III Hilbert Spaces

Subsections:

1. Scalar product spaces and Hilbert spaces

1.1 Sesquilinear form, scalar product, induced norm

Let X be a $\mathbb{K}$-linear space and $\varphi : X \times X \to \mathbb{K}$.

Definition

- φ is a sesquilinear form*) iff it satisfies:

  (i) $\forall_{\substack{x_1, x_2, y \in X \\ \lambda_1, \lambda_2 \in \mathbb{K}}} \quad \varphi(\lambda_1 x_1 + \lambda_2 x_2, y) = \lambda_1 \varphi(x_1, y) + \lambda_2 \varphi(x_2, y)$

  (i.e.: first-variable linearity)

  (ii) $\forall_{\substack{x_1, x_2, y \in X \\ \lambda_1, \lambda_2 \in \mathbb{K}}} \quad \varphi(y, \lambda_1 x_1 + \lambda_2 x_2) = \overline{\lambda_1} \varphi(y, x_1) + \overline{\lambda_2} \varphi(y, x_2)$

  (i.e.: second-variable „conjugate linearity").

Note, that $\overline{\lambda} = \lambda$ for $\mathbb{K} = \mathbb{R} \ni \lambda$, so „sesquilinear" = „bilinear" (i.e. both first- and second-variable linear) in real space case.

- φ is a scalar product**) iff it is a sesquilinear form and:

---

*) Pol.: półtoraliniowa ; in Latin: sesqui = $1\frac{1}{2}$.

**) Pol.: iloczyn skalarny; a second English name: inner product.
We say also scalar/inner product in X.

(iii) $\forall_{x,y\in X}$ $\varphi(y,x)=\overline{\varphi(x,y)}$ (i.e., „conjugate symmetry");

(iv) $\forall_{x\in X}$ $\varphi(x,x)\geqslant 0$ *) (i.e., non-negativity);

(v) $\forall_{x\in X}$ $\left(\varphi(x,x)=0 \overset{**)}{\Longrightarrow} x=0\right)$ (i.e, nondegeneracy).

Note, that having (i) and (iii), we get (ii), so we can omit it in the definition of scalar product. For scalar products we shall often use the symbol $(\cdot,\cdot)$, instead of $\varphi(\cdot,\cdot)$.

Examples

The following „$(\cdot,\cdot)$"-s are scalar products ($\longrightarrow \triangle$):

- in $\mathbb{K}^d$ $\quad (x,y):=\sum_{j=1}^{d} x_j\overline{y_j}$ for $x,y\in\mathbb{K}^d$.

We also use a shorter notation: $\overline{x}:=(\overline{x_1},\dots,\overline{x_d})$ and $xy:=\sum_{j=1}^{d}x_jy_j$, so $(x,y)=x\overline{y}$ (so $(x,y)=xy$, if $\mathbb{K}=\mathbb{R}$). This scalar product is called standard s.p. ***) (there exist many s.p.-s in $\mathbb{K}^d$, obviously).

---

*) Note, that $\forall_{x\in X}$ $\varphi(x,x)\in\mathbb{R}$ by (iii).

**) „$\Leftarrow$" is also true by (i), (ii).

***) s.p. = scalar product.

- in $\ell^2(\mathbb{N})$ (etc.) $\quad (x,y) := \sum_{n\in\mathbb{N}} x(n)\overline{y(n)}$ for $x,y \in \ell^2(\mathbb{N})$

– note, that the serie on the RHS is convergent (moreover – absolutely convergent), because, e.g., $|x(n)\overline{y(n)}| \leqslant |x(n)|^2 + |y(n)|^2$.

Similarly we can define $(\ ,\ )$ in any $\ell^2_w(\Omega)$ by

$$(x,y) := \sum_{t\in\Omega} x(t)\overline{y(t)}\,w(t) = \int_\Omega x\bar{y}w\,d\#.$$

(– see footnotes on p.p PB-28 and OF-60 for the def. of "$\sum_{t\in\Omega}$")

- in $L^2(\Omega,\mu)$ $\quad ([f],[g]) := \int_\Omega f\bar{g}\,d\mu$ for $[f],[g] \in L^2(\Omega,\mu)$

– here, similarly as above, $f\bar{g}$ is integrable (we can use the analogic inequality or just the Hölder inequality) and obviously the RHS does not depend on the member of $[f],[g]$.

- Let $\emptyset \neq G$ be an open connected subset of $\mathbb{C}$, define $A^2(G) := \{f \in \tilde{L}^2(G)^{*)} : f \text{ is analytic}^{**)}\}$. We define $(\cdot,\cdot)$ in $A^2(G)$ by the formula $\quad (f,g) := \int_G f\cdot\bar{g}\,dl_2$ for $f,g \in A^2(G)$.

- By $\mathrm{Trig}(\mathbb{R})$ denote the space of all trigonometric polynomials on $\mathbb{R}$, i.e. $\mathrm{Trig}(\mathbb{R}) \underset{lin}{\subset} \ell(\mathbb{R})$, $\mathrm{Trig}(\mathbb{R}) := \mathrm{lin}\{e_\lambda : \lambda\in\mathbb{R}\}$, where for $\lambda\in\mathbb{R}$ $\quad e_\lambda : \mathbb{R}\to\mathbb{C}$ is given by the formula

---

*) With the usual (2-dimensional) Lebesgue measure (l_G) in $G \subset \mathbb{C} \approx \mathbb{R}^2$ – the restriction of l_2 in $\mathbb{R}^2$ to G.

**) = holomorphic.

$$e_\lambda(t) := e^{i\lambda t}, \quad t \in \mathbb{R}.$$

We define $(f,g) := \lim_{T\to+\infty} \frac{1}{2T} \int_{-T}^{T} f(t)\cdot\overline{g(t)}\,dt$

for $f,g \in \mathrm{Trig}(\mathbb{R})$ — the existence of the limit on the RHS is easy to prove ($\longrightarrow$ △).

Some geometry – orthogonality

When $(\cdot,\cdot)$ is a fixed scalar product in X, we can introduce an important geometrical notion in X: the orthogonality (with respect to $(\cdot,\cdot)$).

Definition

Let $x,y \in X$, then x is orthogonal (or perpendicular)*) to y iff $(x,y)=0$ – we denote it by

$$x \perp y.$$

For $M \subset X$ we write also $x \perp M$ iff $\forall_{y\in M}\ x \perp y$.

Define $M^\perp := \{x \in X : x \perp M\}$,

– we call $M^\perp$ the orthogonal complement**) of M.

Obviously, by (iii) and (i) $x \perp y$ iff $y \perp x$ and $M^\perp$ is

---

*) Pol.: prostopadły

**) Pol.: uzupełnienie ortogonalne (raczej nie używa się tu „prostopadłe").

a linear subspace of X for any subset M of X.

$\boxed{\text{Examples}}$

Consider $X = Trig(\mathbb{R})$. For any $\lambda_1, \lambda_2 \in \mathbb{R}$, if $\lambda_1 \neq \lambda_2$, then we have

$$(e_{\lambda_1}, e_{\lambda_2}) = \lim_{T \to +\infty} \frac{1}{2T} \int_{-T}^{T} e^{i\lambda_1 t} e^{-i\lambda_2 t} dt =$$

$$= \lim_{T\to +\infty} \frac{1}{2T} \int_{-T}^{T} e^{i(\lambda_1-\lambda_2)t} dt = \frac{1}{i(\lambda_1-\lambda_2)} \cdot \lim_{T\to+\infty} \frac{1}{2T} \left(e^{i(\lambda_1-\lambda_2)T} - \right.$$
$$\left. + e^{-i(\lambda_1-\lambda_2)T}\right) = 0,$$

so $e_{\lambda_1} \perp e_{\lambda_2}$. Moreover for $\lambda \in \mathbb{R}$ $(e_\lambda, e_\lambda) = 1 \neq 0$, so it can be easily proved, that

$$\{e_\lambda\}^\perp = lin\{e_\mu : \mu \in \mathbb{R} \setminus \{\lambda\}\}.$$

In $\ell^2(\mathbb{N})$ we have $e_m \perp e_n$ *) for $n, m \in \mathbb{N}$, $n \neq m$,

but $\{e_n\}^\perp = \overline{lin\{e_m : m \in \mathbb{N}\setminus\{n\}\}}^{**)} \neq lin\{e_m : m \in \mathbb{N}\setminus\{n\}\}$

(because ... $\to$ $\triangle$).

---

*) Be carefull: the same notation e_n as e_λ for $\lambda = n$ has here a different meaning!

**) $\overline{}$ is here the closure in the sense of the norm space $\ell^2(\mathbb{N})$ (with $\| \ \|_2$ norm).

$\boxed{\text{HS-6}}$

Polarization identity (ties)

Suppose that φ is a sesquilinear form (in X), then for any $x, y \in X$ and $\lambda \in \mathbb{K}$

$$\varphi(x+\lambda y, x+\lambda y) = \varphi(x,x) + |\lambda|^2 \varphi(y,y) + \lambda\varphi(y,x) + \bar{\lambda}\varphi(x,y). \quad (1)$$

If φ is moreover conjugate symmetric (i.e. (iii) holds), then by (1) we get

$$\mathrm{Re}\left(\bar{\lambda}\varphi(x,y)\right) = \frac{1}{2}\left[\varphi(x+\lambda y, x+\lambda y) - \varphi(x,x) - |\lambda|^2\varphi(y,y)\right] \quad (2)$$

In particular, for $\mathbb{K} = \mathbb{R}$ we get (taking $\lambda = \pm 1$)

$$\left.\begin{aligned} \varphi(x,y) &= \frac{1}{2}\left[\varphi(x+y, x+y) - \varphi(x,x) - \varphi(y,y)\right] \quad (3_+) \\ \varphi(x,y) &= -\frac{1}{2}\left[\varphi(x-y, x-y) - \varphi(x,x) - \varphi(y,y)\right] \quad (3_-) \end{aligned}\right\} \quad (3)$$

and thus also (summing the RHS-s)

$$\varphi(x,y) = \frac{1}{4}\left[\varphi(x+y, x+y) - \varphi(x-y, x-y)\right] \quad (3')$$

Each one of (3) and also (3') are called polarizations formulas / identities **) for φ *). More generally, each formula expressing $\varphi(x,y)$ by „some combinations" of $\varphi(z_1, z_1), \dots \varphi(z_k, z_k)$, where $z_1, \dots, z_k$ are „some combinations" of x and y is called polarisation formula

*) Here: for $\mathbb{K} = \mathbb{R}$ - case only...

**) Pol.: wzór polaryzacyjny

HS-7

Consider now the case $\underline{\mathbb{K} = \mathbb{C}}$:

then using (2) for $\lambda = \pm 1$ we get only

$$\operatorname{Re} \varphi(x,y) = \text{RHS of } (3_+), \qquad (4_+)$$
$$\operatorname{Re} \varphi(x,y) = \text{RHS of } (3_-), \qquad (4_-)$$
$$\operatorname{Re} \varphi(x,y) = \text{RHS of } (3'); \qquad (4')$$

(4_+) and (4_-): (4)

but using (2) for $\lambda = \pm i$ we similarly obtain

$$\pm \operatorname{Im} \varphi(x,y) = \operatorname{Re}\big(\mp i\, \varphi(x,y)\big) = \tfrac{1}{2}\big[\varphi(x \pm iy, x \pm iy) - \varphi(x,x) - \varphi(y,y)\big]. \qquad (5_\pm)$$

So (substracting) also

$$\operatorname{Im} \varphi(x,y) = \tfrac{1}{4}\big[\varphi(x+iy, x+iy) - \varphi(x-iy, x-iy)\big], \qquad (6)$$

and finally we get

$$\varphi(x,y) = \tfrac{1}{4}\Big[\varphi(x+y, x+y) - \varphi(x-y, x-y) + i\big(\varphi(x+iy, x+iy) - \varphi(x-iy, x-iy)\big)\Big], \qquad (6')$$

being a polarization formula for φ in $\mathbb{K} = \mathbb{C}$ case.

One can also obtain different polarization formulas in this case using directly one from $(4_\pm)$ and one from $(5_\pm)$ ($\longrightarrow \triangle_4$), but there exists still much more polarization identities for φ i both $\mathbb{R}$ and $\mathbb{C}$ cases ... $\longrightarrow \triangle_{14}$.

Soon we shall make use of such formulas in the special case, when φ is a scalar product in X.

But observe still one more consequence of (2) for $\lambda = \pm 1$: We get (comparing only the RHS-s)

$$\varphi(x+y,x+y)+\varphi(x-y,x-y)=2\big(\varphi(x,x)+\varphi(y,y)\big), \quad (7)$$

which is called (by some reasons which will be clear later) the parallelogram formula / identity*) (for φ).

The induced norm

Suppose that X is a linear space and $(\cdot,\cdot)$ a fixed scalar product in X. Define $\|\ \|: X \longrightarrow \mathbb{R}$ by

$$\|x\| := (x,x)^{\frac{1}{2}}, \qquad x \in X.$$

We shall prove soon that $\|\ \|$ is a norm in fact — as it is suggested by its notation, but we need first the following result.

Fact ("Schwarz**) inequality")

$\forall_{x,y\in X}$ $\quad |(x,y)| \le \|x\|\cdot\|y\|$. Moreover „<" holds iff x and y are linearly independent.

Proof Suppose first that x and y are linearly

---

*) Pol.: formuła równoległoboku

**) Also names Cauchy and Buniakovsky (or Bunyakovsky, Buniakowski...) are assigned to this fact...

independent. Then for any $\lambda \in \mathbb{K}$ we have $0 < (x+\lambda y, x+\lambda y)$ by (iv) and (v) p. HS-3, so by (1)
$0 < \|x\|^2 + |\lambda|^2\|y\|^2 + \lambda (y,x) + \bar{\lambda}(x,y)$. Now taking
$\lambda := \frac{-(x,y)}{\|y\|^2}$ ($\|y\|^2 \neq 0$ by (V) and fact that $y \neq 0$, which must hold by the linear indep. of x, y) we get by (iii)

$$0 < \|x\|^2 + \frac{|(x,y)|^2}{\|y\|^2} - \frac{|(x,y)|^2}{\|y\|^2} - \frac{|(x,y)|^2}{\|y\|^2} = \|x\|^2 - \frac{|(x,y)|^2}{\|y\|^2}$$

which gives

$$|(x,y)| < \|x\| \, \|y\|.$$

Now suppose, that x, y are linearly dependent. If $x = \alpha y$ for some $\alpha \in \mathbb{K}$, then

$$|(x,y)| = |(\alpha y, y)| = |\alpha| \, |(y,y)| = (\alpha \cdot \bar{\alpha} \, (y,y))^{\frac{1}{2}} (y,y)^{\frac{1}{2}} =$$
$$= (\alpha y, \alpha y)^{\frac{1}{2}} (y,y)^{\frac{1}{2}} = \|x\| \cdot \|y\|.$$

Similarly for $y = \alpha x$. $\blacksquare$

Fact ("On induced norm")

The function $\|\cdot\|$ is a norm in X. Moreover, for any $x, y \in X$ the following formulae holds:

1. the polarization identity:
$$(x,y) = \begin{cases} \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right) & \text{for } \mathbb{K} = \mathbb{R}, \\ \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\right) & \text{for } \mathbb{K} = \mathbb{C}; \end{cases}$$
2. the parallelogram identity: $\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right)$.

$\boxed{\text{Proof}}$

By property (v) of scalar product we obviously have $\|x\|=0 \Longrightarrow x=0$, and (i), (ii) give $\|\lambda x\| = |\lambda|\,\|x\|$ for any $x \in X$, $\lambda \in \mathbb{K}$. For any $x, y \in X$, by (i) and (iii) properties, we have

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

thus by Schwarz inequality

$$\|x+y\|^2 \leq \|x\|^2 + \|y\|^2 + 2\|x\|\,\|y\| = (\|x\| + \|y\|)^2,$$

which proves the triangle inequality for $\|\ \|$.

Now the formulae 1. and 2. holds by (3'), (6'), (7) pp. HS-7,8,9. $\square$

The norm $\|\cdot\|$ defined by the scalar product $(\cdot,\cdot)$ as above is called the <u>norm induced</u>*) by $(\cdot,\cdot)$

Note, that the polarization formula means in particular, that also the scalar product $(\cdot,\cdot)$ is <u>uniquely defined</u> by its induced norm $\|\cdot\|$!

It is also interesting that the following result, which is in some sense „inverse" to the parallelogram identity, is true.

---

*) Pol.: <u>norma indukowana</u>

Theorem ("On parallelogram identity")

Let $\|\cdot\|$ be a certain norm in a linear space X.
Then TFCAE:

(1) $\|\cdot\|$ is induced by some scalar product in X.

(2) $\|\cdot\|$ satisfies the parallelogram identity (2. p. HS-10) for any $x, y \in X$.

"**Proof**"

(1) $\Rightarrow$ (2) – is just proved in Fact "On induced norm".

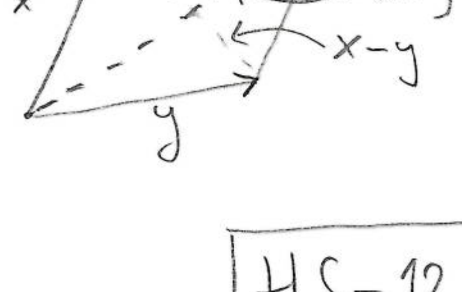
(2) $\Rightarrow$ (1) – a sketch / hint only:

We define $(\cdot,\cdot)$ by the polarization identity using our $\|\cdot\|$ on its RHS. Knowing that $\|\cdot\|$ satisfies the p. id. it is possible to prove (with many "algebraical" steps... $\rightarrow \triangle$) that:

(a) $(\cdot,\cdot)$ is a scalar product in X.

(b) $\|\cdot\|$ is induced by $(\cdot,\cdot)$. $\boxed{\square}$

Note, that the name "parallelogram identity" is related to a well known relation between lengths of edges and diagonals in parallelogram:



1.2. Scalar product / Hilbert spaces

Each linear space X with a fixed scalar product $(\cdot,\cdot)$ in X is called a scalar product or pre-Hilbert or unitary space *); more formally it is a pair $(X, (\cdot,\cdot))$ consisting of a linear space and a scalar product in X. But as we know (by Fact "On induced norm") this scalar product $(\cdot,\cdot)$ defines the induced norm $\|\cdot\|$ for it, and vice versa: the induced norm itself contains all the information on $(\cdot,\cdot)$ — in other words, the scalar product $(\cdot,\cdot)$ can be recovered from $\|\cdot\|$ (by the polarization identity). So, this is why one often change this formal definition of unitary/pre-Hilbert space and says that it is a pair $(X, \|\cdot\|)$, where $\|\cdot\|$ is a norm which is induced by some scalar product in X **). This second approach is more convenient for us, because using it we just treat

---

*) Pol.: przestrzeń z iloczynem skalarnym, prehilbertowska, unitarna.

**) With this changed "formalism" the name scalar product space is rather not used...

each unitary space as a particular case of normed space. So, usually, we shall think on unitary spaces as on some normed space, which additionally possesses a scalar product inducing the norm. And repeat: this scalar product is uniquely determined then. Moreover - similarly as in the general case of normed space - we shall often shortly write/say „ the unitary space X", meaning that there is a fixed norm $\| \ \|$ in X and also the appropriate scalar product $(\cdot,\cdot)$.

Definition

$(X, (\cdot,\cdot))$ or $(X, \|\cdot\|)$ (or simply X) is a **Hilbert space** *) if it is a unitary space which is a Banach space (with respect to the norm $\|\cdot\|$ induced by the scalar product $(\cdot,\cdot)$ in X).

◊ Some examples and constructions

Of course all the examples $\mathbb{K}^d$, $\ell^2(\mathbb{N})$, $\ell^2_\mu(\Omega)$, $L^2(\Omega,\mu)$, $A^2(G)$, $Trig(\mathbb{R})$ are unitary spaces with the scalar products defined on pp. HS-3/4. (both for

*) Pol.: przestrzeń Hilberta.

the case of $\mathbb{K}=\mathbb{R}$ and $\mathbb{C}$, (excluding $\mathrm{Trig}(\mathbb{R})$) — we just should consider the appropriate (spaces of) $\mathbb{K}$-valued functions, vectors, classes of functions. *))

Moreover — and it is very important to us — all the norms in those spaces (excluding $\mathrm{Trig}(\mathbb{R})$), induced by these scalar products are just "our typical" norms: $\|\cdot\|_2$ (in $\mathbb{K}^d$), $\|\cdot\|_2$ in $\ell^2(\mathbb{N})$, $\|\cdot\|_{2,w}$ (in $\ell^2_w(\Omega)$), $\|\cdot\|_2$ (in $L^2(\Omega,\mu)$) and $\|\cdot\|_2$ (restricted to $A^2(G)$ from $L^2(G, \ell_G)$ in $A^2(G)$)**)

So — using our earlier knowledge — we see that $\mathbb{K}^d$ (with $\|\ \|_2$), $\ell^2(\mathbb{N})$, $\ell^2_w(\Omega)$ and $L^2(\Omega,\mu)$ are also Hilbert spaces. The only problem is with the last two spaces: $A^2(G)$ and $\mathrm{Trig}(\mathbb{R})$.

$\boxed{\text{Exercise}}$

Prove that $A^2(G)$ <u>is</u> a Hilbert space and $\mathrm{Trig}(\mathbb{R})$ <u>is not</u>. $\longrightarrow \triangle$.

Hints:

- for $A^2(G)$: prove that the almost uniform convergence follows from the convergence of a sequence of functions from $A^2(G)$ to „a class" $[f] \in L^2(G, \ell_G)$.
- for $\mathrm{Trig}(\mathbb{R})$: consider „an approximation" by elements of $\mathrm{Trig}(\mathbb{R})$ of "an appropriate infinite sum of elements of the form $c e_\lambda$" (with $c \in \mathbb{K}$, $\lambda \in \mathbb{R}$).

---

*) Observe however, that $A^2(G)$ for the case $\mathbb{K}=\mathbb{R}$ is not „a very interesting" space... $\rightarrow \triangle$: why?

**) More precisely, we must identify: $A^2(G) \ni f \longleftrightarrow [f] \in L^2(G, \ell_G)$.

Now – consider our two ± „standard" constructions of norm spaces:

I A subspace of a unitary space

Let X be a unitary space and $Y \underset{lin}{\subset} X$. If we restrict $(\,;\,)$ from $X \times X$ to $Y \times Y$, then $(\cdot\,;\cdot\,)_Y := (\cdot\,,\cdot\,)\big|_{Y\times Y}$ is obviously a scalar product in Y and moreover the subspace norm $\|\ \|_Y$, being the restriction of the norm induced by $(\cdot\,,\cdot\,)$ in X to Y is just the norm induced by $(\cdot\,;\cdot\,)_Y$ in Y.

So – it is natural to call $(Y\ (\,,\,)_Y)$ the unitary subspace of X, and it is compatibile with $(Y, \|\ \|_Y)$ being the unitary space (in the second sense *)). In fact we shall usually say simply "Y is a unitary subspace of X", which means that we consider the norm $\|\ \|_Y$ and the scalar product $(\cdot\,;\cdot\,)_Y$ for Y. Moreover, by Fact "On Banach subspace" p. PB-30, we get directly:

Fact

If X is a Hilbert space and $Y \underset{lin}{\subset} X$, then Y is a Hilbert space iff $Y = \overline{Y}$.

---

*) See the two approaches on pp. HS-13/14.

II A product of two (or more...) unitary spaces

If $X_1, \dots X_k$ are unitary spaces with the appropriate norms $\| \ \|_j$ and scalar products $(\cdot,\cdot)_j$, $j = 1, \dots k$ then our "standard" choice of the norm in $X := X_1 \times \dots \times X_k$ given by $\|x_1\|_1 + \dots + \|x_k\|_k$ (see (1) p. PB – 48) is not a good choice from the point of view of scalar product spaces. But we can choose an equivalent norm $\| \ \|$ in X by the formula (see also Remark p. PB - 50)

$$\| x \| := \| x \|_{[2]} := \Big(\sum_{s=1}^{k} \| x_s \|_s^2 \Big)^{1/2} \quad \text{for } x = (x_1, \dots, x_k)$$

$\in X$. If we define also $(\cdot\,,\cdot)$ for X by

$$(x, y) := \sum_{s=1}^{k} (x_j, y_j)_s$$

then we easily obtain that $\| \ \|$ is exactly the norm induced by $(\, , \,)$, and $(\, , \,)$ is a scalar product in X ($\longrightarrow$)

So X with the norm $\| \ \|$ and the scalar product $(\, , \,)$ is called **the product of unitary spaces** $X_1, \dots X_k$. And similarly as in the subspace case, we get the appropriate "Hilbert" result

$\boxed{\text{Fact}}$

If $X_1, \dots, X_k$ are Hilbert spaces, then the above defined product of unitary spaces $X_1, \dots, X_k$ is also a Hilbert space.

Proof

It follows from Fact "On product" p. PB-49 and from the fact that $\| \ \|_{[2]} \equiv \| \ \|_{[1]}$ in X (see p. PB-50). $\boxed{\text{Fu}}$

Note also, that we could also ask a natural question on the quotient space X/Y in the case when X is a unitary space and $Y \underset{lin}{\subset} X$, $Y = \overline{Y}$. Is X/Y (with its quotient norm) again a unitary space? We shall answer this question later — at least for the case of X being a Hilbert space.

The similar question can be asked about the completion operation: is each completion of a unitary space also a unitary space? — The answer is YES → ⚠ — so it is also a Hilbert space.

◊ Some extra unitary space properties

We prove here some further results concerning unitary spaces, related to the fact, that they are equipped with a norm of special kind. So, assume that X – a unitary space.

Fact („On scalar product continuity")
The scalar product $(\cdot,\cdot) : X \times X \longrightarrow \mathbb{K}$ is a continuous function.

Proof
Let $x_n \longrightarrow x$, $y_n \longrightarrow y$ in X. We have

$$|(x_n, y_n) - (x,y)| = |(x_n, y_n) - (x, y_n) + (x, y_n) - (x,y)|$$
$$\leq |(x_n - x, y_n)| + |(x, y_n - y)|.$$

So, by Schwarz inequality we get

$$|(x_n, y_n) - (x,y)| \leq \|x_n - x\| \cdot \|y_n\| + \|x\| \, \|y_n - y\| \longrightarrow 0$$

(because $\|y_n\| \longrightarrow \|y\|$, so $\{\|y_n\|\}$ is a bounded sequence)
which gives $(x_n, y_n) \longrightarrow (x,y)$. $\boxed{\text{L}}$.

Fact ("Pythagorean formula")
If $x_1, \dots, x_k \in X$ and they are pairwise orthogonal (i.e. $x_i \perp x_j$ for $i \neq j$), then

$$\|\sum_{j=1}^{k} x_j\|^2 = \sum_{j=1}^{k} \|x_j\|^2.$$

Proof

We have by sesquilinearity and orthogonality

$$\|\sum_{j=1}^{k} x_j\|^2 = \left(\sum_{j=1}^{k} x_j , \sum_{j=1}^{k} x_j\right) = \sum_{i=1}^{k}\sum_{j=1}^{k} (x_i , x_j) = \sum_{i=1}^{k} (x_i, x_i) = \sum_{i=1}^{k} \|x_i\|^2 .$$

$\boxed{\text{Ч}}$

Fact ("On isometry of unitary spaces")

Let A be a linear map from X onto Y, where X, Y – unitary spaces. Then TFCAE:

(i) A is an isometry

(ii) $\forall_{x_1, x_2 \in X} \quad (Ax_1 , Ax_2) = (x_1 , x_2)$.

Proof

(ii) $\Rightarrow$ (i) is obvious from the definition of the induced norms (we take $x = x_1 = x_2$).

(i) $\Rightarrow$ (ii) is easy to obtain by the linearity of A and the polarization formulae $\longrightarrow$ △. $\boxed{\text{Ч}}$

Fact ("On $M^\perp$")

For any $M \subset X$ $\quad M^\perp$ is a closed linear subspace of X.

Proof

We know, that $M^\perp \underset{lin}{\subset} X$ (see p.p. HS-5/6). Suppose that $\{x_n\}$ is a sequence with terms in $M^\perp$ and that $x_n \to x \in X$. Then for any $y \in M$ we have $x_n \perp y$, i.e $(x_n, y) = 0$. So $0 = (x_n, y) \longrightarrow (x, y)$, by Fact "On s.p. continuity", thus $0 = (x, y)$, i.e. $x \perp y$. Therefore $x \in M^\perp$. $\boxed{\text{E}}$

Remark

For any $M \subset X$
$$M^\perp = (\operatorname{lin} M)^\perp = (\overline{M})^\perp = (\overline{\operatorname{lin} M})^\perp.$$

Proof

If $M_1 \subset M_2 \subset X$, then $M_1^\perp \supset M_2^\perp$, by the definition of "$\cdot^\perp$".

so
$$M^\perp \supset (\operatorname{lin} M)^\perp \supset (\overline{\operatorname{lin} M})^\perp \text{ and}$$
$$M^\perp \supset (\overline{M})^\perp \supset (\overline{\operatorname{lin} M})^\perp.$$

But if $x \in M^\perp$, i.e. $x \perp M$, then also $x \perp \operatorname{lin} M$, because of the linearity of $(\cdot, x)$, hence $x \in (\operatorname{lin} M)^\perp$, which gives $M^\perp = (\operatorname{lin} M)^\perp$. Analogously, if $x \in M^\perp$, then also $x \perp \overline{M}$, by the continuity of $(\cdot, x)$, hence $M^\perp = (\overline{M})^\perp$. Using this to $\operatorname{lin} M$ we get $(\operatorname{lin} M)^\perp = (\overline{\operatorname{lin} M})^\perp$. $\boxed{\text{E}}$

2. Orthogonal projections

We start here from a specific property of Hilbert spaces, related to parallelogram identity – "an algebraic – geometric" formula. The property itself has rather "analytic – geometric" sense. And it has really crucial consequences for the "simplicity and beauty" of Hilbert spaces...

2.1 Reaching of the minimal distance and the orthogonality

♦ The minimal distance principle

Theorem ("The Minimal Distance Principle")**)

Let $\mathcal{H}$ *) be a Hilbert space, M – a non-empty closed convex subset of $\mathcal{H}$ and $x \in \mathcal{H}$. Then there exists exactly one $y \in M$ such, that $\|x-y\| = \operatorname{dist}(x, M)$.

Proof

Obviously $\tilde{M} := (-x) + M$ is also closed and convex (and $\neq \emptyset$), so it suffices to prove the case $x = 0$ only ($\operatorname{dist}(0, \tilde{M}) = \operatorname{dist}(x, M)$ obviously...). We start from the proof of

---

*) We shall often use the "caligraphic" $\mathcal{H}$ to denote a Hilbert space...

**) = "M.D.P."

uniqueness: suppose that y and y' in M satisfy $\| y \| = \text{dist}(0, M) = \| y' \|$. Then, by the parallelogram identity we have

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

But by the convexity of M also $\frac{y+y'}{2} \in M$, so $\left\| \frac{y+y'}{2} \right\| \geqslant d :=$
$= \text{dist}(0, M)$ and

$$0 \leq \| y - y' \|^2 \leq 2(d^2 + d^2) - 4d^2 = 0,$$

which gives $\| y - y' \| = 0$ and $y = y'$.

Now, we shall use just the same trick to prove the existence part:

by the definition of "dist", there exists $\{y_n\}_{n \geqslant 1}$ in Y such that $\| y_n \| \longrightarrow d$, and we shall prove, that $\{y_n\}_{n \geqslant 1}$ is a Cauchy sequence. Let $\varepsilon > 0$ and choose $N \geqslant 1$ such that $\| y_n \|^2 < d^2 + \varepsilon^2/4$ for any $n \geqslant N$. Then, for $m, n \geqslant N$

$$\| y_n - y_m \|^2 = 2(\| y_n \|^2 + \| y_m \|^2) - 4 \left\| \frac{y_n + y_m}{2} \right\|^2 < 2\left(2d^2 + \frac{\varepsilon^2}{2}\right) - 4d^2$$
$$= \varepsilon^2.$$

So $\{y_n\}_{n \geqslant 1}$ is Cauchy seq., and $\mathcal{H}$ is a Hilbert space, which gives $y_n \longrightarrow y$ for some $y \in \mathcal{H}$. But M is closed and $\| \cdot \|$ is continuous, hence $y \in M$, $d \longleftarrow \| y_n \| \longrightarrow \| y \|$, and $\| y \| = d$. $\boxed{\text{⌐}}$

$\boxed{\text{Note}}$

Compare now the above (very) strong result – namely its $x = 0$ case – to

the general normed space result: the Riesz Lemma (p. PB−13).
Now the remarks p. PB−14/15 on "almost perpendicular/orthogonal" vector seems more clear...

Orthogonal Projection

One of the most important direct applications of M.D.P. concerns M-s being closed linear subspaces, being much more then "convex only". Here, and later on, $\mathcal{H}$ is a Hilbert space.**)

Definition

For any closed linear subspace Y of $\mathcal{H}$ and $x \in \mathcal{H}$ we denote by

$$P_Y x$$

the unique vector $y \in Y$ satisfying $\|x-y\| = \operatorname{dist}(x, Y)$ (see M.D.P.), and we call it the orthogonal projection of x onto Y.*)
Moreover, we call the map $P_Y : X \to Y \subset X$ given by

$$P_Y(x) := P_Y x, \qquad x \in X$$

the orthogonal projection onto Y.*)

The names and notation used above suggest, that P_Y should be something more than "a map" only...

---

*) Pol.: rzut ortogonalny x na Y; rzut ortogonalny na Y (for P_Y)
**) The assumption "$\mathcal{H}$ is a unitary space" will be too weak for us, in general, for most of the forthcoming part...

Theorem ("On orthogonal projection")

Let Y be a closed linear subspace of $\mathcal{H}$.
The orthogonal projection P_Y onto Y has the following properties:

(i) $\forall_{x\in\mathcal{H}} \quad x-P_Y(x) \perp Y$

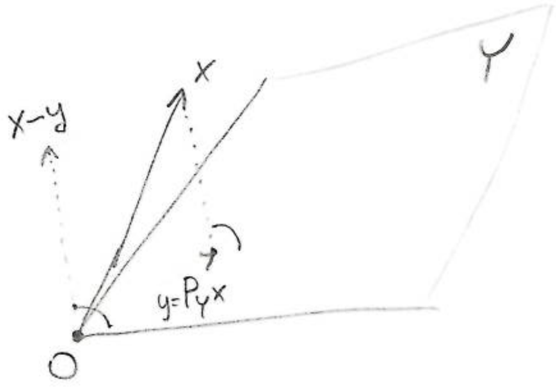
(ii) $P_Y \in B(\mathcal{H}),\ B(\mathcal{H},Y)$ [*)] and $\|P_Y\| = \begin{cases} 1 & \text{if } Y\neq\{0\} \\ 0 & \text{if } Y=\{0\}\end{cases}$

(iii) $(P_Y)^2 = P_Y$, $\operatorname{Ker} P_Y = Y^\perp$, $\operatorname{Ran} P_Y = Y$.

Moreover, for any $x,y\in\mathcal{H}$ TFCAE:

(iv) $\quad y = P_Y x$

(v) $\quad y\in Y$ and $x-y\perp Y$.



Lemma

If X - a unitary space and $v,w\in X$, then TFCAE: (a) $w\perp v$

(b) $\forall_{\lambda\in\mathbb{K}} \quad \|w\|\leq\|w+\lambda v\|$.

Proof (a) $\Rightarrow$ (b) $\quad \|w+\lambda v\|^2 = \|w\|^2 + \|\lambda v\|^2 \geqslant \|w\|^2$

by (the) Pythagorean formula ($\lambda v\perp w$ by (a)).

(b) $\Rightarrow$ (a) When $v=0$ then (a) holds; suppose that $v\neq 0$.

---

*) We can treat P_Y both as a map from $\mathcal{H}$ to $\mathcal{H}$, and as a map from $\mathcal{H}$ to Y.

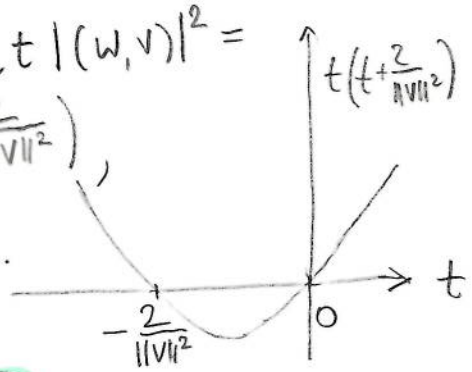
The value of the operator norm means the just the same for both "treatings", obviously ...

We have $\|w+\lambda v\|^2 = \|w\|^2 + |\lambda|^2\|v\|^2 + 2\,\mathrm{Re}\big(\bar{\lambda}(w,v)\big)$, hence by

(5) $\forall_{\lambda\in\mathbb{K}} \quad 0 \le |\lambda|^2\|v\|^2 + 2\,\mathrm{Re}\big(\bar{\lambda}(w,v)\big)$.

W can consider all $\lambda_t := (w,v)\cdot t$ with $t\in\mathbb{R}$, in particular.

Hence: $\forall_{t\in\mathbb{R}} \quad 0 \le |(w,v)|^2\cdot t^2\|v\|^2 + 2t\,|(w,v)|^2 = |(w,v)|^2\,\|v\|^2\, t\left(t + \frac{2}{\|v\|^2}\right)$,



which means, that $|(w,v)|^2 = 0$, so (a) holds.

$\boxed{L}$

Proof (of Theorem)

Fix $x\in\mathcal{H}$. If $y\in Y$, then we have, by Lemma:

$$x-y \perp Y \iff \forall_{y'\in Y}\,(x-y)\perp y' \iff \forall_{\substack{\lambda\in\mathbb{K}\\ y'\in Y}} \|x-y\| \le \|(x-y)+\lambda y'\| = \|x-(y-\lambda y')\| \iff \forall_{y''\in Y} \|x-y\| \le \|x-y''\| \iff \|x-y\| = \mathrm{dist}(x,Y).$$

Hence, by the definition of $P_Y(x)$, the part "Moreover..." holds.

In particular, it gives (i). Let $x_1, x_2\in\mathcal{H}$, $\lambda\in\mathbb{K}$, so $x_j - P_Y(x_j)\in Y^\perp$ for $j=1,2$, hence $(x_1+x_2) - \big(P_Y(x_1)+P_Y(x_2)\big) = (x_1 - P_Y(x_1)) + (x_2 - P_Y(x_2)) \in Y^\perp$ and $(\lambda x_1) - \lambda(P_Y(x_1)) = \lambda(x_1 - P_Y(x_1)) \in Y^\perp$, and $(P_Y(x_1)+P_Y(x_2)),\ \lambda(P_Y(x_1)) \in Y$. Thus, by "(v) $\Rightarrow$ (iv)"

$P_Y(x_1)+P_Y(x_2) = P_Y(x_1+x_2)$ and $\lambda(P_Y(x_1)) = P_Y(\lambda x_1)$ — so $P_Y\in\mathcal{L}(\mathcal{H})$.

If $x\in Y$ then $P_Y x \overset{*)}{=} x$, by definition. Hence $P_Y(P_Y x) = P_Y x$ for any $x\in\mathcal{H}$ (since $P_Y x\in Y$). This gives $\mathrm{Ran}\,P_Y = Y$ and $P_Y^2 = P_Y$.

---

*) We start to use $P_Y x$ in place of $P_Y(x)$, since $P_Y\in\mathcal{L}(\mathcal{H})$ is proved.

Again, by the "(iv) $\Leftrightarrow$ (v)" part, we get $P_Y x = 0 \iff x - 0 \in Y^\perp$, so $\operatorname{Ker} P_Y = Y^\perp$, which finishes the proof of (iii). Of course, if $Y = \{0\}$, then $P_Y = 0$, and $\|P_Y\| = 0$. Suppose that $Y \neq \{0\}$ and let $y \in S_Y(0,1)$. Then $\|P_Y y\| = \|y\| = 1$; moreover by (i) for any $x \in \mathcal{H}$
$x = P_Y x + (x - P_Y x)$ and $P_Y x \perp x - P_Y x$, so by the Pythagorean formula $\|x\|^2 \geq \|P_Y x\|^2$, which gives $P_Y \in B(X,X)$ and $\|P_Y\| = 1$.

$\boxed{\text{Ч}}$

2.2 Projections, direct sums and orthogonal decomposition

♦ Orthogonal decomposition and complement

We shall see now some advantages of possessing orthogonal projections on each closed subspace. First, let us study (recall?...) the general case of projections in a linear case.

Let X be a linear space.

Definition Consider $P \in \mathcal{L}(X)$. P is a projection (linear projection) *) iff

$$P^2 = P.$$

Denote: $\mathcal{LP}(X) := \{P \in \mathcal{L}(X) : P \text{ - projection}\}$.

---

*) Pol.: rzut (rzut liniowy)

Fact ("On linear projection")

0. If $P \in \mathcal{LP}(X)$, then $\operatorname{Ran} P = \{x \in X : Px = x\}$

1. If $P \in \mathcal{LP}(X)$, then $I - P \in \mathcal{LP}(X)$ and $\operatorname{Ker}(I-P) = \operatorname{Ran} P$, $\operatorname{Ran}(I-P) = \operatorname{Ker} P$.

2. If $P \in \mathcal{LP}(X)$, then $X = \operatorname{Ker} P \oplus \operatorname{Ran} P$.

3. If $X_1, X_2 \underset{lin}{\subset} X$ and $X = X_1 \oplus X_2$, then $\exists!\, P \in \mathcal{LP}(X)$ $X_1 = \operatorname{Ran} P$ and $X_2 = \operatorname{Ker} P$.

Proof For parts 0.–2. assume, that $P \in \mathcal{LP}(X)$.

0. „$\supset$" is obvious, since $Px \in \operatorname{Ran} P$ for any $x \in X$. If $x \in \operatorname{Ran} P$, then $x = Py$ for some $y \in X$, so $Px = P(Py) = Py = x$ and „$\subset$" also holds.

1. $(I-P)^2 = I^2 - PI - IP + P^2 = I - 2P + P = I - P$, which means that $I - P \in \mathcal{LP}(X)$. Now: $x \in \operatorname{Ker}(I-P) \iff x - Px = 0 \iff x = Px$ $\iff x \in \operatorname{Ran} P$, by 0. This proves that $\operatorname{Ker}(I-P) = \operatorname{Ran} P$ and so, using this for the projection $I-P$ in place of P, we get $\operatorname{Ker} P = \operatorname{Ker}(I-(I-P)) = \operatorname{Ran}(I-P)$.

2. Let $x \in \operatorname{Ker} P \cap \operatorname{Ran} P$, then by 0. $Px = x$ and $Px = 0$, so $x = 0$. It remains to prove, that $X = \operatorname{Ker} P + \operatorname{Ran} P$. But when $x \in X$, then $x = Px + (I-P)x$ and $(I-P)x \in \operatorname{Ran}(I-P) = \operatorname{Ker} P$, $Px \in \operatorname{Ran} P$.

3. With the assumptions of 3 for any $x \in X$ there exists exactly one pair $(x_1, x_2) \in X_1 \times X_2$ such that $x = x_1 + x_2$, so we define $Px := x_1$, where x_1, x_2 are as above.

Using this uniqueness of the decomposition "$x = x_1 + x_2$" one easily proves both: the linearity of P and the fact, that $P^2 = P$ ($\longrightarrow$) Analogously we check that $X_1 = \operatorname{Ran} P$ and $X_2 = \operatorname{Ker} P$. Thus, it remains to prove the uniqueness of P. Suppose, that P, P' both satisfy the "conditions of 3."
Then for any $x \in X$ we have

$$x = Px + (I-P)x = P'x + (I-P')x \quad \text{and } Px \in \operatorname{Ran} P = X_1,$$
$$P'x \in \operatorname{Ran} P' = X_1, \quad (I-P)x \in \operatorname{Ker} P = X_2, \quad (I-P')x \in \operatorname{Ker} P' = X_2$$

(e.g. by 1.). Hence, since $X = X_1 \oplus X_2$, we get

$$Px = P'x \quad \left(\text{and } (I-P)x = (I-P')x\right), \text{ so } P = P'!$$

Remark

In future (not so far) we shall prove a "similar" result concerning Banach spaces X and bounded linear projections ($P \in B(X) \cap \mathcal{L}P(X)$)...

We shall often use the following simple algebraic lemma

Lemma ("On unicity of direct sum")

Suppose that $X_1, X_2, X_1', X_2' \underset{\text{lin}}{\subset} X$. If $X_1 \oplus X_2 = X = X_1' \oplus X_2'$ and $X_j' \subset X_j$, $j = 1,2$, then $X_j' = X_j$, $j = 1,2$.

Proof

Let $x_1 \in X_1$, choose $x_1' \in X_1'$, $x_2' \in X_2'$ such that $x_1 = x_1' + x_2'$, so we have $x_1 + 0 = x_1' + x_2'$ and $x_1, x_1' \in X_1$, $0, x_2' \in X_2$. Using $X = X_1 \oplus X_2$ we get $x_1 = x_1'$ (and $0 = x_2'$), so $x_1 \in X_1'$, i.e. $X_1 \subset X_1'$. Thus $X_1' = X_1$ and the analogic argument proves that $X_2' = X_2$

Now we are ready to prove our "decomposition" result for orthogonal subspaces in a Hilbert space $\mathcal{H}$.

Theorem "On orthogonal decomposition"

If Y is a closed linear subspace of $\mathcal{H}$ then

(i) $\mathcal{H} = Y \oplus Y^{\perp}$

(ii) $(Y^{\perp})^{\perp} = Y$.

If M is a subset of $\mathcal{H}$, then $(M^{\perp})^{\perp} = \overline{\text{lin} M}$, in particular, when M is a linear subspace of $\mathcal{H}$, then $(M^{\perp})^{\perp} = \overline{M}$.

Proof

Since $P_Y \in \mathcal{LP}(\mathcal{H})$, then by Theorem "On linear proj."

$$\mathcal{H} = \text{Ran} P_Y \oplus \text{Ker} P_Y,$$ so we get (i) by Theorem "On orthogonal proj." p. (iii). But $Y^{\perp}$ is also a closed subspace (see Fact "On $M^{\perp}$" p. HS-20), so using (i) for it we get

$$\mathcal{H} = Y^{\perp} \oplus (Y^{\perp})^{\perp}.$$

On the other hand (i) (for Y) can be also written as

$$\mathcal{H} = Y^{\perp} \oplus Y,$$

and obviously $Y \subset (Y^{\perp})^{\perp}$, by the definition of the orthogonal complement. Hence, by Lemma "On unicity...", (ii) holds.

Now, if $M \subset \mathcal{H}$, then $(M^{\perp})^{\perp} = ((\overline{\text{lin} M})^{\perp})^{\perp} = \overline{\text{lin} M}$ by Remark p. HS-21 and then by (ii) (used to $Y := \overline{\text{lin} M}$).

The "in particular" part uses just $M = \text{lin} M$ for $M \underset{\text{lin}}{\subset} \mathcal{H}$. $\blacksquare$

Corollary 1

If $Y \underset{lin}{\subset} \mathcal{H}$, then $P_{(Y^\perp)} = I - P_{\overline{Y}}$; in particular $P_{(Y^\perp)} = I - P_Y$, when $Y = \overline{Y}$.

Proof

It suffices to prove the part for Y - closed, because $Y^\perp = \overline{Y}^\perp$ (see Rem. p. HS-21), so assume $Y = \overline{Y}$.

If $x \in \mathcal{H}$, then by Theorem "On orth. proj." we have $z := x - P_Y x \in Y^\perp$ and moreover, by Theorem "On orth. dec.", $x - z = P_Y x \in Y = (Y^\perp)^\perp$, i.e. $x - z \perp (Y^\perp)$.

Hence, using again Theorem "On orth. proj." (its "(v) ⇒ (iv)" part) for the subspase $Y^\perp$, we get

$$P_{(Y^\perp)} x = z = x - P_Y x = (I - P_Y) x.$$

∎

Note

The practical meaning of (i) of the above theorem is the following: each $x \in \mathcal{H}$ can be uniquely "decomposed" in the form

$$x = x_1 + x_2,$$

where $x_1 \in Y$ and $x_2 \in Y^\perp$ (if $Y = \overline{Y} \underset{lin}{\subset} \mathcal{H}$).

This explains the name of the theorem...

Corollary 2 ("A linear density criterion")

If $M \subset \mathcal{H}$, then TFCAE:

(i) M is linearly dense, (ii) $M^\perp = \{0\}$.

Proof

(i) $\Rightarrow$ (ii)

We have $M^\perp = (\overline{\operatorname{lin} M})^\perp$, by Remark p. HS-21, so by (i)

$$M^\perp = \mathcal{H}^\perp = \{0\} \quad (\text{if } x \in \mathcal{H}^\perp \text{ then } 0 = (x,x) = \|x\|^2, \text{ so } x = 0).$$

(ii) $\Rightarrow$ (i)

By (ii)

$(M^\perp)^\perp = \{0\}^\perp = \mathcal{H}$, so by Theorem "On orth. decomp."

$\overline{\operatorname{lin} M} = (M^\perp)^\perp = \mathcal{H}$, so (i) holds. $\boxed{}$

The quotient of Hilbert space

Recall our question p. HS-18 :
"Is X/Y a unitary space for X - unitary and Y - a closed subspace of X ?"
As we promissed, we answer it for the case $X = \mathcal{H}$ - - a Hilbert space.

We start from a result being a kind of "inversion" of Fact "On isometry of unitary spaces" (p. HS-20)

Fact
If X is a unitary space and Y is a normed space isometric*) to X, then Y is also a unitary space.

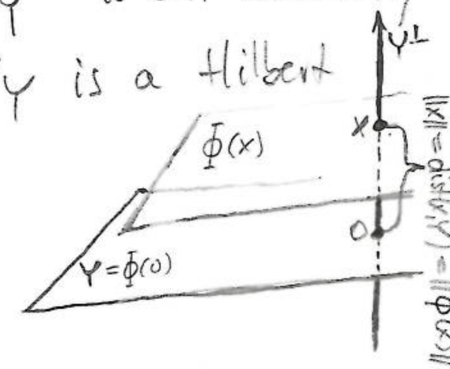
Proof
Let J be an isometry from X onto Y. By Theorem "On parallelogram ident." the norm $\| \ \|_X$ in X satisfies the parallelogram ident. So, $\| \ \|_Y$ - the norm in Y also satisfies it, because J preserves both $+, -$, and the norms ($\longrightarrow$ △ for the details...).
Hence, again by the same theorem, $\| \ \|_Y$ is induced by a scalar product in Y, i.e. Y is a unitary space. $\boxed{\text{I4}}$

---

*) Recall, that "isometric" refers always to linear isometriticy in this AFI lecture!

Theorem ("On the quotient of Hilbert space")

If Y is a closed linear subspace of a Hibert space $\mathcal{H}$ then $\Phi := \pi|_{Y^\perp} : Y^\perp \to \mathcal{H}/Y$ is an isometry between $Y^\perp$ and $\mathcal{H}/Y$. In particular, $\mathcal{H}/Y$ is a Hilbert space.*)



Proof

Obviously $\Phi \in \mathcal{L}(Y^\perp, \mathcal{H}/Y)$, because π is linear.

If $x \in \mathcal{H}$ then by Theorem "On ort. decomp."

$$x = y + z \quad \text{for some } y \in Y,\ z \in Y^\perp,$$

so $\quad [x] = [y+z] = [z] = \pi(z) = \Phi(z)$, which means that Φ is onto $\mathcal{H}/Y$. Recall that for $x \in X$

$\|[x]\| = \mathrm{dist}(x, Y)$. On the other hand, by the definition of $P_Y x$ (see p. HS-24), we have

$$\|x - P_Y x\| = \mathrm{dist}(x, Y) = \|[x]\|. \qquad (1)$$

But by Corollary p. HS-31 $x - P_Y x = P_{Y^\perp} x$, and thus for $x \in Y^\perp$ we obtain (by (1) and by $x = P_{Y^\perp} x \ldots$)

$$\|\Phi(x)\| = \|[x]\| = \|P_{Y^\perp} x\| = \|x\|.$$

---

*) With its quotient norm, of course.

So Φ is an isometry from $Y^\perp$ onto $\mathcal{H}/Y$.
By Fact p. HS-31 $\mathcal{H}/Y$ is thus a unitary space, because $Y^\perp$ is unitary space, as a linear subspace of the Hilbert space $\mathcal{H}$. Moreover $\mathcal{H}/Y$ is Banach space, because $\mathcal{H}$ is Hilbert. Hence $\mathcal{H}/Y$ is a unitary Banach space $=$
$=$ Hilbert space. $\boxed{\;}$

Remark

The above result means that the space $\mathcal{H}/Y$ can be identified with $Y^\perp$ in a natural way (namely, by: $Y^\perp \ni x \leftrightsquigarrow [x] \in \mathcal{H}/Y$).

3. Orthogonal / orthonormal systems and bases

Our main goal here is to study orthogonal (and orthonormal) bases in Hilbert spaces. They play the rôle somewhat similar to the rôle of bases (linear, i.e. of Hamel) in linear spaces. However, they are much more related to the analytic (the norm...) and geometric (the orthogonality...) structure of Hilbert space.

3.1. Orthogonal systems and sets

Let $\mathcal{A}$ be a set (we shall treat it as a set of "indices") and let $\{x_\alpha\}_{\alpha\in\mathcal{A}}$ be an indexed system*) of vectors in a unitary space X. And let $S\subset X$.

Definition

- $\{x_\alpha\}_{\alpha\in\mathcal{A}}$ is an orthogonal system**) iff $\forall_{\substack{\alpha,\beta\in\mathcal{A}\\ \alpha\neq\beta}} \ x_\alpha \perp x_\beta$;
- S is an orthogonal set**) iff $\forall_{\substack{x,y\in S\\ x\neq y}} \ x\perp y$.
- $\{x_\alpha\}_{\alpha\in\mathcal{A}}$ is an orthonormal system***) iff it is an orthogonal system, and $\forall_{\alpha\in\mathcal{A}} \ \|x_\alpha\|=1$.
- S is an orthonormal set***) iff S is an orthogonal set and $\forall_{x\in S} \ \|x\|=1$.

---

*) i.e., $\{x_\alpha\}_{\alpha\in\mathcal{A}}$ is a function from $\mathcal{A}$ to $\mathcal{H}$, from the formal point of view.

**) Pol.: układ / zbiór ortogonalny. Some people assume "extra" that $\forall_{\alpha\in\mathcal{A}} \ x_\alpha\neq 0$ ($\forall_{x\in S} \ x\neq 0$, for orth. set), but we DON'T! – see also the definition of n-d. orth. system/set, p. HS-47.

***) Pol.: ortonormalny

We shall use often the abbreviation o.s. for orthogonal system, and o.set (for orthogonal set.)
In some parts of subsection 3 we shall use rather orthogonal systems than sets, because of some notation convenience. Note however, that having an orthogonal set S we can "canonically" get an orthogonal system $\{x_s\}_{s\in S}$ just defining $x_s := s$ for $s \in S$. But the "inverse", i.e., taking $S := \{x_\alpha : \alpha \in \mathcal{A}\}$ for an o.s. $\{x_\alpha\}_{\alpha\in\mathcal{A}}$ can be sometimes not a good idea: if there is "a lot of" zero-vectors x_α among all x_α-s (which can be important for us), we get "only one" $0 \in S$...
Thus, for $\{x_\alpha\}_{\alpha\in\mathcal{A}}$ - an o.s. denote:

$$\operatorname{supp}\{x_\alpha\}_{\alpha\in\mathcal{A}} := \{\alpha \in \mathcal{A} : x_\alpha \neq 0\}$$

(the support*) of $\{x_\alpha\}_{\alpha\in\mathcal{A}}$). We shall call $\{x_\alpha\}_{\alpha\in\mathcal{A}}$ quasi-countable**) iff $\operatorname{supp}\{x_\alpha\}_{\alpha\in\mathcal{A}}$ is at most countable set.
E.g., if $\mathcal{A}$ is a finite or countable set, then $\{x_\alpha\}_{\alpha\in\mathcal{A}}$ is quasi-countable. Denote by QCO (or by QCO(X), QCO($\mathcal{A}$,X), when we should be more precise) the set of all quasi-countable orthogonal systems for the set $\mathcal{A}$ and the unitary space X.

The sum of a "quasi-countable" orthogonal system

We shall need a notion of the convergence and of the sum for "quasi-countable" orthogonal systems. The idea is quite

---

*) Pol.: nośnik.
**) Pol.: quasi-przeliczalny

obvious – when "summing", we (can) restrict the summation only to the nonzero terms x_α. – When their number is finite, the definition is trivial, and when it is countable – we use just the notion of the sum of a serie in a normed space. However, to make it properly, we should be carefull **), and we must perform several "formal steps"...
We start from the most crucial result for our definition. It concerns orthogonal sequences, i.e. O.S. with $A = \mathbb{N}$ (or $A = \mathbb{N}_{n_0}$ for some n_0 *)).

$\boxed{\text{Fact}}$ ("On orthogonal sequences")

Let $\{x_n\}_{n\in\mathbb{N}_1}$ be an O.S. in a Hilbert space $\mathcal{H}$.

Then:

(1) $\sum_{n=1}^{+\infty} x_n$ is convergent iff $\sum_{n=1}^{+\infty} \|x_n\|^2 < +\infty$

(2) If $\sum_{n=1}^{+\infty} x_n$ is convergent, then

$$\left\|\sum_{n=1}^{+\infty} x_n\right\|^2 = \sum_{n=1}^{+\infty} \|x_n\|^2, \qquad \text{(GPF)}$$

and $\sum_{n=1}^{+\infty} x_n$ is unconditionally convergent, i.e., for any bijection (permutation) p of $\mathbb{N}_1$ the series $\sum_{n=1}^{+\infty} x_{p(n)}$ is convergent, and

$$\sum_{n=1}^{+\infty} x_{p(n)} = \sum_{n=1}^{+\infty} x_n \quad \bullet$$

---

*) Recall: $\mathbb{N}_{n_0} := \{n\in\mathbb{Z} : n \geqslant n_0\}$ for $n_0 \in \mathbb{Z}$.

**) Why to be carefull? And why we do this only for O.S., and not just for any quasi-countable systems in Banach spaces, where the sums of series are "properly" defined? – The main reason is the problem of independance of the "order of summation", as we shall see very soon...

Remarks

1. The part (1) + the first part of (2) are sometimes jointly called "generalized/infinite Pythagorean formula" ("GPF").

2. As we shall see below, if we assume that $\{x_n\}_{n\in\mathbb{N}_1}$ is an o.s. in a unitary space X only (instead of the Hilbert $\mathcal{H}$), then we have "$\Longrightarrow$" in (1) and the first part of (2) is also true.

Proof (of Fact)

(1) Since both $\mathcal{H}$ and $\mathbb{R}$ are complete spaces, thus the convergences of both series in (1) are equivalent to the appropriate Cauchy conditions. But by the Pythagorean formula, if $n \geq m$ then

$$\left\|\sum_{k=m}^{n} x_k\right\|^2 = \sum_{k=m}^{n} \|x_k\|^2,$$

so those Cauchy conditions are eqivalent one to the other).

(2) Suppose, that $\sum_{n=1}^{+\infty} x_n$ is convergent. Then, again by Pyth. form.,

$$\left\|\sum_{n=1}^{\infty} x_n\right\|^2 = \lim_{n\to+\infty} \left\|\sum_{k=1}^{n} x_k\right\|^2 = \lim_{n\to+\infty} \sum_{k=1}^{n} \|x_k\|^2 = \sum_{n=1}^{+\infty} \|x_n\|^2,$$

so (GPF) holds. Now, by (1) and the fact that the scalar series $\sum_{n=1}^{+\infty} \|x_n\|^2$ is absolutely convergent (and so – also unconditionally) we get the convergence of $\sum_{n=1}^{+\infty} x_{p(n)}$ for any permutation $p: \mathbb{N}_1 \longrightarrow \mathbb{N}_1$. Fix now the permutation p and denote for $n \in \mathbb{N}_1$

$$L(n) := \{ p(j) : 1 \leq j \leq n \}.$$

Using the fact, that p is onto $\mathbb{N}_1$ let us choose for any $k \in \mathbb{N}_1$ such $n(k) \in \mathbb{N}_1$ that $\{1, \dots, k\} \subset L(n(k))$.
We have $k \leq \# L(n(k)) = n(k)$ - by the fact that p is injective. Hence

$$\lim_{k \to +\infty} n(k) = +\infty. \qquad (2)$$

Denote $S_k := \sum_{j=1}^{k} x_j$, $\tilde{S}_k := \sum_{j=1}^{k} x_{p(j)}$ and let
$S := \sum_{j=1}^{+\infty} x_j$, $\tilde{S} := \sum_{j=1}^{+\infty} x_{p(j)}$.

Now, using once again the Pythagorean formula and the injectivity of p we get

$$\| \tilde{S}_{n(k)} - S_k \|^2 = \Big\| \sum_{m \in L(n(k)) \setminus \{1, \dots, k\}} x_m \Big\|^2 = \sum_{m \in L(n(k)) \setminus \{1, \dots, k\}} \| x_m \|^2 \leq$$

$$\leq \sum_{m=k+1}^{+\infty} \| x_m \|^2 \xrightarrow[k \to +\infty]{} 0, \text{ by } (1).$$

Hence, $\left(\tilde{S}_{n(k)} - S_k \right) \xrightarrow[k \to +\infty]{} 0$, but by (2) also

$$\left(\tilde{S}_{n(k)} - S_k \right) \xrightarrow[k \to +\infty]{} \tilde{S} - S,$$

i.e., $\tilde{S} = S$. $\boxed{\text{q}}$

Suppose now, that $\{x_\alpha\}_{\alpha \in A}$ is an (quasi-countable) O.S. in a Hilbert space $\mathcal{H}$. We shell define:

1°. the notion: $\sum_{\alpha \in A} x_\alpha$ is convergent / divergent

2°. for the case of $\sum_{\alpha \in A} x_\alpha$ convergent the notion of $\sum_{\alpha \in A} x_\alpha$ (— the vector of $\mathcal{H}$) called also the sum (of) $\sum_{\alpha \in A} x_\alpha$ or the sum of $\{x_\alpha\}_{\alpha \in A}$

Denote $\tilde{A} := \operatorname{supp}\{x_\alpha\}_{\alpha \in A}$. If $\tilde{A}$ is finite, then we always say convergent (and never-divergent) about $\sum_{\alpha \in A} x_\alpha$. Moreover we define $\sum_{\alpha \in A} x_\alpha := \sum_{\alpha \in \tilde{A}} x_\alpha$, where the RHS means the usual finite sum (the inductive definition with respect to $\# \tilde{A}$, obviously independent of the way of numbering).*)

Now, to make 1° for $\{x_\alpha\}_{\alpha \in A} \in QCO$ and $\tilde{A}$-infinite we consider all the set of all "numberings" for $\{x_\alpha\}_{\alpha \in A}$:

$$\mathrm{Num}\ \left(= \mathrm{Num}(\{x_\alpha\}_{\alpha \in A}\right) := \{\gamma : \mathbb{N}_1 \to \tilde{A} : \gamma \text{ is a bijection}\}.$$

Observe that by standard properties of non-negative scalar series we have:

$\boxed{\text{Fact 1}}$

The value (in $[0; +\infty]$) of $\sum_{n=1}^{+\infty} \|x_{\gamma(n)}\|^2$ does not depend on $\gamma \in \mathrm{Num}$.

---

*) And obviously, the fact that $\{x_\alpha\}_{\alpha \in A}$ is an O.S. is not important for $\tilde{A}$-finite.

Obviously $Num \neq \emptyset$, because $\# \tilde{A} = \# \mathbb{N}_1$ in this case, so we can define the convergence as follows:

Definition

$\sum_{\alpha \in A} x_\alpha$ is convergent iff for any (equivalently: "some" – by Fact 1) $\gamma \in Num$ the series $\sum_{n=1}^{+\infty} \|x_{\gamma(n)}\|^2$ is convergent.
In the opposite case $\sum_{\alpha \in A} x_\alpha$ is divergent.

Now, to make 2° we need the following result related to Fact "On orthogonal sequences".

Fact 2

If $\sum_{\alpha \in A} x_\alpha$ is convergent, then the series $\sum_{n=1}^{+\infty} x_{\gamma(n)}$ converges for any $\gamma \in Num$ and (the sum of) $\sum_{n=1}^{+\infty} x_{\gamma(n)}$ does not depend on $\gamma \in Num$.

Proof

The convergence of $\sum_{n=1}^{+\infty} x_{\gamma(n)}$ for each $\gamma \in Num$ follows from Fact "On orth. seq" part (1). Suppose that $\gamma, \gamma' \in Num$, then $p := \gamma^{-1} \circ \gamma'$ is a permutation of $\mathbb{N}_1$, so by the same fact, part (2) used for O.S. $\{y_n\}_{n \in \mathbb{N}_1} := \{x_{\gamma(n)}\}_{n \in \mathbb{N}_1}$ we get

$$\sum_{n=1}^{+\infty} x_{\gamma(n)} = \sum_{n=1}^{+\infty} y_n = \sum_{n=1}^{+\infty} y_{p(n)} = \sum_{n=1}^{+\infty} x_{\gamma(p(n))} = \sum_{n=1}^{+\infty} x_{(\gamma \circ \gamma^{-1} \circ \gamma')(n)} = \sum_{n=1}^{+\infty} x_{\gamma'(n)}.$$

$\boxed{\text{h}}$

Definition

If $\sum_{\alpha\in A} x_\alpha$ is convergent, then the sum of $\sum_{\alpha\in A} x_\alpha$, denoted also by $\sum_{\alpha\in A} x_\alpha$ is the sum of the series $\sum_{n=1}^{+\infty} x_{\gamma(n)}$ for any (equivalently: "some" – by Fact 2) $\gamma \in Num$.

Remark

Note, that there is a possible risk of confusion related to the above two definitions. In the case of A - finite it is simpler – the symbol $\sum_{\alpha\in A} x_\alpha$ has two meanings:

(1) the "usual" finite sum of vectors

(2) the above (p. HS-41) defined $\sum_{\alpha\in A} x_\alpha := \sum_{\alpha\in \hat{A}} x_\alpha$.

– Fortunately, both meanings give the same value (the difference is just finite sum of zeros...).

But the problem exists also for the case of $A = \mathbb{N}_{n_0}$ – it is not so "strict", because the symbols

(i) $\sum_{n\in \mathbb{N}_{n_0}} x_n$ and (ii) $\sum_{n=n_0}^{+\infty} x_n$ do not coincide,

however they look very similar! - Does the both values coincide? – YES, but it should be proved...

The same question, also with positive answer, concerns the problem of the convergence \ divergence of (i) and (ii).

The proofs follow immediately from the simple fact below ($\longrightarrow$)

Lemma

Let X be a normed space and $\{x_n\}_{n\in\mathbb{N}_1}$ – a sequence in X with infinitely many non-zero terms. Let $k_n \in \mathbb{N}_1$ be the index of the n-th non-zero term of $\{x_n\}_{n\in\mathbb{N}_1}$ *) for any $n \in \mathbb{N}_1$. Then $\sum_{n=1}^{+\infty} x_n$ is convergent iff $\sum_{n=1}^{+\infty} x_{k_n}$ is convergent, and if those series are convergent, then they converge to the same vector of X.

Proof $\longrightarrow$

---

*) i.e., $\{k_n\}_{n\in\mathbb{N}_1}$ is strictly increasing and $m \in \mathbb{N}_1$ has the form k_n for some $n \in \mathbb{N}_1$ iff $x_m \neq 0$.

Bessel's Lemma and the orthogonal projection formula

Lemma ("The Bessel lemma")

Suppose that $\{x_\alpha\}_{\alpha \in A}$ is an O.S. *) in a unitary space X and $x \in X$. Then:

(i) ("the Bessel inequality") if $\tilde{x}_\alpha := \begin{cases} 0 & \text{if } x_\alpha = 0 \\ x_\alpha / \|x_\alpha\| & \text{if } x_\alpha \neq 0 \end{cases}$ for any $\alpha \in A$, then

$$\sum_{\alpha \in A} |(x, \tilde{x}_\alpha)|^2 \leq \|x\|^2, \qquad (1)$$

(ii) $\{\alpha \in A : (x, x_\alpha) \neq 0\}$ is at most countable.

Proof

(i) By the definition of a finite sum of non-negative numbers, to get (1) it is sufficient to prove

$$\sum_{\alpha \in A'} |(x, \tilde{x}_\alpha)|^2 \leq \|x\|^2 \qquad (1')$$

for any finite subset A' of A. So, fix such A' and consider $\mathcal{H} := \mathrm{lin}(\{\tilde{x}_\alpha : \alpha \in A'\} \cup \{x\})$ and

$$Y := \mathrm{lin}\{\tilde{x}_\alpha : \alpha \in A'\}.$$

Since the dimension of $\mathcal{H}$ is finite, $\mathcal{H}$ is a Hilbert space (a subspace of X) and Y is its closed linear subspace.

*) It is NOT assumed, that it is in QCO!

Let $y := \sum_{\alpha \in A'} (x, \tilde{x}_\alpha)\tilde{x}_\alpha$. We have $y \in Y$ and for any $\beta \in A'$, by ortogonality of $\{\tilde{x}_\alpha\}_{\alpha \in A}$ we also have

$$(x-y, \tilde{x}_\beta) = (x, \tilde{x}_\beta) - \sum_{\alpha \in A'} (x, \tilde{x}_\alpha)(\tilde{x}_\alpha, \tilde{x}_\beta) =$$

$$= (x, \tilde{x}_\beta) - (x, \tilde{x}_\beta)(\tilde{x}_\beta, \tilde{x}_\beta),$$

but $\tilde{x}_\beta = 0$ or $(\tilde{x}_\beta, \tilde{x}_\beta) = \|\tilde{x}_\beta\|^2 = 1$, hence

$$(x-y, \tilde{x}_\beta) = 0.$$

Thus, by the definition of Y we get $x - y \perp Y$, so $y = P_Y x$ (where P_Y denotes the orthogonal projection onto Y in $\mathcal{H}$) by Theorem "On orthogonal projection". Therefore $\|y\| = \|P_Y x\| \leq \|x\|$ by the same theorem, and $\|y\|^2 = \sum_{\alpha \in A'} \|(x, \tilde{x}_\alpha)\tilde{x}_\alpha\|^2 = \sum_{\alpha \in A'} |(x, \tilde{x}_\alpha)|^2$ by the Pythagorean formula, so $(1')$ holds.

(ii) Note first, that $\{\alpha \in A : (x, x_\alpha) \neq 0\} = \{\alpha \in A : (x, \tilde{x}_\alpha) \neq 0\} = \bigcup_{n \in \mathbb{N}_1} A_n$, where $A_n := \{\alpha \in A : |(x, \tilde{x}_\alpha)| > \frac{1}{n}\}$. So it sufices to prove that each A_n is finite. Suppose, that it is infinite for some $n \in \mathbb{N}_1$. Then there exists an <u>injective</u> sequence $\{\alpha_k\}_{k \in \mathbb{N}_1}$ of elements of A_n, so $\{x_{\alpha_k}\}_{k \in \mathbb{N}_1}$ is also an o.s. in X. Thus, using (i) for this o.s., we get:

$$\|x\|^2 \geq \sum_{k \in \mathbb{N}_1} |(x, \tilde{x}_{\alpha_k})|^2 > \sum_{k \in \mathbb{N}_1} \frac{1}{n^2} = +\infty \quad \text{– a contradiction!}$$

$\boxed{\text{Ы}}$

Having Bessel's lemma we can easily obtain formulae for scalar products and norms of sums of quasi-countable orthogonal systems.

Recall first the notion of the sum for $f : \mathcal{A} \to \mathbb{K}$ satisfying $\sum_{\alpha \in \mathcal{A}} |f(\alpha)| < +\infty$. The definition of the sum $\sum_{\alpha \in \mathcal{A}} f(\alpha)$ via the definitions of such sums for non-negative functions was recalled in *) p. OF-60, and it can be equivalently defined by the integral with respect to the measure $\#$ for $\mathcal{A}$, i.e., we have

$$\sum_{\alpha \in \mathcal{A}} f(\alpha) = \int_{\mathcal{A}} f \, d\#$$

(and in particular $\int_{\mathcal{A}} |f| \, d\# = \sum_{\alpha \in \mathcal{A}} |f(\alpha)| < +\infty$).

The following result can be easily proved by the standard use of the (Monotone or) Dominated Convergence Theorem for integrals.

$\boxed{\text{Fact}}$

Suppose that $f : \mathcal{A} \to \mathbb{K}$ and $\gamma : \mathbb{N}_1 \to \mathcal{A}$ satisfy

(i) γ is injective (ii) $\operatorname{supp} f \subset \gamma(\mathbb{N}_1)$.

Then:

(a)
$$\sum_{\alpha \in \mathcal{A}} |f(\alpha)| = \sum_{n=1}^{+\infty} |f(\gamma(n))|,$$

(b) if $\sum_{\alpha \in \mathcal{A}} |f(\alpha)| < +\infty$, then

$$\sum_{\alpha \in \mathcal{A}} f(\alpha) = \sum_{n=1}^{+\infty} f(\gamma(n)).$$

$\boxed{\text{Proof}}$ $\longrightarrow$

Now we are ready to prove the following formulae for sums of orthogonal systems.

Fact

Suppose that $\mathcal{H}$ is a Hilbert space and $\{x_\alpha\}_{\alpha\in\mathcal{A}} \in QCO(\mathcal{H})$. Then:

(i) if $\sum_{\alpha\in\mathcal{A}} x_\alpha$ is convergent and $x = \sum_{\alpha\in\mathcal{A}} x_\alpha$, then for any $y \in \mathcal{H}$ $\sum_{\alpha\in\mathcal{A}} |(y, x_\alpha)| \leq \|y\|\cdot\|x\|$ and

$$(y, x) = \sum_{\alpha\in\mathcal{A}} (y, x_\alpha),$$

(ii) $\sum_{\alpha\in\mathcal{A}} x_\alpha$ is convergent iff $\sum_{\alpha\in\mathcal{A}} \|x_\alpha\|^2 < +\infty$ and if $\sum_{\alpha\in\mathcal{A}} \|x_\alpha\|^2 < +\infty$, then $\|\sum_{\alpha\in\mathcal{A}} x_\alpha\|^2 = \sum_{\alpha\in\mathcal{A}} \|x_\alpha\|^2$.

Proof

Let $\tilde{\mathcal{A}} = \operatorname{supp}\{x_\alpha\}_{\alpha\in\mathcal{A}}$. By the definition of $\sum_{\alpha\in\mathcal{A}} x_\alpha$ it is convenient to study two cases: $\tilde{\mathcal{A}}$ - finite and $\tilde{\mathcal{A}}$ - countable. The proof of (i) and (ii) is obvious for the first case. For $\tilde{\mathcal{A}}$ - countable consider any $\gamma \in Num(\{x_\alpha\}_{\alpha\in\mathcal{A}})$ and define $\{\tilde{x}_\alpha\}_{\alpha\in\mathcal{A}}$ as in Bessel's lemma. We have

$$\forall_{\alpha\in\mathcal{A}} \quad x_\alpha = \|x_\alpha\|\cdot\tilde{x}_\alpha,$$

thus, by the Shwarz inequality*) and then by Bessel's inequality

$$\sum_{\alpha\in\mathcal{A}} |(y, x_\alpha)| = \sum_{\alpha\in\mathcal{A}} \|x_\alpha\|\,|(y, \tilde{x}_\alpha)| \leq \Big(\sum_{\alpha\in\mathcal{A}} \|x_\alpha\|^2\Big)^{\frac12} \Big(\sum_{\alpha\in\mathcal{A}} |(y, \tilde{x}_\alpha)|^2\Big)^{\frac12} \leq$$

---

*) in $\ell^2(\mathcal{A})$ or just by Hölder inequality for $p=2$ for the integrals with respect the measure $\#$ on $\mathcal{A}$.

$$\le \left(\sum_{\alpha\in A}\|x_\alpha\|^2\right)^{1/2}\cdot\|y\|.$$ (1)

Observe, that (by the definition of the convergence) $\sum_{\alpha\in A} x_\alpha$ is convergent iff $\sum_{n=1}^{+\infty}\|x_{\gamma(n)}\|^2 < +\infty$. On the other hand

$\sum_{n=1}^{+\infty}\|x_{\gamma(n)}\|^2 = \sum_{\alpha\in A}\|x_\alpha\|^2$, by Fact p. HS-47 (a) – this gives the first part of (ii).

So, if $\sum_{\alpha\in A}\|x_\alpha\|^2 < \infty$, then $\sum_{\alpha\in A} x_\alpha$ is convergent and using again (the above equality and the def. of the sum of o.s.) we get

$$\Big\|\sum_{\alpha\in A} x_\alpha\Big\|^2 = \Big\|\sum_{n=1}^{+\infty} x_{\gamma(n)}\Big\|^2 = \sum_{\alpha\in A}\|x_\alpha\|^2,$$

which proves the second part of (ii). Now, by (1) and (ii) we have

$$\sum_{\alpha\in A}|(y,x_\alpha)| \le \|x\|\cdot\|y\| \text{ under the assumptions of (i).}$$

To finish the proof of (i) we can use Fact p. 47 (b), the continuity of $(\ ,\ )$, and the fact that $x = \sum_{\alpha\in A} x_\alpha = \sum_{n=1}^{+\infty} x_{\gamma(n)}$:

$$\sum_{\alpha\in A}(y,x_\alpha) = \sum_{n=1}^{+\infty}(y,x_{\gamma(n)}) = \Big(y,\sum_{n=1}^{+\infty}x_{\gamma(n)}\Big) = (y,x). \qquad \boxed{\text{KC}}$$

This result allows us to obtain quite easily a very important formula for the orthogonal projection onto a space being a "closed span" of any o. s. (not necessary belonging to QCO).

Theorem ("The orthogonal projection formula I")

Suppose that $\{x_\alpha\}_{\alpha\in A}$ is an o.s. in a Hilbert space $\mathcal{H}$. Let

$$Y := \overline{\mathrm{lin}\{x_\alpha : \alpha \in A\}}$$

and denote $\tilde{x}_\alpha := \begin{cases} 0 & \text{if } x_\alpha = 0, \\ \frac{x_\alpha}{\|x_\alpha\|} & \text{if } x_\alpha \neq 0. \end{cases}$

Then for any $x \in \mathcal{H}$ $\{(x,\tilde{x}_\alpha)\cdot\tilde{x}_\alpha\}_{\alpha\in A} \in QCO$ and

$$P_Y x = \sum_{\alpha\in A} (x,\tilde{x}_\alpha)\tilde{x}_\alpha ; \qquad (2)$$

in particular the sum $\sum_{\alpha\in A}(x,\tilde{x}_\alpha)\tilde{x}_\alpha$ is convergent and

$$\Big\|\sum_{\alpha\in A}(x,\tilde{x}_\alpha)\tilde{x}_\alpha\Big\|^2 = \sum_{\alpha\in A}|(x,\tilde{x}_\alpha)|^2 \leq \|x\|^2. \qquad (3)$$

Proof

By Bessel lemma we get $\{(x,\tilde{x}_\alpha)\cdot\tilde{x}_\alpha\}_{\alpha\in A} \in QCO$ by its part (ii). So, we can use the Fact p. HS-48 for it, and we get (the) convergence of $\sum_{\alpha\in A}(x,\tilde{x}_\alpha)\tilde{x}_\alpha$, because $\sum_{\alpha\in A}\|(x,\tilde{x}_\alpha)\tilde{x}_\alpha\|^2 = \sum_{\alpha\in A}|(x,\tilde{x}_\alpha)|^2 \leq \|x\|^2$, by the Bessel inequality; moreover we get also (3) by the part (ii) of the above Fact. To prove (2) observe first, that the RHS of (2) belongs to Y, by the definition of the sum of o.s., because

$\tilde{x}_\alpha \in Y$ for any $\alpha \in A$. So, by Theorem "On orth. projection", to prove (2) is suffices to check, that

$$z := \Big(x - \sum_{\alpha \in A} (x, \tilde{x}_\alpha)\tilde{x}_\alpha \Big) \in Y^\perp. \qquad (4)$$

But if $\beta \in A$, then by part (i) of the Fact p. HS-48 we have

$$(\tilde{x}_\beta, z) = (\tilde{x}_\beta, x) - \Big(\tilde{x}_\beta, \Big(\sum_{\alpha \in A}(x, \tilde{x}_\alpha)\tilde{x}_\alpha\Big)\Big) =$$

$$= (\tilde{x}_\beta, x) - \sum_{\alpha \in A} \overline{(x, \tilde{x}_\alpha)}(\tilde{x}_\beta, \tilde{x}_\alpha) = (\tilde{x}_\beta, x) - (\tilde{x}_\beta, x)\|\tilde{x}_\beta\|^2$$

(we use the ortogonality of $\{\tilde{x}_\alpha\}_{\alpha \in A}$), hence

$$(\tilde{x}_\beta, z) = \left\{ \begin{matrix} 0 - 0 & \text{for } x_\beta = 0 \\ (\tilde{x}_\beta, x) - (\tilde{x}_\beta, x)\cdot 1 & \text{for } x_\beta \neq 0 \end{matrix} \right\} = 0.$$

This means, that $\forall_{\beta \in A}\ x_\beta \perp z$ i.e. $z \in \{x_\alpha : \alpha \in A\}^\perp$.

But $Y^\perp = \{x_\alpha : \alpha \in A\}^\perp$, by Remark p. HS-21, so (4) holds.

$\boxed{\text{M}}$

$\boxed{\text{Corollary}}$ ("The orthogonal projection formula II")

If $\{x_\alpha\}_{\alpha \in A}$ is an ortho<u>normal</u> system in a Hilbert space $\mathcal{H}$ and $Y = \overline{\operatorname{lin}\{x_\alpha : \alpha \in A\}}$, then for any $x \in \mathcal{H}$ the system $\{(x, x_\alpha)x_\alpha\}$ is a quasi-countable o.s., the sum $\sum_{\alpha \in A}(x, x_\alpha)x_\alpha$ is convergent and

$$P_Y x = \sum_{\alpha \in A} (x, x_\alpha)x_\alpha, \quad \|P_Y x\|^2 = \sum_{\alpha \in A} |(x, x_\alpha)|^2 \leq \|x\|^2. \qquad (5)$$

$\boxed{\text{Proof}}$ We use the previous theorem, so the assertion is obvious by $\tilde{x}_\alpha = x_\alpha$. $\boxed{\text{M}}$

3.2. Orthogonal bases

We distinguish here special kinds of orthogonal sets / systems – the orthogonal bases. – To do this, we first restrict our attention to sligtly smaller family of orth. sets/systems than the studied before ones.

Non-degenerated orthogonal systems

Definition *)

An o.s. $\{x_\alpha\}_{\alpha\in A}$ in a unitary space X is non-degenerated iff $\forall_{\alpha\in A}\ x_\alpha \neq 0$. An o. set $S \subset X$ is non-degenerated iff $0 \notin S$.

We use the abbreviation n-d. for "non-degenerated" – in particular – we shall use: n-d. o.s. and n-d. o. set.

---

*) See **) p. HS-36.

Remarks

1. If $\{x_\alpha\}_{\alpha\in A}$ is n-d. o.s., then $A \ni \alpha \mapsto x_\alpha \in S := \{x_\beta : \beta \in A\}$ is a bijection, because $x_\alpha \perp x_\beta$ and $x_\alpha = x_\beta$ give $(x_\alpha, x_\alpha) = \|x_\alpha\|^2 = 0$, i.e., $x_\alpha = 0$.

So, the notions of n-d. o.s. and of n-d o. set are "equivalent"/"almost the same" *), in contrast to the pair of notions: o.s and o. set – see p. HS-37.

2. If $\{x_\alpha\}_{\alpha\in A}$ is o.s, and $\tilde{A} := \operatorname{supp}\{x_\alpha\}_{\alpha\in A}$, then $\{x_\alpha\}_{\alpha\in\tilde{A}}$ is n-d. o.s. Analogously: if S is o. set, then $\tilde{S} := S\setminus\{0\}$ is n-d. o.s. Moreover, obviously,

$$\operatorname{lin}\{x_\alpha : \alpha\in\tilde{A}\} = \operatorname{lin}\{x_\alpha : \alpha\in A\},\ \operatorname{lin}\tilde{S} = \operatorname{lin} S.$$

3. Despite 1. and 2. above, we needed the general notion of o.s. (without the "n-d." restriction) to define the convergence ("summability") and the sum for some quasi-countable o.s.-s.
They could often be not n-d. o.s., as we saw e.g. in "The ort. proj. form." (see also the Fourier expansion, p. HS-58)
Namely, by Bessel's lemma, for any <u>orthonormal</u> system $\{x_\alpha\}_{\alpha\in A}$ and $x \in X$ the "new" system

$$\{(x, x_\alpha)x_\alpha\}_{\alpha\in A}$$

is still an orthogonal, but often <u>no</u> n-d. orth. sytem (e.g., it is <u>never</u> n-d.o.s., if A is incountable...) •

---

*) More precisely: $\{x_\alpha\}_{\alpha\in A}$ is n-d. o.s. $\Leftrightarrow$ ($S := \{x_\alpha : \alpha\in A\}$ is n-d. o. set <u>and</u> $A \ni \alpha \mapsto x_\alpha \in S$ is a bijection).
($\to \Delta$)

The Remark 1. can be even strengthened.

$\boxed{\text{Fact}}$

If $\{x_\alpha\}_{\alpha\in A}$ is n-d. o. s. / S is n-d. o. set, then it is a linearly independent system / set.

$\boxed{\text{Proof}}$

Suppose that $\{x_\alpha\}_{\alpha\in A}$ is n-d. o.s. and for some finite $A'\subset A$ $\sum_{\alpha\in A'}\lambda_\alpha x_\alpha = 0$, with some $\lambda_\alpha\in\mathbb{K}$, $\alpha\in A'$. Then for any $\beta\in A'$ we have $0=\left(\sum_{\alpha\in A'}\lambda_\alpha x_\alpha, x_\beta\right)=\lambda_\beta(x_\beta,x_\beta)=$ $=\lambda_\beta\|x_\beta\|^2$, and $\lambda_\beta=0$, since $x_\beta\neq 0$. For the o. set the proof is anclogic. $\boxed{\text{Ч}}$

Orthogonal base – the existence

Let X be a unitary space, and let $S \subset X$, $\{x_\alpha\}_{\alpha\in A}$ – (an indexed) system of vectors in X. We shall define the same name "orthogonal base" for both set and system cases.

Definition

- S is an orthogonal base iff it is a maximal element in the family of all non-degenerated orthogonal sets in X, with respect to the $\subset$ (inclusion) order.
- $\{x_\alpha\}_{\alpha\in A}$ is an orthogonal base iff $S := \{x_\alpha \in X : \alpha\in A\}$ is an orthogonal base (in the above defined sense) and $A \ni \alpha \longmapsto x_\alpha \in S$ is an injection.
- $S \,/\, \{x_\alpha\}_{\alpha\in A}$ is an orthonormal base iff it is an orthogonal base and $S \subset S(0,1) \,/\, \forall_{\alpha\in A} \|x_\alpha\| = 1$.

Be shall often say "... base for X" or "... base in X" in all such cases.

Warning! "Usually" an orthogonal base is NOT a base ... (in the linear base sense).

The existence result (and its proof...), is similar, as for the (linear) base case:

Theorem ("On existence of orthogonal base")

For any unitary space X there exists an orthogonal base B for X. Moreover, for any n-d. orthogonal set $S \subset X$ there exists an orthogonal base B for X, such that $S \subset B$.

Proof

Denote by $NDOS(X)$ - the family of the all n-d. o.sets in X, and let $\mathcal{F} \subset NDOS(X)$ be a chain (with respect to $\subset$). To prove the ("moreover") assertion it suffices to find some $U \in NDOS(X)$ such that U is an upper bound for $\mathcal{F}$, due to the Kuratowski-Zorn Lemma.
Consider $U := \bigcup \mathcal{F}$ - surely it is an upper bound for $\mathcal{F}$ in the $\subset$ sense as a set, but we must check, that $U \in NDOS(X)$.
Let $x, y \in U$, then $x \in S_1$ for some $S_1 \in \mathcal{F}$ and $y \in S_2$ for some $S_2 \in \mathcal{F}$. But $\mathcal{F}$ is a chain, i.e. $S_1 \subset S_2$ or $S_2 \subset S_1$ and in particular both $x, y \in S_i$ for $i = 1$ or 2, so
$(x,y) = \begin{cases} 0 & \text{if } x \neq y \\ \|x\|^2 \neq 0 & \text{if } x = y \end{cases}$, because S_i is a n-d. o.set.

This proves the "moreover" part, and to get the first part it suffices to observe, that $NDOS(X)$ is always non-empty, because $\emptyset \in NDOS(X)$ ... $\blacksquare$

Remarks

1. The appropriate existance result (please, give the strict formulation $\rightarrow$) in terms of "orthogonal bases - systems" follows from this, obviously – one can always consider the
system $\{x\}_{x \in B}$ for B – an "orthogonal base – set,"
and it is "also" an orthogonal base.

2. Having an orthogonal base (in both set/system cases) we can always normalize it, to obtain an appropriate orthonormal base. More preisely, the normalized set/system

$$\left\{ \frac{x}{\|x\|} : x \in S \right\} \Big/ \left\{ \frac{x_\alpha}{\|x_\alpha\|} \right\}_{\alpha \in A}$$

is an orthonormal base, if $S \big/ \{x_\alpha\}_{\alpha \in A}$ is an orthogonal one.

Some equivalent conditions and fundamental properties

Both – the definition and the existence theorem for orthogonal base work in any unitary space, but to get a really important notion, we rather need Hilbert spaces.
So, assume here, that $\mathcal{H}$ is a Hilbert space.

We prove now an important result, which can be treated both as several options of "alternative definitions" (equivalent ones) of orthonormal base, and as a set of some fundamental properties of orthonormal bases.

Theorem ("On orthonormal base")

Let $\{x_\alpha\}_{\alpha\in A}$ be an orthonormal system in $\mathcal{H}$.

TFCAE:

(i) $\{x_\alpha\}_{\alpha\in A}$ is an orthonormal base for $\mathcal{H}$;

(ii) $\{x_\alpha : \alpha\in A\}$ is linearly dense in $\mathcal{H}$ *);

(iii) ("the Fourier expansion") for any $x\in\mathcal{H}$

$$x=\sum_{\alpha\in A}(x,x_\alpha)\,x_\alpha;$$

(iv) ("the Parseval equality") for any $x\in\mathcal{H}$

$$\|x\|^2=\sum_{\alpha\in A}|(x,x_\alpha)|^2.$$

Proof

(i) $\Rightarrow$ (ii)

From (i) we get: $M:=\{x_\alpha : \alpha\in A\}$ is an orthogonal base (as the set). Suppose that M is not linearly dense, then by "A linar density criterion" there exists $0\neq z\in M^\perp$, so $M\cup\{z\}$ is a n-d. o. set, and $z\notin M$ (since $M\cap M^\perp=\{0\}$), i.e. $M\subsetneq M\cup\{z\}$ – a contradiction with the maximality of M.

---

*) The property (ii) is sometimes called "completeness" (of o.s. $\{x_\alpha\}_{\alpha\in A}$ or o.set $\{x_\alpha:\alpha\in A\}$)

(ii) $\Rightarrow$ (iii)

If (ii) holds, then $X = \overline{\mathrm{lin}\{x_\alpha : \alpha \in A\}}$, so by "The orth. projection formula II" we have

$$x = P_X x = \sum_{\alpha \in A} (x, x_\alpha) x_\alpha .$$

(iii) $\Rightarrow$ (iv)

If (iii) holds, then using again The orth. proj. formula II we get

$$\|x\|^2 = \Big\| \sum_{\alpha \in A} (x, x_\alpha) x_\alpha \Big\|^2 = \sum_{\alpha \in A} |(x, x_\alpha)|^2$$

(the second "=" we obtain from (5))

(iv) $\Rightarrow$ (i)

Suppose that (iv) holds, but $\{x_\alpha\}_{\alpha \in A}$ is not an orth. base. Since $\{x_\alpha\}_{\alpha \in A}$ is an orthonormal system, thus $A \ni \alpha \mapsto x_\alpha$ should be injective (see Remark 1. p. HS-53), thus, by definition p. HS-55 $S := \{x_\alpha : \alpha \in A\}$ is not an orthogonal base, so there exists $\tilde{S}$ - some n-d.o. set larger than S. In particular, there exits $0 \neq x \in \tilde{S} \setminus S$ such that $x \in S^\perp$. Hence, by (iv)

$$0 \neq \|x\|^2 = \sum_{\alpha \in A} 0 = 0 \quad - \text{ a contradiction.} \qquad \blacksquare$$

Corollary ("On orthogonal base")

Let $\{x_\alpha\}$ be a n-d. o. s. in $\mathcal{H}$.

TFCAE:

(i) $\{x_\alpha\}_{\alpha\in\mathcal{A}}$ is an orthogonal base for $\mathcal{H}$;

(ii) $\{x_\alpha\}_{\alpha\in\mathcal{A}}$ is linearly dense;

(iii) ("the Fourier expansion") for any $x \in \mathcal{H}$

$$x = \sum_{\alpha\in\mathcal{A}}' \frac{(x, x_\alpha)}{\|x_\alpha\|^2} x_\alpha ;$$

(iv) ("the Parseval equality") for any $x \in \mathcal{H}$

$$\|x\|^2 = \sum_{\alpha\in\mathcal{A}}' \left(\frac{|(x, x_\alpha)|}{\|x_\alpha\|} \right)^2 .$$

Proof

It suffices to use the previous theorem for the orthonormal system $\left\{ \frac{x_\alpha}{\|x_\alpha\|} \right\}_{\alpha\in\mathcal{A}}$ ($\rightarrow$ △ - please check those "calculations") $\boxed{\text{LL}}$

Remark

Both above results can be easily *) reformulated for orthonormal / orthogonal sets. Recall that S is an o. set / orthogon. base / n-d. o. set / orthonormal base iff $\{x\}_{x\in S}$ is an o. s. / orth. base / n-d. o. system / orth. base.

---

*) $\longrightarrow$ △.

Exercise 1 $\longrightarrow$ $\triangle$

Each orthogonal base ($\{x_n\}_{n \in \mathbb{N}_{n_0}}$) indexed by $\mathbb{N}_{n_0}$ ($n_0 \in \mathbb{Z}$) is a Schauder base.

Warning! It is very easy, but not so obvious, as it seems at first glance ("$A \subset B \Rightarrow (\exists!_{a \in A} \dots \Rightarrow \exists!_{a \in B} \dots)$" ??).

Fact

A finite orthogonal base (in a Hilbert space) is a base.

Proof

If $\{x_n\}_{n=1,\dots,d}$, ($d \in \mathbb{N}_0$) *) (the choice of a d-dimensional index set is not important...) is an orthogonal base in $\mathcal{H}$, then by Th. "On orth. base" $\mathcal{H} = \overline{\operatorname{lin}\{x_1, \dots, x_d\}}$, but $\operatorname{lin}\{x_1, \dots, x_d\}$ is finite dimensional, so it is closed (see Remark p. PB-15), hence $\mathcal{H} = \operatorname{lin}\{x_1, \dots, x_d\}$. But also $\{x_n\}_{n=1,\dots d}$ is a linearly independent system (Fact p. HS-54), so it is a (linear) base. $\boxed{\square}$

This result can be slightly strengthen:

Exercise 2 $\longrightarrow$ $\triangle$

Let X – a unitary space, B – its orthogonal base. TFCAE:
(i) $\dim X < +\infty$; (ii) $\operatorname{card} B < +\infty$; (iii) B is a base.

---

*) We get $\emptyset$ if $d = 0$...

The Hilbert dimension and separable Hilbert spaces

Copying the idea of the dimension notion for linear spaces (called also "linear dimension") we would like to introduce the following:

Definition

The **Hilbert dimension** of a Hilbert space $\mathcal{H}$ is the cardinality of an orthogonal base**) of $\mathcal{H}$. We denote it by

$$\dim_H(\mathcal{H}).$$

To make this definition reasonable "in 100%" *) we need the unicity result.

Theorem ("On Hilbert dimension")

Any two orthogonal bases of a Hilbert space $\mathcal{H}$ have the same cardinality.

Proof

Let us consider orth. bases $\mathcal{B}$ and $\tilde{\mathcal{B}}$ of $\mathcal{H}$ (it suffices to consider "the set case" – see **)) For any $x \in \mathcal{B}$ define:
$\tilde{C}(x) := \{\tilde{x} \in \tilde{\mathcal{B}} : (x, \tilde{x}) \neq 0\} \subset \tilde{\mathcal{B}}$. Observe, that:

---

*) We can say that it is already "50% reasonable", because of Theorem "On existance of orthogonal base"...

**) In the "system" case of orth. base $\{x_\alpha\}_{\alpha \in \mathcal{A}}$ we mean here the cardinality of the set $\{x_\alpha : \alpha \in \mathcal{A}\}$. Note, that it is equal to $\operatorname{card} \mathcal{A}$, because "$\alpha \mapsto x_\alpha$" is injective.

(i) $\forall_{x\in B}$ $\quad \mathrm{card}\left(\tilde{C}(x)\right) \leq \aleph_0 \;\left(:= \mathrm{card}(\mathbb{N})\right)$;

(ii) $\tilde{B} = \bigcup_{x\in B} \tilde{C}(x)$.

To prove (ii) observe that $\bigcup_{x\in B} \tilde{C}(x) \subset \tilde{B}$, but

$$\tilde{B} \setminus \bigcup_{x\in B} \tilde{C}(x) = \bigcap_{x\in B}\left(\tilde{B}\setminus\tilde{C}(x)\right) = \left\{\tilde{x}\in\tilde{B} : \forall_{x\in B} (\tilde{x},x)=0\right\} =$$

$$= \tilde{B} \cap B^{\perp} \underset{(*)}{=} \tilde{B} \cap \left(\overline{\mathrm{lin} B}\right)^{\perp} \underset{(**)}{=} \tilde{B} \cap \mathcal{H}^{\perp} = \tilde{B} \cap \{0\} = \emptyset,$$

where $\underset{(*)}{=}$ follows from Remark p. HS-21 and $\underset{(**)}{=}$ from Theorem "On orth. base".

And (i) follows from the Bessel lemma.

If B or $\tilde{B}$ is finite, then by Fact p. HS-61 $\mathcal{H}$ is finite dimensional, so by Fact p. HS-54, the both B and $\tilde{B}$ must be finite, as linearly indep. sets. Hence both are bases, and $\mathrm{card}\, B = \mathrm{card}\, \tilde{B}$ (see - Linear Algebra ...).

It remains only to study the case of B and $\tilde{B}$ being infinite. In this case, by (i) and (ii)

$\mathrm{card}\, \tilde{B} \leq \mathrm{card}\, B \cdot \aleph_0$. But $\mathrm{card}\, B \cdot \aleph_0 = \mathrm{card}\, B$, because $\mathrm{card}\, B$ is infinite (see Set Theory...). So, finally $\mathrm{card}\, \tilde{B} \leq \mathrm{card}\, B$ and, by symmetry, $\mathrm{card}\, B \leq \mathrm{card}\, \tilde{B}$ which gives (see Set Theory again...) $\mathrm{card}\, B = \mathrm{card}\, \tilde{B}$. $\boxed{\angle}$

Now the $\dim_H(\mathcal{H})$ is well-defined!

Example

Let A be <u>an arbitrary</u> set and consider the Hilbert space $\ell^2(A) \equiv L^2(A, \#)$. For any $\alpha \in A$ define $e_\alpha \in \ell^2(A)$ by $e_\alpha(\beta) = \begin{cases} 1 & \alpha = \beta \\ 0 & \alpha \neq \beta \end{cases}$.

Obviously $\{e_\alpha\}_{\alpha \in A}$ is an orthonormal system. Moreover if $f \in \ell^2(A)$ and $\forall_{\alpha \in A}\ f \perp e_\alpha$, then

$$\forall_{\alpha \in A} \quad 0 = (f, e_\alpha) = \sum_{\beta \in A} f(\beta) e_\alpha(\beta) = f(\alpha), \text{ i.e. } f = 0.$$

This means, that $\{e_\alpha\}_{\alpha \in A}$ is linearly dense, and hence - - it is an orthonormal base! It means, that $\dim_{\mathcal{H}}(\ell^2(A)) = \operatorname{card}(A)$ - so there exists a Hilbert space of <u>any</u> Hilbert dimension!

It can be very surprising for you, that any Hilbert space is in fact, up to an isometry, the space of the form from the above example, for some set A.

In Hilbert spaces context the notion "isometry" is somewhat less popular then in general normed spaces context.

Terminology / Definition

Unitary map (transformation, operator etc) is an isometry (linear) between Hilbert spaces. The isometric Hilbert spaces are also called **unitary equivalent**.

Recall here, that unitary map preserves not only the linear operations and norm, but also the scalar product (see Fact "On isometry..." p. HS-20).

Theorem ("On L^2 (ℓ^2...) representation of Hilbert space")

Any two Hilbert spaces of the same Hilbert dimension are unitary equivalent. *) Moreover, each Hilbert space $\mathcal{H}$ is unitary equivalent to any $\ell^2(\mathcal{A})$ **), such that the cardinality of the set $\mathcal{A}$ equals to $\dim_H(\mathcal{H})$. More precisely, if $\{x_\alpha\}_{\alpha\in\mathcal{A}}$ is an orthonormal base in $\mathcal{H}$, then the map $\Phi : \ell^2(\mathcal{A}) \longrightarrow \mathcal{H}$ given by

$$\Phi(f) := \sum_{\alpha\in\mathcal{A}} f(\alpha)\, x_\alpha \quad , \qquad f \in \ell^2(\mathcal{A}) \qquad (1)$$

is a unitary transformation (from $\ell^2(\mathcal{A})$ onto $\mathcal{H}$).

Proof

By Theory "On existence of orth. base" and Remark 2 p. HS-57 $\mathcal{H}$ posesses an orthonormal base, and by the definition of $\dim_H(\mathcal{H})$, there exits an orthonormal base of the form $\{x_\alpha\}_{\alpha\in\mathcal{A}}$ for any $\mathcal{A}$ with $\operatorname{card}\mathcal{A} = \dim_H(\mathcal{H})$. Hence, it suffices to prove that Φ given by (1)

---

*) Warning! Beware of misunderstanding! This theorem (as many others of "similar type") works separately for $\mathbb{K} = \mathbb{R}$ and for $\mathbb{K} = \mathbb{C}$.

**) Where we take $\ell^2(\mathcal{A})$ as the space of functions from $\mathcal{A}$ into $\mathbb{K}$ with $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ appropriately!

is in fact a unitary map. ($\longrightarrow \triangle$ - for "it suff." ...)

Observe first, that formula (1) is correct, in the sense, that the RHS sum is convergent. Indeed, $\{f(\alpha)x_\alpha\}_{\alpha\in A} \in QCO$ because of Bessel's lemma and the fact that $\forall_{\alpha\in A}\ f(\alpha) = (f, e_\alpha)$ (see Example p. HS-64); moreover

$$\sum_{\alpha\in A} \|f(\alpha)x_\alpha\|^2 = \sum_{\alpha\in A}|f(\alpha)|^2 = \|f\|_2 < +\infty, \quad (2)$$

which gives the convergence of the RHS, by Fact (ii) p. HS-48. Observe, that Φ is linear – also by the above Fact, part (i): if $f, g \in \ell^2(A)$, then for any $\beta\in A$

$$(x_\beta, \Phi(f+g) - \Phi(f) - \Phi(g)) = (x_\beta, \Phi(f+g)) - (x_\beta, \Phi(f)) - (x_\beta, \Phi(g)) =$$

$$= \sum_{\alpha\in A}(x_\beta, (f(\alpha)+g(\alpha))x_\alpha) - \sum_{\alpha\in A}(x_\beta, f(\alpha)x_\alpha) - \sum_{\alpha\in A}(x_\beta, g(\alpha)x_\alpha) =$$

$$= (x_\beta, (f(\beta)+g(\beta))x_\beta) - (x_\beta, f(\beta)x_\beta) - (x_\beta, g(\beta)x_\beta) = 0,$$

hence $\Phi(f+g) = \Phi(f) + \Phi(g)$, by the fact that $\{x_\beta : \beta\in A\}$ is linearly dense (Thm "On orth. base" + "A linear density criterion"); analogously ($\longrightarrow \triangle$) we prove that Φ is homogenous. Now, again using Fact (ii), we see that

$$\|\Phi(f)\|^2 = \sum_{\alpha\in A}\|f(\alpha)x_\alpha\|^2,$$

so $\|\Phi(f)\|^2 = \|f\|^2$ for any $f\in\ell^2(A)$. This means, that Φ is an isometry from $\ell^2(A)$ onto its image $Y := \operatorname{Ran}\Phi$. So, Y is a Banach space, hence it is a closed linear subspace of $\mathcal{H}$.

On the other hand $\Phi(e_\beta) = \sum_{\alpha\in A} e_\beta(\alpha) x_\alpha = x_\beta$
for any $\beta \in A$, so

$$\mathcal{H} = \overline{\operatorname{lin}\{x_\beta : \beta\in A\}} \subset Y,$$

i.e., $Y = \mathcal{H}$. $\boxed{\text{Ш}}$

$\boxed{\text{Corollary 1}}$

Two Hilbert spaces*) $\mathcal{H}_1$ and $\mathcal{H}_2$ are unitary equivalent iff $\dim_H \mathcal{H}_1 = \dim_H \mathcal{H}_2$.

$\boxed{\text{Proof}}$

"$\Leftarrow$" Follows just from the theorem above.

"$\Rightarrow$" If $U : \mathcal{H}_1 \to \mathcal{H}_2$ is a unitary transformation and B is an orthogonal base in $\mathcal{H}_1$, then $U(B)$ is an orthogonal base in $\mathcal{H}_2$. But U is injective, hence $\dim_H \mathcal{H}_1 = \operatorname{card} B = \operatorname{card}(U(B)) = \dim_H \mathcal{H}_2$. $\boxed{\text{Ш}}$

$\boxed{\text{Corollary 2}}$

Let $\mathcal{H}$ be a Hilbert space. $\mathcal{H}$ is separable iff $\dim_H \mathcal{H}$ is at most countable. If $\dim_H \mathcal{H} = d \in \mathbb{N}_0$, then $\mathcal{H}$ is unitary eqivalent to $\mathbb{K}^d$ **), if $\dim_H \mathcal{H} = \aleph_0$, then $\mathcal{H}$ is unitary eqivalent to $\ell^2(\mathbb{N})$. In particular, any two separable, infinite dimensional***) Hilbert spaces*) are unitary equivalent.

---

*) Over the same scalar field $\mathbb{K}$...

**) $\mathbb{K}^0 = \{0\}$. In general d-case we take the Euclidean norm in $\mathbb{K}^d$.

***) Note, that the choice of the dimension kind (Hilbert, linear) is not important here, since $\dim \mathcal{H} = +\infty \Leftrightarrow \dim_H \mathcal{H} = +\infty \to \triangle$.

Proof

If $\dim_H \mathcal{H} \leq \aleph_0$ then by Thm "On orth. base" (ii) $\mathcal{H}$ has an at most countable linearly dense subset. This gives separability for $\mathcal{H}$, by Fact "On separability from linear density" (p. PB-63). Suppose that $\mathcal{H}$ is separable and let $\{x_\alpha\}_{\alpha\in A}$ be an orthonormal base of $\mathcal{H}$. Suppose, that $\operatorname{card} A > \aleph_0$. Then the incountable family of balls $\{K(x_\alpha, \frac{1}{2})\}_{\alpha\in A}$ is pairwise disjoint, because $\|x_\alpha - x_\beta\| = \sqrt{2}$, by the Pythagorean formula for $\alpha\neq\beta$. Hence $\mathcal{H}$ would be not separable. This finishes the proof of the first sentence. Now, the case of $\dim_H \mathcal{H} = d$ follows from Fact p. HS-61 ($\dim \mathcal{H} = \dim_H \mathcal{H} = d$ then, or we can use Thm p. HS-65 and the fact that $\mathbb{K}^d$ is unitary eqiuivalent to $\ell^2(\{1,\dots,d\})$). If $\dim_H \mathcal{H} = \aleph_0$, then $\mathcal{H}$ is un. equiv. to $\ell^2(\mathbb{N})$ by Theorem "On ℓ^2 representation ...". The last sentence follows by the previous two. $\boxed{\text{ck}}$

The last several results show that the variety of Hilbert spaces is much less than the variety of all the Banach spaces. Especialy for the most popular - infinite dimensional, but separable, Hilbert spaces! - They are all "almost the same" - being unitary eqivalent to the well known $\ell^2(\mathbb{N})$ space ...

The Gram - Schmidt ortogonalization process *)

In many practical cases of Hilbert spaces an orthogonal base can be in fact constructed, which seem somewhat better, than the abstract existence result "on existence", obtained via Kuratowski - Zorn Lemma (see Th. p. HS-56).
This construction is usually based on the so-called Gram - Schmidt ortogonalization process. It has several variants - we describe here all of them, but the most popular is the constmction of an orthogonal / orthonormal system "starting" from a linearly independent one.

Let $\{v_n\}_{n\in \mathbb{I}}$ be a system of vectors in a unitary space X, where $\mathbb{I}$ is a set of integers of the form $\{1,\dots,N\}$, $N\in\mathbb{N}_1$ or $\mathbb{I}=\mathbb{N}$.

We define recursively two new systems $\{x_n\}_{n\in\mathbb{I}}$, $\{\tilde{x}_n\}_{n\in\mathbb{I}}$ in X:

$$(G\text{-}S)\quad \begin{cases} x_1 := v_1, \quad \tilde{x}_1 := (x_1)_{nor} \\ \forall_{n\in\mathbb{I}} \left[n+1\in\mathbb{I} \Rightarrow \left(x_{n+1} := (I-P_n)v_{n+1},\ \tilde{x}_{n+1} := (x_{n+1})_{nor} \right) \right] \end{cases}$$

---

*) Pol.: ortogonalizacja Grama - Schmidta.

where we denote: $(y)_{nor} := \begin{cases} 0 & \text{for } y=0 \\ y/\|y\| & \text{for } y\neq 0 \end{cases}$ and

P_n is the orthogonal projection in the space V_{n+1} onto V_n, where $V_k := \text{lin}\{v_1, \dots, v_k\}$ for any $k \in \mathbb{I}$. Denote also $V := \text{lin}\{v_j : j \in \mathbb{I}\}$.

Theorem ("On Gram-Schmidt process")

The systems $\{x_n\}_{n\in\mathbb{I}}$, $\{\tilde{x}_n\}_{n\in\mathbb{I}}$ have the following properties:

(i) $\forall_{n\in\mathbb{I}}$ $\text{lin}\{x_1, \dots, x_n\} = \text{lin}\{\tilde{x}_1, \dots, \tilde{x}_n\} = V_n$,

(i') $\text{lin}\{x_j : j\in\mathbb{I}\} = \text{lin}\{\tilde{x}_j : j\in\mathbb{I}\} = V$

(ii) $\{x_n\}_{n\in\mathbb{I}}$ is an o.s.

(iii) $\forall_{n\in\mathbb{I}} \left[n+1\in\mathbb{I} \Rightarrow x_{n+1} = v_{n+1} - \sum_{j=1}^{n} (v_{n+1}, \tilde{x}_j)\tilde{x}_j \right]$

(iv) if $\{v_n\}_{n\in\mathbb{I}}$ is linearly independent, then $\{x_n\}_{n\in\mathbb{I}}$ is a n-d. o.s and $\{\tilde{x}_n\}_{n\in\mathbb{I}}$ is an orthonormal system and moreover

$$\forall_{n\in\mathbb{I}} \left[n+1\in\mathbb{I} \Rightarrow x_{n+1} = v_{n+1} - \sum_{j=1}^{n} \frac{(v_{n+1}, x_j)}{\|x_j\|^2} x_j \right] \quad (1)$$

(v) if $\tilde{\mathbb{I}} := \{n\in\mathbb{I} : x_n \neq 0\}$ then $\{\tilde{x}_n\}_{n\in\tilde{\mathbb{I}}}$ is an orthonormal syst., $\{x_n\}_{n\in\mathbb{I}}$ is a) n-d. o.s. and $\{\tilde{x}_j : j\in\tilde{\mathbb{I}}\} = V = \text{lin}\{x_j : j\in\tilde{\mathbb{I}}\}$

(vi) if X is a Hilbert space, then $\{x_n\}_{n\in\tilde{\mathbb{I}}}$ is an orthogonal base of $\overline{V}$ and $\{\tilde{x}_n\}_{n\in\tilde{\mathbb{I}}}$ – an orthonormal base of $\overline{V}$, with $\tilde{\mathbb{I}}$ as in (v). In particular they are orthogonal/orthonormal bases of X, respectively if X is Hilbert and $\{v_n\}_{n\in\mathbb{I}}$ is linearly dense*)

---

*) i.e. $\{v_n : n\in\mathbb{I}\}$ is linearly dense.

(vii) if X is a Hilbert space, $\{v_n\}_{n\in\mathbb{I}}$ is linearly independent and linearly dense, then $\{x_n\}_{n\in\mathbb{I}}$ is an orthogonal base of X and $\{\tilde{x}_n\}_{n\in\mathbb{I}}$ is an orthonormal base of X.

Proof

Observe first, that (iii) holds directly from the definition of x_{n+1} and from The orthogonal projection formula I *) (p. HS-50). Thus (i) can be easily proved by induction ($\longrightarrow$ △ ; (i) could also be obtained directly from the definition if we use the fact that $P_n x \in V_n$ for any $x \in X$ and $n \in \mathbb{I}$ $\longrightarrow$ △). (i') follows directly from (i). To get (ii) it suffices to prove that for any $n \in \mathbb{I}$ if $n+1 \in \mathbb{I}$, then $x_{n+1} \perp V_n = \mathrm{lin}\{x_1, \dots, x_n\}$ (by (i)). But this is obvious from the definition of x_{n+1}, since $I - P_n$ is the orth. projection onto $V_n^{\perp}$ (in V_{n+1}). Hence we get (v) by (i') and (ii) and this gives also (vi) by Corollary "On orthogonal base" (p. HS-60). Let us prove (iv) now, so assume, that $\{v_n\}_{n\in\mathbb{I}}$ is linearly independent. Hence $x_1 = v_1 \neq 0$. Moreover if $n \in \mathbb{I}$ and $(n+1) \in \mathbb{I}$, then $v_{n+1} \notin V_n$. Suppose that $x_{n+1} = 0$, then by (i) $v_{n+1} \in V_{n+1} = \mathrm{lin}\{x_1, \dots, x_n, x_{n+1}\} = \mathrm{lin}\{x_1, \dots x_n\} = V_n$ – a contradiction. So, $x_{n+1} \neq 0$, which

*) In "the simplest" – finite $\mathcal{A}$ case...

proves that $\{x_n\}_{n\in\mathbb{I}}$ is n-d., by induction. Hence, $\{\tilde{x}_n\}_{n\in\mathbb{I}}$ is an orthonormal syst., and (1) holds by (iii). So (iv) is proved and now we get (vii) directly from (iv) and (vi).

$\square$

Remark

Observe that using the second part of (vi) we get the proof (the constructive one!) of existence of an orthonormal base for any separable Hilbert space! It suffices to take any $\{v_n\}_{\in\mathbb{N}_1}$ being dense sequence (i.e. with $\{v_n : n\in\mathbb{N}_1\}$ being dense), which is possible by separability.

Exercise →

Suppose that the system $\{v_n\}_{n\in\mathbb{I}}$ is linearly independent. and $\{y_n\}_{n\in\mathbb{I}}$ is a system in X. Then TFCAE:

(a) $\{y_n\}_{n\in\mathbb{I}} = \{\tilde{x}_n\}_{n\in\mathbb{I}}$ (i.e. $\{y_n\}_{n\in\mathbb{I}}$ is the orthonormal system obtained from $\{v_n\}_{n\in\mathbb{I}}$ by the G.-S. orthonormalization process);

(b) the following properties (A), (B), (C) hold:

(A) $\{y_n\}_{n\in\mathbb{I}}$ is an orthonormal system

(B) $\mathrm{lin}\{y_1, \dots y_n\} = \mathrm{lin}\{v_1, \dots v_n\}$ for any $n\in\mathbb{I}$

(C) for any $n\in\mathbb{I}$ (if $y_n = \sum_{j=1}^{n} \lambda_{jn} v_j$ for some $\lambda_{1n}, \lambda_{nn}\in\mathbb{K}$, then $\lambda_{nn} > 0$).

Examples of orthogonal bases

We give here several examples of orthogonal bases in several Hilbert spaces (some in "set", some in "system" form). We use here the abbreviations:

o.g.b. and o.n.b. for orthogonal and orthonormal base, respectively. We mostly don't give the proofs here – only the formulae for the o.g.b / o.n.b, so there is a lot of "$\to \triangle$-s" (sometimes, the proof of linear density is not so easy yet – then try to find the orthogonality/orthonormality proof at least ...)

Example 1 For any $\Omega \neq \emptyset$ the system $\{e_t\}_{t\in\Omega}$ given by $e_t(s) = \begin{cases} 1 & s = t \\ 0 & s \neq t \end{cases}$, $s, t \in \Omega$ is an o.n.b. in $l^2(\Omega)$ (see Example p. HS-64).

Example 2 ("The ("exp") Fourier o.n. base")

For any $n \in \mathbb{Z}$ define $e_n : I_a \to \mathbb{C}$, where $a \in \mathbb{R}$ ("fixed") and $I := [a; a+2\pi]$, by the formula

$$e_n(s) := \frac{1}{\sqrt{2\pi}} e^{ins}, \quad s \in I_a.$$

The system $\{[e_n]\}_{n\in\mathbb{Z}}$ is an o.n.b. in $L^2(I_a)$ (with the standard Lebesgue measure, as usual)

$\longrightarrow \triangle$ (**Hint**: use the abstract Stone-Weierstrass theorem "for $C(I_a)$" to get the lin-density of $\{e_n\}_{n\in\mathbb{Z}}$ in $C(I_a)$ first...).

$\boxed{\text{Example } 2\frac{1}{2}}$ ("The ("sin/cos") Fourier o.n. base")

(a, I_a as in Ex. 2) Define $c_n, s_n : I_a \to \mathbb{R}$ for $n \in \mathbb{N}_1$, by

$c_n(s) := \cos(ns)$, $s_n(s) := \sin(ns)$, $s \in I_a$, and let $\mathbb{1}$ be the constant 1 function.

The set $\{[\mathbb{1}]\} \cup \{[c_n] : n \in \mathbb{N}_1\} \cup \{[s_n] : n \in \mathbb{N}_1\}$ is an o.g.b. in $L^2(I_a)$ $\longrightarrow \triangle$.

The set $\{(2\pi)^{-1}[\mathbb{1}]\} \cup \{\pi^{-1}[c_n]\} \cup \{\pi^{-1}[s_n]\}$ is an o.n.b. in $L^2(I_a)$. $\longrightarrow \triangle$.

(Here both $\mathbb{K} = \mathbb{R}$ and $\mathbb{K} = \mathbb{C}$ can be considered).

$\boxed{\text{Example } 2\frac{1}{1001}}$ ("The unit circle Fourier o.n. base")

Let $\mathbb{T}$ be a unite circle in $\mathbb{R}^2 \equiv \mathbb{C}$, i.e. $\mathbb{T} := \{z \in \mathbb{C} : |z| = 1\}$ and consider the normalized*) "manifold" Lebesgue measure (i.e. on Borel $\mathbb{T}$ subsets) μ on $\mathbb{T}$

*) i.e. the Lebesgue measure multiplied by $\frac{1}{2\pi}$, to get $\mu(\mathbb{T}) = 1$.

Consider $\mathbb{Z}: \mathbb{T} \to \mathbb{C}$ – the identity function ($\mathbb{Z}(z) = z$ for $z \in \mathbb{T}$).

$\{[\mathbb{Z}^n]\}_{n \in \mathbb{Z}}$ is an o.n.b. in $L^2(\mathbb{T}, \mu) \longrightarrow \triangle$

($\mathbb{Z}^n(z) := (\mathbb{Z}(z))^n = z^n$ for $z \in \mathbb{T}$, $n \in \mathbb{Z}$).

(Hint: Find an (appropriate) unitary transformation between $L^2(\mathbb{T}, \mu)$ and $L^2(I_0)$ (or I_π)...) .

$\boxed{\text{Example 3}}$ ("The Legendre Polynomials o.n. base")

Let $a, b \in \mathbb{R}$, $a < b$. Define the Legendre polynomials L_n for $[a;b]$, $n \in \mathbb{N}_0$ by

$L_n : [a;b] \longrightarrow \mathbb{R}$

$$L_n := g_n^{[n]} \quad \text{(the n-th derivative), where}$$

$$g_n(x) := (x-a)^n (x-b)^n, \quad x \in [a;b]. \quad *)$$

Let $c_n := \sqrt{\frac{2n+1}{b-a}}\,(b-a)^{-n} \cdot (n!)^{-1}$.

$\{[L_n]\}_{n \in \mathbb{N}_0}$ is an o.g.b. and moreover $\{c_n \cdot [L_n]\}_{n \in \mathbb{N}_0}$ is an o.n.b. in $L^2([a;b])$. $\longrightarrow \triangle$

---

*) Here „$0^0 := 1$", so $g_0(x) = 1$ also for $x = a, b$...

(Hint:*) to calculate the orthonormality use the "multiple" integration by parts...).

Example 4

("The Hermite o.n. base")

Define the Hermite functions $H_n : \mathbb{R} \longrightarrow \mathbb{R}$ for $n \in \mathbb{N}_0$ by $H_n(x) := e^{(x^2/2)} \cdot h^{[n]}(x)$, $x \in \mathbb{R}$,
where $h : \mathbb{R} \longrightarrow \mathbb{R}$ is given by $h(x) := e^{(-x^2)}$, $x \in \mathbb{R}$.

Let $d_n := (-1)^n (2^n n!)^{-\frac{1}{2}} (\pi)^{-\frac{1}{4}}$, $n \in \mathbb{N}_0$.

$\{[H_n]\}_{n\in\mathbb{N}_0}$ is an o.g.b. and $\{d_n [H_n]\}_{n\in\mathbb{N}_0}$ an o.n.b. in $L^2(\mathbb{R})$. $\longrightarrow$ (Hint*): as above...).

Exercise $\longrightarrow$

For $\{[L_n]\}_{n\in\mathbb{N}_0}$ and $\{[H_n]\}_{n\in\mathbb{N}_0}$ from the above two examples find such scalar sequences $\{z_n\}_{n\in\mathbb{N}_0}$ that $\{z_n[L_n]\}_{n\in\mathbb{N}_0}$ / $\{z_n [H_n]\}_{n\in\mathbb{N}_0}$ is the orthonormal system obtained from $\{x^n\}_{n\in\mathbb{N}_0}$ / $\{x^n e^{(-x^2/2)}\}_{n\in\mathbb{N}_0}$ by the G.-S. orthonorm. process (see Exercise p. HS-72), respectively.

---

*) And an extra hint is given below - in the Exercise ...

$\boxed{\text{Example 5}}$ ("The Chebyshev*) polynomials")

Consider the sequence $\{p_n\}_{n\in\mathbb{N}_0}$ of functions $p_n : \mathbb{R} \to \mathbb{R}$ given by the following recurrent formula:

$$(Ch) \quad \begin{cases} p_0 = \mathbb{1}, \quad p_1 = \mathbb{X} \\ p_{n+1} + p_{n-1} = \mathbb{X} \cdot p_n \qquad n \geq 1. \end{cases}$$

It can be easily proved by the induction, that p_n is an n-order polynomial for each $n \in \mathbb{N}_0$.

The polynomials U_n given by

$$U_n(x) = p_n(2x), \quad n \in \mathbb{N}_0, \; x \in \mathbb{R}$$

are called Chebyshev polynomials (of the second kind)

Thus the polynomials p_n are sometimes called "rescaled Chebyshev polynomials (of t. s. k.)". It can be proved**) that $\{[p_n|_{[-2;2]}]\}_{n\in\mathbb{N}_0}$ is an o. n. b. in the space $L^2([-2;2], \mu)$, where $\mu = g\,dx$ with the density function $g : [-2;2] \to \mathbb{R}$ given by

$$g(x) := \frac{1}{\pi}\sqrt{1 - (x/2)^2}, \quad x \in [-2;2].$$

---

*) Pol.: Czebyszew, wielomiany Czebyszewa.

**) But not easily, now...

Example 6 ("Wavelets*) and Haar orthonormal bases")

The general notion "wavelet" (as well as "wave") has no unique meaning (usually it is ± a special kind of "oscillation", which does not mean anything very concrete too...) Also the mathematical notion "wavelet / wavelets" has several meanings, depending on the context. We shall concentrate here on one of those meaning:

a system $\{\psi_\alpha\}_{\alpha\in A}$ of scalar functions from $\tilde{L}^2(\mathbb{R}^d)$ or $\{[\psi_\alpha]\}_{\alpha\in A}$ in $L^2(\mathbb{R}^d)$ which is parametrized by $\alpha \in A$ in such a way, that for each $\alpha \in A$ $\psi_\alpha : \mathbb{R}^d \to \mathbb{C}$ has a form

$$\psi_\alpha(x) := \sqrt[2d]{a_\alpha}\,\psi(a_\alpha x - b_\alpha), \qquad x \in \mathbb{R}^d \qquad (W)$$

where:

1) $\forall_{\alpha\in A}$ $a_\alpha \in (0; +\infty)$, $b_\alpha \in \mathbb{R}^d$ are so-called "scale" (a_α)**) and "shift" (b_α)**) parameters

2) ψ is a fixed $\tilde{L}^2(\mathbb{R}^d)$ function, so-called "mother wavelet" function

---

*) Pol.: falki ("wavelet" = falka — od "wave" = fala ...).

**) but often the name scale is given to a_α^{-1} and shift to $a_\alpha^{-1} b_\alpha$ ...

Additionally also two extra conditions on Ψ are often assumed:

2 a) (the "normalization" cond.) $\int_{\mathbb{R}^d} |\Psi(x)|^2 \, dx = 1$ (i.e. $\|[\Psi]\|_2 = 1$).

2 b) (the "oscillation" cond.) $\int_{\mathbb{R}^d} |\Psi(x)| \, dx < +\infty$ and $\int_{\mathbb{R}^d} \Psi(x) \, dx = 0$.

Such kind of a system $\{\Psi_\alpha\}_{\alpha \in A}$ or $\{[\Psi_\alpha]\}_{\alpha \in A}$ (both: with or without 2a)-b) ...) is also called mother wavelet system. The functions Ψ_α or classes $[\Psi_\alpha]$ are called "child wavelets"...

$\boxed{\text{Observation}}$ ("Disjoint support mother wavelet systems")

If Ψ, $\{a_\alpha\}_{\alpha \in A}$ and $\{b_\alpha\}_{\alpha \in A}$ are such, that 2a) holds and

3) $\forall_{\substack{\alpha, \beta \in A \\ \alpha \neq \beta}} \quad (\operatorname{supp} \Psi_\alpha) \cap (\operatorname{supp} \Psi_\beta) = \emptyset$, *)

then the mother wavelet system $\{[\Psi_\alpha]\}_{\alpha \in A}$ defined by (W) is an orthonormal system in $L^2(\mathbb{R}^d)$.

$\boxed{\text{Proof}} \longrightarrow \triangle$ (an easy calculation...).

---

*) It suffices that the LHS has measure 0.

One of the simplest (and the oldest) mother wavelet systems is so-called Haar wavelet system which is defined by the following "mother" and "parameters":

$$\psi : \mathbb{R} \longrightarrow \mathbb{R} \qquad (d=1) \quad \psi(x) := \begin{cases} 1 & x \in [0; \frac{1}{2}) \\ -1 & x \in [\frac{1}{2}; 1) \\ 0 & x \in \mathbb{R} \setminus [0;1) \end{cases}$$

$$\mathcal{A} := \mathbb{Z}^2, \quad a_\alpha := 2^n, \quad b_\alpha := k \quad \text{for } \alpha = (n,k) \in \mathcal{A}$$

i.e. $\quad \Psi_{(n,k)}(x) := \sqrt{2^n}\, \psi(2^n x - k), \quad x \in \mathbb{R}$

So, we have:

$$\operatorname{supp} \Psi_{(n,k)} = 2^{-n}\big(k + [0;1)\big) = [2^{-n} \cdot k \,;\, 2^{-n}(k+1)),$$

and in particular 2a), 2b) and 3) hold in this case.

Hence, Haar wavelet system is a disjoint support system.

Moreover it is an orthonormal base (the $\{[\Psi_\alpha]\}_{\alpha \in \mathcal{A}}$ case) in $L^2(\mathbb{R})$. $\longrightarrow$ ⚠. It is also caled the Haar orthonormal base in $L^2(\mathbb{R})$.

There is also another famous system related to this one.

Consider, for $\alpha \in \mathcal{A}$, the functions $\varphi_\alpha : [0;1] \longrightarrow \mathbb{R}$ given by

$$\varphi_\alpha := \Psi_\alpha \big|_{[0;1]}$$

and let $\mathcal{A}_1 := \{(n,k) \in \mathcal{A} : n \in \mathbb{N}_0,\ 0 \leq k \leq 2^n - 1\}$.

So, for $\alpha \in \mathcal{A}_1$ $\quad \operatorname{supp} \Psi_\alpha \subset [0;1]$, hence $\operatorname{supp} \varphi_\alpha = \operatorname{supp} \Psi_\alpha$, which means that $\{\varphi_\alpha\}_{\alpha \in \mathcal{A}_1}$ consists of disjoint support functions. Therefore $\{[\varphi_\alpha]\}_{\alpha \in \mathcal{A}_1}$ is an orthonormal system in $L^2([0;1])$. Moreover $\{[\varphi_\alpha] : \alpha \in \mathcal{A}_1\}^{\perp} = \mathrm{lin}[\mathbb{1}]$, where $\mathbb{1}$ is a constant 1 function on $[0;1]$. $\longrightarrow \triangle$.

This proves*), that the system $\{[\varphi_\alpha]\}_{\alpha \in \mathcal{A}_1 \cup \{*\}}$, where $*$ is any object such that $* \notin \mathcal{A}_1$ and $\varphi_* = \mathbb{1}$, is an orthonormal base in $L^2([0;1])$. This system is also called <u>the Haar orthonormal base in $L^2([0;1])$</u>.

---

*) Why? $\longrightarrow \triangle$ ...