Tema 11

Espacios de Hilbert

Este último tema está dedicado a 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 cada espacio de Hilbert con su dual.

Además, del teorema de la proyección ortogonal se deduce que todo subespacio cerrado de un espacio de Hilbert está complementado, una propiedad que obviamente se conserva por isomorfismos. Hoy día se sabe que, recíprocamente, un espacio de Banach en el que todos los subespacios cerrados estén complementados, ha de ser isomorfo a un espacio de Hilbert.

11.1. 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 cada $y \in X$ 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 cada $x \in X$ 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 λ. 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 φ 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 φ es una forma bilineal simétrica.

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

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

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

y por tanto, φ 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

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

En el caso real, está ya bien claro que φ queda determinada por Q.

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

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

luego usando (1) deducimos que

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

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 φ 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.

11.2. 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 φ una forma sexquilineal hermítica en X, cuya forma cuadrática asociada Q es positiva. Entonces:*

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

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

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

o lo que es lo mismo,

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

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

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

Pero si $Q(y) = 0$, se ha de tener $\operatorname{Re} \varphi\left(x, y\right) = 0$, pues en otro caso podríamos tomar $t \in \mathbb{R}$ de forma que $2\,t \operatorname{Re} \varphi\left(x, y\right) > Q(x)$, 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 $\left|\varphi\left(x, y\right)\right| = \alpha\, \varphi\left(x, y\right) = \varphi\left(\alpha x, y\right)$. Es claro entonces que $\varphi\left(\alpha x, y\right) \in \mathbb{R}_0^+$, así como que $Q(\alpha x) = Q(x)$, luego usando (4) con αx en lugar de x, obtenemos

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

y 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

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

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 φ en X, cuya forma cuadrática asociada es definida positiva. Habitualmente, un producto escalar se denota por $(x,y) \mapsto (x|y)$, para $x,y \in X$, y se dice que $(x|y)$ 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 $(x,y) \mapsto (x|y)$, de $X \times X \to \mathbb{K}$, es un producto escalar en X si, y sólo si, verifica las tres condiciones siguientes:

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

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

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

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

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

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:

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

En el caso complejo, basta observar que $\operatorname{Im} (x|y) = \operatorname{Re} (x|iy)$ 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, $\big(T(x)\big|T(z)\big) = (x|z)$ 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

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

que nos permite obtener la siguiente consecuencia.

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

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

$$(x|y) - (u|v) = (x-u|y) + (u|y-v)$$

y la desigualdad de Cauchy-Schwartz nos dice que

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

lo que implica claramente que el producto escalar es continuo en el punto (u,v). ■

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:*

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

(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

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

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:

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

Entonces, $\|x+y\| = \|x\| + \|y\|$ equivale a que se tenga $\mathrm{Re}(x|y) = |(x|y)| = \|x\|\,\|y\|$. 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, basta recordar el producto escalar usual en $\mathbb{R}^N$, cuya versión en 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:*

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

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:

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

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*

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

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. Para $f,g \in L_2$, la desigualdad integral de Hölder nos dice que $f\,\overline{g} \in L_1$, y esto permite claramente definir un producto escalar en L_2 que da lugar a su norma:

- *El espacio de Lebesgue L_2 es un espacio de Hilbert, con el producto escalar dado por*

$$\big(f\,\big|\,g\big) = \int_0^1 f(t)\,\overline{g(t)}\,dt \qquad \forall f,g \in L_2$$

Nótese que la desigualdad de Cauchy-Schwartz en el espacio de Hilbert L_2 no es más que la desigualdad integral de Hölder para $p = 2$. Puesto que $C[0,1]$ puede verse como subespacio denso en L_2, que no es cerrado, tenemos que $C[0,1]$, no con su norma natural que es la del máximo, sino con la inducida por L_2, es otro ejemplo de espacio pre-hilbertiano no completo.

11.3. 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

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

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

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

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 φ, mientras que en el caso complejo, φ 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

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

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 φ 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 φ 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 φ es definida positiva, y φ 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 ψ ha de ser el producto escalar que buscamos. Las propiedades que ψ debe verificar, se deducirán de las de φ, 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 ψ 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 ψ es una forma sexquilineal hermítica. Finalmente, para $x \in X$, usando (10) y el hecho de que φ 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 ψ es definida positiva, luego ψ es el producto escalar que buscábamos. ■

El teorema anterior tiene varias consecuencias inmediatas que conviene destacar. Para la primera de ellas, dado un espacio normado X, usaremos el espacio real subyacente $X_{\mathbb{R}}$, que obviamente es también un espacio normado, 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 φ definida en (6), mientras que el de X es la aplicación ψ, 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. Equivalentemente, la completación de un espacio pre-hilbertiano es un espacio de Hilbert.*

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 $\|e_1 \pm e_2\|_p = 2^{1/p}$ 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_∞^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_∞^2 es isométricamente isomorfo a un subespacio de c_0, luego c_0 no es un espacio de Hilbert, y por tanto l_∞ 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.*

En cuanto a los espacios de Lebesgue, para $1 \leqslant p \leqslant \infty$ sabemos que l_p es isométricamente isomorfo a un subespacio de L_p. Por tanto:

- *Para $1 \leqslant p \leqslant \infty$ se tiene que L_p es un espacio de Hilbert si, y sólo si, $p = 2$.*

Si Γ es un conjunto no numerable, está claro que el espacio de funciones acotadas $l_\infty(\Gamma)$ contiene un subespacio isométricamente isomorfo a l_∞^2, por lo que $l_\infty(\Gamma)$ no es un espacio de Hilbert, como ocurría cuando Γ es numerable.

Comentemos finalmente que el espacio de Banach $C[0,1]$, con la norma del máximo, no es un espacio de Hilbert. Para verlo, basta considerar por ejemplo las funciones $f, g \in C[0,1]$ definidas por $f(t) = 1$ y $g(t) = t$ para todo $t \in [0,1]$. Se tiene claramente que

$$\| f+g \|^2 + \| f-g \|^2 = 5 \neq 4 = 2 \| f \|^2 + 2 \| g \|^2$$

luego $C[0,1]$ no verifica la identidad del paralelogramo. De hecho, no es difícil comprobar que, si K es un espacio topológico compacto y de Hausdorff, que no se reduce a un punto, entonces el espacio de Banach $C(K)$ no es un espacio de Hilbert.

11.4. 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. A partir de ahora trabajaremos siempre en un espacio de Hilbert, con lo que casi nunca se pierde generalidad, pues en el caso de un espacio pre-hilbertiano no completo, siempre podemos considerar su completación.

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

Demostración. Fijado $x \in H$, para $y, z \in M$, la identidad del paralelogramo nos dice que

$$\| z-y \|^2 = \| (x-y) - (x-z) \|^2 = 2 \| x-y \|^2 + 2 \| x-z \|^2 - 4 \left\| x - \frac{y+z}{2} \right\|^2$$

Por ser M convexo, tenemos $(y+z)/2 \in M$, y de la igualdad anterior deducimos que

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

desigualdad que será la clave de la demostración.

Fijamos una sucesión $\{y_n\}$ de puntos de M tal que $\{\|x-y_n\|\} \to d(x,M)$. Dado $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que, para $k \geqslant n_0$ se tiene $\|x-y_k\|^2 < d(x,M)^2 + (\varepsilon^2/4)$. Para $n,m \geqslant n_0$, podemos entonces usar (11), con $y = y_m$ y $z = y_n$, para obtener que

$$\|y_n - y_m\|^2 \leqslant 2\|x-y_m\|^2 + 2\|x-y_n\|^2 - 4\,d(x,M)^2 < \varepsilon^2$$

luego $\{y_n\}$ es una sucesión de Cauchy. Como H es completo y M cerrado, tenemos por tanto que $\{y_n\} \to y \in M$. La continuidad de la norma de H nos dice entonces que

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

de modo que y es una mejor aproximación de x en M. Es la única, pues si $z \in M$ también verifica que $\|x-z\| = d(x,M)$, la desigualdad (11) 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 M, es decir, para que M sea proximinal en H, es necesario que M sea cerrado, luego esta hipótesis no se puede suprimir.

Más importante es observar que la convexidad de M tampoco se puede suprimir. Así pues, a diferencia de lo que ocurría en los espacios normados de dimensión finita, un subconjunto cerrado A de un espacio de Hilbert H puede no ser proximinal en H, como veremos con el siguiente ejemplo.

- *Sea $\{e_n\}$ la sucesión de vectores unidad en l_2, y sea $y_n = (n+1)\,e_n/n$ para todo $n \in \mathbb{N}$. Entonces el conjunto $A = \{y_n : n \in \mathbb{N}\}$ es cerrado, pero no es proximinal en l_2.*

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 toda sucesión convergente de puntos de A es constante a partir de un término en adelante, luego converge a un punto de A.

Por otra parte, comprobamos que 0 no tiene ninguna mejor aproximación en A, pues basta observar que $d(0,A) = \inf\{\|y_n\| : n \in \mathbb{N}\} = 1$, pero $\|y_n\| > 1$ para todo $n \in \mathbb{N}$. ■

En resumen, para asegurar la existencia de mejores aproximaciones, no podemos suprimir ninguna de las hipótesis del teorema anterior. 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 M es convexo, luego no se precisa que M sea cerrado. Por tanto, si M es un subconjunto no vacío y convexo de un espacio de Hilbert H, entonces cada $x \in H$ tiene a lo sumo una mejor aproximación en M. Finalmente está bien claro que la convexidad de M no se puede suprimir, pues basta pensar lo que ocurre cuando M tiene exactamente dos puntos.

11.5. 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 de Hilbert H. Entonces $y_0 \in M$ es una mejor aproximación de un $x \in H$ 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 H. 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 de Hilbert H. Para $x \in H$ e $y_0 \in M$, se tiene:*

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

Si M es de hecho un subespacio de H, se tiene también:

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

Para interpretar geométricamente estas equivalencias, usamos una nomenclatura, motivada por la segunda, que a partir de ahora será muy conveniente. Si H es un espacio de Hilbert, para $x,y \in H$ 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 H, definimos

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

Por la linealidad y continuidad del producto escalar en la primera variable, vemos que $E^{\perp}$ es un subespacio cerrado de H. Se dice que $E^{\perp}$ es el **subespacio ortogonal** a E. La notación es la misma que usábamos para el anulador de un subconjunto de un espacio normado X, que era un subespacio cerrado del espacio dual X^*. Pronto veremos que es perfectamente coherente emplear la misma notación en ambas situaciones.

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 H, luego $\overline{\operatorname{Lin} E} \subset E^{\perp\perp}$.

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 H es un espacio de Hilbert, para $x,y \in H$ se tiene

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

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

- *Si H es un espacio de Hilbert, para $x,y \in H$ se tiene:*
  (i) *En caso real:* $\quad x \perp y \iff \big\|x+y\big\|^2 = \big\|x\big\|^2 + \big\|y\big\|^2$
  (ii) *En caso complejo:* $\quad x \perp y \iff \big\|x+y\big\|^2 = \big\|x+iy\big\|^2 = \big\|x\big\|^2 + \big\|y\big\|^2$

La caracterización de las mejores aproximaciones tiene ya una interpretación geométrica muy intuitiva. Si M es un subespacio del espacio de Hilbert H, vimos que $y_0 \in M$ es la mejor aproximación de $x \in H$ en M si, y sólo si, $x - y_0 \in M^{\perp}$. Esto equivale a que $x - y_0$ sea perpendicular al subespacio M, o que el punto y_0 sea el pie de la perpendicular a M que pasa por el punto x, lo que está totalmente de acuerdo con la intuición geométrica.

Para un conjunto convexo $M \subset H$, la interpretación es también intuitiva, aunque no tan sencilla. La hacemos en caso real, pues si H es complejo, basta pensar que M es un subconjunto convexo de $H_{\mathbb{R}}$. Fijado $x \in H$, vimos que $y_0 \in M$ es la mejor aproximación de x en M si, y sólo si, $\big(x - y_0\,\big|\,y - y_0\big) \leqslant 0$ para todo $y \in M$. Tomando $\alpha_0 = \big(x - y_0\,\big|\,y_0\big) \in \mathbb{R}$, esto equivale a $\big(x - y_0\,\big|\,y\big) \leqslant \alpha_0$ para todo $y \in M$. Como $\big(x - y_0\,\big|\,x - y_0\big) > 0$, tenemos $\big(x - y_0\,\big|\,x\big) \geqslant \alpha_0$. Por tanto, el conjunto $Z = \big\{ z \in H : \big(x - y_0\,\big|\,z\big) = \alpha_0 \big\}$ es un *hiperplano afín*, que pasa por el punto y_0, ya que $y_0 \in Z$, deja el punto x a un lado, y el conjunto M al otro. En particular, el funcional lineal y continuo $z \mapsto \big(x - y_0\,\big|\,z\big)$, o si se quiere el hiperplano Z, separa x de M.

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.

11.6. 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 H 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 $x \in H$, 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 L_2^* se identifica con L_2. 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

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

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 Ψ 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, Ψ es lineal.

En cualquier caso, sabemos que Ψ 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 Ψ 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 Ψ es una biyección conjugado-lineal e isométrica de H sobre H^*.

Demostración. Sólo hay que comprobar la primera afirmación: la sobreyectividad de Ψ. 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 $\ker f \neq H$, existe $u \in (\ker f)^{\perp}$ tal que $\|u\| = 1$. Para $x \in H$ se tiene que $f(x)u - f(u)x \in \ker f$, de donde

$$0 = \big(f(x)u - f(u)x \,\big|\, u\big) = f(x)\|u\|^2 - f(u)\big(x \,\big|\, u\big) = f(x) - \big(x \,\big|\, \overline{f(u)}\,u\big)$$

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

Vemos que todo espacio de Hilbert real H se identifica con su espacio dual H^*, pues Ψ es un isomorfismo isométrico de H sobre H^*. En cambio, cuando H es complejo, Ψ no es lineal, sino conjugado-lineal. Aunque no vamos a hacerlo, se puede probar que, también en caso complejo, existe un isomorfismo isométrico Φ de H sobre H^*. Sin embargo, mientras que la aplicación Ψ es canónica, porque su definición sólo involucra el producto escalar de H, el isomorfismo Φ no puede serlo.

La existencia del isomorfismo isométrico Φ es fácil de comprobar en los casos particulares que conocemos. En el caso $H = l_2^N$ con $N \in \mathbb{N}$, para todo $y \in l_2^N$ se define $\Phi(y) = \Psi(\overline{y})$, donde $\overline{y}(k) = \overline{y(k)}$ para todo $k \in \{1, 2, \ldots, N\}$. Análogamente, cuando $H = l_2$, 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, en el caso $H = L_2$, para toda $g \in L_2$ se define $\Phi(g) = \Psi(\overline{g})$ donde $\overline{g}(t) = \overline{g(t)}$ p.c.t. $t \in [0,1]$. En los tres casos vemos claramente que Φ es un isomorfismo isométrico de H sobre H^*, que fue exactamente el que usamos en su momento para identificar estos dos espacios.

Conviene resaltar que, en el caso $H = L_2$, hemos probado que Ψ y Φ son sobreyectivas, gracias al teorema anterior, sin usar ningún resultado de Teoría de la Medida.

Dado un subconjunto no vacío E de un espacio de Hilbert H, aclaremos ahora la cuestión de notación planteada al definir el subespacio ortogonal $E^{\perp}$. Cuando identificamos H con H^* mediante Ψ, vemos que $E^{\perp}$ se identifica con el subespacio cerrado de H^* dado por

$$\Psi\big(E^{\perp}\big) = \big\{\widetilde{y} : y \in E^{\perp}\big\} = \big\{f \in H^* : f(x) = 0 \ \ \forall x \in E\big\}$$

que es precisamente el anulador de E, tal y como se definió, para subconjuntos de cualquier espacio normado, en el estudio de la dualidad. Así pues, al identificar H con H^* hacemos coincidir las dos definiciones de $E^{\perp}$, luego al usar la misma notación para ambas nociones, no hay peligro de confusión.

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 un espacio de Banach X es 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.*