Tema 6

# Espacios de Hilbert

Tras los espacios de dimensión finita, vamos a estudiar ahora los de Hilbert, que son los espacios de Banach más perfectos desde un punto de vista geométrico, pues verifican todos los postulados de la geometría euclídea. Con el estudio, por parte de David Hilbert (1862-1943) y su escuela, de las formas cuadráticas en infinitas variables, aparecen los primeros espacios de Hilbert de dimensión infinita, y puede decirse que arranca, en los albores del siglo XX, la prehistoria del Análisis Funcional.

Como primer resultado relevante, caracterizamos los espacios de Hilbert, entre los espacios de Banach, mediante la *identidad del paralelogramo*, para pasar inmediatamente al estudio del resultado sin duda más importante: el *teorema de la proyección ortogonal*. Como consecuencia más destacable, probamos el *teorema de Riesz-Fréchet*, que permite identificar en cierto modo cada espacio de Hilbert con su dual.

## 6.1. Formas sexquilineales y formas cuadráticas

Necesitamos algunas nociones básicas bien sencillas, que nos llevarán a la definición de los espacios de Hilbert. En lo que sigue, para evitar repeticiones, fijamos un espacio vectorial $X$.

Una **forma sexquilineal** en $X$ es una aplicación $\varphi : X \times X \to \mathbb{K}$ que verifica las siguientes dos condiciones:

$(i)$ Es lineal en la primera variable, es decir, para todo $y \in X$ se tiene:

$$\varphi\big(\lambda x + z, y\big) = \lambda\,\varphi\big(x, y\big) + \varphi\big(z, y\big) \qquad \forall x,z \in X,\ \forall \lambda \in \mathbb{K}$$

$(ii)$ Es **conjugado-lineal** en la segunda variable, es decir, para todo $x \in X$ se tiene:

$$\varphi\big(x, \lambda y + z\big) = \overline{\lambda}\,\varphi\big(x, y\big) + \varphi\big(x, z\big) \qquad \forall y,z \in X,\ \forall \lambda \in \mathbb{K}$$

Como es habitual, para $\lambda \in \mathbb{K}$, estamos denotando por $\overline{\lambda}$ al escalar conjugado de $\lambda$. En el caso real, se tiene $\overline{\lambda} = \lambda$ para todo $\lambda \in \mathbb{R}$, luego una forma sexquilineal no es otra cosa que una forma bilineal, es decir, lineal en cada variable.

Se dice ahora que la forma sexquilineal $\varphi$ es **hermítica**, cuando verifica que:

$$\varphi\big(y,x\big) = \overline{\varphi\big(x,y\big)} \qquad \forall x,y \in X$$

En el caso real, esto es tanto como decir que $\varphi$ es una forma bilineal simétrica.

Si $\varphi$ es una forma sexquilineal hermítica, es claro que $\varphi\big(x,x\big) \in \mathbb{R}$ para todo $x \in X$. Se dice entonces que la aplicación

$$Q : X \to \mathbb{R}, \qquad Q(x) = \varphi\big(x,x\big) \quad \forall x \in X$$

es la **forma cuadrática** asociada a $\varphi$. Se entiende por tanto que, una aplicación $Q : X \to \mathbb{R}$ es una forma cuadrática en $X$, cuando existe una forma sexquilineal hermítica $\varphi$ en $X$, verificando que $Q(x) = \varphi\big(x,x\big)$ para todo $x \in X$. Vamos a comprobar que entonces $\varphi$ es única y se obtiene fácilmente a partir de $Q$.

**Identidad de polarización.** *Sea $Q$ la forma cuadrática asociada a una forma sexquilineal hermítica $\varphi$. Se tiene entonces*

$$4\,\mathrm{Re}\,\varphi\big(x,y\big) = Q(x+y) - Q(x-y) \qquad \forall x,y \in X \tag{1}$$

*y por tanto, $\varphi$ es la única forma sexquilineal hermítica cuya forma cuadrática asociada es $Q$.*

**Demostración.** La igualdad $(1)$ es inmediata, pues para $x,y \in X$ se tiene

$$\begin{aligned} Q(x+y) - Q(x-y) &= \varphi\big(x+y,x+y\big) - \varphi\big(x-y,x-y\big) \\ &= 2\varphi\big(x,y\big) + 2\varphi\big(y,x\big) = 2\varphi\big(x,y\big) + 2\overline{\varphi\big(x,y\big)} \\ &= 4\,\mathrm{Re}\,\varphi\big(x,y\big) \end{aligned}$$

En el caso real, está ya bien claro que $\varphi$ está determinada por $Q$.

En el caso complejo, también para $x,y \in X$ se tiene

$$\mathrm{Im}\,\varphi\big(x,y\big) = \mathrm{Re}\,\big[-i\varphi\big(x,y\big)\big] = \mathrm{Re}\,\varphi\big(x,iy\big)$$

luego usando $(1)$ deducimos que

$$4\,\mathrm{Im}\,\varphi\big(x,y\big) = Q(x+iy) - Q(x-iy) \qquad \forall x,y \in X$$

Se tiene por tanto,

$$4\,\varphi\big(x,y\big) = Q(x+y) - Q(x-y) + i\,Q(x+iy) - i\,Q(x-iy) \qquad \forall x,y \in X$$

y esto prueba que $\varphi$ queda determinada por $Q$, también en el caso complejo. ■

De esta forma, hemos obtenido una correspondencia biunívoca entre formas sexquilineales hermíticas y formas cuadráticas. Usaremos solamente un tipo particular de formas cuadráticas, que enseguida vamos a definir.

## 6.2. Productos escalares y espacios de Hilbert

Seguimos manteniendo fijo el espacio vectorial $X$ en el que trabajamos. Se dice que una forma cuadrática $Q : X \to \mathbb{R}$ es **positiva**, cuando $Q(x) \in \mathbb{R}_0^+$ para todo $x \in X$. La siguiente propiedad clave de las formas cuadráticas positivas es el punto de partida obligado para el estudio de los espacios de Hilbert.

**Desigualdad de Cauchy-Schwartz.** *Sea $\varphi$ una forma sexquilineal hermítica en $X$, cuya forma cuadrática asociada $Q$ es positiva. Entonces:*

$$\big|\varphi\big(x, y\big)\big| \leqslant Q(x)^{1/2}\, Q(y)^{1/2} \qquad \forall x, y \in X \tag{2}$$

*Como consecuencia, la función $x \mapsto Q(x)^{1/2}$ es una seminorma en $X$.*

**Demostración.** Fijados $x, y \in X$, para todo $t \in \mathbb{R}$ tenemos claramente

$$\begin{aligned} 0 \leqslant Q(x - ty) &= \varphi\big(x, x\big) - t\,\varphi\big(x, y\big) - t\,\varphi\big(y, x\big) + t^2\varphi\big(y, y\big) \\ &= Q(x) - 2\,t\,\mathrm{Re}\,\varphi\big(x, y\big) + t^2 Q(y) \end{aligned}$$

o lo que es lo mismo,

$$2\,t\,\mathrm{Re}\,\varphi\big(x, y\big) \leqslant Q(x) + t^2 Q(y) \qquad \forall t \in \mathbb{R} \tag{3}$$

Si $Q(y) > 0$, tomando $t = \mathrm{Re}\,\varphi\big(x, y\big)/Q(y)$, obtenemos claramente

$$\big|\,\mathrm{Re}\,\varphi\big(x, y\big)\big| \leqslant Q(x)^{1/2}\, Q(y)^{1/2} \tag{4}$$

Pero si $Q(y) = 0$, se ha de tener $\mathrm{Re}\,\varphi\big(x, y\big) = 0$, pues en otro caso podríamos tomar $t \in \mathbb{R}$ de forma que $2\,t\,\mathrm{Re}\,\varphi\big(x, y\big) > Q(x)$, lo que contradice $(3)$. Así pues, se verifica $(4)$ para cualesquiera $x, y \in X$, luego hemos probado $(2)$ en caso real.

En caso complejo, fijando de nuevo $x, y \in X$, tomamos $\alpha \in \mathbb{C}$ con $|\alpha| = 1$ de forma que se tenga $\big|\varphi\big(x, y\big)\big| = \alpha\,\varphi\big(x, y\big) = \varphi\big(\alpha x, y\big)$. Es claro entonces que $\varphi\big(\alpha x, y\big) \in \mathbb{R}_0^+$, así como que $Q(\alpha x) = Q(x)$, luego usando $(4)$ con $\alpha x$ en lugar de $x$, obtenemos

$$\big|\varphi\big(x, y\big)\big| = \varphi\big(\alpha x, y\big) = \big|\,\mathrm{Re}\,\varphi\big(\alpha x, y\big)\big| \leqslant Q(\alpha x)^{1/2}\, Q(y)^{1/2} = Q(x)^{1/2}\, Q(y)^{1/2}$$

con lo que hemos probado $(2)$ en caso complejo.

Para ver que la función $x \mapsto Q(x)^{1/2}$ es una seminorma en $X$, la desigualdad triangular es ya inmediata, pues para cualesquiera $x, y \in X$ se tiene

$$\begin{aligned} Q(x+y) &= Q(x) + Q(y) + 2\,\mathrm{Re}\,\varphi\big(x, y\big) \leqslant Q(x) + Q(y) + 2\big|\varphi\big(x, y\big)\big| \\ &\leqslant Q(x) + Q(y) + 2\,Q(x)^{1/2}\, Q(y)^{1/2} = \big(Q(x)^{1/2} + Q(y)^{1/2}\big)^2 \end{aligned}$$

Finalmente es claro que, para $x \in X$ y $\alpha \in \mathbb{K}$, se tiene $Q(\alpha x) = |\alpha|^2\, Q(x)$. ■

Está bien clara la condición que nos permite obtener una norma en $X$. Se dice que una forma cuadrática $Q$ es **definida positiva**, cuando $Q(x) \in \mathbb{R}^+$ para todo $x \in X \setminus \{0\}$. Como toda forma cuadrática $Q$ verifica que $Q(0) = 0$, si $Q$ es definida positiva, vemos que $Q$ es positiva.

Pues bien, un **producto escalar** en $X$ es una forma sexquilineal hermítica $\varphi$ en $X$, cuya forma cuadrática asociada es definida positiva. Habitualmente, un producto escalar se denota por $(x,y) \mapsto \big(x\,\big|\,y\big)$, para $x,y \in X$, y se dice que $\big(x\,\big|\,y\big)$ es el producto escalar de $x$ por $y$. La anterior definición de producto escalar es redundante, pues al ser una forma lineal en la primera variable, la condición para ser hermítica ya hace que sea conjugado-lineal en la segunda variable. Vemos por tanto que una aplicación $(x,y) \mapsto \big(x\,\big|\,y\big)$, de $X \times X \to \mathbb{K}$, es un producto escalar en $X$ si, y sólo si, verifica las tres condiciones siguientes:

**(a)** $\big(\lambda x + z\,\big|\,y\big) = \lambda\,\big(x\,\big|\,y\big) + \big(z\,\big|\,y\big) \qquad \forall x,y,z \in X\,,\ \ \forall \lambda \in \mathbb{K}$

**(b)** $\big(y\,\big|\,x\big) \,=\, \overline{\big(x\,\big|\,y\big)} \qquad \forall x,y \in X$

**(c)** $\big(x\,\big|\,x\big) \,\in\, \mathbb{R}^+ \qquad \forall x \in X \setminus \{0\}$

Un **espacio pre-hilbertiano** es un espacio vectorial $X$, dotado de un producto escalar. Se considera automáticamente a $X$ como un espacio normado, cuya norma viene dada por

$$\|x\| \,=\, \big(x\,\big|\,x\big)^{1/2} \qquad \forall x \in X$$

Así pues, un espacio pre-hilbertiano no es más que un espacio normado $X$, cuya norma procede de un producto escalar, mediante la igualdad anterior. A su vez, el producto escalar de $X$ viene determinado por su norma, mediante la identidad de polarización, que toma la forma:

$$4\,\mathrm{Re}\,\big(x\,\big|\,y\big) \,=\, \|x+y\|^2 \,-\, \|x-y\|^2 \qquad \forall x,y \in X$$

En el caso complejo, basta observar que $\mathrm{Im}\,\big(x\,\big|\,y\big) \,=\, \mathrm{Re}\,\big(x\,\big|\,iy\big)$ para cualesquiera $x,y \in X$.

Deducimos claramente que, si $Y$ es otro espacio pre-hilbertiano, cuyos producto escalar y norma denotamos como los de $X$, y $T : X \to Y$ es un operador lineal isométrico, entonces $T$ preserva el producto escalar, es decir, $\big(T(x)\,\big|\,T(z)\big) \,=\, \big(x\,\big|\,z\big)$ para cualesquiera $x,z \in X$. En particular, si $T$ es un isomorfismo isométrico, entonces $T$ identifica totalmente $X$ con $Y$, no sólo como espacios normados, sino también como espacios pre-hilbertianos.

En un espacio pre-hilbertiano $X$, la desigualdad de Cauchy-Schwartz toma la forma

$$\big|\,\big(x\,\big|\,y\big)\,\big| \,\leqslant\, \big\|x\big\|\,\big\|y\big\| \qquad \forall x,y \in X$$

que nos permite obtener la siguiente consecuencia.

- *El producto escalar de un espacio pre-hilbertiano $X$ es una función continua.*

Basta observar que, para cualesquiera $x,y,u,v \in X$, se tiene

$$\big(x\,\big|\,y\big) - \big(u\,\big|\,v\big) = \big(x-u\,\big|\,y\big) + \big(u\,\big|\,y-v\big) = \big(x-u\,\big|\,y-v\big) + \big(x-u\,\big|\,v\big) + \big(u\,\big|\,y-v\big)$$

y la desigualdad de Cauchy-Schwartz nos dice que

$$\big|\,\big(x\,\big|\,y\big) \,-\, \big(u\,\big|\,v\big)\,\big| \,\leqslant\, \big\|x-u\big\|\,\big\|y-v\big\| \,+\, \big\|x-u\big\|\,\big\|v\big\| \,+\, \big\|u\big\|\,\big\|y-v\big\|$$

de donde se deduce que el producto escalar es continuo en cada punto $(u,v) \in X \times X$. ■

Por otra parte, para cualquier espacio pre-hilbertiano, vamos a discutir la posibilidad de que se tenga la igualdad en la desigualdad de Cauchy-Schwartz, o en la triangular.

- *Sea $X$ un espacio pre-hilbertiano y $x,y \in X$ con $y \neq 0$. Entonces:*

$$(i) \quad |(x|y)| = \|x\| \, \|y\| \quad \Longleftrightarrow \quad \exists \, \lambda \in \mathbb{K} \, : \, x = \lambda y$$
$$(ii) \quad \|x+y\| = \|x\| + \|y\| \quad \Longleftrightarrow \quad \exists \, \rho \in \mathbb{R}_0^+ \, : \, x = \rho y$$

$(i)$ Suponiendo que $|(x|y)| = \|x\| \, \|y\|$, basta tomar $\lambda = (x|y)/\|y\|^2$. En efecto, se tiene $|\lambda|^2 = |(x|y)|^2/\|y\|^4 = \|x\|^2/\|y\|^2$ y también $\overline{\lambda}(x|y) = \|x\|^2$, de donde

$$\|x - \lambda y\|^2 = \|x\|^2 + |\lambda|^2 \, \|y\|^2 - 2 \, \text{Re} \left( \overline{\lambda}(x,|y) \right) = 0$$

luego $x = \lambda y$ como se quería. El recíproco es aún más fácil, pues si $x = \lambda y$ con $\lambda \in \mathbb{K}$, se tiene $\|x\| = |\lambda| \, \|y\|$, de donde $|(x|y)| = |(\lambda y|y)| = |\lambda| \, \|y\|^2 = \|x\| \, \|y\|$.

$(ii)$ Recordamos la prueba de la desigualdad triangular. basada en la de Cauchy-Schwartz:

$$\begin{aligned} \|x+y\|^2 &= \|x\|^2 + \|y\|^2 + 2 \, \text{Re}(x|y) \leqslant \|x\|^2 + \|y\|^2 + 2 \, \big|(x|y)\big| \\ &\leqslant \|x\|^2 + \|y\|^2 + 2 \, \|x\| \, \|y\| = \big(\|x\| + \|y\|\big)^2 \end{aligned}$$

Entonces, $\|x+y\| = \|x\| + \|y\|$ equivale a que se tenga $\text{Re}(x|y) = |(x|y)| = \|x\| \, \|y\|$. Usando $(i)$ la segunda igualdad equivale a $x = \rho y$ con $\rho \in \mathbb{K}$, pero entonces $(x|y) = \rho \, \|y\|^2$ y la primera igualdad equivale a que se tenga $\rho \, \|y\|^2 \in \mathbb{R}_0^+$, es decir, $\rho \in \mathbb{R}_0^+$. ■

Presentamos ya los espacios que más nos interesan. Un **espacio de Hilbert** es un espacio pre-hilbertiano cuya norma es completa, o lo que es lo mismo, un espacio de Banach cuya norma procede de un producto escalar. Seguidamente vamos a mostrar varios ejemplos concretos de espacios de Hilbert. Para el primero de ellos basta recordar el producto escalar usual en $\mathbb{R}^N$ cuya versión en el caso complejo es fácil de adivinar. Obtenemos claramente el siguiente resultado:

- *Para todo $N \in \mathbb{N}$, se tiene que $l_2^N$ es un espacio de Hilbert, cuyo producto escalar es:*

$$(x|y) = \sum_{k=1}^{N} x(k) \, \overline{y(k)} \qquad \forall x,y \in l_2^N$$

Obsérvese que la desigualdad de Cauchy-Schwartz en el espacio $l_2^N$ es caso particular de la de Hölder, pues la segunda nos permite escribir:

$$\big|(x|y)\big| \leqslant \sum_{k=1}^{N} |x(k)| \, |\overline{y(k)}| = \sum_{k=1}^{N} |x(k)| \, |y(k)| \leqslant \|x\|_2 \, \|y\|_2 \qquad \forall x,y \in \mathbb{K}^N$$

Razonando con sucesiones de forma análoga a como lo hemos hecho en $\mathbb{K}^N$, encontramos el primer ejemplo de espacio de Hilbert de dimensión infinita. Para $x,y \in l_2$, sabemos que la serie $\sum_{n \geqslant 1} x(n) \, \overline{y(n)}$ es absolutamente convergente, lo que permite claramente definir un producto escalar en $l_2$ que da lugar a su norma:

- *El espacio de sucesiones $l_2$ es un espacio de Hilbert, con el producto escalar dado por*

$$(x|y) = \sum_{n=1}^{\infty} x(n) \, \overline{y(n)} \qquad \forall x,y \in l_2$$

Recordemos que el subespacio $\mathbb{K}^{(\mathbb{N})} \subset l_2$, formado por las sucesiones de soporte finito, es denso en $l_2$ pero no es cerrado, lo que nos da el primer ejemplo de un espacio pre-hilbertiano que no es completo.

Como se puede ya adivinar, nuestro tercer ejemplo de espacio de Hilbert es el espacio de Lebesgue $L_2(\Omega)$, donde $\Omega$ es un abierto no vacío de $\mathbb{R}^N$, con $N \in \mathbb{N}$. Para $f, g \in L_2(\Omega)$, la desigualdad integral de Hölder nos dice que $f\,\overline{g} \in L_1(\Omega)$, y esto permite claramente definir un producto escalar en $L_2(\Omega)$ que da lugar a su norma:

- *Si $N \in \mathbb{N}$ y $\Omega$ es un abierto no vacío de $\mathbb{R}^N$, el espacio de Lebesgue $L_2(\Omega)$ es un espacio de Hilbert, cuyo producto escalar viene dado por*

$$\big(f \,|\, g\big) = \int_\Omega f(t)\, \overline{g(t)}\, dt \qquad \forall f, g \in L_2(\Omega)$$

Nótese que la desigualdad de Cauchy-Schwartz en el espacio de Hilbert $L_2(\Omega)$ no es más que la desigualdad integral de Hölder para $p = 2$.

Recordando que $C_{00}(\Omega)$ siempre puede verse como un subespacio denso en $L_2(\Omega)$, que no es cerrado, tenemos que $C_{00}(\Omega)$, no con su norma natural, sino con la inducida por $L_2(\Omega)$, es otro ejemplo de espacio pre-hilbertiano no completo.

## 6.3. La identidad del paralelogramo

Es natural buscar una caracterización de los espacios pre-hilbertianos entre los espacios normados, o lo que es lo mismo, una caracterización de las normas que proceden de un producto escalar. Entre las muchas respuestas que se conocen a esta pregunta, veremos la más clásica.

Si $X$ es un espacio pre-hilbertiano, se tiene claramente

$$\|x+y\|^2 = \big(x+y \,|\, x+y\big) = \|x\|^2 + \|y\|^2 + 2\ \mathrm{Re}\,\big(x \,|\, y\big) \qquad \forall x, y \in X$$

Hacemos desaparecer el producto escalar, sumando miembro a miembro esta igualdad con la que se obtiene al sustituir $y$ por $-y$, con lo que tenemos

$$\|x+y\|^2 + \|x-y\|^2 = 2\,\|x\|^2 + 2\,\|y\|^2 \qquad \forall x, y \in X \tag{5}$$

Esta propiedad recibe el nombre de **identidad del paralelogramo**, por su obvia interpretación geométrica: en todo paralelogramo, la suma de los cuadrados de las diagonales coincide con la suma de los cuadrados de los lados.

Pero lo importante de la igualdad $(5)$ es que en ella no aparece el producto escalar de $X$, sino solamente su norma. Podemos decir que un espacio normado $X$ verifica la identidad del paralelogramo cuando se cumple $(5)$. Tenemos así una condición necesaria para que una norma proceda de un producto escalar. Como la forma de obtenerla ha sido tan sencilla, no parece que tal condición pueda ser suficiente, pero vamos a probar que sí lo es.

**Teorema (Jordan-Von Neumann, 1935).** *Un espacio normado es pre-hilbertiano si, y sólo si, verifica la identidad del paralelogramo.*

**Demostración.** Como ya conocemos una implicación, dado un espacio normado $X$, que verifica la identidad del paralelogramo, se trata de probar que la norma de $X$ procede de un producto escalar. En vista de la identidad de polarización, está claro que debemos considerar la aplicación $\varphi : X \times X \to \mathbb{R}$ definida por

$$4\,\varphi\big(x\,,y\big) = \|x+y\|^2 - \|x-y\|^2 \qquad \forall x,y \in X \tag{6}$$

pues en el caso real, el producto escalar que buscamos no puede ser otro que $\varphi$, mientras que en el caso complejo, $\varphi$ ha de ser la parte real de dicho producto escalar.

Dados $u,v,y \in X$ la identidad del paralelogramo nos dice que

$$\begin{aligned} 4\varphi\big(u+v\,,y\big) &= \|u+v+y\|^2 - \|u+v-y\|^2 \\ &= 2\,\|u+y\|^2 + 2\,\|v\|^2 - \|u+y-v\|^2 - 2\,\|u\|^2 - 2\,\|v-y\|^2 + \|u-v+y\|^2 \\ &= 2\,\|u+y\|^2 + 2\,\|v\|^2 - 2\,\|u\|^2 - 2\|v-y\|^2 \end{aligned}$$

Sumando esta igualdad con la que se obtiene al intercambiar $u$ y $v$, vemos que

$$8\varphi\big(u+v\,,y\big) = 2\,\|u+y\|^2 - 2\,\|u-y\|^2 + 2\,\|v+y\|^2 - 2\,\|v-y\|^2$$

El segundo miembro es, por definición, $8\,\varphi\big(u\,,y\big) + 8\,\varphi\big(v\,,y\big)$, luego hemos probado que

$$\varphi\big(u+v\,,y\big) = \varphi\big(u\,,y\big) + \varphi\big(v\,,y\big) \qquad \forall u,v,y \in X \tag{7}$$

Consideramos ahora el conjunto

$$E = \big\{\alpha \in \mathbb{R} \,:\, \varphi(\alpha x\,,y\big) = \alpha\varphi\big(x\,,y\big) \ \forall x,y \in X\big\}$$

que no es vacío, pues $0,1 \in E$. Dados $\alpha,\beta \in E$ y $x,y \in X$, usando $(7)$ se tiene

$$\varphi\big((\alpha-\beta)x\,,y\big) + \beta\,\varphi\big(x\,,y\big) = \varphi\big((\alpha-\beta)x\,,y\big) + \varphi\big(\beta x\,,y\big) = \varphi\big(\alpha x\,,y\big) = \alpha\,\varphi\big(x\,,y\big)$$

de donde se deduce que $\alpha-\beta \in E$. Pero además, si $\beta \neq 0$, observamos que

$$\beta\,\varphi\big((\alpha/\beta)x\,,y\big) = \varphi\big(\beta(\alpha/\beta)x\,,y\big) = \varphi\big(\alpha x\,,y\big) = \alpha\,\varphi\big(x\,,y\big)$$

y deducimos que $\alpha/\beta \in E$. En resumen, $E$ es un subcuerpo de $\mathbb{R}$, luego se tiene $\mathbb{Q} \subset E$.

Fijados $x,y \in X$, usamos la función $f_{xy} : \mathbb{R} \to \mathbb{R}$ definida, para todo $\alpha \in \mathbb{R}$ por

$$f_{xy}(\alpha) = \big(\alpha x\,\big|\,y\big) - \alpha\big(x\,\big|\,y\big) = (1/4)\|\,\alpha x + y\,\|^2 - (1/4)\|\,\alpha x - y\,\|^2 - \alpha\big(x\,\big|\,y\big)$$

La continuidad de la suma, el producto por escalares y la norma de $X$ nos dicen que $f_{xy}$ es una función continua, luego el conjunto $E_{xy} = f_{xy}^{-1}\big(\{0\}\big)$ es cerrado en $\mathbb{R}$. Ahora bien, $E$ es la intersección de todos los conjuntos de la forma $E_{xy}$ con $x,y \in X$, luego $E$ también es cerrado. Como $\mathbb{Q} \subset E$, concluimos que $E = \mathbb{R}$. Esto significa que

$$\varphi\big(\alpha x\,,y\big) = \alpha\,\varphi\big(x\,,y\big) \qquad \forall \alpha \in \mathbb{R},\ \forall x,y \in X \tag{8}$$

En el caso real, es ya fácil completar la demostración. En (7) y (8) tenemos que $\varphi$ es lineal en la primera variable, pero en (6) vemos claramente que $\varphi(x,y) = \varphi(y,x)$ para cualesquiera $x,y \in X$, luego $\varphi$ es una forma bilineal simétrica en $X$. Pero de nuevo (6) nos dice que $\varphi(x,x) = \|x\|^2$ para todo $x \in X$, luego la forma cuadrática asociada a $\varphi$ es definida positiva, y $\varphi$ es el producto escalar que buscábamos.

En el caso complejo, debemos considerar la aplicación $\psi : X \times X \to \mathbb{C}$ dada por

$$\psi(x,y) = \varphi(x,y) + i\varphi(x,iy) \qquad \forall x,y \in \mathbb{R} \tag{9}$$

pues $\psi$ ha de ser el producto escalar que buscamos. Las propiedades que $\psi$ debe verificar, se deducirán de las de $\varphi$, con una observación clave: para $x,y \in X$, como $X$ es un espacio normado complejo, se tiene $\|x+iy\| = \|ix-y\|$, así como $\|x-iy\| = \|ix+y\|$, luego

$$\varphi(x,iy) = -\varphi(ix,y) \qquad \forall x,y \in X \tag{10}$$

Veamos entonces que $\psi$ es lineal en la primera variable, sin olvidar que ahora $\mathbb{K} = \mathbb{C}$. Para cualesquiera $x,y,z \in X$ y $\alpha \in \mathbb{R}$, de (7) y (8) deducimos claramente que

$$\psi(x+z,y) = \psi(x,y) + \psi(z,y) \qquad \text{y} \qquad \psi(\alpha x,y) = \alpha\,\psi(x,y)$$

luego sólo queda comprobar que $\psi(ix,y) = i\,\psi(x,y)$, pero esto se deduce de (10), ya que

$$\begin{aligned}\psi(ix,y) &= \varphi(ix,y) + i\,\varphi(ix,iy) = -\varphi(x,iy) - i\,\varphi(x,-y)\\ &= i\left[i\,\varphi(x,iy) + \varphi(x,y)\right] = i\,\psi(x,y)\end{aligned}$$

Por otra parte, para $x,y \in X$, también deducimos de (10) que

$$\psi(y,x) = \varphi(y,x) + i\,\varphi(y,ix) = \varphi(x,y) - i\,\varphi(x,iy) = \overline{\psi(x,y)}$$

Como ya teníamos linealidad en la primera variable, la última igualdad nos dice que $\psi$ es una forma sexquilineal hermítica. Finalmente, para $x \in X$, usando (10) y el hecho de que $\varphi$ es simétrica, obtenemos $\varphi(x,ix) = -\varphi(ix,x) = -\varphi(x,ix)$, luego $\varphi(x,ix) = 0$, de donde

$$\psi(x,x) = \varphi(x,x) + i\,\varphi(x,ix) = \|x\|^2 \qquad \forall x \in X$$

Por tanto, la forma cuadrática asociada a $\psi$ es definida positiva, luego $\psi$ es el producto escalar que buscábamos. ■

El teorema anterior tiene varias consecuencias inmediatas que conviene destacar. Para la primera de ellas, usaremos que todo espacio normado complejo se puede ver también como un espacio normado real, una idea que volverá a aparecer varias veces, por lo que conviene formalizarla con detalle.

Si $X$ es un espacio vectorial complejo, su producto por escalares es una aplicación de $\mathbb{C} \times X$ en $X$, que puede restringirse a $\mathbb{R} \times X$, con lo que obviamente $X$ se convierte en un espacio vectorial real, al que llamaremos **espacio real subyacente** a $X$ y denotaremos por $X_{\mathbb{R}}$.

Es evidente que toda norma en $X$ es también una norma en $X_{\mathbb{R}}$, pero el recíproco no es cierto. Por ejemplo, para $X = \mathbb{C}$ se tiene $X_{\mathbb{R}} = \mathbb{R}^2$, y una norma en $\mathbb{R}^2$ sólo es una norma en $\mathbb{C}$ cuando es proporcional a la euclídea. En cualquier caso, si $X$ es un espacio normado complejo, vemos siempre a $X$ como espacio normado real, con la misma norma que $X$. Como la identidad del paralelogramo no involucra el producto por escalares, se tiene:

- *Un espacio normado complejo $X$ es pre-hilbertiano si, y sólo si, lo es $X_{\mathbb{R}}$.*

Conviene aclarar que, cuando $X$ y $X_{\mathbb{R}}$ son espacios pre-hilbertianos, aunque ambos tienen la misma norma, sus productos escalares no pueden coincidir. Esto puede verse en la demostración del teorema anterior, pues el producto escalar de $X_{\mathbb{R}}$ es la aplicación $\varphi$ definida en $(6)$, mientras que el de $X$ es la aplicación $\psi$, dada por $(9)$. Por tanto, la relación entre ambos productos escalares está bien clara: para cualesquiera $x, y \in X$ se tiene

$$\psi(x,y) = \varphi(x,y) + i\,\varphi(x,iy), \qquad \text{o bien,} \qquad \varphi(x,y) = \operatorname{Re}\psi(x,y)$$

Por otra parte, como la identidad del paralelogramo sólo involucra dos vectores de nuestro espacio normado, deducimos lo siguiente:

- *Un espacio normado es pre-hilbertiano si, y sólo si, lo son todos sus subespacios de dimensión $2$.*

Uniendo las dos observaciones anteriores, para saber si un espacio normado $X$, es o no un espacio pre-hilbertiano, sólo tenemos que examinar los subespacios bidimensionales de $X_{\mathbb{R}}$.

Para una tercera observación, sea $M$ un subespacio de un espacio normado $X$, con la norma inducida. Si $X$ es pre-hilbertiano, está claro que $M$ también lo es, pues al restringir a $M \times M$ el producto escalar de $X$, se tiene un producto escalar en $M$ que obviamente da lugar a la norma inducida. El teorema anterior nos permite comprobar que el recíproco es cierto, siempre que $M$ sea denso en $X$. Para ello basta ver los dos miembros de la identidad del paralelogramo como funciones continuas de dos variables, $x, y \in X$, definidas en $X \times X$, con lo que el conjunto $E$ de los pares $(x,y) \in X \times X$ que verifican dicha identidad es cerrado. Como $M$ es pre-hilbertiano, tenemos $M \times M \subset E$, pero por ser $M$ denso en $X$, también $M \times M$ es denso en $X \times X$ y concluimos que $E = X \times X$. Hemos probado:

- *Sea $X$ un espacio normado y $M$ un subespacio denso en $X$. Si $M$, con la norma inducida por $X$, es un espacio pre-hilbertiano, entonces $X$ también lo es.*

Podemos ahora averiguar fácilmente, cuáles de los espacios de Banach presentados en su momento como ejemplos, son espacios de Hilbert. Para $N \in \mathbb{N}$ con $N > 1$ y $1 \leqslant p \leqslant \infty$, usemos los dos primeros vectores, $e_1$ y $e_2$, en la base usual de $\mathbb{K}^N$. Es claro que $\|e_1\|_p = \|e_2\|_p = 1$, mientras que $\|e_1 \pm e_2\| = 2^{1/p}$ si $p < \infty$, y $\|e_1 \pm e_2\|_\infty = 1$. Vemos entonces que

$$\big(\|e_1 + e_2\|_\infty\big)^2 + \big(\|e_1 - e_2\|_\infty\big)^2 = 2 \neq 4 = 2\big(\|e_1\|_\infty\big)^2 + 2\big(\|e_2\|_\infty\big)^2$$

luego $l_\infty^N$ no verifica la identidad del paralelogramo. Si $1 \leqslant p < \infty$ y $l_p^N$ la verifica, tenemos

$$2 \cdot 2^{2/p} = \big(\|e_1 + e_2\|_p\big)^2 + \big(\|e_1 - e_2\|_p\big)^2 = 2\big(\|e_1\|_p\big)^2 + 2\big(\|e_2\|_p\big)^2 = 4$$

de donde se deduce que $p = 2$ y hemos obtenido el resultado que sigue.

- *Para* $N \in \mathbb{N}$ *con* $N \geqslant 2$ *y* $1 \leqslant p \leqslant \infty$, $l_p^N$ *es un espacio de Hilbert si, y sólo si,* $p = 2$.

Observemos ahora que para $1 \leqslant p < \infty$, el espacio de sucesiones $l_p$ contiene un subespacio isométricamente isomorfo a $l_p^2$, luego $l_p$ sólo puede ser un espacio de Hilbert para $p = 2$. En el caso $p = \infty$, vemos también que $l_\infty^2$ es isométricamente isomorfo a un subespacio de $c_0$, luego $c_0$ no es un espacio de Hilbert, y por tanto $l_\infty$ tampoco puede serlo. En resumen, para los espacios clásicos de sucesiones tenemos:

- *Para* $1 \leqslant p \leqslant \infty$, *se tiene que* $l_p$ *es un espacio de Hilbert si, y sólo si,* $p = 2$. *Por su parte,* $c_0$ *tampoco es un espacio de Hilbert.*

Para los espacios de Lebesgue, si $\Omega$ es un abierto no vacío de $\mathbb{R}^N$ y $1 \leqslant p \leqslant \infty$, entonces sabemos que $L_p(\Omega)$ contiene un subespacio isométricamente isomorfo a $l_p$. Por tanto:

- *Si* $\Omega$ *es un abierto no vacío de* $\mathbb{R}^N$, *con* $N \in \mathbb{N}$ *arbitrario, y* $1 \leqslant p \leqslant \infty$, *entonces* $L_p(\Omega)$ *es un espacio de Hilbert si, y sólo si,* $p = 2$.

Si $\Gamma$ es un conjunto no numerable, está claro que el espacio de funciones acotadas $l_\infty(\Gamma)$ contiene un subespacio isométricamente isomorfo a $l_\infty^2$, luego $l_\infty(\Gamma)$ no es un espacio de Hilbert, igual que cuando $\Gamma$ era numerable.

Queda estudiar los espacios de funciones continuas, cosa que tampoco es difícil. Sea pues $\Omega$ un espacio topológico de Hausdorff, localmente compacto, que contenga al menos dos puntos distintos $s, t \in \Omega$. Tomamos entonces dos abiertos $U, V \subset \Omega$ tales que $s \in U$, $t \in V$ y $U \cap V = \emptyset$ y sabemos que existen $f, g \in C_{00}(\Omega)$ que verifican:

$$f(\Omega) \cup g(\Omega) \subset [0,1], \quad f(s) = g(t) = 1, \quad \operatorname{sop} f \subset U, \quad \operatorname{sop} g \subset V$$

deducimos claramente que $\|f \pm g\|_\infty = \|f\|_\infty = \|g\|_\infty = 1$, luego $C_{00}(\Omega)$, ahora con su norma natural $\|\cdot\|_\infty$, no verifica la identidad del paralelogramo. Obtenemos así la siguiente conclusión:

- *Si* $\Omega$ *es un espacio topológico de Hausdorff, localmente compacto, que no se reduce a un punto, entonces* $C_{00}(\Omega)$ *no es un espacio pre-hilbertiano. Por tanto, los espacios de Banach* $C_0(\Omega)$ *y* $C_b(\Omega)$ *no son espacios de Hilbert.*

## 6.4. Mejor aproximación en espacios de Hilbert

La identidad del paralelogramo permite obtener muy fácilmente un resultado clave, sobre existencia y unicidad de mejores aproximaciones en espacios de Hilbert, que nos llevará después al teorema más importante referente a dichos espacios. Recordemos que un subconjunto $A$ de un espacio vectorial $X$ es **convexo**, cuando para cualesquiera $x, y \in A$, se tiene que el segmento de extremos $x$ e $y$ está contenido en $A$, es decir, $(1-t)x + ty \in A$ para todo $t \in [0,1]$.

**Teorema de aproximación óptima.** *Sea $A$ un subconjunto no vacío, convexo y cerrado, de un espacio de Hilbert $H$. Entonces, cada $x \in X$ tiene una única mejor aproximación en $A$, es decir: para cada $x \in H$ existe un único $y \in A$ verificando que $\|x - y\| = d(x,A)$.*

**Demostración.** Fijado $x \in H$, para $a, b \in A$, por ser $A$ convexo se tiene que $(a+b)/2 \in A$ de donde, usando la identidad del paralelogramo deducimos que

$$4\,d(x,A)^2 \leqslant 4\left\|x - \frac{a+b}{2}\right\|^2 = \|(x-a)+(x-b)\|^2 = 2\,\|x-a\|^2 + 2\,\|x-b\|^2 - \|b-a\|^2$$

o lo que es lo mismo,

$$\|b-a\|^2 \leqslant 2\,\|x-a\|^2 + 2\,\|x-b\|^2 - 4\,d(x,A)^2 \qquad \forall a,b \in A \tag{11}$$

y esta desigualdad es la clave de la demostración.

Tomamos una sucesión $\{y_n\}$ de puntos de $A$ tal que $\big\{\|x - y_n\|\big\} \to d(x,A)$. Dado $\varepsilon > 0$, existe entonces $n_0 \in \mathbb{N}$ tal que, para $k \geqslant n_0$ se tiene $\|x - y_k\|^2 < d(x,A)^2 + \big(\varepsilon^2/4\big)$.

Para $n,m \in \mathbb{N}$ con $n,m \geqslant n_0$, podemos entonces usar $(11)$ con $a = y_m$ y $b = y_n$ para obtener que $\|y_n - y_m\|^2 \leqslant 2\,\|x - y_m\|^2 + 2\,\|x - y_n\|^2 - 4\,d(x,A)^2 < \varepsilon^2$, luego $\{y_n\}$ es una sucesión de Cauchy de puntos de $A$. Como $H$ es completo y $A$ cerrado, tenemos $\{y_n\} \to y \in A$. La continuidad de la norma de $H$ nos dice entonces que

$$\|x - y\| = \lim_{n\to\infty} \|x - y_n\| = d(x,A)$$

de modo que $y$ es una mejor aproximación de $x$ en $A$. Es la única, pues si $z \in A$ también verifica que $\|x - z\| = d(x,A)$, la desigualdad $(11)$ con $a = z$ y $b = y$ nos dice que $z = y$. ■

Merece la pena discutir brevemente el papel que juegan las hipótesis del teorema anterior, por una parte para la existencia, y por otra para la unicidad, de las mejores aproximaciones.

Empezando por la existencia, para que todo $x \in H$ tenga al menos una mejor aproximación en $A$, es decir, para que $A$ sea proximinal en $H$, es necesario que $A$ sea cerrado, luego esta hipótesis no se puede suprimir. Con respecto a la complitud de $H$, en la demostración sólo se usa para asegurarse la complitud de $A$. Por tanto, el teorema es cierto para un subconjunto no vacío, convexo y completo, de cualquier espacio pre-hilbertiano, pero esto tiene poca utilidad.

Más importante es comprobar que la convexidad de $M$ no se puede suprimir, es decir, que un subconjunto cerrado $A$ de un espacio de Hilbert $H$ puede no ser proximinal en $H$, como veremos con un ejemplo.

Sea $\{e_n\}$ la sucesión de vectores unidad en el espacio de Hilbert $l_2$, y consideremos el conjunto $A = \big\{y_n : n \in \mathbb{N}\big\}$ donde $y_n = (n+1)\,e_n/n$ para todo $n \in \mathbb{N}$.

Para $n,m \in \mathbb{N}$ con $n \neq m$ se tiene $\|y_n - y_m\| \geqslant \sqrt{2}$, de donde se deduce que $A$ es cerrado, pues una sucesión convergente de puntos de $A$ ha de ser constante a partir de un término en adelante, luego converge a un punto de $A$. Por otra parte, vemos claramente que $0$ no tiene una mejor aproximación en $A$, ya que $d(0,A) = \inf\big\{\|y_n\| : n \in \mathbb{N}\big\} = 1$, pero $\|y_n\| > 1$ para todo $n \in \mathbb{N}$. Así pues, $A$ es un subconjunto cerrado del espacio de Hilbert $l_2$, pero $A$ no es proximinal en $l_2$.

En resumen, ninguna de las hipótesis del teorema se puede suprimir, si queremos asegurar la existencia de mejores aproximaciones. Para la unicidad la situación es completamente diferente, como vamos a ver.

En la demostración del teorema, la unicidad se deduce de la desigualdad $(11)$, obtenida usando solamente que $A$ es convexo, luego no se precisa que $H$ sea completo, ni que $A$ sea cerrado. Por tanto, si $A$ es un subconjunto convexo de un espacio pre-hilbertiano $H$, entonces cada $x \in H$ tiene a lo sumo una mejor aproximación en $A$. Finalmente está bien claro que la convexidad de $A$ no se puede suprimir, pues basta pensar lo que ocurre cuando $A$ tiene exactamente dos puntos.

## 6.5. Ortogonalidad

Para llegar al principal resultado en la teoría de los espacios de Hilbert, sólo nos queda obtener una caracterización de las mejores aproximaciones, que es independiente del teorema anterior, y acorde con la intuición geométrica.

Sea $M$ un subconjunto no vacío de un espacio pre-hilbertiano $X$. Entonces $y_0 \in M$ es una mejor aproximación de un $x \in X$ en $M$ si, y sólo si, verifica que

$$\|x - y_0\|^2 \leqslant \|x - z\|^2 \qquad \forall z \in M \tag{12}$$

Ahora bien, para todo $z \in M$ se tiene

$$\|x-z\|^2 = \|(x-y_0)-(z-y_0)\|^2 = \|x-y_0\|^2 + \|z-y_0\|^2 - 2\,\mathrm{Re}\,\big(x-y_0\,\big|\,z-y_0\big)$$

luego la condición $(12)$ equivale a

$$2\,\mathrm{Re}\,\big(x-y_0\,\big|\,z-y_0\big) \leqslant \big\|z-y_0\big\|^2 \qquad \forall z \in M \tag{13}$$

Si $M$ es convexo, dados $y \in M$ y $t \in ]0,1[$, usamos $(13)$ con $z = (1-t)y_0 + t\,y \in M$, con lo que $z - y_0 = t\,(y - y_0)$, y obtenemos

$$2t\,\mathrm{Re}\,\big(x-y_0\,\big|\,y-y_0\big) \leqslant t^2\big\|y-y_0\big\|^2$$

Dividiendo ambos miembros por $t > 0$, y haciendo $t \to 0$, deducimos que

$$\mathrm{Re}\,\big(x-y_0\,\big|\,y-y_0\big) \leqslant 0 \qquad \forall y \in M \tag{14}$$

Recíprocamente, de $(14)$ se deduce $(13)$, o lo que es lo mismo $(12)$, sin más que tomar $y = z$.

La condición $(14)$ se simplifica notablemente cuando $M$ es un subespacio de $X$. De hecho, para todo $z \in M$, usando $(14)$ con $y = y_0 \pm z \in M$, obtenemos que $\mathrm{Re}\,\big(x-y_0\,\big|\,z\big) = 0$. En el caso complejo, tenemos también $\mathrm{Im}\,\big(x-y_0\,\big|\,z\big) = \mathrm{Re}\,\big(x-y_0\,\big|\,iz\big) = 0$ para todo $z \in M$. Así pues, en cualquier caso, de $(14)$ hemos deducido que

$$\big(x-y_0\,\big|\,z\big) = 0 \qquad \forall z \in M \tag{15}$$

Recíprocamente, de $(15)$ deducimos $(14)$ tomando $z = y - y_0$. El siguiente enunciado recoge las equivalencias recién probadas.

- **Caracterización de las mejores aproximaciones.** *Sea $M$ un subconjunto convexo de un espacio pre-hilbertiano $X$. Para $x \in X$ e $y_0 \in M$, se tiene:*

$$\|x - y_0\| \leqslant \|x - y\| \quad \forall y \in M \quad \Longleftrightarrow \quad \mathrm{Re}\,\big(x - y_0 \,\big|\, y - y_0\big) \leqslant 0 \quad \forall y \in M$$

*Si $M$ es de hecho un subespacio de $X$, se tiene también:*

$$\|x - y_0\| \leqslant \|x - y\| \quad \forall y \in M \quad \Longleftrightarrow \quad \big(x - y_0 \,\big|\, y\big) = 0 \quad \forall y \in M$$

Para interpretar geométricamente estas equivalencias, usamos una nomenclatura, motivada por la segunda, que a partir de ahora será muy conveniente. Si $X$ es un espacio pre-hilbertiano, dados $x, y \in X$, decimos que $x$ es **ortogonal** a $y$, cuando se tiene que $\big(x \,\big|\, y\big) = 0$, en cuyo caso escribimos $x \perp y$. También podemos decir que $x$ e $y$ son ortogonales, pues se trata de una relación simétrica: $x \perp y \Leftrightarrow y \perp x$. Si ahora $E$, es un subconjunto no vacío de $X$, definimos

$$E^{\perp} = \big\{ x \in X : x \perp y \;\; \forall y \in E \big\} = \big\{ x \in X : \big(x \,\big|\, y\big) = 0 \;\; \forall y \in E \big\}$$

Por la linealidad y continuidad del producto escalar en la primera variable, vemos que $E^{\perp}$ es un subespacio cerrado de $X$. Se dice que $E^{\perp}$ es el **subespacio ortogonal** a $E$.

Iterando el paso de $E$ a $E^{\perp}$, es natural escribir $E^{\perp\perp} = \big(E^{\perp}\big)^{\perp}$. Es claro que $E \subset E^{\perp\perp}$, pero $E^{\perp\perp}$ es un subespacio cerrado de $X$, luego $\overline{\mathrm{Lin}\,E} \subset E^{\perp\perp}$.

Para entender geométricamente la ortogonalidad, basta pensar en el teorema de Pitágoras: dos vectores $x, y$ son *perpendiculares* cuando $\|x + y\|^2 = \|x\|^2 + \|y\|^2$. Si $X$ es un espacio pre-hilbertiano, para $x, y \in X$ se tiene

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\,\mathrm{Re}\,\big(x \,\big|\, y\big)$$

luego $x$ e $y$ son perpendiculares si, y sólo si, $\mathrm{Re}\,\big(x \,\big|\, y\big) = 0$, lo que en caso real, equivale a que $x$ e $y$ sean ortogonales. En caso complejo, tenemos $\mathrm{Im}\,\big(x \,\big|\, y\big) = \mathrm{Re}\,\big(x \,\big|\, i y\big)$, luego $x$ e $y$ son ortogonales cuando $x$ es perpendicular tanto a $y$ como a $i y$, es decir, cuando el vector $x$ es perpendicular al plano (real) determinado por $y$ e $i y$. Por tanto, la interpretación geométrica de la ortogonalidad queda resumida en el enunciado que sigue.

- *Si $X$ es un espacio pre-hilbertiano, para $x, y \in X$ se tiene:*

  $(i)$ *En caso real:* $\quad x \perp y \iff \|x + y\|^2 = \|x\|^2 + \|y\|^2$

  $(ii)$ *En caso complejo:* $\quad x \perp y \iff \|x + y\|^2 = \|x + iy\|^2 = \|x\|^2 + \|y\|^2$

La caracterización de las mejores aproximaciones tiene ya una interpretación geométrica muy intuitiva. Si $M$ es un subespacio del espacio pre-hilbertiano $X$, vimos que $y_0 \in M$ es la mejor aproximación de $x \in X$ en $M$ si, y sólo si, $x - y_0 \in M^{\perp}$. En términos geométricos, esto significa que el vector $x - y_0$ es perpendicular al subespacio $M$, o que el punto $y_0$ es el pie de la recta perpendicular a $M$ que pasa por el punto $x$, lo que está totalmente de acuerdo con la intuición geométrica.

En el caso de un conjunto convexo $M \subset X$, la interpretación es también intuitiva, aunque no tan sencilla. La hacemos en caso real, pues si $X$ es complejo, basta pensar que $M$ es un subconjunto convexo de $X_{\mathbb{R}}$ y que las distancias en $X$ son las mismas que en $X_{\mathbb{R}}$.

Vimos que $y_0 \in M$ es la mejor aproximación de $x \in X$ en $M$ cuando $\big(x - y_0 \,\big|\, y - y_0\big) \leqslant 0$ para todo $y \in M$. Si $\alpha_0 = \big(x - y_0 \,\big|\, y_0\big) \in \mathbb{R}$, esto equivale a $\big(x - y_0 \,\big|\, y\big) \leqslant \alpha_0$ para todo $y \in M$. Descartando el caso trivial $x = y_0$, tenemos $\big(x - y_0 \,\big|\, x - y_0\big) > 0$, luego $\big(x - y_0 \,\big|\, x\big) > \alpha_0$. Entonces, $Z = \big\{ z \in X : \big(x - y_0 \,\big|\, z\big) = \alpha_0 \big\}$ es un *hiperplano afín* que *pasa* por el punto $y_0$, ya que $y_0 \in Z$, y *separa* el punto $x$ del conjunto $M$, es decir, deja $x$ a un lado y $M$ al otro.

Conviene describir mejor el hiperplano $Z$. Para $z_1, z_2 \in Z$ se tiene $\big(x - y_0 \,\big|\, z_1 - z_2\big) = 0$, así que $x - y_0$ es perpendicular a todas las rectas contenidas en $Z$, es un *vector normal* al hiperplano $Z$. En resumen, el único hiperplano que pasa por $y_0$ y tiene a $x - y_0$ como vector normal, separa $x$ de $M$. Que esta condición caracterice a la mejor aproximación de $x$ en $M$, vuelve a estar de acuerdo con la intuición geométrica.

## 6.6. Teorema de la proyección ortogonal

Enlazando el teorema de aproximación óptima, que garantiza la existencia y unicidad de mejores aproximaciones, con la caracterización de las mismas en términos de ortogonalidad, obtenemos el resultado más importante en el estudio de los espacios de Hilbert:

**Teorema de la proyección ortogonal.** *Si $M$ es un subespacio cerrado de un espacio de Hilbert $H$, se tiene $H = M \oplus M^{\perp}$, y esta suma es topológico-directa.*

*De hecho, si $P_M$ es la única proyección lineal de $X$ sobre $M$ cuyo núcleo es $M^{\perp}$, se tiene:*

$$\big\| x \big\|^2 = \big\| P_M(x) \big\|^2 + \big\| x - P_M(x) \big\|^2 \qquad \forall x \in X \tag{16}$$

*luego $P_M$ es continua y $\|P_M\| = 1$, a menos que $M = \{0\}$. Además, para cada $x \in X$, se tiene que $P_M(x)$ es la única mejor aproximación de $x$ en $M$. Finalmente, se verifica que $M^{\perp\perp} = M$.*

**Demostración.** Cada $x \in H$ tiene una única mejor aproximación en $M$ que, adelantando acontecimientos, denotamos por $P_M(x)$. Sabemos que $P_M(x) \in M$ se caracteriza por verificar que $x - P_M(x) \in M^{\perp}$.

Se tiene entonces que $x = P_M(x) + \big(x - P_M(x)\big) \in M + M^{\perp}$ para todo $x \in H$. Además, vemos que $M \cap M^{\perp} = \{0\}$, pues para todo $y \in M \cap M^{\perp}$ se tiene $y \perp y$, luego $y = 0$. Por tanto tenemos $H = M \oplus M^{\perp}$ y efectivamente, $P_M$ es la única proyección lineal de $X$ sobre $M$ cuyo núcleo es $M^{\perp}$.

La igualdad (16) se debe a que $P_M(x) \perp \big(x - P_M(x)\big)$ para todo $x \in H$, y de ella deducimos que $\|P_M(x)\| \leqslant \|x\|$ para todo $x \in H$, luego $P_M$ es continua con $\|P_M\| \leqslant 1$. Si $M \neq \{0\}$, al ser $P_M(y) = y$ para todo $y \in M$, se ha de tener $\|P_M\| = 1$. Sabemos que la continuidad de $P_M$ equivale a que la suma directa $H = M \oplus M^{\perp}$ sea topológico-directa.

Finalmente, para $z \in M^{\perp\perp}$ se tiene $P_M(z) \in M \subset M^{\perp\perp}$, de donde $z - P_M(z) \in M^{\perp\perp}$, pero también $z - P_M(z) \in M^{\perp}$, así que $z - P_M(z) \in M^{\perp\perp} \cap M^{\perp} = \{0\}$. Por tanto $z = P_M(z) \in M$ y hemos probado que $M^{\perp\perp} \subset M$, pero la otra inclusión es sabida. ■

Con la notación del teorema anterior, se dice que $P_M$ es la **proyección ortogonal** de $H$ sobre $M$, cuya interpretación geométrica está bien clara. Nótese que $M^\perp$ está en la misma situación que $M$, pues también es subespacio cerrado de $H$. Tenemos por tanto las proyecciones ortogonales $P_M$ y $P_{M^\perp}$, de $H$ sobre $M$ y $M^\perp$ respectivamente. Como $\ker P_{M^\perp} = M^{\perp\perp} = M$, dichas proyecciones son complementarias: $P_M(x) + P_{M^\perp}(x) = x$ para todo $x \in H$.

Antes de estudiar la principal consecuencia del teorema anterior, generalizamos su última afirmación, sustituyendo el subespacio cerrado $M$ por un conjunto no vacío arbitrario:

- *Si $E$ es un subconjunto no vacío de un espacio de Hilbert $H$, se tiene que $E^{\perp\perp} = \overline{\operatorname{Lin} E}$. En particular, si $Y$ es un subespacio de $H$, se tiene $\overline{Y} = Y^{\perp\perp}$, luego $Y$ es denso en $H$ si, y sólo si, $Y^\perp = \{0\}$.*

Tomando $M = \overline{\operatorname{Lin} E}$, tenemos $M \subset E^{\perp\perp}$. Ahora bien, de $E \subset M$ se deduce $M^\perp \subset E^\perp$, luego $E^{\perp\perp} \subset M^{\perp\perp}$, pero como $M$ es un subespacio cerrado de $H$, el teorema anterior nos dice que $M^{\perp\perp} = M$, luego $E^{\perp\perp} \subset M$. Por tanto, $E^{\perp\perp} = M$, como se quería.

Si $Y$ es subespacio de $H$, se tiene obviamente $Y^{\perp\perp} = \overline{\operatorname{Lin} Y} = \overline{Y}$. Por tanto, si $Y^\perp = \{0\}$ tendremos $\overline{Y} = \{0\}^\perp = H$, luego $Y$ es denso en $H$. Recíprocamente, si $Y$ es denso en $H$, tenemos $Y^{\perp\perp} = H$, de donde $Y^\perp = Y^\perp \cap Y^{\perp\perp} = \{0\}$. ■

El teorema de la proyección ortogonal se debe a la escuela de Hilbert, aparece ya en un trabajo de E. Schmidt, su más directo colaborador, publicado en 1908. Como una consecuencia muy relevante, F. Riesz y M. Fréchet probaron poco después la autodualidad de los espacios de Hilbert, resultado que ahora vamos a estudiar.

Recordemos que todos los espacios de Hilbert, que hemos presentado como ejemplos, se identifican con su espacio dual. Para $N \in \mathbb{N}$ sabemos que $\left(l_2^N\right)^*$ es isométricamente isomorfo a $l_2^N$, y análogamente, $l_2^*$ se identifica con $l_2$. Aunque no dimos una demostración completa, también sabemos que, si $\Omega$ es un abierto no vacío de $\mathbb{R}^N$ con $N \in \mathbb{N}$, entonces el espacio de Lebesgue $L_2(\Omega)$ es isométricamente isomorfo a su espacio dual $L_2(\Omega)^*$. Pues bien, estos resultados no son más que casos particulares de un teorema general que permite en cierto modo identificar todo espacio de Hilbert $H$ con su dual $H^*$. Veamos paso a paso cómo se consigue dicha identificación.

Si $X$ es un espacio pre-hilbertiano, fijado $y \in X$, podemos definir

$$\widetilde{y} : X \to \mathbb{K}, \qquad \widetilde{y}(x) = \big(x \,\big|\, y\big) \quad \forall x \in X$$

La linealidad y continuidad del producto escalar en la primera variable, significan que $\widetilde{y} \in X^*$, y vamos a calcular fácilmente la norma dual de $\widetilde{y}$. La desigualdad de Cauchy-Schwartz nos dice que $\big|\widetilde{y}(x)\big| \leqslant \|x\|\,\|y\|$ para todo $x \in X$, es decir, $\|\widetilde{y}\| \leqslant \|y\|$, pero de hecho tenemos la igualdad, ya que $\|y\|^2 = \widetilde{y}(y) \leqslant \|\widetilde{y}\|\,\|y\|$. Consideramos ahora la aplicación

$$\Psi : X \to X^*, \qquad \Psi(y) = \widetilde{y} \quad \forall y \in X$$

Como el producto escalar es conjugado-lineal en la segunda variable, vemos que $\Psi$ también es conjugado-lineal, es decir, para $y, z \in X$ y $\lambda \in \mathbb{K}$, se tiene que $\Psi(\lambda y + z) = \overline{\lambda}\Psi(y) + \Psi(z)$. Claro está que, en el caso real, $\Psi$ es lineal.

En cualquier caso, sabemos que $\Psi$ preserva la norma, luego es isométrica: para $y, z \in X$, se tiene que $\|\Psi(y) - \Psi(z)\| = \|\Psi(y-z)\| = \|y-z\|$. Para que $\Psi$ sea sobreyectiva, $X$ ha de ser completo, puesto que $X^*$ lo es. Pero lo más interesante es el recíproco.

**Teorema de Riesz-Fréchet**. *Si $H$ un espacio de Hilbert y $f \in H^*$, existe $y \in H$ tal que*

$$f(x) = \big(x \,|\, y\big) \qquad \forall x \in H$$

*Por tanto, si para $y \in H$ escribimos $\widetilde{y}(x) = (x \,|\, y)$ para todo $x \in H$, definiendo $\Psi(y) = \widetilde{y}$ para todo $y \in H$, se tiene que $\Psi$ es una biyección conjugado-lineal e isométrica de $H$ sobre $H^*$.*

**Demostración.** Sólo hay que comprobar la primera afirmación, la sobreyectividad de $\Psi$. Dado $f \in H^* \setminus \{0\}$, tenemos que $\ker f$ es un subespacio cerrado de $H$, al que podemos aplicar el teorema de la proyección ortogonal. Como $\ker f \neq X$, existe $y_0 \in (\ker f)^\perp$ tal que $\|y_0\| = 1$. Para $x \in H$ se tiene que $f(x)\,y_0 - f(y_0)\,x \in \ker f$, de donde

$$0 = \big(f(x)\,y_0 - f(y_0)\,x \,\big|\, y_0\big) = f(x) - f(y_0)\,\big(x \,\big|\, y_0\big) = f(x) - \big(x \,\big|\, \overline{f(y_0)}\,y_0\big) \qquad \forall x \in H$$

y tomando $y = \overline{f(y_0)}\,y_0$, se tiene $f(x) = (x \,|\, y)$ para todo $x \in H$, como se quería. ■

Así pues, todo espacio de Hilbert real $H$ se identifica con su espacio dual $H^*$, pues $\Psi$ es un isomorfismo isométrico. Cuando $H$ es complejo, $\Psi$ no es lineal sino conjugado-lineal, pero veremos más adelante que existe una biyección conjugado-lineal e isométrica de $H$ sobre sí mismo, cuya composición con $\Psi$ nos da un isomorfismo isométrico de $H$ sobre $H^*$. Esto es fácil de comprobar en los casos particulares que conocemos bien. Si $H = l_2^N$ con $N \in \mathbb{N}$, para todo $y \in l_2^N$, definimos $\Phi(y) = \Psi(\overline{y})$, donde $\overline{y}(k) = \overline{y(k)}$ para todo $k \in \{1, 2, \ldots, N\}$. Análogamente, para $y \in l_2$ se define $\Phi(y) = \Psi(\overline{y})$ donde $\overline{y} \in l_2$ viene dado por $\overline{y}(n) = \overline{y(n)}$ para todo $n \in \mathbb{N}$. Por último, si $g \in H = L_2(\Omega)$ donde $\Omega$ es un abierto no vacío de $\mathbb{R}^N$, para toda $g \in L_2(\Omega)$ se toma $\Phi(g) = \Psi(\overline{g})$ donde $\overline{g} \in L_2(\Omega)$ verifica que $\overline{g}(t) = \overline{g(t)}$ p.c.t. $t \in \Omega$. En los tres casos vemos claramente que $\Phi$ es un isomorfismo isométrico de $H$ sobre $H^*$, que fue exactamente el que usamos en su momento para identificar ambos espacios.

Es importante resaltar que, en el caso $H = L_2(\Omega)$, hemos probado que tanto $\Psi$ en el caso real, como $\Phi$ en el caso complejo, son aplicaciones sobreyectivas, gracias al teorema anterior, sin usar ningún resultado de Teoría de la Medida.

Comentemos finalmente la consecuencia más clara del teorema de la proyección ortogonal. Si $M$ es un subespacio cerrado de un espacio de Hilbert $H$, entonces $M^\perp$ es un complemento topológico canónico de $M$ y, consecuentemente el espacio cociente $H/M$ carece de interés. En particular, en cualquier espacio de Hilbert $H$, todo subespacio cerrado está complementado, y es claro que esta propiedad de $H$ se conserva por isomorfismos:

- *Si $X$ es un espacio de Banach isomorfo a un espacio de Hilbert, todo subespacio cerrado de $X$ está complementado en $X$.*

En 1971, J. Lindenstrauss y L. Tzafriri probaron el recíproco del resultado anterior: *si todo subespacio cerrado de un espacio de Banach $X$ está complementado en $X$, entonces $X$ es isomorfo a un espacio de Hilbert.* Resolvieron así un problema que había permanecido abierto durante varias décadas, conocido como el *problema del subespacio complementado.*