TEMA 2º: ESPACIOS DE HILBERT

1. Espacios prehilbertianos o espacios dotados de un producto escalar o interior.

DEFINICION: Sea E un espacio vectorial sobre $\mathbb{K}$ ($\mathbb{R}$ o $\mathbb{C}$). Un producto escalar o interior en E es una aplicación

$$(\ldots|\ldots): (x,y)\in E\times E \longmapsto (x|y)\in \mathbb{K}$$

que verifica los axiomas siguientes:

i) $\forall \lambda,\mu\in\mathbb{K}, \forall x,y,z\in E,\ (\lambda x+\mu y|z)=\lambda(x|z)+\mu(y|z)$

ii) $\forall x,y\in E,\ (x|y)=\overline{(y|x)}$ $\quad (*)$

iii) $\forall x\in E-\{0\},\ (x|x)>0$

Observar que $\forall x\in E-\{0\}$, $(x|x)\in\mathbb{R}$, pues por ii), $(x|x)=\overline{(x|x)}$, y por ello podemos escribir $(x|x)>0$.

1.1. PROPOSICION: Si $(\ldots|\ldots)$ es un producto escalar en E, se verifica

iv) $\forall x\in E,\ (x|0)=(0|x)=0$.

v) $\forall\lambda,\mu\in\mathbb{K}, \forall x,y,z\in E,\ (x|\lambda y+\mu z)=\bar\lambda(x|y)+\bar\mu(x|z)$ $\quad (**)$

vi) Desigualdad de Cauchy-Schwartz: $\forall x,y\in E,\ |(x|y)|\le (x|x)^{1/2}(y|y)^{1/2}$

vii) Desigualdad de Minkowsky: $\forall x,y\in E,\ (x+y|x+y)^{1/2}\le (x|x)^{1/2}+(y|y)^{1/2}$

Demostr: iv) $(0|x)=(x-x|x)=(x|x)-(x|x)=0$. Que $(x|0)=0$ se deduce de esto y de ii).

v) $(x|\lambda y+\mu z)=\overline{(\lambda y+\mu z|x)}=\overline{\lambda(y|x)+\mu(z|x)}=$
$=\bar\lambda\,\overline{(y|x)}+\bar\mu\,\overline{(z|x)}=\bar\lambda(x|y)+\bar\mu(x|z)$

vi) $\forall x,y\in E, \forall\lambda\in\mathbb{K},\ 0\le(x+\lambda y|x+\lambda y)=(x|x)+(x|\lambda y)+$
$+(\lambda y|x)+(\lambda y|\lambda y)=(x|x)+\bar\lambda(x|y)+\lambda(y|x)+\lambda\bar\lambda(y|y)=$
$=(x|x)+\bar\lambda(x|y)+\lambda\overline{(x|y)}+|\lambda|^2(y|y)$

ya que $\lambda\bar\lambda=|\lambda|^2$. Si, en particular, hacemos $\lambda=t(x|y)$, $t\in\mathbb{R}$ tenemos que:

$$0\le(x|x)+t\overline{(x|y)}(x|y)+t(x|y)\overline{(x|y)}+t^2|(x|y)|^2(y|y)=$$
$$=(x|x)+2|(x|y)|^2t+|(x|y)|^2(y|y)t^2$$

y esto cualquiera que sea $t\in\mathbb{R}$. Pero esta desigualdad se verifica si el discriminante del trinomio es menor o igual que cero, es decir si:

$$4|(x|y)|^4-4|(x|y)|^2(y|y)(x|x)\le 0$$

lo cual es equivalente a lo que queremos probar (si fuese $(x|y)=0$

La desigualdad es trivial).

vii) $(x+y|x+y) = (x|x) + (x|y) + (y|x) + (y|y) =$
$= (x|x) + (x|y) + \overline{(x|y)} + (y|y) =$
$= (x|x) + 2\,\mathrm{Re}\,(x|y) + (y|y) \leq$
$\leq (x|x) + 2\,|(x|y)| + (y|y) \overset{(1)}{\leq} (x|x) + 2\,(x|x)^{1/2}(y|y)^{1/2} + (y|y) =$
$= [(x|x)^{1/2} + (y|y)^{1/2}]^2$

La desigualdad (1) es cierta por la desigualdad de Cauchy-Schwartz.
Luego, $\forall x,y \in E$, $(x+y|x+y)^{1/2} \leq (x|x)^{1/2} + (y|y)^{1/2}$. csqd.

DEFINICIONES: 1) Las aplicaciones $(\cdot|\cdot)$ que verifican i) y v) se llaman sesquilineales (bilineales en el caso $\mathbb{K}=\mathbb{R}$).
2) Una aplicación sesquilineal que verifica ii) se llama hermitiana.
3) Una aplicación que verifique iii) se dice que es definida positiva.
Por tanto, un producto escalar en E es una aplicación hermitiana de $E \times E$ en $\mathbb{K}$ definida positiva.

1.2. TEOREMA: (de Jordan-Von Neumann).

Si la aplicación $(\ldots|\ldots): (x,y) \in E\times E \longmapsto (x|y) \in \mathbb{K}$ es un producto escalar en E, entonces la aplicación $\|\ldots\|: x \in E \longmapsto (x|x)^{1/2} \in \mathbb{R}$ es una norma en E que verifica la "igualdad del paralelogramo":
$$\forall x,y \in E,\quad \|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2)$$
Recíprocamente, si $\|\cdot\|$ es una norma en E que verifica la igualdad del paralelogramo, dicha norma "está inducida" por el producto escalar siguiente:
si $\mathbb{K}=\mathbb{R}$, $(x|y) = \frac{1}{4}[\|x+y\|^2 - \|x-y\|^2] \in \mathbb{R}$
si $\mathbb{K}=\mathbb{C}$, $(x|y) = \frac{1}{4}[\|x+y\|^2 - \|x-y\|^2 + i(\|x+iy\|^2 - \|x-iy\|^2)] \in \mathbb{C}$

Demostr: 1ª PARTE: Probemos que $\|\cdot\|$ verifica los axiomas de las normas.
N1) $(x|x)^{1/2} = 0 \iff (x|x) = 0 \iff x = 0$, en virtud de iii) y iv)
N2) $\|x+y\| = (x+y|x+y)^{1/2} \leq (x|x)^{1/2} + (y|y)^{1/2} = \|x\| + \|y\|$, por la desigualdad de Minkowsky.
N.3) $\|\lambda x\| = (\lambda x|\lambda x)^{1/2} = [\lambda\bar{\lambda}(x|x)]^{1/2} = [|\lambda|^2(x|x)]^{1/2} = |\lambda|\cdot\|x\|$.
Probemos que $\|\cdot\|$ verifica la igualdad del paralelogramo.
$\forall x,y \in E$, $\|x+y\|^2 + \|x-y\|^2 = (x+y|x+y) + (x-y|x-y) =$
$= (x|x) + (x|y) + (y|x) + (y|y) + (x|x) - (x|y) - (y|x) + (y|y) =$
$= 2[(x|x) + (y|y)] = 2(\|x\|^2 + \|y\|^2)$.

2ª PARTE (Recíproco). Lo probaremos solamente para el caso $\mathbb{K}=\mathbb{R}$ (*)
Definiendo $(x|y) = \frac{1}{4}[\|x+y\|^2 - \|x-y\|^2]$, es claro que

cida por el producto escalar $(\cdot|\cdot)$, si es que probamos que, efectivamente, $(\cdot|\cdot)$ es un producto escalar en E. Veamos que $(\cdot|\cdot)$ verifica los axiomas de producto escalar para $\mathbb{K}=\mathbb{R}$.

ii) $\forall x,y\in E$, $(x|y)=\frac{1}{4}\left[\|x+y\|^2-\|x-y\|^2\right]=\frac{1}{4}\left[\|y+x\|^2-\|y-x\|^2\right]=$
$=(y|x)$.

iii) $\forall x\in E-\{0\}$, $(x|x)=\|x\|^2>0$

i) Probemos que $\forall x,y,z\in E$, $(x+y|z)=(x|z)+(y|z)$.

Consideremos la función

$$\phi : E^3 \longrightarrow \mathbb{R}$$
$$(x,y,z)\longmapsto \phi(x,y,z)=4\left[(x+y|z)-(x|z)-(y|z)\right]$$

Si probamos que ϕ es la función nula sobre E^3 quedará probado que $(x+y|z)=(x|z)+(y|z)$. Por definición de $(\cdot|\cdot)$,

$$\phi(x,y,z)=4\left[(x+y|z)-(x|z)-(y|z)\right]=$$
$$=\|x+y+z\|^2-\|x+y-z\|^2-\|x+z\|^2+\|x-z\|^2-\|y+z\|^2+\|y-z\|^2 \qquad (\mathrm{I})$$

En virtud de la ley del paralelogramo tenemos:

$$\|x+z+y\|^2=2\|x+z\|^2+2\|y\|^2-\|x+z-y\|^2$$
$$\|x-z+y\|^2=2\|x-z\|^2+2\|y\|^2-\|x-z-y\|^2$$

Sustituyendo en (I) estas expresiones obtenemos:

$$\phi(x,y,z)=-\|x+z-y\|^2+\|x-z-y\|^2+\|x+z\|^2-\|x-z\|^2-\|y+z\|^2+$$
$$+\|y-z\|^2 \qquad (\mathrm{II})$$

Sumando miembro a miembro las expresiones (I) y (II) y dividiendo por dos, se tiene que:

$$\phi(x,y,z)=\frac{1}{2}\left[\|y+z+x\|^2+\|y+z-x\|^2\right]-\frac{1}{2}\left[\|y-z+x\|^2+\|y-z-x\|^2\right]-$$
$$-\|y+z\|^2+\|y-z\|^2. \qquad (\mathrm{III})$$

Por la ley del paralelogramo: $\frac{1}{2}\left[\|y+z+x\|^2+\|y+z-x\|^2\right]=$
$=\|y+z\|^2+\|x\|^2$, $-\frac{1}{2}\left[\|y-z+x\|^2+\|y-z-x\|^2\right]=-\|y-z\|^2-\|x\|^2$.

Luego, sustituyendo en (III), deducimos que $\phi(x,y,z)=0$ y, en consecuencia, $\forall x,y,z\in E$, $(x+y|z)=(x|z)+(y|z)$.

Probemos ahora que $\forall\lambda\in\mathbb{R}$, $\forall x,y\in E$, $(\lambda x|y)=\lambda(x|y)$

* La propiedad es cierta para todo natural n:
  - Es cierta, para $n=2$ pues $(2x|y)=(x+x|y)=(x|y)+(x|y)=2(x|y)$.
  - Supuesto que es cierta para $K\in\mathbb{N}$, $((K+1)x|y)=(Kx+x|y)=(Kx|y)+(x|y)=$
  $=K(x|y)+(x|y)=(K+1)(x|y)$.
* La propiedad también es cierta para todo entero negativo:
  - Es cierta para -1: $(-x|y)=\frac{1}{4}\left[\|-x+y\|^2-\|-x-y\|^2\right]=-\frac{1}{4}\left[\|x+y\|^2-\|x-y\|^2\right]=-(x|y)$.

- Es cierta para todo $m \in \mathbb{Z}^-$: $(mx|y) = (-(-m)x|y) = -(-mx|y) = -(-m)(x|y) = m(x|y)$, pues $-m \in \mathbb{N}$.

* La propiedad es cierta para todo racional de la forma $\frac{1}{n}$, $n \in \mathbb{N}$:
$\forall n \in \mathbb{N}$, $(x|y) = (n\frac{1}{n}x|y) = n(\frac{1}{n}x|y) \Rightarrow \frac{1}{n}(x|y) = (\frac{1}{n}x|y)$.

* La propiedad es cierta para todo número racional, que es de la forma $\frac{m}{n}$, $m \in \mathbb{Z}$ y $n \in \mathbb{N}$:

$$\left(\frac{m}{n}x|y\right) = m\left(\frac{1}{n}x|y\right) = \frac{m}{n}(x|y).$$

Dados $x, y \in E$, las aplicaciones $\lambda \in \mathbb{R} \longmapsto (\lambda x|y) \in \mathbb{R}$ y $\lambda \in \mathbb{R} \longmapsto \lambda(x|y) \in \mathbb{R}$ son continuas (según la definición de $(x|y)$). Además, coinciden sobre $\mathbb{Q}$, que es denso en $\mathbb{R}$. Luego, también coinciden en $\mathbb{R}$ (Topología I: Principio de prolongación de identidades, Tema 7º), como queríamos probar ■

<u>DEFINICION</u>: Llamamos espacios prehilbertianos, o espacios pH, a los espacios vectoriales sobre $\mathbb{K}$ ($\mathbb{R}$ o $\mathbb{C}$) dotados de un producto escalar.

Llamamos espacios de Hilbert a los espacios prehilbertianos completos.

<u>EJEMPLOS DE ESPACIOS PREHILBERTIANOS</u>:

① Los productos escalares en $\mathbb{K}$ son las aplicaciones de la forma

$$(x,y) \in \mathbb{K}\times\mathbb{K} \longmapsto \lambda x \bar{y} \in \mathbb{K}, \quad \lambda \in \mathbb{R}^+$$

pues los funcionales bilineales son de esta forma.

② Los productos escalares en $\mathbb{R}^2$ son las aplicaciones de la forma

$$[(x_1,x_2),(y_1,y_2)] \in \mathbb{R}^2\times\mathbb{R}^2 \longmapsto (x_1,x_2)\begin{pmatrix} a & b \\ b & c \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbb{R}$$

con $a>0$ y $ac-b^2>0$, pues los funcionales bilineales sobre $\mathbb{R}^2$ son de la forma

$$f[(x_1,x_2),(y_1,y_2)] = (x_1,x_2)\begin{pmatrix} a & b \\ d & c \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

Los que son simétricos son aquellos para los que $b=d$, y los definidos positivos son aquellos para los que $a>0$ y $ac-b^2>0$.

③ Los productos escalares en $\mathbb{R}^n$ son las aplicaciones de la forma

$$[(x_1,\dots,x_n),(y_1,\dots,y_n)] \in \mathbb{R}^n\times\mathbb{R}^n \longmapsto (x_1 \dots x_n)\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

tales que $a_{ij} = a_{ji}$ y $\det \Delta_j > 0$ siendo $\Delta_j = \begin{pmatrix} a_{11} & \cdots & a_{1j} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jj} \end{pmatrix}$, $(j=1,2,\dots,n)$.

En el caso $n=2$ (ejemplo ②), la esfera unidad es $S = \{(x_1,x_2) / \|(x_1,x_2)\| = 1\}$:

$\|(x_1,x_2)\| = \left[(x_1,x_2)\begin{pmatrix} a & b \\ b & c \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right]^{1/2} = 1 \iff ax_1^2 + 2bx_1x_2 + cx_2^2 = 1$, que es una elip-
se centrada en $(0,0)$. En general, los espacios $\mathbb{R}^n$ dotados de un producto

escalar son aquellos espacios $\mathbb{R}^n$, y solo aquellos, cuya esfera unidad es un hiperelipsoide.

④ Sea ℓ^2 el espacio vectorial de las sucesiones (x_n) de $\mathbb{K}$ de cuadrado de módulo sumable, es decir, tales que $\sum_{n=1}^{\infty} |x_n|^2 < +\infty$. Un producto escalar en ℓ^2 es:

$$((x_n)|(y_n)) = \sum_{n=1}^{\infty} x_n \overline{y_n}.$$

Probemos que si $(x_n),(y_n) \in \ell^2$, la serie $\sum_{n=1}^{\infty} x_n \overline{y_n}$ es convergente. Es suficiente ver que es absolutamente convergente; siendo $|x_n \overline{y_n}| = |x_n| \cdot |\overline{y_n}| = |x_n| \cdot |y_n| = |x_n y_n|$, tenemos que ver que $\sum_{n=1}^{\infty} |x_n y_n|$ es convergente.

Según la desigualdad numérica de Cauchy-Schwartz:

$$\forall m \in \mathbb{N}, \sum_{n=1}^{m} |x_n||y_n| \leq \left[\sum_{n=1}^{m} |x_n|^2\right]^{1/2} \left[\sum_{n=1}^{m} |y_n|^2\right]^{1/2}$$

Luego $\sum_{n=1}^{\infty} |x_n y_n| \leq \left[\sum_{n=1}^{\infty} |x_n|^2\right]^{1/2} \left[\sum_{n=1}^{\infty} |y_n|^2\right]^{1/2}$. Luego $((x_n)|(y_n)) \in \mathbb{R}$

La norma inducida por este producto escalar es:

$$\|(x_n)\| = \left[\sum_{n=1}^{\infty} |x_n \overline{x_n}|\right]^{1/2} = \left[\sum_{n=1}^{\infty} |x_n|^2\right]^{1/2}$$

⑤ En $C(I)$ es un producto escalar el definido por: $(f|g) = \int_I fg$.

Los dos primeros axiomas se verifican por las propiedades de la integral. Además, si $f \neq 0$, $(f|f) = \int_I f^2 > 0$, pues sabemos que la integral en I de una función continua no nula y no negativa es estrictamente positiva.

La norma inducida por este producto escalar es:

$$\|f\|_2 = (f|f)^{1/2} = \left[\int_I f^2\right]^{1/2}$$

$C(I)$ con la norma $\|\cdot\|_2$ no es completo. La compleción de $(C(I), \|\cdot\|_2)$ se denota por $L^2(I)$. $L^2(I)$, ℓ^2 y $(\mathbb{K}^n, \|\cdot\|_2)$ son espacios de Hilbert. Estos tres espacios, junto con $C_\infty(I)$ son los espacios fundamentales del Análisis actual.

<u>OBSERVACION</u>: Todo espacio prehilbertiano de dimensión finita es ~~de~~ de Hilbert, pues todo espacio normado de dimensión finita es completo.

2. Ortogonalidad en un espacio prehilbertiano.

<u>DEFINICION</u>: Sea H un espacio prehilbertiano. Se dice que dos vectores $x, y \in H$ son ortogonales, y escribiremos $x \perp y$, si $(x|y) = 0$.

<u>PROPIEDADES DE LA ORTOGONALIDAD</u>: Sea H un espacio prehilbertiano.

P.1) $\forall x \in H$, $x \perp 0$, pues $\forall x \in H$, $(x|0) = 0$.

P.2) $x \perp y \Rightarrow y \perp x$ (propiedad simétrica).

P.3) Homogeneidad: $x \perp y \Rightarrow \forall \lambda, \mu \in \mathbb{K}$, $\lambda x \perp \mu y$.

P.4) Aditividad: $x \perp y \wedge x \perp z \Rightarrow x \perp y+z$.

P.5) Si dos vectores x e y no nulos son ortogonales, entonces son linealmente independientes

Demostr.: Si x e y fuesen linealmente dependientes, existe $\lambda \neq 0$ tal que $y = \lambda x$. Si $x \perp y \Rightarrow x \perp \lambda x \Rightarrow (\lambda x | x) = 0 \Rightarrow$ $\Rightarrow \lambda (x|x) = 0$. Luego sería $(x|x)=0$, pues $\lambda \neq 0$, lo cual es absurdo, pues $x \neq 0 \Rightarrow (x|x) > 0$.

P.6.) Teorema de Pitágoras: Si $x \perp y$, entonces

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2$$

El recíproco es cierto para $\mathbb{K} = \mathbb{R}$.

Demostr: $\|x+y\|^2 = (x+y|x+y) = \|x\|^2 + \|y\|^2 + (x|y) + (y|x)$.

Si $x \perp y$, $(x|y) = (y|x) = 0$. csqd.

P.7). Propiedad del isósceles: Si $x \perp y$, entonces $\|x+y\| = \|x-y\|$.

En general, $\forall \lambda \in \mathbb{K}$, $\|x+\lambda y\| = \|x - \lambda y\|$.

P.8) $\forall x,y \in H - \{0\}$, $\exists \alpha \in \mathbb{K} \,/\, x \perp \alpha x + y$

Demostr: Sea $\alpha = -\dfrac{(y|x)}{(x|x)}$. Entonces $x \perp \alpha x + y$. Además, es única trivialmente.

P.9) Dados $x,y \in H$, x es ortogonal a y si y solo si se verifica que $\forall \lambda \in \mathbb{K}$, $\|x\| \leq \|x + \lambda y\|$.

Demostr.: $\boxed{\Rightarrow}$ $\forall \lambda \in \mathbb{K}$, $\|x+\lambda y\|^2 = (x+\lambda y | x + \lambda y) = (x|x) + \lambda (y|x) +$ $+ \bar{\lambda}(x|y) + |\lambda|^2 (y|y) = \|x\|^2 + \lambda(y|x) + \bar{\lambda}(x|y) + |\lambda|^2 \|y\|^2 =$ $= \|x\|^2 + |\lambda|^2 \|y\|^2 \geq \|x\|^2$. Luego $\|x\| \leq \|x+\lambda y\|$, $\forall \lambda \in \mathbb{K}$.

$\boxed{\Leftarrow}$ Por hipótesis, $\forall \lambda \in \mathbb{K}$, $\|x\| \leq \|x+\lambda y\|$. Luego

$(x|x) \leq (x+\lambda y | x + \lambda y) = (x|x) + \lambda (y|x) + \bar{\lambda}(x|y) + \lambda\bar{\lambda}(y|y)$, $\forall \lambda \in \mathbb{K}$.

En particular es cierto para $\lambda = -\dfrac{(x|y)}{(y|y)}$, supuesto que $y \neq 0$ (*)

Entonces $(x|x) \leq (x|x) - \dfrac{|(x|y)|^2}{(y|y)} - \dfrac{|(x|y)|^2}{(y|y)} + \dfrac{|(x|y)|^2}{(y|y)} = (x|x) - \dfrac{|(x|y)|^2}{(y|y)}$

Luego $\dfrac{|(x|y)|^2}{(y|y)} \leq 0$. Pero $\dfrac{|(x|y)|^2}{(y|y)}$ es positivo. Luego, debe ser $\dfrac{|(x|y)|^2}{(y|y)} = 0$ y, por tanto, $(x|y) = 0$. csqd.

3. Aproximación óptima. Proyección ortogonal.

DEFINICION: Sea H un espacio prehilbertiano y L un subespacio vectorial de H. Dado un vector x de H, decimos que $y_0 \in L$ es aproxima ción óptima de x relativa a L si $\|x-y_0\| \leq \|x-y\|$, $\forall y \in L$, es decir, si $d(x,y_0) = \inf_{y\in L} d(x,y) = d(x,L)$

OBSERVACIONES: ① Decir que $\forall y \in L$, $\|x-y_0\| \leq \|x-y\|$ es equivalente a decir $\forall y \in L$, $\forall \lambda \in \mathbb{K}$, $\|x-y_0\| \leq \|x-y_0+\lambda y\|$, ya que $y_0-\lambda y$ recorre L cuando λ recorre $\mathbb{K}$ e y recorre L, y esto significa, de acuerdo con la propiedad P.9) de la ortogonalidad, que $\forall y \in L$, $x-y_0 \perp y$, y esto equivale, por definición, a decir que $x-y_0 \perp L$.

② El concepto de aproximación óptima, como se deduce de la definición, solo "utiliza" la estructura métrica de H, y se puede definir, por tanto, en un espacio métrico y relativa a un subconjunto del mismo. En el caso de que H sea espacio prehilbertiano decimos que y_0 es una aproximación óptima de x relativa a L si y solo si, por definición, y_0 es la proyección ortogonal de x sobre L. Se dice "la" proyección ortogonal pues, como veremos despues, es única, si L es completo.

3.1. TEOREMA: (de la proyección ortogonal).

Sea H un espacio prehilbertiano y L un subespacio vectorial de H. Si L es completo (de Hilbert), entonces todo punto $x \in H$ tiene una y solo una proyección ortogonal sobre L o aproximación óptima relativa a L.

Demostr.: Existencia: Si $x \in L$, el teorema es trivial, pues la aproximación óptima de x relativa a L sería x, y es único.

Supongamos que $x \in H - L$. Siendo L completo, es cerrado y, en consecuencia, $\alpha = d(x,L) = \inf_{y \in L} \|x-y\| > 0$.

Por definición de α, existe en L una sucesión $(y_n)_n$ tal que $(\|x-y_n\|)_n \to \alpha$. Veamos que $(y_n)_n$ es de Cauchy, con lo cual será convergente, por ser L completo, y su límite será "una" aproximación óptima de x relativa a L, pues $\alpha = \lim_{n\to\infty} \|x-y_n\| = \|x - \lim_n y_n\|$

En virtud de la igualdad del paralelogramo:

$$2(\|x-y_p\|^2 + \|x-y_q\|^2) = \|2x-y_p-y_q\|^2 + \|y_p-y_q\|^2 =$$
$$= 4\,\|x-\tfrac{1}{2}(y_p+y_q)\|^2 + \|y_p-y_q\|^2. \text{ Luego:}$$

$$\|y_p-y_q\|^2 = 2(\|x-y_p\|^2 + \|x-y_q\|^2) - 4\,\|x-\tfrac{1}{2}(y_p+y_q)\|^2$$

Como $\frac{1}{2}(y_p+y_q) \in L$, pues L es subespacio de H, por definición de α, tenemos que $\|x-\frac{1}{2}(y_p+y_q)\| \geq \alpha \Rightarrow -4\,\|x-\frac{1}{2}(y_p+y_q)\|^2 \leq -4\alpha^2$

Luego:

$$\|y_p - y_q\|^2 \leq 2(\|x-y_p\|^2 + \|x-y_q\|^2) - 4\alpha^2.$$

Ahora bien, por hipótesis, $(\|x-y_n\|)_n \to \alpha$. Luego $(\|x-y_n\|^2)_n \to \alpha^2$. Por tanto, $\forall \varepsilon>0, \exists \nu \in \mathbb{N} / n>\nu \Rightarrow \|x-y_n\|^2 < \alpha^2 + \frac{\varepsilon^2}{4}$.

Luego, $\forall \varepsilon>0, \exists \nu \in \mathbb{N} / p,q>\nu \Rightarrow \|y_p - y_q\|^2 \leq 2(\|x-y_p\|^2 + \|x-y_q\|^2) - 4\alpha^2 <$

$$< 2\left(\alpha^2 + \frac{\varepsilon^2}{4} + \alpha^2 + \frac{\varepsilon^2}{4}\right) - 4\alpha^2 = \varepsilon^2.$$

Luego si $p,q>\nu \Rightarrow \|y_p - y_q\| < \varepsilon$. Por tanto, $(y_n)_n$ es de Cauchy en L y, como se indicó anteriormente, existe "una" proyección ortogonal de x sobre L.

Unicidad: Supongamos que y_0 e y_0' son aproximaciones óptimas de x relativas a L. En virtud de la igualdad del paralelogramo tenemos que

$$\|y_0 - y_0'\|^2 = 2(\|x-y_0\|^2 + \|x-y_0'\|^2) - 4\left\|x - \frac{1}{2}(y_0+y_0')\right\|^2.$$

Como $\|x-y_0\| = \|x-y_0'\| = \alpha$ y $4\left\|x-\frac{1}{2}(y_0+y_0')\right\|^2 \geq 4\alpha^2$, pues $\frac{1}{2}(y_0+y_0') \in L$, se tiene que $\|y_0 - y_0'\|^2 \leq 2(\alpha^2+\alpha^2) - 4\alpha^2 = 0$.

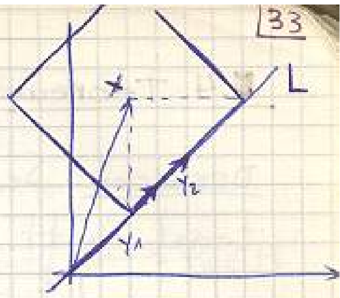
Luego $y_0 = y_0'$. csqd.

OBSERVACIONES: ① En la demostración del teorema solo hemos utilizado que el "punto medio" de dos puntos de L está en L y no se utiliza que L sea subespacio, sino para esto. Como los conjuntos convexos son los que verifican que el punto medio de dos puntos pertenezca al conjunto podemos generalizar un poco más imponiendo la condición más débil de que L sea convexo.

② Si L no es completo, la proyección ortogonal puede existir o no, pero de existir, será única, pues en la demostración de la unicidad no se ha utilizado la completitud de L.

③ Consideremos el espacio normado $C_2(I)$ y L el subespacio de las funciones polinómicas en I. L es denso en $C_2(I)$, pues es denso en $C_\infty(I)$ y la topología de $C_\infty(I)$ es más fina que la de $C_2(I)$. Sea $f \in C_2(I) - L$. Entonces f no tiene aproximación óptima relativa a L, pues $d(f,L)=0$, (ya que L es denso en $C_2(I)$ y hay polinomios que distan tan poco como se desee de f) y, sin embargo, no existe $g \in L$ tal que $d(f,g)=0$, pues en este caso, sería $f=g$ y f no es un polinomio.

④ Consideremos el espacio normado $(\mathbb{R}^2, \|\cdot\|_1)$ y sea $L = \{(x,x) / x \in \mathbb{R}\}$



Geométricamente, se ve que la proyección ortogonal de un vector x sobre L no es única (y_1 e y_2 son proyecciones ortogonales de x).

* Sea H un espacio prehilbertiano y L un subespacio vectorial completo de H. Entonces podemos definir una aplicación de H en L, que llamaremos proyección ortogonal y denotaremos por P_L, que a cada $x \in H$ asocia su proyección ortogonal sobre L, $P_L(x)$. Entonces:

3.2. PROPOSICION: La aplicación P_L es una proyección lineal de norma 1.(*)

Demostr: – $P_L \circ P_L = P_L$, pues $\forall x \in H$, $P_L(x) \in L$ y $\forall y \in L$, $P_L(y) = y$.
Luego, $\forall x \in H$, $(P_L \circ P_L)(x) = P_L(P_L(x)) = P_L(x)$.
– P_L es lineal: En virtud de la linealidad de la ortogonalidad y de la definición de proyección ortogonal de x sobre L ($x - P_L(x) \perp L, \forall x \in H$) se tiene: $\forall x, x' \in H, \forall \lambda, \lambda' \in \mathbb{K}$, $x - P_L(x) \perp L$ y $x' - P_L(x') \perp L \Rightarrow$
$\Rightarrow \lambda [x - P_L(x)] + \lambda' [x' - P_L(x')] \perp L \Rightarrow$
$\Rightarrow (\lambda x + \lambda' x') - (\lambda P_L(x) + \lambda' P_L(x')) \perp L$. Luego $\lambda P_L(x) + \lambda' P_L(x')$ es una proyección ortogonal de $\lambda x + \lambda' x'$. En virtud del teorema 3.1 la proyección ortogonal de un vector cualquiera de H existe y es única.
Luego, debe ser: $P_L(\lambda x + \lambda' x') = \lambda P_L(x) + \lambda' P_L(x')$.
– P_L es de norma 1: Por definición, $\|P_L\| = \sup_{x \neq 0} \frac{\|P_L(x)\|}{\|x\|}$
Entonces, $\|P_L\| \geq 1$, pues si $x \in L - \{0\}$, $P_L(x) = x$ y se tiene que $\frac{\|P_L(x)\|}{\|x\|} = \frac{\|x\|}{\|x\|} = 1$. Si probamos que $\forall x \in H$, $\|P_L(x)\| \leq \|x\|$ quedará visto que $\|P_L\| = 1$, pues 1 será una cota superior alcanzable.
$\forall x \in H$, $x - P_L(x) \perp L$. En particular, $\forall x \in H$, $P_L(x) \perp x - P_L(x)$ pues $P_L(x) \in L$. Entonces, por la propiedad P.9) de la ortogonalidad:
$\forall \lambda \in \mathbb{K}$, $\|P_L(x)\| \leq \|P_L(x) + \lambda(x - P_L(x))\|$. En particular, para $\lambda = 1$ tenemos que $\|P_L(x)\| \leq \|x\|$, $\forall x \in H$. Luego $\|P_L\| = 1$. c.s.q.d.

3.3. TEOREMA: (de Kakutani)

Sea E espacio normado de dimensión mayor o igual que 3. Entonces E es prehilbertiano si y solo si existen proyecciones lineales de norma 1 sobre todo subespacio bidimensional de E.

La demostración de este teorema es muy complicada y no la haremos aquí.

4. Teorema de F. Riesz.

DEFINICION: Sea H un espacio prehilbertiano y L un subespacio vectorial de H. Se llama complemento ortogonal de L al conjunto

$$L^\perp = \{x \in H \,/\, x \perp L\}.$$

Se demuestra que cualquiera que sea el subespacio L, $L^\perp$ es cerrado en H. Se prueba también que si H es de Hilbert y L es cerrado, entonces $(L^\perp)^\perp = L$.

Por el teorema de la proyección ortogonal, si L es completo todo elemento x de H tiene una y solo una proyección ortogonal $P_L(x)$. Como $\forall x \in H$, $x = P_L(x) + (x - P_L(x))$ con $P_L(x) \in L$ y $x - P_L(x) \in L^\perp$, y esta descomposición de x como suma de un elemento de L y otro de $L^\perp$ es única, se tiene que

$$H = L \oplus L^\perp$$

4.1. PROPOSICION: Sea H un espacio prehilbertiano y a un elemento de H. Entonces la aplicación $\varphi_a : x \in H \longmapsto (x|a) \in \mathbb{K}$ es lineal y continua. Además $\|\varphi_a\| = \|a\|$.

Demostr.: – φ_a es lineal, trivialmente.

– φ_a es continua si y solo si es acotado el conjunto $\{|\varphi_a(x)| \,/\, \|x\|=1\} = \{|(x|a)| \,/\, \|x\|=1\}$.

Por la desigualdad de Cauchy-Schwartz, $\forall x \in H$, $|(x|a)| \leq \|x\| \cdot \|a\|$.

Si $\|x\| = 1$, se verifica que $|(x|a)| \leq \|a\|$.

Luego φ_a es continua.

– $\|\varphi_a\| = \|a\|$: Siendo $\|\varphi_a\| = \sup\{|(x|a)| \,/\, \|x\|=1\}$ se deduce de lo anterior que $\|\varphi_a\| \leq \|a\|$. Veamos que esta cota superior es alcanzable con lo cual será $\|\varphi_a\| = \|a\|$.

Si $a = 0$, es trivial que $\|\varphi_a\| = \|a\| = 0$.

Supongamos que $a \neq 0$, entonces $\left\|\frac{a}{\|a\|}\right\| = 1$ y

$$\left|\varphi_a\left(\frac{a}{\|a\|}\right)\right| = \left|\left(\frac{a}{\|a\|} \,\Big|\, a\right)\right| = \frac{|(a,a)|}{\|a\|} = \frac{\|a\|^2}{\|a\|} = \|a\|$$

Luego $\|\varphi_a\| = \|a\|$. c.s.q.d.

4.2. COROLARIO: La aplicación $f : a \in H \longmapsto \varphi_a \in H'$ es antilineal e isométrica y, a fortiori, inyectiva.

Demostr.: $f(\lambda a + \mu b) = \varphi_{\lambda a + \mu b} = \bar{\lambda}\varphi_a + \bar{\mu}\varphi_b$, pues

$\varphi_{\lambda a + \mu b}(x) = (x | \lambda a + \mu b) = \bar{\lambda}(x|a) + \bar{\mu}(x|b) = \bar{\lambda}\varphi_a(x) + \bar{\mu}\varphi_b(x) \quad \forall x \in H$

Si f es isométrica es inyectiva, pues $a \neq b \Leftrightarrow \|a-b\| \neq 0 \Leftrightarrow$
$\Leftrightarrow \|\varphi_{a-b}\| \neq 0 \Leftrightarrow \|\varphi_a - \varphi_b\| \neq 0 \Leftrightarrow \varphi_a \neq \varphi_b$. c.s.q.d.

4.3. TEOREMA: (de F. Riesz)

Si H es un espacio de Hilbert, la aplicación $f: a \in H \longmapsto \varphi_a \in H'$ es suprayectiva, es decir, todo funcional lineal continuo en H es de la forma $x \longmapsto (x|a)$, para algún $a \in H$.

Demostr.: Sea φ un funcional lineal continuo sobre H. Queremos ver que existe $a \in H$ tal que $\varphi = \varphi_a$, es decir, tal que
$$\varphi(x) = (x|a), \ \forall x \in H.$$
Dado $\varphi \in H'$, el conjunto $L = \{y \in H / \varphi(y) = 0\}$ es un hiperplano de H. Además L es cerrado, pues $L = \varphi^{-1}(\{0\})$, φ es continua y $\{0\}$ es un cerrado de $\mathbb{K}$. Siendo H completo y L un subespacio cerrado de H, entonces L es completo.

Por el teorema de la proyección ortogonal, existe un punto $b \in H - L$ tal que $b \perp L$, basta tomar un $c \in H - L$ y hacer $b = c - P_L(c)$, pues $P_L(c)$ existe ya que L es completo. Entonces, $\forall y \in L$, $\varphi_b(y) = (y|b) = 0$. Luego, $L = \{y \in H / \varphi_b(y) = 0\}$. Por tanto, $\operatorname{Ker} \varphi = \operatorname{Ker} \varphi_b$; entonces φ es proporcional a φ_b (ver GEOMETRIA I, TEMA 9, Teorema 6.2), es decir, $\exists \lambda \in \mathbb{K} / \varphi = \lambda \varphi_b$. Pero $\lambda \varphi_b = \varphi_{\bar{\lambda} b}$. Luego $\exists a = \bar{\lambda} b \in H / \varphi = \varphi_a$. c.s.q.d.

De todo lo anterior deducimos que: "si H es un espacio de Hilbert, H y H' son "identificables" por medio de una aplicación biyectiva, antilineal e isométrica".

*EJEMPLO: Sea H un espacio prehilbertiano y $L = \langle v_1, \ldots, v_n \rangle$ un subespacio vectorial de H de dimensión n. Vamos a calcular la aproximación óptima de un vector x de H relativa a L.

Dicha aproximación óptima existe y es única, pues, al ser L de dimensión finita, es completo. Decir que $P_L(x) = \mu_1 v_1 + \ldots + \mu_n v_n$ equivale a decir que $x - (\mu_1 v_1 + \ldots + \mu_n v_n) \perp L$, y esto equivale a decir que $x - (\mu_1 v_1 + \ldots + \mu_n v_n) \perp \lambda_1 v_1 + \ldots + \lambda_n v_n$, $\forall \lambda_1, \ldots, \lambda_n \in \mathbb{K}$ y esto equivale, por la linealidad de la ortogonalidad, a que $x - (\mu_1 v_1 + \ldots + \mu_n v_n) \perp v_i$, $i \in \{1, \ldots, n\}$. Esto significa que:
$$\mu_1 (v_1|v_1) + \ldots + \mu_n (v_n|v_1) = (x|v_1), \ \ldots, \ \mu_1 (v_1|v_n) + \ldots + \mu_n (v_n|v_n) = (x|v_n)$$
pues $(x - (\mu_1 v_1 + \ldots + \mu_n v_n) | v_i) = 0 \Rightarrow (x|v_i) - (\mu_1 v_1 + \ldots + \mu_n v_n | v_i) = 0$.

Se ha formado el siguiente sistema de n ecuaciones con n incógnitas

$$\left\{\begin{array}{l}\mu_1 (v_1|v_1) + \dots + \mu_n (v_n|v_1) = (x|v_1) \\ \dots\dots\dots\dots\dots\dots \\ \mu_1 (v_1|v_n) + \dots + \mu_n (v_n|v_n) = (x|v_n)\end{array}\right. \qquad (I)$$

cuya solución nos da las componentes de $P_L(x)$ en la base $\{v_1, \dots, v_n\}$, y esta solución, como sabemos de antemano, existe y es única.

De este razonamiento deducimos lo siguiente: "Si $\{v_1, \dots, v_n\}$ es un sistema libre de un espacio prehilbertiano H, entonces, la matriz de Gram

$$\begin{pmatrix} (v_1|v_1) & \dots & (v_n|v_1) \\ \dots & \dots & \dots \\ (v_1|v_n) & \dots & (v_n|v_n) \end{pmatrix}$$

es regular (su determinante es distinto de cero)."

* SISTEMA ORTONORMAL: Si el sistema $\{v_1, \dots, v_n\}$ es ortonormal, es decir, si $(v_i|v_j) = \delta_{ij} = \begin{cases} = 1 & \text{si } i = j \\ = 0 & \text{si } i \neq j \end{cases}$, es decir, si la matriz de Gram es la matriz unidad, el sistema (I) está resuelto trivialmente, y se tiene $\mu_k = (x|v_k)$, $k = 1, \dots, n$. Es decir:

$$P_L(x) = (x|v_1)\, v_1 + \dots + (x|v_n)\, v_n.$$

* PROCESO DE ORTONORMALIZACION DE GRAM-SCHMIDT:

"Sea H un espacio prehilbertiano y $\{v_1, \dots, v_n\}$ un sistema libre en H. Entonces existe un sistema $\{u_1, \dots, u_n\}$ ortonormal en H y tal que $\langle u_1, \dots, u_k\rangle = \langle v_1, \dots, v_k\rangle$". Lo mismo se puede decir para una colección numerable y libre en H.

Demostr.: Sea $u_1 = \dfrac{v_1}{\|v_1\|}$.

Para determinar u_2, hallamos primero u_2' de modo que $u_2' \perp u_1$ y $\langle u_1, u_2'\rangle = \langle v_1, v_2\rangle$. Evidentemente, $\langle v_1, v_2\rangle = \langle u_1, v_2\rangle$. Si ha de ser $\langle u_1, u_2'\rangle = \langle v_1, v_2\rangle$ podemos escribir $u_2' = \lambda u_1 + v_2$. Si $u_2' \perp u_1$ entonces $0 = (\lambda u_1 + v_2 | u_1) = \lambda (u_1|u_1) + (v_2|u_1) = \lambda + (v_2|u_1) \Rightarrow \lambda = -(v_2|u_1)$.

Tenemos, por tanto, determinado u_2'. Hacemos entonces $u_2 = \dfrac{u_2'}{\|u_2'\|}$, con lo cual, $u_2 \perp u_1$, $\langle u_1, u_2\rangle = \langle v_1, v_2\rangle$ y $\|u_2\| = 1$.

Fácilmente se termina la demostración de esta proposición por inducción ∎

* SISTEMAS ORTONORMALES TOTALES: Sea H un espacio prehilbertiano. Se dice que un sistema de vectores de H $\{x_i / i \in I\}$ es ortonormal si $(x_i|x_j) = \delta_{ij}$ (delta de Krönecker).

Ordenando por inclusión el conjunto de los sistemas ortonormales de H resulta, trivialmente, que toda cadena o parte totalmente ordenade Zorn garantiza la existencia de sistemas ortonormales maximales en H.

DEF: Sea H espacio prehilbertiano y $\{x_i / i\in I\}$ un sistema ortonormal de H. Decimos que $\{x_i / i\in I\}$ es total si se verifica la implicación siguiente:

$$[(x|x_i)=0, \forall i\in I] \Rightarrow [x=0].$$

4.4. PROPOSICION: Es condición suficiente para que un sistema ortonormal sea total que el subespacio vectorial de H engendrado por él sea denso en H.

Demostr.: Sea $\{x_i / i\in I\}$ un sistema ortonormal de H. El subespacio L engendrado por este sistema es, como sabemos, el conjunto de las combinaciones lineales finitas de vectores del sistema.

Sea entonces $x\in H$ de modo que $(x|x_i)=0, \forall i\in I$. Por definición de L, y por la linealidad de la ortogonalidad se verifica que $x\perp L$, y esto equivale a decir que $0\in P_L(x)$, pues $x-0\perp L$. (*)

Por otra parte, siendo L denso, $\overline{L}=H$ y, en consecuencia, $x\in\overline{L}$. Luego: $d(x,L)=0$; es decir, $\inf\limits_{y\in L} d(x,y)=0$, y también, como $d(x,P_L(x)) = \inf\limits_{y\in L} d(x,y)$ y $0\in P_L(x)$, se tiene que $d(x,0)=\inf\limits_{y\in L} d(x,y)=0$. Luego $x=0$ y, en consecuencia, el sistema ortonormal $\{x_i / i\in I\}$ es total. csqd.

4.5. PROPOSICION: Si H es de Hilbert entonces es condición necesaria para que un sistema ortonormal de H sea total que el subespacio vectorial de H engendrado por él sea denso en H.

Demostr.: Supongamos que $L=\langle\{x_i / i\in I\}\rangle$ no es denso y probemos que, entonces, el sistema ortonormal $\{x_i / i\in I\}$ no es total.

Como $\overline{L}$ es un cerrado en H y H es completo se verifica que $\overline{L}$ es completo, pues todo subespacio cerrado de un espacio completo es completo.

Sea $x\in H-\overline{L}$, que existe pues suponemos que L no es denso. Entonces, por el teorema de la proyección ortogonal existe, y es única, la proyección ortogonal de x relativa a $\overline{L}$. Si hacemos $z=x-P_{\overline{L}}(x)$ se tiene que $z\neq 0$ pues $x\notin\overline{L}$ y $z\perp\overline{L}$ y, por tanto, $z\perp L$, pues $L\subset\overline{L}$. En definitiva, $\exists z\in H-\{0\} / (z|x_i)=0, \forall i\in I$, lo cual contradice que $\{x_i / i\in I\}$ es total. Luego L debe ser denso. csqd.

4.6. PROPOSICION: Si H es un espacio prehilbertiano separable, es decir, que admite un subconjunto denso y numerable, entonces existe en H un sistema ortonormal que engendra un subespacio denso y, por tanto, dicho sistema ortonormal es total.

Demostr.: Sea $\{x_n \mid n\in\mathbb{N}\}$ un conjunto denso y numerable en H. De este conjunto podemos obtener, por el proceso de ortonormalización de Gram-Schmidt, un sistema ortonormal de modo que el subespacio engendrado por él contenga a $\{x_n \mid n\in\mathbb{N}\}$. Por tanto, dicho subespacio será denso y, en consecuencia, el sistema ortonormal es total. csqd.

Todo lo anterior lo resumimos en los siguientes puntos:

- En todo espacio de Hilbert o prehilbertiano separable existe un sistema ortonormal que engendra un subespacio denso y, por tanto, dicho sistema es total.
- En espacios de Hilbert es equivalente el que un sistema ortonormal sea total y que engendre subespacios densos.
- En todo espacio pH hay sistemas ortonormales maximales.
- Un sistema ortonormal es total si y solo si es maximal, pues si $\{x_i \mid i\in I\}$ es total debe ser maximal pues si pudiesemos añadir un vector $x\neq x_i, \forall i\in I$ de modo que $\{x_i \mid i\in I\}\cup\{x\}$ fuese ortonormal se verificaría que $(x|x_i)=0, \forall i\in I$ y $x\neq 0$ en contra de que $\{x_i \mid i\in I\}$ es total y, recíprocamente, si $\{x_i \mid i\in I\}$ es maximal y $(x|x_i)=0, \forall i\in I$ se tiene que $x=0$, pues de lo contrario el sistema $\{x_i \mid i\in I\}\cup\{\frac{x}{\|x\|}\}$ sería ortonormal en contra de que $\{x_i \mid i\in I\}$ es maximal; luego $\{x_i \mid i\in I\}$ es total.