{\mathscr{B}'}(Q) = \begin{pmatrix} a_1 & 0 & \ldots & 0 \\ 0 & a_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & a_n \end{pmatrix} \in M_n(\mathbb{R}),$$

donde $r = \mathrm{rg}(Q)$, y donde $a_1, \ldots, a_r \neq 0$ y $a_{r+1} = \ldots = a_n = 0$.

Demostración. Al fijar la base $\mathscr{B} = (e_1, \ldots, e_n)$ podemos suponer sin pérdida de generalidad que $V = \mathbb{C}^n$ y que

$$Q(x_1, \ldots, x_n) = \sum_{1 \le j,k \le n} h_{jk} x_j \overline{x_k}.$$

Demostraremos por inducción en n que Q se descompone de la forma deseada $Q = \sum_{j=1}^{r} a_j |\ell_j|^2$.

El primer caso a considerar es cuando Q admite un *término cuadrado* no-nulo. Por ejemplo, supongamos que $h_{11}x_1\overline{x_1} = h_{11}|x_1|^2$ es no-nulo. En este caso, el método de Gauss consiste en reagrupar todos los términos que contengan la variable x_1 factorizando el coeficiente $h_{11} \neq 0$ (que es real, pues Q es hermitiana). Obtenemos así

$$h_{11}\left(x_1\overline{x_1} + \frac{h_{12}}{h_{11}} x_1\overline{x_2} + \ldots + \frac{h_{1n}}{h_{11}} x_1\overline{x_n} + \frac{h_{21}}{h_{11}} x_2\overline{x_1} + \ldots + \frac{h_{n1}}{h_{11}} x_n\overline{x_1} \right).$$

[9] Por ejemplo, agregando coordenadas $\ell_j(x) = x_{k_j}$ que no hayan sido utilizadas en el método de reducción de Gauss.

Teniendo en cuenta las relaciones $h_{jk} = \overline{h_{kj}}$ dada por la simetría hermitiana, esto se factoriza como

$$h_{11}\left(x_1 + \frac{h_{21}}{h_{11}}x_2 + \ldots + \frac{h_{n1}}{h_{11}}x_n\right)\overline{\left(x_1 + \frac{h_{21}}{h_{11}}x_2 + \ldots + \frac{h_{n1}}{h_{11}}x_n\right)} + Q'(x_2, \ldots, x_n),$$

donde $Q'(x_2, \ldots, x_n)$ es una forma cuadrática hermitiana que sólo depende de las variables $x_2, \ldots, x_n$. Si definimos la forma lineal

$$\ell_1(x_1, \ldots, x_n) = x_1 + \frac{h_{21}}{h_{11}}x_2 + \ldots + \frac{h_{n1}}{h_{11}}x_n,$$

entonces $Q(x_1, \ldots, x_n) = h_{11}|\ell_1(x_1, \ldots, x_n)|^2 + Q'(x_2, \ldots, x_n)$. Luego, la hipótesis de inducción aplicada a Q' implica que $Q' = \sum_{j=2}^r a_j|\ell_j|^2$ donde $a_j \in \mathbb{R}\setminus\{0\}$ y $(\ell_2, \ldots, \ell_r)$ son formas lineales independientes en V^* en las variables $(x_2, \ldots, x_n)$. Dado que ℓ_1 depende de la variable x_1, concluímos así que $\ell_1, \ldots, \ell_r \in V^*$ son independientes.

El segundo caso a considerar es cuando Q no admite términos cuadrados no-nulos, pero sólo admite *términos cruzados* no-nulos. Por ejemplo, supongamos que $h_{12}x_1\overline{x_2} + h_{21}x_2\overline{x_1}$ es no-nulo. En este caso, el método de Gauss consiste en reagrupar todos los términos que contengan las variables x_1 y x_2 y factorizarlo en un producto de la forma

$$h_{12}(x_1 + \ldots)\overline{(x_2 + \ldots)} + \text{conjugado}.$$

Para evitar escribir dos veces cada término y su conjugado, es que nos limitaremos a escribir sólamente los términos de la forma $x_j\overline{x_k}$ conteniendo x_j (no-conjugado) y $\overline{x_k}$ (conjugado). Obtenemos así

$$\begin{aligned}
&Q(x_1, \ldots, x_n)\\
&= h_{12}\left(x_1\overline{x_2} + \frac{h_{13}}{h_{12}}x_1\overline{x_3} + \ldots + \frac{h_{1n}}{h_{12}}x_1\overline{x_n} + \frac{h_{32}}{h_{12}}x_3\overline{x_2} + \ldots + \frac{h_{n2}}{h_{12}}x_n\overline{x_2}\right) + \text{conjugado} + Q'(x_3, \ldots, x_n)\\
&= h_{12}\left(x_1 + \frac{h_{32}}{h_{12}}x_3 + \ldots + \frac{h_{n2}}{h_{12}}x_n\right)\overline{\left(x_2 + \frac{h_{31}}{h_{21}}x_3 + \ldots + \frac{h_{n1}}{h_{21}}x_n\right)} + \text{conjugado} + Q'(x_3, \ldots, x_n),
\end{aligned}$$

donde $Q'(x_3, \ldots, x_n)$ es una forma cuadrática hermitiana que sólo depende de las variables $x_3, \ldots, x_n$. Luego,

$$Q(x_1, \ldots, x_n) = f\overline{g} + g\overline{f} + Q'(x_3, \ldots, x_n),$$

donde

$$f(x_1, \ldots, x_n) = h_{12}\left(x_1 + \frac{h_{32}}{h_{12}}x_3 + \ldots + \frac{h_{n2}}{h_{12}}x_n\right) \quad \text{y} \quad g(x_1, \ldots, x_n) = x_2 + \frac{h_{31}}{h_{21}}x_3 + \ldots + \frac{h_{n1}}{h_{21}}x_n.$$

Para escribir Q como combinación lineal de cuadrados, usamos el hecho que

$$f\overline{g} + g\overline{f} = \frac{1}{2}\left((f+g)\overline{(f+g)} - (f-g)\overline{(f-g)}\right) = \frac{1}{2}|f+g|^2 - \frac{1}{2}|f-g|^2.$$

Si definimos las formas lineales

$$\ell_1(x_1, \ldots, x_n) = f + g \quad \text{y} \quad \ell_2(x_1, \ldots, x_n) = f - g$$

entonces ellas son linealmente independientes, puesto que $h_{12} \neq 0$. Entonces,

$$Q(x_1, \ldots, x_n) = \frac{1}{2}|\ell_1(x_1, \ldots, x_n)|^2 - \frac{1}{2}|\ell_2(x_1, \ldots, x_n)|^2 + Q'(x_3, \ldots, x_n).$$

Finalmente, la hipótesis de inducción aplicada a Q' implica que $Q' = \sum_{j=3}^r a_j|\ell_j|^2$ donde $a_j \in \mathbb{R}\setminus\{0\}$ y $(\ell_3, \ldots, \ell_r)$ son formas lineales independientes en V^* en las variables $(x_3, \ldots, x_n)$. Dado que ℓ_1 y ℓ_2 dependen de las variables x_1 y x_2, y dado que ℓ_1 y ℓ_2 son linealmente independientes, concluímos así que $\ell_1, \ldots, \ell_r \in V^*$ son independientes. □

Ejercicio 7.3. Considerar las siguientes formas cuadráticas hermitianas $Q : \mathbb{C}^3 \to \mathbb{R}$.

a) $Q(x_1, x_2, x_3) = x_1\overline{x_1} - 2ix_1\overline{x_2} + 2ix_2\overline{x_1} + ix_1\overline{x_3} - ix_3\overline{x_1} + 2x_2\overline{x_2} + 2x_3\overline{x_3} - 2ix_2\overline{x_3} + 2ix_3\overline{x_2}$.

b) $Q(x_1, x_2, x_3) = x_1\overline{x_2} + x_2\overline{x_1} + ix_1\overline{x_3} - ix_3\overline{x_1} + (1+i)x_2\overline{x_3} + (1-i)x_3\overline{x_2}$.

En cada caso, utilizando el método de reducción de Gauss, descomponer Q como combinación lineal de suma de cuadrados de formas lineales independientes. Determinar la signatura y rango de Q, así como una base de $\mathbb{C}^3$ que sea ortogonal respecto a Q.

8. Apéndice B: Forma normal de matrices ortogonales

El presente apéndice tiene por objetivo aplicar el teorema espectral para endomorfismos normales en espacios hermitianos para dar una descripción sencilla de matrices ortogonales reales (i.e., elementos de $\mathbf{O}(n)$). Esta es una de las tantas instancias en que métodos complejos nos permiten atacar problemas reales.

Proposición 8.1. *Sea $V \cong \mathbb{R}^n$ espacio vectorial real de dimensión $n \geq 1$ y sea $u : V \to V$ un endomorfismo. Entonces u posee una recta estable L o un plano estable Π.*

Demostración. Si u posee un valor propio real $\lambda \in \mathbb{R}$, entonces existe un vector propio $v \neq 0$ asociado a λ y luego la recta $L = \mathrm{Vect}_{\mathbb{R}}(v)$ generada por v es estable por u, i.e., $u(L) \subseteq L$. Si u no posee valores propios reales, entonces el teorema fundamental del álgebra asegura que u posee al menos un valor propio complejo $\lambda = \alpha + i\beta \in \mathbb{C}$. Si consideramos una base $\mathscr{B}$ de V y escribimos $A = \mathrm{Mat}_{\mathscr{B}}(u)$ entonces la ecuación $AZ = \lambda Z$ tiene una solución *compleja* no-nula $Z \in \mathbb{C}^n$.

Si escribimos $Z = X + iY$, con $X, Y \in \mathbb{R}^n$. Entonces la ecuación $AZ = \lambda Z$ se reduce a

$$AX + iAY = (\alpha + i\beta)(X + iY) \Leftrightarrow AX = \alpha X - \beta Y,\ AY = \beta X + \alpha Y.$$

Si $\lambda \notin \mathbb{R}$, los vectores columnas X e Y no pueden ser $\mathbb{R}$-linealmente dependientes (colineales): en caso contrario tendríamos por ejemplo (si $X \neq 0$) que $Y = tX$ con $t \in \mathbb{R}$ y luego $Z = X + itX = (1 + it)X$. Esto último implicaría que $X \in \mathbb{R}^n$ también sería un vector propio *real* de la matriz real A asociada al valor propio complejo λ, pero esto último es una contradicción con la relación $AX = \lambda X$. El razonamiento es el mismo si $Y \neq 0$ y $X = tY$.

Así, si consideramos el plano $\Pi = \mathrm{Vect}_{\mathbb{R}}(v_1, v_2)$ generado por los vectores $\mathbb{R}$-linealmente independientes v_1 y v_2, de coordenadas $X \in \mathbb{R}^n$ e $Y \in \mathbb{R}^n$ respecto a la base $\mathscr{B}$, entonces tenemos que

$$u(v_1) = \alpha v_1 - \beta v_2 \in \Pi \quad \text{y} \quad u(v_2) = \beta v_1 + \alpha v_2 \in \Pi,$$

por lo que Π es estable por u, i.e., $u(\Pi) \subseteq \Pi$. $\square$

Teorema 8.2 (teorema espectral, caso ortogonal)**.** *Sea $V \cong \mathbb{R}^n$ un espacio euclideano, y sea $u : V \to V$ un endomorfismo. Entonces, u es* ortogonal (i.e., $u \in \mathbf{O}(n)$) *si y sólo si V admite una descomposición en suma directa ortogonal*

$$V = \Pi_1 \oplus \Pi_2 \oplus \ldots \oplus \Pi_s \oplus L_1 \oplus \ldots \oplus L_t$$

formada de planos $\Pi_j \cong \mathbb{R}^2$ y de rectas $L_j \cong \mathbb{R}$ estables por u, tales que (al orientar cada Π_j) la restricción $u|_{\Pi_j}$ es una rotación $r_{\theta_j} \in \mathbf{SO}(2)$ de ángulo $\theta_j \in \mathbb{R}$ y donde $u|_{L_j} = \pm \mathrm{Id}_{L_j}$.

Más aún, eligiendo una base ortonormal directa (a_j, b_j) de Π_j y un vector director unitario v_j de L_j, entonces (reordenando las rectas L_j si fuese necesario) la matriz de u respecto a la base ortonormal $\mathscr{B} = (a_1, b_1, \ldots, a_s, b_s, v_1, \ldots, v_t)$ de V está dada por

$$A = \mathrm{Mat}_{\mathscr{B}}(u) = \left(\begin{array}{c|c|cc|c|cc}
\mathrm{I}_p & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ \hline
0 & -\mathrm{I}_q & 0 & 0 & \cdots & \cdots & 0 \\ \hline
\vdots & 0 & \cos(\theta_1) & -\sin(\theta_1) & 0 & \cdots & 0 \\
\vdots & 0 & \sin(\theta_1) & \cos(\theta_1) & 0 & \cdots & 0 \\ \hline
\vdots & \vdots & 0 & 0 & \ddots & 0 & 0 \\ \hline
\vdots & \vdots & \vdots & \vdots & 0 & \cos(\theta_s) & -\sin(\theta_s) \\
0 & 0 & 0 & 0 & 0 & \sin(\theta_s) & \cos(\theta_s)
\end{array}\right)$$

donde $p + q = t$, y la cual es llamada la forma normal *del endomorfismo ortogonal $u \in \mathbf{O}(n)$.*

Demostración. La demostración es por inducción en la dimensión $n = \dim_{\mathbb{R}}(V)$. Si $n = 1$ el resultado se obtiene del hecho que los únicos endomorfismos de un espacio euclideano V de dimensión 1 son las homotecias $u = \lambda\,\mathrm{Id}_V$, por lo que u es ortogonal si y sólo si $\lambda = \pm 1$, i.e., $u = \pm\,\mathrm{Id}_V$.

Si $n = 2$ esto se reduce al estudio matricial de elementos de $\mathbf{O}(2)$, i.e., las isometrías de un plano euclideano. Recordemos este cálculo brevemente: sea (e_1, e_2) una base ortonormal de $V \cong \mathbb{R}^2$ y sea $A = \mathrm{Mat}_{\mathscr{B}}(u)$. Entonces, u es ortogonal si y sólo si $(u(e_1), u(e_2))$ es una base ortonormal de V. Una vez que escogemos una orientación del plano euclideano $V \cong \mathbb{R}^2$, tenemos que existe un único ángulo $\theta \in]-\pi, \pi]$ tal que $u(e_1) = \cos(\theta)e_1 + \sin(\theta)e_2$. El vector $u(e_2)$ es ortogonal a $u(e_1)$ y es de norma 1, por lo que las únicas posibilidades para $u(e_2)$ son

$$u(e_2) = -\sin(\theta)e_1 + \cos(\theta)e_2 \quad \text{o bien} \quad u(e_2) = \sin(\theta)e_1 - \cos(\theta)e_2,$$

o en otras palabras, las únicas posibilidades para A son

$$A = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \quad \text{o bien} \quad A = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}.$$

En el primer caso tenemos que $u = r_\theta$ es una rotación de ángulo θ (y los valores propios de A son los números complejos conjugados $\lambda = e^{i\theta}$ y $\overline{\lambda} = e^{-i\theta}$), mientras que en el segundo caso calculamos directamente que

$$A = P^{-1}SP, \; P = \begin{pmatrix} \cos(\theta/2) & -\sin(\theta/2) \\ \sin(\theta/2) & \cos(\theta/2) \end{pmatrix}, \; S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

por lo que $u = s_L$ es una simetría ortogonal (reflexión) respecto a la recta $L = \ker(u - \mathrm{Id}_V)$ (cuyo ángulo polar es $\theta/2$) y los valores propios de u son $\lambda = 1$ y $\mu = -1$. En ambos casos, existen bases ortonormales en las cuales la matriz de u tiene la forma normal del enunciado.

Supongamos ahora que el resultado es cierto para dimensiones $\leq n-1$, con $n \geq 3$. Sabemos entonces que u posee al menos una recta estable L (si u posee un valor propio real) o un plano estable Π (si posee un valor propio complejo no-real). Sea U dicho sub-espacio estable (una recta o un plano), y consideremos la descomposición ortogonal $V = U \oplus U^\perp$. Dado que U es estable por u, tenemos que $U^\perp$ es también estable por u, por lo que la restricción $u|_{U^\perp}$ es un endomorfismo ortogonal de $U^\perp$. Finalmente, la hipótesis de inducción implica la existencia de una descomposición ortogonal en $U^\perp$ verificando la conclusión del teorema, mientras que en U aplicamos las observaciones antes hechas en dimensión 1 y 2. $\square$

Observación 8.3.

1. *En dimensión 2, escoger una orientación del plano euclideano $V \cong \mathbb{R}^2$ (i.e., al fijar una base ortonormal $\mathscr{B}_0$ de V y declararla una base directa) determina de manera única el ángulo $\theta \in]-\pi, \pi]$ de una rotación u en $\mathbf{SO}(2)$. En particular, para toda base ortonormal directa $\mathscr{B}$ (i.e., que verifica $\det_{\mathscr{B}_0}(\mathscr{B}) = 1$), tenemos $\mathrm{Mat}_{\mathscr{B}}(u) = R_\theta$ es la matriz de rotación de ángulo θ.*

2. *En dimensión 3, si fijamos una orientación de $\mathbb{R}^3$ dada por la base canónica $\mathscr{C}$, entonces todo elemento $u \in \mathbf{SO}(3)$ tal que $u \neq \mathrm{Id}_V$ admite un único* eje de rotación *dado por la recta (espacio propio) $L = \ker(u - \mathrm{Id}_V)$. Si orientamos la recta L escogiendo un vector director unitario $v \in L$, entonces el ángulo de rotación θ está únicamente determinado por la* regla de la mano derecha*: para toda base ortonormal (v_1, v_2) del plano ortogonal $\Pi = L^\perp$ tal que (v_1, v_2, v) sea una base directa de $\mathbb{R}^3$, tenemos que $\mathrm{Mat}_{(v_1,v_2)}(u|_\Pi) = R_\theta$.*

3. ¡Atención! *En dimensión par ≥ 4, un elemento $u \in \mathbf{SO}(n)$ no posee necesariamente un eje de rotación, i.e., puede ocurrir que $\ker(u - \mathrm{Id}_V) = \{0\}$. Este es el caso por ejemplo para la matriz*

$$A = \left(\begin{array}{c:c} R_{\theta_1} & 0 \\ \hdashline 0 & R_{\theta_2} \end{array}\right) \in \mathbf{SO}(4)$$

donde $\theta_1, \theta_2 \in]-\pi, \pi[$ son no-nulos.

Aplicación Física: Concluímos esta sección observando que, en Física, el movimiento de un sólido está caracterizado por la variación en el tiempo de un sistema coordenado ortonormal $(M_0(t), \mathscr{B}(t))$, considerando el desplazamiento de un punto fijo M_0 del sólido y la variación de un sistema de coordenadas ortonormales asociado al sólido. Para ilustrar esto último, consideremos por ejemplo la variación de la base $\mathscr{B}(t)$ en función del tiempo.

Escogamos $\mathscr{B}(0) = (e_1, \ldots, e_n)$ una base de referencia inicial. Entonces, $\mathscr{B}(t)$ está dada por una matriz de cambio de base ortogonal $P(t) \in \mathbf{O}(n)$ tal que $P(0) = \mathrm{I}_n$. El hecho que la matriz $P(t)$ sea ortogonal se obtiene gracias a la hipótesis de que el objeto en cuestión es un sólido: las distancias entre sus puntos no cambian a lo largo del tiempo, por lo que la transformación debe preservar la norma de los vectores.

Ahora, por razones físicas, la función $t \mapsto P(t)$ con valores matriciales debe ser continua. Esto implica que $t \mapsto \det(P(t))$ es una función continua también. Dado que $\det(P(t)) = \pm 1$ y dado que no pueden haber saltos en el valor del determinante (por continuidad), tenemos que $\det(P(t)) = 1$ para todo tiempo t, puesto que $\det(P(0)) = \det(\mathrm{Id}_V) = 1$. En otras palabras, esto implica que el movimiento de sólidos sólo se puede efectuar a través de transformaciones $P(t) \in \mathbf{SO}(n)$ en el grupo especial ortogonal.

Recíprocamente, toda matriz $A \in \mathbf{SO}(n)$ puede obtenerse a través de un movimiento continuo $t \mapsto P(t)$ en un intervalo de tiempo $[0,1]$, i.e., tal que $P(0) = \mathrm{I}_n$ y $P(1) = A$. En efecto, si escribimos

$$A = Q^{-1} R_{(\theta_1, \ldots, \theta_s; p)} Q,$$

donde $Q \in \mathbf{O}(n)$ es una matriz ortogonal y donde $R_{(\theta_1, \ldots, \theta_s; p)}$ es la forma normal de A formada por s bloques de rotaciones planas de ángulos $\theta_1, \ldots, \theta_s$ y de p valores propios $+1$ (dado que $A \in \mathbf{SO}(n)$, hay un número par de valores propios -1, que podemos reagrupar en rotaciones planas de ángulo $\theta = \pi$). Si definimos

$$P(t) = Q^{-1} R_{(t\theta_1, \ldots, t\theta_s; p)} Q$$

obtenemos una aplicación matricial $t \mapsto P(t)$ que es continua (incluso diferenciable) para $t \in [0,1]$ y que cumple $P(0) = \mathrm{I}_n$ y $P(1) = A$.

Índice alfabético