Lectures 16-17

6. Hilbert spaces

In this section we examine a special type of Banach spaces.

Definition. Let $\mathbb{K}$ be one of the fields $\mathbb{R}$ or $\mathbb{C}$, and let $\mathcal{X}$ be a vector space over $\mathbb{K}$. An *inner product on* $\mathcal{X}$ is a map

$$\mathcal{X} \times \mathcal{X} \ni (\xi, \eta) \longmapsto \big(\xi \,\big|\, \eta\big) \in \mathbb{K},$$

with the following properties:

- $\big(\xi \,\big|\, \xi\big) \geq 0$, $\forall\, \xi \in \mathcal{X}$;
- if $\xi \in \mathcal{X}$ satisfies $\big(\xi \,\big|\, \xi\big) = 0$, then $\xi = 0$;
- for any $\xi \in \mathcal{X}$, the map $\mathcal{X} \ni \eta \longmapsto \big(\xi \,\big|\, \eta\big) \in \mathbb{K}$ is $\mathbb{K}$-linear;
- $\big(\eta \,\big|\, \xi\big) = \overline{\big(\xi \,\big|\, \eta\big)}$, $\forall\, xi, \eta \in \mathcal{X}$.

Comments. Combining the last two properties, one gets

$$\big(\xi \,\big|\, \lambda\eta_1 + \eta_2\big) = \lambda\big(\xi \,\big|\, \eta_1\big) + \big(\xi \,\big|\, \eta_2\big),\ \ \forall\, \xi, \eta_1, \eta_2 \in \mathcal{X},\ \lambda \in \mathbb{K};$$
$$\big(\lambda\xi_1 + \xi_2 \,\big|\, \eta\big) = \overline{\lambda}\big(\xi_1 \,\big|\, \eta\big) + \big(\xi_2 \,\big|\, \eta\big),\ \ \forall\, \xi_1, \xi_2, \eta \in \mathcal{X},\ \lambda \in \mathbb{K}.$$

In particular, one has

$$(1)\qquad \big(\lambda\xi \,\big|\, \lambda\xi\big) = \overline{\lambda}\lambda\big(\xi \,\big|\, \xi\big) = |\lambda|^2 \cdot \big(\xi \,\big|\, \xi\big),\ \ \forall\, \xi \in \mathcal{X},\ \lambda \in \mathbb{K}.$$

Proposition 6.1 (Cauchy-Bunyakowski-Schwartz Inequality). *Let* $\big(\,\cdot\,\big|\,\cdot\,\big)$ *be an inner product on the* $\mathbb{K}$*-vector space* $\mathcal{X}$. *Then*

$$(2)\qquad \big|\big(\xi \,\big|\, \eta\big)\big|^2 \leq \big(\xi \,\big|\, \xi\big) \cdot \big(\eta \,\big|\, \eta\big),\ \ \forall\, \xi, \eta \in \mathcal{X}.$$

Moreover, if equality holds then ξ *and* η *are proportional, in the sense that either* $\xi = 0$*, or* $\eta = 0$*, or* $\xi = \lambda\eta$.

Proof. Fix $\xi, \eta \in \mathcal{X}$. Assume $\eta \neq 0$. (In the case when $\eta = 0$, both statements are trivial). Choose a number $\lambda \in \mathbb{K}$, with $|\lambda| = 1$, such that

$$\big|\big(\xi \,\big|\, \eta\big)\big| = \lambda\big(\xi \,\big|\, \eta\big) = \big(\xi \,\big|\, \lambda\eta\big).$$

Define the map $F : \mathbb{K} \to \mathbb{K}$ by

$$F(z) = \big(z\lambda\eta + \xi \,\big|\, z\lambda\eta + \xi\big),\ \ \forall\, z \in \mathbb{K}.$$

A simple computation gives

$$\begin{aligned} F(z) &= z\lambda\overline{z}\overline{\lambda}\big(\eta \,\big|\, \eta\big) + z\lambda\big(\xi \,\big|\, \eta\big) + \overline{z}\overline{\lambda}\big(\eta \,\big|\, \xi\big) + \big(\xi \,\big|\, \xi\big) = \\ &= |z|^2|\lambda|^2\big(\eta \,\big|\, \eta\big) + z\lambda\big(\xi \,\big|\, \eta\big) + \overline{z}\overline{\lambda\big(\xi \,\big|\, \eta\big)} + \big(\xi \,\big|\, \xi\big) = \\ &= |z|^2\big(\eta \,\big|\, \eta\big) + z\big|\big(\xi \,\big|\, \eta\big)\big| + \overline{z}\big|\big(\xi \,\big|\, \eta\big)\big| + \big(\xi \,\big|\, \xi\big),\ \ \forall\, z \in \mathbb{R}. \end{aligned}$$

In particular, when we restrict F to $\mathbb{R}$, it becomes a quadratic function:

$$F(t) = at^2 + bt + c, \ \ \forall\, t \in \mathbb{R},$$

where $a = \big(\,\eta\,\big|\,\eta\,\big) > 0$, $b = 2\big|\big(\,\xi\,\big|\,\eta\,\big)\big|$, $c = \big(\,\xi\,\big|\,\xi\,\big)$. Notice that we have

$$F(t) \geq 0, \ \ \forall\, t \in \mathbb{R}.$$

This forces $b^2 - 4ac \leq 0$. This last inequality gives

$$4\big|\big(\,\xi\,\big|\,\eta\,\big)\big|^2 - 4\big(\,\xi\,\big|\,\xi\,\big)\cdot\big(\,\eta\,\big|\,\eta\,\big) \leq 0,$$

so we get

$$\big|\big(\,\xi\,\big|\,\eta\,\big)\big|^2 \leq \big(\,\xi\,\big|\,\xi\,\big)\cdot\big(\,\eta\,\big|\,\eta\,\big),$$

and the inequality (2) is proven. Let us examine now when we have equality. The equality in (2) gives $b^2 - 4ac = 0$, which in terms of quadratic equations says that the equation

$$F(t) = at^2 + bt + c = 0$$

has a $\big($unique$\big)$ solution t_0. This will give

$$\big(\,t_0\lambda\eta + \xi\,\big|\,t_0\lambda\eta + \xi\,\big) = F(t_0) = 0,$$

which forces $t_0\lambda\eta + \xi = 0$, i.e. $\xi = (-t_0\lambda)\eta$. □

Corollary 6.1. *Let $\big(\,\cdot\,\big|\,\cdot\,\big)$ be an inner product on the $\mathbb{K}$-vector space $\mathcal{X}$. Then the map*

$$\mathcal{X} \ni \xi \longmapsto \sqrt{\big(\xi|\xi\big)} \in [0, \infty)$$

is a norm on $\mathcal{X}$.

Proof. Denote $\sqrt{\big(\xi|\xi\big)}$ simply by $\|\xi\|$. The fact that $\|\xi\|$ is non-negative is clear. The implication $\|\xi\| = 0 \Rightarrow \xi = 0$ is also clear. Using (1) we have

$$\|\lambda\xi\| = \sqrt{\big(\lambda\xi|\lambda\xi\big)} = \sqrt{|\lambda|^2\big(\xi|\xi\big)} = |\lambda|\cdot\sqrt{\big(\xi|\xi\big)} = |\lambda|\cdot\|\xi\|, \ \ \forall\,\xi \in \mathcal{X},\ \lambda \in \mathbb{K}.$$

Finally, for $\xi, \eta \in \mathcal{X}$, we have

$$\begin{aligned}\|\xi + \eta\|^2 = \big(\,\xi + \eta\,\big|\,\xi + \eta\,\big) &= \big(\,\xi\,\big|\,\xi\,\big) + \big(\,\eta\,\big|\,\eta\,\big) + \big(\,\xi\,\big|\,\eta\,\big) + \big(\,\eta\,\big|\,\xi\,\big) = \\ &= \|\xi\|^2 + \|\eta\|^2 + \big(\,\xi\,\big|\,\eta\,\big) + \overline{\big(\,\xi\,\big|\,\eta\,\big)} = \|\xi\|^2 + \|\eta\|^2 + 2\mathrm{Re}\big(\,\xi\,\big|\,\eta\,\big).\end{aligned}$$

We now use the C-B-S inequality, which reads

$$\big|\big(\,\xi\,\big|\,\eta\,\big)\big| \leq \|\xi\|\cdot\|\eta\|, \tag{3}$$

so the above computation gives

$$\begin{aligned}\|\xi + \eta\|^2 = \|\xi\|^2 + \|\eta\|^2 + 2\mathrm{Re}\big(\xi|\eta\big) &\leq \|\xi\|^2 + \|\eta\|^2 + 2\big|\big(\xi|\eta\big)\big| \leq \\ &\leq \|\xi\|^2 + \|\eta\|^2 + 2\|\xi\|\cdot\|\eta\| = \big(\|\xi\| + \|\eta\|\big)^2,\end{aligned}$$

so we immediately get $\|\xi + \eta\| \leq \|\xi\| + \|\eta\|$. □

Definition. The norm constructed in the above result is called the *norm defined by the inner product* $\big(\,\cdot\,\big|\,\cdot\,\big)$.

Exercise 1. Use the above notations, and assume we have two vectors $\xi, \eta \neq 0$, such that $\|\xi + \eta\| = \|\xi\| + \|\eta\|$. Prove that there exists some $\lambda > 0$ such that $\xi = \lambda\eta$.

Lemma 6.1. *Let $\mathcal{X}$ be a $\mathbb{K}$-vector space, equipped with an inner product.*

(ii) [Parallelogram Law] $\|\xi + \eta\|^2 + \|\xi - \eta\|^2 = 2\big(\|\xi\|^2 + \|\eta\|^2\big)$.

(i) [Polarization Identities]
(a) *If* $\mathbb{K} = \mathbb{R}$, *then*

$$\big(\,\xi\,\big|\,\eta\,\big) = \frac{1}{4}\big[\|\xi+\eta\|^2 - \|\xi-\eta\|\big],\ \ \forall\,\xi,\eta\in\mathcal{X}.$$

(b) *If* $\mathbb{K} = \mathbb{R}$, *then*

$$\big(\,\xi\,\big|\,\eta\,\big) = \frac{1}{4}\sum_{k=0}^{3} i^{-k}\|\xi + i^k\eta\|^2,\ \ \forall\,\xi,\eta\in\mathcal{X}.$$

Proof. (i). This is obvious, since (since the computations from the proof of Corollary ??)

$$\|\xi\pm\eta\|^2 = \|\xi\|^2 + \|\eta\|^2 \pm 2\mathrm{Re}\big(\,\xi\,\big|\,\eta\,\big).$$

(ii).(a). In the real case, the above identity gives

$$\|\xi\pm\eta\|^2 = \|\xi\|^2 + \|\eta\|^2 \pm 2\big(\,\xi\,\big|\,\eta\,\big),$$

so we immediately get

$$\|\xi+\eta\|^2 - \|\xi-\eta\|^2 = 4\big(\,\xi\,\big|\,\eta\,\big).$$

(b). For every $k\in\{0,1,2,3\}$, we have

$$\|\xi + i^k\eta\|^2 = \|\xi\|^2 + \|\eta\|^2 + 2\mathrm{Re}\big(\,\xi\,\big|\,i^k\eta\,\big) = \|\xi\|^2 + \|\eta\|^2 + i^k\big(\,\xi\,\big|\,\eta\,\big) + i^{-k}\big(\,\eta\,\big|\,\xi\,\big).$$

Then, when we sum up, we have

$$\sum_{k=0}^{3} i^{-k}\|\xi + i^k\eta\|^2 = \big(\|\xi\|^2+\|\eta\|^2\big)\sum_{k=0}^{3} i^{-k} + 4\big(\,\xi\,\big|\,\eta\,\big) + \big(\,\eta\,\big|\,\xi\,\big)\sum_{k=0}^{3} i^{-2k}.$$

Since

$$\sum_{k=0}^{3} i^{-k} = \sum_{k=0}^{3} i^{-2k} = 0,$$

the above computation proves that we indeed have

$$\sum_{k=0}^{3} i^{-k}\|\xi + i^k\eta\|^2 = 4\big(\,\xi\,\big|\,\eta\,\big).$$

□

Corollary 6.2. *Let* $\mathcal{X}$ *be a* $\mathbb{K}$*-vector space equipped with an inner product* $\big(\,\cdot\,\big|\,\cdot\,\big)$. *Then the map*

$$\mathcal{X}\times\mathcal{X}\ni(\xi,\eta)\longmapsto\big(\,\xi\,\big|\,\eta\,\big)\in\mathbb{K}$$

is continuous, with respect to the product topologies.

Proof. Immediate from the polarization identities. □

Corollary 6.3. *Let* $\mathcal{X}$ *and* $\mathcal{Y}$ *be two* $\mathbb{K}$*-vector spaces equipped with inner products* $\big(\,\cdot\,\big|\,\cdot\,\big)_{\mathcal{X}}$ *and* $\big(\,\cdot\,\big|\,\cdot\,\big)_{\mathcal{Y}}$. *If* $T:\mathcal{X}\to\mathbb{K}$ *is an isometric linear map, then*

$$\big(\,T\xi\,\big|\,T\eta\,\big)_{\mathcal{Y}} = \big(\,\xi\,\big|\,\eta\,\big)_{\mathcal{X}},\ \ \forall\,\xi,\eta\in\mathcal{X}.$$

Proof. Immediate from the polarization identities. □

Exercise 2. Let $\mathcal{X}$ be a normed $\mathbb{K}$-vector space. Assume the norm satisfies the Parallelogram Law. Prove that there exists an inner product $(\,\cdot\,|\,\cdot\,)$ on $\mathcal{X}$, such that

$$\|\xi\| = \sqrt{(\,\xi\,|\,\xi\,)},\ \ \forall\,\xi \in \mathcal{X}.$$

HINT: *Define* the inner product by the Polarization Identity, and then prove that it is indeed an inner product.

PROPOSITION 6.2. *Let $\mathcal{X}$ be a $\mathbb{K}$-vector space, equipped with an inner product $(\,\cdot\,|\,\cdot\,)_{\mathcal{X}}$. Let $\mathcal{Z}$ be the completion of $\mathcal{X}$ with respect to the norm defined by the inner product. Then $\mathcal{Z}$ carries a unique inner product $(\,\cdot\,|\,\cdot\,)_{\mathcal{Z}}$, so that the norm on $\mathcal{Z}$ is defined by $(\,\cdot\,|\,\cdot\,)_{\mathcal{Z}}$. Moreover, this inner product extends $(\,\cdot\,|\,\cdot\,)_{\mathcal{X}}$, in the sense that*

$$(\,\langle\xi\rangle\,|\,\langle\eta\rangle\,)_{\mathcal{Z}} = (\,\xi\,|\,\eta\,)_{\mathcal{X}},\ \ \forall\,\xi,\eta \in \mathcal{X}.$$

PROOF. It is obvious that the norm on $\mathcal{Z}$ satisfies the Parallelogram Law. We then apply Exercise 2. □

DEFINITIONS. Let $\mathbb{K}$ be one of the fields $\mathbb{R}$ or $\mathbb{C}$. A *Hilbert space over* $\mathbb{K}$ is a $\mathbb{K}$-vector space, equipped with an inner product, which is complete with respect to the norm defined by the inner product. Some textbooks use the term *Euclidean* for real Hilbert spaces, and reserve the term *Hilbert* only for the complex case.

EXAMPLES 6.1. For I a non-empty set, the space $\ell^2_{\mathbb{K}}(I)$ is a Hilbert space. We know that this is a Banach space. The inner product defining the norm is

$$(\,\alpha\,|\,\beta\,) = \sum_{j\in I}\overline{\alpha(j)}\beta(j),\ \ \forall\,\alpha,\beta \in \ell^2_{\mathbb{K}}(I).$$

The fact that the function $\overline{\alpha}\beta : I \to \mathbb{K}$ is summable is a consequence of Hölder's inequality.

More generally, a Banach space whose norm satisfies the Parallelogram Law is a Hilbert space.

DEFINITIONS. Let $\mathcal{X}$ be a $\mathbb{K}$-vector space, equipped with an inner product $(\,\cdot\,|\,\cdot\,)$. Two vectors $\xi,\eta \in \mathcal{X}$ are said to be *orthogonal*, if $(\,\xi\,|\,\eta\,) = 0$. In this case we write $\xi \perp \eta$. Given a set $\mathcal{M} \subset \mathcal{X}$, and a vector $\xi \in \mathcal{X}$, we write $\xi \perp \mathcal{M}$, if

$$\xi \perp \eta,\ \ \forall\,\eta \in \mathcal{M}.$$

Finally, two subsets $\mathcal{M},\mathcal{N} \subset \mathcal{X}$ are said to be orthogonal, and we write $\mathcal{M} \perp \mathcal{N}$, if

$$\xi \perp \eta,\ \ \forall\,\xi \in \mathcal{M},\ \eta \in \mathcal{N}.$$

NOTATION. Let $\mathcal{X}$ be a vector space equipped with an inner product. For a subset $\mathcal{M} \subset \mathcal{X}$, we define the set

$$\mathcal{M}^{\perp} = \{\xi \in \mathcal{X} : \xi \perp \mathcal{M}\}.$$

REMARKS 6.1. Let $\mathcal{X}$ be a $\mathbb{K}$-vector space equipped with an inner product.

A. The relation $\perp$ is symmetric.

B. If $\xi,\eta \in \mathcal{X}$ satisfy $\xi \perp \eta$, then one has the Pythagorean Theorem:

$$\|\xi+\eta\|^2 = \|\xi\|^2 + \|\eta\|^2.$$

This is a consequence of the equality $\|\xi+\eta\|^2 = \|\xi\|^2 + \|\eta\|^2 + 2\mathrm{Re}(\,\xi\,|\,\eta\,)$.

C. If $\mathcal{M} \subset \mathcal{X}$ is an arbitrary subset, then $\mathcal{M}^\perp$ is a closed linear subspace of $\mathcal{X}$. This follows from the linearity of the inner product in the second variable, and from the continuity.

D. For sets $\mathcal{M} \subset \mathcal{N} \subset \mathcal{X}$, one has

$$\mathcal{M}^\perp \supset \mathcal{N}^\perp.$$

E. For any set $\mathcal{M} \subset \mathcal{X}$, one has

$$\mathcal{M}^\perp = \left(\overline{\text{Span}\,\mathcal{M}}\right)^\perp,$$

where $\overline{\text{Span}\,\mathcal{M}}$ denotes the norm closure of the linear span of $\mathcal{M}$. The inclusion

$$\mathcal{M}^\perp \supset \left(\overline{\text{Span}\,\mathcal{M}}\right)^\perp$$

is trivial, since we have $\mathcal{M} \subset \overline{\text{Span}\,\mathcal{M}}$. Conversely, if $\xi \in \mathcal{M}^\perp$, then $\mathcal{M} \subset \{\xi\}^\perp$. But since $\{\xi\}^\perp$ is a closed linear subspace, this gives

$$\overline{\text{Span}\,\mathcal{M}} \subset \{\xi\}^\perp,$$

i.e. $\xi \in \left(\overline{\text{Span}\,\mathcal{M}}\right)^\perp$.

The following result gives a very interesting property of Hilbert spaces.

PROPOSITION 6.3. *Let $\mathcal{H}$ be a Hilbert space, let $\mathcal{C} \subset \mathcal{H}$ be a non-empty closed convex set. For every $\xi \in \mathcal{H}$, there exists a unique vector $\xi' \in \mathcal{C}$, such that*

$$\|\xi - \xi'\| = \text{dist}(\xi, \mathcal{C}).$$

PROOF. Denote $\text{dist}(\xi, \mathcal{C})$ simply by d. By definition, we have

$$\delta = \inf_{\eta \in \mathcal{C}} \|\xi - \eta\|.$$

Choose a sequence $(\eta_n)_{n\geq 1} \subset \mathcal{C}$, such that $\lim_{n\to\infty} \|\xi - \eta_n\| = \delta$.

Claim: *One has the inequality*

$$\|\eta_m - \eta_n\|^2 \leq 2\|\xi - \eta_m\|^2 + 2\|\xi - \eta_n\|^2 - 4\delta^2, \ \ \forall\, m, n \geq 1.$$

Use the Parallelogram Law

$$2\|\xi - \eta_m\|^2 + 2\|\xi - \eta_n\|^2 = \|2\xi - \eta_m - \eta_n\|^2 + \|\eta_m - \eta_n\|^2. \tag{4}$$

We notice that, since $\frac{1}{2}(\eta_m + \eta_n) \in \mathcal{C}$, we have

$$\|\xi - \tfrac{1}{2}(\eta_m + \eta_n)\| \geq \delta,$$

so we get

$$\|2\xi - \eta_m - \eta_n\|^2 = 4\|\xi - \tfrac{1}{2}(\eta_m + \eta_n)\|^2 \geq 4\delta^2,$$

so if we go back to (4) we get

$$2\|\xi - \eta_m\|^2 + 2\|\xi - \eta_n\|^2 = \|2\xi - \eta_m - \eta_n\|^2 + \|\eta_m - \eta_n\|^2 \geq 4\delta^2 + \|\eta_m - \eta_n\|^2,$$

and the Claim follows.

Having proven the Claim, we now notice that, since $\lim_{n\to\infty} \|\xi - \eta_n\| = \delta$, we immediately get the fact that *the sequence $(\eta_n)_{n\geq 1}$ is Cauchy.* Since $\mathcal{H}$ is complete, it follows that the sequence is convergent to some point ξ'. Since $\mathcal{C}$ is closed, it follows that $\xi' \in \mathcal{C}$. So far we have

$$\|\xi - \xi'\| = \lim_{n\to\infty} \|\xi - \eta_n\| = \delta = \text{dist}(\xi, \mathcal{C}),$$

thus proving the existence.

Let us prove now the uniqueness. Assume $\xi'' \in \mathcal{C}$ is another point such that $\|\xi - \xi''\| = \delta$. Using the Parallelogram Law, we have

$$4\delta^2 = 2\|\xi - \xi'\|^2 + \|\xi - \xi''\|^2 = \|2\xi - \xi' - \xi''\|^2 + \|\xi' - \xi''\|^2.$$

If $\xi' \neq \xi''$, then we will have

$$4\delta^2 > \|2\xi - \xi' - \xi''\|^2 = 4\|\xi - \tfrac{1}{2}(\xi' + \xi'')\|^2,$$

so we have a new vector $\eta = \frac{1}{2}(\xi' + \xi'') \in \mathcal{C}$, such that

$$\|\xi - \eta\| < \delta,$$

thus contracting the definition of δ. □

Definition. Let $\mathcal{H}$ be a Hilbert space, and let $\mathcal{X} \subset \mathcal{H}$ be a closed linear subspace. For every $\xi \in \mathcal{H}$, using the above result, we let $P_{\mathcal{X}}\xi \in \mathcal{X}$ denote the unique vector in $\mathcal{X}$ with the property

$$\|\xi - P_{\mathcal{X}}\xi\| = \text{dist}(\xi, \mathcal{X}).$$

This way we have constructed a map $P_{\mathcal{X}} : \mathcal{H} \to \mathcal{H}$, which is called the *orthogonal projection ont* $\mathcal{X}$.

The properties of the orthogonal projection are summarized in the following result.

Proposition 6.4. *Let $\mathcal{H}$ be a Hilbert space, and let $\mathcal{X} \subset \mathcal{H}$ be a closed linear subspace.*

(i) *For vectors $\xi \in \mathcal{H}$ and $\zeta \in \mathcal{X}$ one has the equivalence*

$$\zeta = P_{\mathcal{X}}\xi \Longleftrightarrow (\xi - \zeta) \perp \mathcal{X}.$$

(ii) $P_{\mathcal{X}}\big|_{\mathcal{X}} = \text{Id}_{\mathcal{X}}$.

(iii) *The map $P_{\mathcal{X}} : \mathcal{H} \to \mathcal{X}$ is linear, continuous. If $\mathcal{X} \neq \{0\}$, then $\|P_{\mathcal{X}}\| = 1$.*

(iv) $\text{Ran}\, P_{\mathcal{X}} = \mathcal{X}$ *and* $\text{Ker}\, P_{\mathcal{X}} = \mathcal{X}^{\perp}$.

Proof. (i). "⇒." Assume $\zeta = P_{\mathcal{X}}\xi$. Fix an arbitrary vector $\eta \in \mathcal{X} \smallsetminus \{0\}$, and choose a number $\lambda \in \mathbb{K}$, with $|\lambda| = 1$, such that

$$\lambda\big(\,\xi - \zeta\,\big|\,\eta\,\big) = \big|\big(\,\xi - \zeta\,\big|\,\eta\,\big)\big|.$$

In particular, we have

$$\big|\big(\,\xi - \zeta\,\big|\,\eta\,\big)\big| = \text{Re}\big(\,\xi - \zeta\,\big|\,\lambda\eta\,\big).$$

Define the map $F : \mathbb{R} \to \mathbb{R}$ by

$$F(t) = \|\xi - \zeta - t\lambda\eta\|^2 - \|\xi - \zeta\|^2.$$

By the definition of $\zeta = P_{\mathcal{X}}\xi$, we have

(5) $$F(t) > 0, \ \ \forall\, t \in \mathbb{R} \smallsetminus \{0\}.$$

Notice that $F(t) = at^2 + bt$, $\forall\, t \in \mathbb{R}$, where $a = \big(\,\lambda\eta\,\big|\,\lambda\eta\,\big) = \|\eta\|^2$, and $b = 2\text{Re}\big(\,\xi - \zeta\,\big|\,\lambda\eta\,\big) = 2\big|\big(\,\xi - \zeta\,\big|\,\eta\,\big)\big|$. Of course, the property

$$at^2 + bt > 0, \ \ \forall\, t \in \mathbb{R} \smallsetminus \{0\}$$

forces $b = 0$, so we indeed get $\big(\,\xi - \zeta\,\big|\,\eta\,\big) = 0$.

"⇐." Assume $(\xi - \zeta) \perp \mathcal{X}$. For any $\eta \in \mathcal{X}$, we have $(\xi - \zeta) \perp (\zeta - \eta)$, so using the Pythagorean Theorem, we get

$$\|\xi - \eta\|^2 = \|\xi - \zeta\|^2 + \|\zeta - \eta\|^2,$$

which forces
$$\|\xi - \eta\| \geq \|\xi - \zeta\|, \ \forall \eta \in \mathcal{X}.$$
This proves that
$$\|\xi - \zeta\| = \text{dist}(\xi, \mathcal{X}),$$
i.e. $\zeta = P_{\mathcal{X}}\xi$.

(ii). This property is pretty clear. If $\xi \in \mathcal{X}$, then $0 = \xi - \xi$ is orthogonal to $\mathcal{X}$, so by (i) we get $\xi = P_{\mathcal{X}}\xi$.

(iii). We prove the linearity of $P_{\mathcal{X}}$. Start with vectors $\xi_1, \xi - 2 \in \mathcal{H}$ and a scalar $\lambda \in \mathbb{K}$. Take $\zeta_1 = P_{\mathcal{X}}\xi_1$ and $\zeta_2 = P_{\mathcal{X}}\xi_2$. Consider the vector $\zeta = \lambda\zeta_1 + \zeta_2$. For any $\eta \in \mathcal{X}$, we have
$$\big(\lambda\xi_1 + \xi_2 - \zeta \,\big|\, \eta\big) = \big((\lambda\xi_1 - \lambda\zeta_1) + (\xi_2 - \zeta_2) \,\big|\, \eta\big) =$$
$$= \big(\lambda\xi_1 - \lambda\zeta_1 \,\big|\, \eta\big) + \big(\xi_2 - \zeta_2 \,\big|\, \eta\big) = \overline{\lambda}\big(\xi_1 - \zeta_1 \,\big|\, \eta\big) + \big(\xi_2 - \zeta_2 \,\big|\, \eta\big) = 0.$$
By (i) we have $(\xi_1 - \zeta_1) \perp \mathcal{X}$ and $(\xi_1 - \zeta_1) \perp \mathcal{X}$, so the above computation proves that
$$(\lambda\xi_1 + \xi_2 - \zeta) \perp \mathcal{X},$$
so using (i) we get
$$P_{\mathcal{X}}(\lambda\xi_1 + \xi_2) = \zeta = \lambda\zeta_1 + \zeta_2 = \lambda P_{\mathcal{X}}\xi_1 + P_{\mathcal{X}}\xi_2,$$
so $P_{\mathcal{X}}$ is indeed linear.

To prove the continuity, we start with an arbitrary vector $\xi \in \mathcal{H}$ and we use the fact that $(\xi - P_{\mathcal{X}}\xi) \perp P_{\mathcal{X}}\xi$. By the Pythagorean Theorem we then have
$$\|\xi\|^2 = \|(\xi - P_{\mathcal{X}}\xi) + P_{\mathcal{X}}\xi\|^2 = \|\xi - P_{\mathcal{X}}\xi\|^2 + \|P_{\mathcal{X}}\xi\|^2 \geq \|P_{\mathcal{X}}\xi\|^2.$$
In other words, we have
$$\|P_{\mathcal{X}}\xi\| \leq \|\xi\|, \ \forall \xi \in \mathcal{H},$$
so $P_{\mathcal{X}}$ is indeed continuous, and we have $\|P_{\mathcal{X}}\| \leq 1$. Using (ii) we immediately get that, when $\mathcal{X} \neq \{0\}$, we have $\|P_{\mathcal{X}}\| = 1$.

(iv). The equality $\text{Ran}\, P_{\mathcal{X}} = \mathcal{X}$ is trivial by the construction of $P_{\mathcal{X}}$ and by (ii). If $\xi \in \text{Ker}\, P_{\mathcal{X}}$, then by (i), we have $\xi \in \mathcal{X}^{\perp}$. Conversely, if $\xi \perp \mathcal{X}$, then $\zeta = 0$ satisfies the condition in (i), i.e. $P_{\mathcal{X}}\xi = 0$. ☐

Corollary 6.4. *If $\mathcal{H}$ is a Hilbert space, and $\mathcal{X} \subset \mathcal{H}$ is a closed linear subspace, then*
$$\mathcal{X} + \mathcal{X}^{\perp} = \mathcal{H} \text{ and } \mathcal{X} \cap \mathcal{X}^{\perp}.$$
In other words the map
$$\mathcal{X} \times \mathcal{X}^{\perp} \ni (\eta, \zeta) \longmapsto \eta + \zeta \in \mathcal{H} \tag{6}$$
is a linear isomorphism.

Proof. If $\xi \in \mathcal{H}$ then $P_{\mathcal{X}}\xi \in \mathcal{X}$, and $\xi - P_{\mathcal{X}}\xi \in \mathcal{X}^{\perp}$, and then the equality
$$\xi = P_{\mathcal{X}}\xi + (\xi - P_{\mathcal{X}}\xi)$$
proves that $\xi \in \mathcal{X} + \mathcal{X}^{\perp}$. The equality $\mathcal{X} \cap \mathcal{X}^{\perp} = \{0\}$ is trivial, since for $\zeta \in \mathcal{X} \cap \mathcal{X}^{\perp}$, we must have $\zeta \perp \zeta$, which forces $\zeta = 0$. ☐

Exercise 3. Let $\mathcal{H}$ be a Hilbert space.

(i) Prove that, for any closed subspace $\mathcal{X} \subset \mathcal{H}$, one has the equality
$$P_{\mathcal{X}^{\perp}} = I - P_{\mathcal{X}}.$$

(ii) Prove that two closed subspaces $\mathcal{X}, \mathcal{Y} \subset \mathcal{H}$, the following are equivalent:
- $\mathcal{X} \perp \mathcal{Y}$;
- $P_{\mathcal{X}} P_{\mathcal{Y}} = 0$;
- $P_{\mathcal{Y}} P_{\mathcal{X}} = 0$.

(iii) Prove that two closed subspaces $\mathcal{X}, \mathcal{Y} \subset \mathcal{H}$, the following are equivalent:
- $\mathcal{X} \subset \mathcal{Y}$;
- $P_{\mathcal{X}} P_{\mathcal{Y}} = P_{\mathcal{X}}$;
- $P_{\mathcal{Y}} P_{\mathcal{X}} = P_{\mathcal{X}}$.

(iv) Let $\mathcal{X}, \mathcal{Y} \subset \mathcal{H}$ are closed subspaces, such that $\mathcal{X} \perp \mathcal{Y}$, then
- $\mathcal{X} + \mathcal{Y}$ is c closed linear subspace of $\mathcal{H}$;
- $P_{\mathcal{X}+\mathcal{Y}} = P_{\mathcal{X}} + P_{\mathcal{Y}}$.

Corollary 6.5. *Let $\mathcal{H}$ be a Hilbert space, and let $\mathcal{X} \subset \mathcal{H}$ be a linear (not necessarily closed) subspace. Then on has the equality*

$$\overline{\mathcal{X}} = \left(\mathcal{X}^{\perp}\right)^{\perp}.$$

Proof. Denote the closed subspace $\left(\mathcal{X}^{\perp}\right)^{\perp}$ by $\mathcal{Z}$. Since $\mathcal{X}^{\perp} = \overline{\mathcal{X}}^{\perp}$, by the previous exercise we have

$$P_{\mathcal{Z}} = I - P_{\mathcal{X}^{\perp}} = I - P_{\overline{\mathcal{X}}^{\perp}} = I - (I - P_{\overline{\mathcal{X}}}) = P_{\overline{\mathcal{X}}},$$

which forces

$$\mathcal{Z} = \operatorname{Ran} P_{\mathcal{Z}} = \operatorname{Ran} P_{\overline{\mathcal{X}}} = \overline{\mathcal{X}}.$$

□

Theorem 6.1 (Riesz' Representation Theorem). *Let $\mathcal{H}$ be a Hilbert space over $\mathbb{K}$, and let $\phi : \mathcal{H} \to \mathbb{K}$ be a linear continuous map. Then there exists a unique vector $\xi \in \mathcal{H}$, such that*

$$\phi(\eta) = \left(\xi \,\middle|\, \eta\right), \ \ \forall\, \eta \in \mathcal{H}.$$

Moreover one has $\|\xi\| = \|\phi\|$.

Proof. First we show the existence. If $\phi = 0$, we simply take $\xi = 0$. Assume $\phi \neq 0$. Define the subspace $\mathcal{X} = \operatorname{Ker} \phi$. Notice that $\mathcal{X}$ is *closed*. Using the linear isomorphism (6) we see that the composition

$$\mathcal{X}^{\perp} \hookrightarrow \mathcal{H} \xrightarrow{\text{quotient map}} \mathcal{H}/\mathcal{X}$$

is a linear isomorphism. Since

$$\mathcal{H}/\mathcal{X} = \mathcal{H}/\operatorname{Ker} \phi \simeq \operatorname{Ran} \phi = \mathbb{K},$$

it follows that $\dim(\mathcal{X}^{\perp}) = 1$. In other words, there exists $\xi_0 \in \mathcal{X}^{\perp}$, $\xi_0 \neq 0$, such that

$$\mathcal{X}^{\perp} = \mathbb{K}\xi.$$

Start now with some arbitrary vector $\eta \in \mathbb{H}$. On the one hand, using the equality $\mathbb{K}\xi_0 + \mathcal{X} = \mathcal{H}$, there exists $\lambda \in \mathbb{K}$ and $\zeta \in \mathcal{X}$, such that

$$\eta = \lambda \xi_0 + \zeta,$$

and since $\zeta \in \mathcal{X} = \operatorname{Ker} \phi$, we get

$$\phi(\eta) = \phi(\lambda \xi_0) = \lambda \phi(\xi_0).$$

On the other hand, we have

$$\left(\xi_0 \,\middle|\, \eta\right) = \left(\xi_0 \,\middle|\, \lambda \xi_0\right) + \left(\xi_0 \,\middle|\, \zeta\right) = \lambda \|\xi_0\|^2,$$

so if we define $\xi = \overline{\phi(\xi_0)}\|\xi_0\|^{-2}$ we will have

$$\left(\xi \,\middle|\, \eta \right) = \left(\overline{\phi(\xi_0)}\|\xi_0\|^{-2}\xi_0 \,|\, \eta \right) = \phi(\xi_0)\|\xi_0\|^{-2}\left(\xi_0 \,\middle|\, \eta \right) = \lambda\phi(\xi_0) = \phi(\eta).$$

To prove uniqueness, assume $\xi' \in \mathcal{H}$ is another vector with

$$\phi(\eta) = \left(\xi' \,\middle|\, \eta \right), \ \ \forall\, \eta \in \mathcal{H}.$$

In particular, we have

$$\|\xi - \xi'\|^2 = \left(\xi - \xi' \,|\, \xi - \xi' \right) = \left(\xi \,|\, \xi - \xi' \right) - \left(\xi' \,|\, \xi - \xi' \right) = \phi(\xi - \xi') - \phi(\xi - \xi') = 0,$$

which forces $\xi = \xi'$.

Finally, to prove the norm equality, we first observe that when $\xi = 0$, the equality is trivial. If $\xi \neq 0$, then on the one hand, using C-B-S inequality we have

$$|\phi(\eta)| = \left|\left(\xi \,\middle|\, \eta \right)\right| \leq \|\xi\| \cdot \|\eta\|, \ \ \forall\, \eta \in \mathcal{H},$$

so we immediately get $\|\phi\| \leq \|\xi\|$. If we take the vector $\zeta = \|\xi\|^{-1}\xi$, then $\|\zeta\| = 1$, and

$$\phi(\zeta) = \left(\xi \,\middle|\, \|\xi\|^{-1}\xi \right) = \|\xi\|,$$

so we also have $\|\phi\| \geq \|\xi\|$. □

In the remainder of this section we discuss a Hilbert space notion of linear independence. This should be thought as a "rigid" linear independence.

Definition. Let $\mathcal{X}$ be a $\mathbb{K}$-vector space, equipped with an inner product. A set $\mathcal{F} \subset \mathcal{X}$ is said to be *orthogonal*, if $0 \notin \mathcal{F}$, and

$$\xi \perp \eta, \ \ \forall\, \xi, \eta \in \mathcal{F}, \text{ with } \xi \neq \eta.$$

A set $\mathcal{F} \subset \mathcal{X}$ is said to be *orthonormal*, if it is orthogonal, but it also satisfies:

$$\|\xi\| = 1, \ \ \forall\, \xi \in \mathcal{F}.$$

Remark that, if one starts with an orthogonal set $\mathcal{F} \subset \mathcal{X}$, then the set

$$\mathcal{F}^{(1)} = \left\{ \|\xi\|^{-1}\xi \,:\, \xi \in \mathcal{F} \right\}$$

is orthonormal.

Proposition 6.5. *Let $\mathcal{X}$ be a $\mathbb{K}$-vector space equipped with an inner product. Any orthogonal set $\mathcal{F} \subset \mathcal{X}$ is linearly independent.*

Proof. Indeed, if one starts with a vanishing linear combination

$$\lambda_1\xi_1 + \cdots + \lambda_n\xi_n = 0,$$

with $\lambda_1, \dots, \lambda_n \in \mathbb{K}$, $\xi_1, \dots, \xi_n \in \mathcal{X}$, such that $\xi_k \neq \xi_\ell$, for all $k, \ell \in \{1, \dots, n\}$ with $k \neq \ell$, then for each $k \in \{1, \dots, n\}$ we clearly have

$$\lambda_k\|\xi_k\|^2 = \left(\xi_k \,\middle|\, \lambda - 1\xi_1 + \cdots + \lambda_n\xi_n \right) = 0,$$

and since $\xi_k \neq 0$, we get $\lambda_k = 0$. □

Lemma 6.2. *Let $\mathcal{X}$ be a $\mathbb{K}$-vector space equipped with an inner product, and let $\mathcal{F} \subset \mathcal{X}$ be an orthogonal set. Then there exists a maximal (with respect to inclusion) orthogonal set $\mathcal{G} \subset \mathcal{X}$ with $\mathcal{F} \subset \mathcal{G}$.*

PROOF. Consider the sets

$$\mathfrak{A} = \{\mathcal{G} : \mathcal{G} \text{ orthogonal subset of } \mathcal{X}\},$$
$$\mathfrak{A}_{\mathcal{F}} = \{\mathcal{G} \in \mathfrak{A} : \mathcal{G} \supset \mathcal{F}\},$$

ordered with the inclusion. We are going to apply Zorn's Lemma to $\mathfrak{A}_{\mathcal{F}}$. Let $\mathfrak{T} \subset \mathfrak{A}_{\mathcal{F}}$ be a subcollection, which is totally ordered, i.e. for any $\mathcal{G}_1, \mathcal{G}_2 \in \mathfrak{T}$ one has $\mathcal{G}_1 \subset \mathcal{G}_2$ or $\mathcal{G}_1 \supset \mathcal{G}_2$. Define the set

$$\mathcal{M} = \bigcup_{\mathcal{G} \in \mathfrak{T}} \mathcal{G}.$$

Since $\mathcal{G} \subset \mathcal{X} \smallsetminus \{0\}$, for all $\mathcal{G} \in \mathfrak{T}$, it is clear that $\mathcal{M} \subset \mathcal{X} \smallsetminus \{0\}$. If $\xi_1, \xi_2 \in \mathcal{M}$ are vectors with $\xi_1 \neq \xi_2$, then we can find $\mathcal{G}_1, \mathcal{G}_2 \in \mathfrak{T}$ with $\xi_1 \in \mathcal{G}_1$ and $\xi_2 \in \mathcal{G}_2$. Using the fact that $\mathfrak{T}$ is totally ordered, it follows that there is $k \in \{1, 2\}$ such that $\xi_1, \xi_2 \in \mathcal{G}_k$, so we indeed get $\xi_1 \perp \xi_2$. It is now clear that $\mathcal{M} \in \mathfrak{A}_{\mathcal{F}}$, and $\mathcal{M} \supset \mathcal{G}$, for all $\mathcal{G} \in \mathfrak{T}$. In other words, we have shown that *every totally ordered subset of* $\mathfrak{A}_{\mathcal{F}}$ *has an upper bound, in* $\mathfrak{A}_{\mathcal{F}}$. By Zorn's Lemma, $\mathfrak{A}_{\mathcal{F}}$ has a maximal element. Finally, it is clear that any maximal element for $\mathfrak{A}_{\mathcal{F}}$ is also a maximal element in $\mathfrak{A}$. □

REMARK 6.2. Using the notations from the proof above, given an *orthonormal* set $\mathcal{M} \subset \mathcal{X}$, the following are equivalent:

(i) $\mathcal{M}$ is maximal in $\mathfrak{A}$;
(ii) $\mathcal{M}$ is maximal in

$$\mathfrak{A}^{(1)} = \{\mathcal{G} : \mathcal{G} \text{ orthonormal subset of } \mathcal{X}\}.$$

The implication $(i) \Rightarrow (ii)$ is trivial. Conversely, if $\mathcal{M}$ is maximal in $\mathfrak{A}^{(1)}$, we use the Lemma to find a maximal $\mathcal{N} \in \mathfrak{A}$ with $\mathcal{N} \supset \mathcal{M}$. But then $\mathcal{N}^{(1)}$ is orthonromal, and $\mathcal{N}^{(1)} \supset \mathcal{M}$, which by the maximality of $\mathcal{M}$ in $\mathfrak{A}^{(1)}$ will force $\mathcal{N}^{(1)} = \mathcal{M}$. Since $\mathcal{N}$ is linearly independent, the relations

$$\mathcal{N}^{(1)} = \mathcal{M} \subset \mathcal{N},$$

will force $\mathcal{N} = \mathcal{N}^{(1)} = \mathcal{M}$.

COMMENT. In linear algebra we know that a linearly independent set is maximal, if and only if it spans the whole space. In the case of orthogonal sets, this statement has a version described by the following result.

THEOREM 6.2. *Let $\mathcal{H}$ be a Hilbert space, and let $\mathcal{F}$ be an orthogonal set in $\mathcal{H}$. The following are equivalent:*

(i) *$\mathcal{F}$ is maximal among all orthogonal subsets of $\mathcal{H}$;*
(ii) Span $\mathcal{F}$ *is dense in $\mathcal{H}$ in the norm topology.*

PROOF. $(i) \Rightarrow (ii)$. Assume $\mathcal{F}$ is maximal. We are going to show that Span $\mathcal{F}$ is dense in $\mathcal{H}$, by contradiction. Denote the closure $\overline{\text{Span}\,\mathcal{F}}$ simply by $\mathcal{X}$, and assume $\mathcal{X} \subsetneq \mathcal{H}$. Since

$$\mathcal{X} = \left(\mathcal{X}^{\perp}\right)^{\perp},$$

we see that, the strict inclusion $\mathcal{X} \subsetneq \mathcal{H}$ forces $\mathcal{X}^{\perp} \neq \{0\}$. But now if we take a non-zero vector $\xi \in \mathcal{X}^{\perp}$, we immediately see that the set $\mathcal{F} \cup \{\xi\}$ is still orthogonal, thus contradicting the maximality of $\mathcal{F}$.

$(ii) \Rightarrow (i)$. Assume $\text{Span}\,\mathcal{F}$ is dense in $\mathcal{H}$, and let us prove that $\mathcal{F}$ is maximal. We do this by contardiction. If $\mathcal{F}$ is not maximal, then there exists $\xi \in \mathcal{H} \smallsetminus \mathcal{F}$, such that $\mathcal{F} \cup \{\xi\}$ is still orthogonal. This would force $\xi \perp \mathcal{F}$, so we will also have

$$\xi \perp \overline{\text{Span}\,\mathcal{F}}.$$

But since $\text{Span}\,\mathcal{F}$ is dense in $\mathcal{H}$, this will give $\xi \perp \mathcal{H}$. In particular we have $\xi \perp \xi$, which would force $\xi = 0$, thus contradicting the fact that $\mathcal{F} \cup \{\xi\}$ is orthigonal. (Recall that all elements of an orthigonal set are non-zero.) □

Definition. Let $\mathcal{H}$ be a Hilbert space An orthonormal set $\mathcal{B} \subset \mathcal{H}$, which is maximal among all orthogonal (or orthonormal) subsets of $\mathcal{H}$, is called an *orthonormal basis for* $\mathcal{H}$.

By Lemma ??, we know that given any orthonormal set $\mathcal{F} \subset \mathcal{H}$, there exists an orthonormal basis $\mathcal{B} \supset \mathcal{F}$.

By the above result, an orthonormal set $\mathcal{B} \subset \mathcal{H}$ is an orthonormal basis for $\mathcal{H}$, if and only if $\text{Span}\,\mathcal{B}$ is dense in $\mathcal{H}$.

Example 6.2. Let I be a non-empty set. Consider the Hilbert space $\ell^2_{\mathbb{K}}(I)$. Consider (see section II.2) the set

$$\mathcal{B} = \{\delta^i \,:\, i \in I\}.$$

Then

$$\text{Span}\,\mathcal{B} = \mathit{fin}_{\mathbb{K}}(I),$$

which is dense in $\ell^2_{\mathbb{K}}(I)$. The above result then says that $\mathcal{B}$ *is an orthonormal basis for* $\ell^2_{\mathbb{K}}(I)$.

The following exercise will be useful in the discussion of another interesting example.

Exercise 4. Equipp the space $C([0,1])$ with the inner product

$$\left(f \,\middle|\, g \right) = \int_0^1 \overline{f(t)} g(t)\, dt, \quad f, g \in C([0,1]).$$

The norm defined by this inner product is

$$\|f\|_2 = \left(\int_0^1 |f(t)|^2 dt \right)^{\frac{1}{2}}, \quad f \in C([0,1]).$$

Define the maps $e_n : [0,1] \ni t \longmapsto \exp(2n\pi i t) \in \mathbb{T}$, $n \in \mathbb{Z}$. (Here $\mathbb{T}$ denotes the unit circle in $\mathbb{C}$.) Prove that the set

$$\mathcal{B} = \{e_n \,:\, n \in \mathbb{Z}\}$$

is orthonormal in $C([0,1])$, and $\text{Span}\,\mathcal{B}$ is dense in $C([0,1])$ in the topology defined by the norm $\|\cdot\|_2$.

Hints: Define the space

$$\mathcal{P} = \{f \in C([0,1]) \,:\, f(0) = f(1)\}.$$

Prove that $\mathcal{P}$ is dense in $C([0,1])$ in the topology defined by the norm $\|\cdot\|_2$.

Prove that the map

$$\Phi : C(\mathbb{T}) \ni F \longmapsto F \circ e \in \mathcal{P}$$

is a linear isomorphism, which is isometric with respect to the *uniform* norms.

In order to prove that $\text{Span}\,\mathcal{B}$ is dense in $C([0,1])$ with respect to $\|\cdot\|_2$, it suffices to show that $\text{Span}\,\mathcal{B}$ is dense in $\mathcal{P}$ in the *uniform* norm. Equivalently, it suffices to show that

$$\Phi^{-1}(\text{Span}\,\mathcal{B})$$

is dense in $C(\mathbb{T})$, with respect to the uniform norm. To get this density use Stone-Weierstrass Theorem, plus the fact that the functions $\zeta_n = \Phi^{-1}(e_n) \in C(\mathbb{T})$ are defined by

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

EXAMPLE 6.3. We define $L^2([0,1])$ to be the completion of $C([0,1])$ with respect to the norm $\|\cdot\|_2$. Regard $C([0,1])$ as a dense linear subspace in $L^2([0,1])$, so we also regard

$$\mathcal{B} = \{e_n : n \in \mathbb{Z}\}$$

as a subset in $L^2([0,1])$. Then $\operatorname{Span}\mathcal{B}$ is dense in $L^2([0,1])$, so $\mathcal{B}$ is an *orthonormal basis for* $L^2([0,1])$.

LEMMA 6.3. *Let $\mathcal{B}$ be an orthonormal basis for the Hilbert space $\mathcal{H}$, and let $\mathcal{F} \subsetneq \mathcal{B}$ be an arbitrary non-empty subset.*

(i) *$\mathcal{F}$ is an orthonormal basis for the Hilbert space $\overline{\operatorname{Span}\mathcal{F}}$.*
(ii) $\big(\overline{\operatorname{Span}\mathcal{F}}\big)^{\perp} = \overline{\operatorname{Span}(\mathcal{B} \smallsetminus \mathcal{F})}$.

PROOF. (i). This is clear, since $\mathcal{F}$ is orthonormal and has dense span.

(ii). Denote for simplicity $\overline{\operatorname{Span}\mathcal{F}} = \mathcal{X}$ and $\overline{\operatorname{Span}(\mathcal{B} \smallsetminus \mathcal{F})} = \mathcal{Y}$. Since

$$\xi \perp \eta, \ \forall \xi \in \mathcal{F},\, \eta \in \mathcal{B} \smallsetminus \mathcal{F},$$

it is pretty obvious that $\mathcal{X} \perp \mathcal{Y}$. Since $\mathcal{X} + \mathcal{Y}$ clearly contains $\operatorname{Span}\mathcal{B}$, it follows that $\mathcal{X} + \mathcal{Y}$ is dense in $\mathcal{H}$. We know howver that $\mathcal{X} + \mathcal{Y}$ is closed, so we have in fact the equality

$$\mathcal{X} + \mathcal{Y} = \mathcal{H}.$$

This will then give

$$I = P_{\mathcal{H}} = P_{\mathcal{X}} + P_{\mathcal{Y}},$$

so we get

$$P_{\mathcal{Y}} = I - P_{\mathcal{X}} = P_{\mathcal{X}^{\perp}},$$

so

$$\mathcal{X}^{\perp} = \operatorname{Ran} P_{\mathcal{X}^{\perp}} = \operatorname{Ran} P_{\mathcal{Y}} = \mathcal{Y}.$$

□

THEOREM 6.3. *Let $\mathcal{H}$ be a Hilbert space, and let $\mathcal{B}$ be an orthonormal basis for $\mathcal{H}$, labelled*[1] *as $\mathcal{B} = \{\xi_j : j \in I\}$. For every vector $\eta \in \mathcal{H}$, let $\alpha^\eta : I \to \mathbb{K}$ be the map defined by*

$$\alpha^\eta(j) = \big(\xi_j \,\big|\, \eta\big), \ \forall j \in I.$$

(i) *For every $\eta \in \mathcal{H}$, the map α^η belongs to $\ell^2_{\mathbb{K}}(I)$.*
(ii) *The map*

$$T : \mathcal{H} \ni \eta \longmapsto \alpha^\eta \in \ell^2_{\mathbb{K}}(I)$$

is an isometric linear isomorphism.

PROOF. (i). Fix for the moment $\eta \in \mathcal{H}$. We must show that

$$\sup\Big\{\sum_{j \in F} |\alpha^\eta(i)|^2 : F \subset I, \text{ finite}\Big\} < \infty.$$

For any non-empty finite subset $F \subset I$, we define the subspace

$$\mathcal{H}_F = \operatorname{Span}\{\xi_j : j \in F\},$$

[1] This notation implicitly assumes that $\xi_j \neq \xi_k$, for all $j, k \in I$ with $j \neq k$.

and define the vector

$$\eta^F = \sum_{j\in F} big(\xi_j \,|\, \eta\,) \cdot \xi_j.$$

Claim: *For every finite set $F \subset I$, one has the equality*

$$\eta^F = P_{\mathcal{H}_F}\eta.$$

It suffices to prove that

$$(\eta - \eta^F) \perp \mathcal{H}_F.$$

But this is obvious, since if we start with some $k \in F$, then using the fact that $\big(\,\xi_k \,\big|\, \xi_j\,\big) = 0$, for all $j \in F \smallsetminus \{k\}$, together with the equality $\|\xi_k\| = 1$, we get

$$\big(\,\xi_k \,\big|\, \eta - \eta_F\,\big) = \big(\,\xi_k \,\big|\, \eta\,\big) - \sum_{j\in F} \big(\,\xi_j \,\big|\, \eta\,\big)\cdot\big(\,\xi_k \,\big|\, \xi_j\,\big) = \big(\,\xi_k \,\big|\, \eta\,\big) - \big(\,\xi_k \,\big|\, \eta\,\big)\cdot\big(\,\xi_k \,\big|\, \xi_k\,\big) = 0.$$

Having proven the Claim, let us observe that, since the terms in the sum that defines η^F are all orthogonal, we get

$$\|\eta^F\|^2 = \sum_{j\in F} \big\|\big(\,\xi_j \,\big|\, \eta\,\big)\cdot\xi_j\big\|^2 = \sum_{j\in F} \big|\big(\,\xi_j \,\big|\, \eta\,\big)\big|^2 \cdot \|\xi_j\|^2 = \sum_{j\in F} |\alpha^\eta(j)|^2.$$

Combining this computation with the Claim, we now have

$$\sum_{j\in F} |\alpha^\eta(j)|^2 = \|\eta_F\|^2 = \|P_{\mathcal{H}_F}\eta\|^2 \le \|\eta\|^2,$$

which proves that

$$\sup\Big\{\sum_{j\in F} |\alpha^\eta(i)|^2 \,:\, F \subset I,\ \text{finite}\Big\} < \|\eta\|.$$

(ii). The linearlity of T is obvious. The above inequality actually proves that

$$\|T\eta\| \le \|\eta\|,\ \ \forall\,\eta \in \mathbb{H}.$$

We now prove that in fact T is isometric. Since T is linear and continuous, it suffices to prove that $T\big|_{\operatorname{Span}\mathcal{B}}$ is isometric. Start with some vector $\eta \in \operatorname{Span}\mathcal{B}$, which means that there exists some finite set $F \subset I$, and scalars $(\lambda_k)_{k\in F} \subset \mathbb{K}$, such that $\eta = \sum_{k\in F} \lambda_k \xi_k$. Remark that

$$\big(\,\xi_j \,|\, \eta\,\big) = \sum_{k\in F} \lambda_k \big(\,\xi_j \,|\, \xi_j\,\big) = \begin{cases} \lambda_k & \text{if } k \in F \\ 0 & \text{if } k \in F \end{cases}$$

so the element $\alpha^\eta = T\eta \in \ell^2_{\mathbb{K}}(I)$ is defined by

$$\alpha^\eta(k) = \begin{cases} \lambda_k & \text{if } k \in F \\ 0 & \text{if } k \in F \end{cases}$$

This gives

$$\|\eta\|^2 = \sum_{j,k\in F} \lambda_j \overline{\lambda}_k \big(\,\xi_j \,\big|\, \xi_k\,\big) = \sum_{k\in F} |\lambda_k|^2 = \sum_{k\in F} |\alpha^\eta(k)|^2 = \|\alpha^\eta\|^2,$$

so we indeed get

$$\|\eta\| = \|T\eta\|,\ \ \forall\,\eta \in \operatorname{Span}\mathcal{B}.$$

Let us prove that T is surjective. Notice that, the above computation, applied to singleton sets $F = \{k\}$, $k \in I$, proves that

$$T\xi_k = \delta^k,\ \ \forall\,k \in I.$$

In particular, we have

$$\operatorname{Ran} T \supset T\big(\operatorname{Span} \mathcal{B}\big) = \operatorname{Span} T(\mathcal{B}) =$$
$$= \operatorname{Span}\{T\xi_k \,:\, k \in I\} = \operatorname{Span}\{\delta^k \,:\, k \in I\} = \mathit{fin}_{\mathbb{K}}(I),$$

which proves that $\operatorname{Ran} T$ is dense in $\ell^2_{\mathbb{K}}(I)$. We know however that T is isometric, so $\operatorname{Ran} T \subset \ell^2_{\mathbb{K}}(I)$ is closed. This forces $\operatorname{Ran} T = \ell^2_{\mathbb{K}}(I)$. □

COROLLARY 6.6 (Parseval Identity). *Let $\mathcal{H}$ be a Hilbert space, and let $\mathcal{B} = \{\xi_j \,:\, j \in I\}$ be an orthonormal basis for $\mathcal{H}$. One has:*

$$\big(\,\zeta \,\big|\, \eta\,\big) = \sum_{j\in I} \big(\,\zeta \,\big|\, \xi_j\,\big) \cdot \big(\,\xi_j \,\big|\, \eta\,\big), \;\; \forall\, \zeta, \eta \in \mathcal{H}.$$

PROOF. If we define $\alpha(j) = \big(\,\xi_j \,\big|\, \zeta\,\big)$ and $\big(\,\xi_j \,\big|\, \eta\,\big)$, $\forall\, j \in I$, then by construction we have $\alpha = T\zeta$ and $\beta = T\eta$. Using the fact that T is isometric, the right hand side of the above equality is the equal to

$$\sum_{j\in I} \overline{\alpha(j)}\beta(j) = \big(\,\alpha \,\big|\, \beta\,\big) = \big(\,T\zeta \,\big|\, T\eta\,\big) = \big(\,\zeta \,\big|\, \eta\,\big).$$

□

NOTATION. Let $\mathcal{H}$ be a Hilbert space, let $\mathcal{B} = \{\xi_j \,:\, j \in I\}$ be an orthonormal basis for $\mathcal{H}$, and let $T : \mathcal{H} \to \ell^2_{\mathbb{K}}(I)$ be the isometric linear isomorphism defined in the previous theorem. Given an element $\alpha \in \ell^2_{\mathbb{K}}(I)$, we denote the vector $T^{-1}\alpha \in \mathcal{H}$ by

$$\sum_{j\in I} \alpha(j)\xi_j.$$

The summation notation is justified by the following fact.

PROPOSITION 6.6. *With the above notations, for every $\varepsilon > 0$, there exists some finite subset $F_\varepsilon \subset I$, such that*

$$\Big\| \sum_{j\in I} \alpha(j)\xi_j - \sum_{k\in F} \alpha(k)\xi_k \Big\|^2 < \varepsilon, \textit{ for all finite sets } F \subset I \textit{ with } F \supset F_\varepsilon.$$

PROOF. Define the vector $\eta = \sum_{j\in I} \alpha(j)\xi_j$. By construction we have $T\eta = \alpha$. Likewise, if we define, for each finite set $F \subset I$, the element $\alpha_F \in \ell^2_{\mathbb{K}}(I)$ by

$$\alpha_F(k) = \begin{cases} \alpha(k) & \text{if } k \in F \\ 0 \text{ if } k \in I \smallsetminus F \end{cases}$$

then $T^{-1}\alpha_F = \sum_{k\in F} \alpha(k)\xi_k$. Using the fact that T is an isometry, we have

$$\|\eta - T^{-1}\alpha_F\| = \|T\eta - \alpha_F\| = \|\alpha - \alpha_F\|,$$

and the desired property follows from the well-known properties of $\ell^2_{\mathbb{K}}(I)$. □

Exercise 5. Let $\mathcal{H}$ be a Hilbert space, let $\mathcal{F} = \{\xi_j \,:\, j \in J\}$ be an orthonormal set. Define the closed linear subspace $\mathcal{H}_{\mathcal{F}} = \overline{\operatorname{Span} \mathcal{F}}$. Prove that the orthogonal projection $P_{\mathcal{H}_{\mathcal{F}}}$ is defined by

$$P_{\mathcal{H}_{\mathcal{F}}}\eta = \sum_{j\in J} \big(\,\xi_j \,\big|\, \eta\,\big)\xi_j, \;\; \forall\, \eta \in \mathcal{H}.$$

HINTS: Extend $\mathcal{F}$ to an orthonormal basis $\mathcal{B}$. Let $\mathcal{B}$ be labelled as $\{\xi_i : i \in I\}$ for some set $I \supset J$. First prove that for any $\eta \in \mathcal{H}$, the map $\beta^\eta = T\eta\big|_J$ belongs to $\ell^2_{\mathbb{K}}(J)$. In particular, the sum

$$\eta_{\mathcal{F}} = \sum_{j\in J} (\, \xi_j \,|\, \eta \,)\xi_j$$

is "legitimate" and defines an element in $\mathcal{H}_{\mathcal{F}}$ (use the fact that $\mathcal{F}$ is an orthonormal basis for $\mathcal{H}_{\mathcal{F}}$). Finally, prove that $(\eta - \eta_{\mathcal{F}}) \perp \mathcal{F}$, using Parseval Identity.

EXAMPLE 6.4. Let us analyze the space $L^2([0,1])$. Use the orthonormal basis $\{e_n : n \in \mathbb{Z}\}$ defined by

$$e_n(t) = \exp(2n\pi i t), \ \ \forall\, t \in [0,1],\, n \in \mathbb{Z}.$$

For any $f \in C([0,1])$ we define

$$\hat{f}(n) = \int_0^1 \exp(-2n\pi i t) f(t)\, dt = \big(\, e_n \,\big|\, f \,\big).$$

We then know that

$$f = \sum_{n\in\mathbb{Z}} \hat{f}(n) e_n.$$

One can think the right hand side as a series, but the reader should be aware of the fact that this series is *convergent only in the norm* $\|\cdot\|_2$. One can define for example for any $N \geq 1$, a partial sum $f_N : [0,1] \to \mathbb{C}$ by

$$f_N(t) = \sum_{n=-N}^{N} \hat{f}(n)\exp(2n\pi i t), \ \ t \in [0,1].$$

We will have

$$\lim_{N\to\infty} \|f - f_N\|_2 = 0,$$

but in general there are (many) values of $t \in [0,1]$ for which the limit $\lim_{N\to\infty} f_N(t)$ does not exist. One can consider a formal infinite series

$$\sum_{n=-\infty}^{\infty} \hat{f}(n)\exp(2n\pi i t). \tag{7}$$

Although this series is not convergent (pointwise) for all $t \in [0,1]$, it plays an important role in analysis. The series (7) is called the *complex Fourier series of* f.

Note that Parseval's Identity gives

$$\int_0^1 \overline{f(t)} g(t)\, dt = \sum_{n=-\infty}^{\infty} \overline{\hat{f}(n)} \hat{g}(n).$$

One can construct another orthonormal basis for $L^2([0,1])$, by taking real and imaginary parts of e_n. More explicitly, we define the sequences of functions $(g_n)_{n=0}^{\infty}$ and $(h_n)_{n=1}^{\infty}$ by

$$g_0(t) = 1, \ \ \forall\, t \in [0,1];$$
$$g_n(t) = \sqrt{2}\cos(2n\pi t), \ \ \forall\, t \in [0,1],\, n \geq 1;$$
$$h_n(t) = \sqrt{2}\sin(2n\pi t), \ \ \forall\, t \in [0,1],\, n \geq 1.$$

Then $\mathcal{B}' = \{g_n : n \geq 0\} \cup \{h_n : n \geq 1\}$ is again an orthonormal basis for $L^2([0,1])$. (It is clear that $\mathcal{B}'$ is orthonormal, and $\operatorname{Span} \mathcal{B}' \ni e_n$, $\forall\, n \in \mathbb{Z}$, so $\operatorname{Span} \mathcal{B}'$ is dense in $L^2([0,1])$.) For $f \in C([0,1])$ one can then define its *real Fourier series*

$$\hat{f}(0) + \sum_{n=1}^{\infty} \big[a_n \cos(2n\pi t) + b_n \sin(2n\pi t)\big],$$

where

$$a_n = \sqrt{2} \int_0^1 f(t) \cos(2n\pi t)\, dt \text{ and } b_n = \sqrt{2} \int_0^1 f(t) \sin(2n\pi t)\, dt, \ \ \forall\, n \geq 1.$$

Note that

$$a_n = \frac{\sqrt{2}}{2} [\hat{f}(-n) + \hat{f}(n)] \text{ and } b_n = \frac{\sqrt{2}}{2i} [\hat{f}(-n) - \hat{f}(n)], \ \ \forall\, n \geq 1.$$

The next result discusses the appropriate notion of dimension for Hilbert spaces.

THEOREM 6.4. *Let $\mathcal{H}$ be a Hilbert space. Then any two orthonormal bases of $\mathcal{H}$ have the same cardinality.*

PROOF. Fix two orthonormal bases $\mathcal{B}$ and $\mathcal{B}'$. There are two possible cases.

CASE I: *One of the sets $\mathcal{B}$ or $\mathcal{B}'$ is finite.*

In this case $\mathcal{H}$ is finite dimensional, since the linear span of a finite set is automatically closed. Since both $\mathcal{B}$ and $\mathcal{B}'$ are linearly independent, it follows that *both* $\mathcal{B}$ and $\mathcal{B}'$ are finite, hence their linear spans are both closed. It follows that

$$\operatorname{Span} \mathcal{B} = \operatorname{Span} \mathcal{B}' = \mathcal{H},$$

so $\mathcal{B}$ and $\mathcal{B}'$ are in fact linear bases for $\mathcal{H}$, and then we get

$$\operatorname{Card} \mathcal{B} = \operatorname{Card} \mathcal{B}' = \dim \mathcal{H}.$$

CASE II: *Both $\mathcal{B}$ and $\mathcal{B}'$ are infinite.*

The key step we need in this case is the following.

Claim 1: *There exists a dense subset $\mathcal{Z} \subset \mathcal{H}$, with*

$$\operatorname{Card} \mathcal{Z} = \operatorname{Card} \mathcal{B}'.$$

To prove this fact, we define the set

$$\mathcal{X} = \operatorname{Span}_{\mathbb{Q}} \mathcal{B}'.$$

It is clear that

$$\operatorname{Card} \mathcal{X} = \operatorname{Card} \mathcal{B}'.$$

Notice that $\mathcal{X}$ is dense in $\operatorname{Span}_{\mathbb{R}} \mathcal{B}'$. If we work over $\mathbb{K} = \mathbb{R}$, then we are done. If we work over $\mathbb{K} = \mathbb{C}$, we define

$$\mathcal{Z} = \mathcal{X} + i\mathcal{X},$$

and we will still have

$$\operatorname{Card} \mathcal{Z} = \operatorname{Card} \mathcal{X} = \operatorname{Card} \mathcal{B}'.$$

Now we are done, since clearly $\mathcal{Z}$ is dense in $\operatorname{Span}_{\mathbb{C}} \mathcal{B}'$.

Choose $\mathcal{Z}$ as in Claim 1. For every $\xi \in \mathcal{B}$ we choose a vector $\zeta_\xi \in \mathcal{Z}$, such that

$$\|\xi - \zeta_\xi\| \leq \frac{\sqrt{2} - 1}{2}.$$

Claim 2: *The map $\mathcal{B} \ni \xi \longmapsto \zeta_\xi \in \mathcal{Z}$ is injective.*

Start with two vectors $\xi_1, \xi_2 \in \mathcal{B}$, such that $\xi_1 \neq \xi_2$. In particular, $\xi_1 \perp \xi_2$, so we also have $\xi_1 \perp (-\xi_2)$, and using the Pythogorean Theorem we get

$$\|\xi_1 - \xi_2\|^2 = \|\xi_2\|^2 + \| - \xi_2\|^2 = 2,$$

which gives

$$\|\xi_1 - \xi_2\| = \sqrt{2}.$$

Using the triangle inequality, we now have

$$\sqrt{2} = \|\xi_1 - \xi_2\| \leq \|\xi_1 - \zeta_{\xi_1}\| + \|\xi_2 - \zeta_{\xi_2}\| + \|\zeta_{\xi_1} - \zeta_{\xi_2}\| \leq \sqrt{2} - 1 + \|\zeta_{\xi_1} - \zeta_{\xi_2}\|.$$

This gives

$$\|\zeta_{\xi_1} - \zeta_{\xi_2}\| \geq 1,$$

which forces $\zeta_{\xi_1} \neq \zeta_{\xi_2}$.

Using Claim 2, we have constructed an injective map $\mathcal{B} \to \mathcal{Z}$. In particular, using Claim 1 and the cardinal arithmetic rules, we get

$$\text{Card}\,\mathcal{B} \leq \text{Card}\,\mathcal{Z} = \text{Card}\,\mathcal{B}'.$$

By symmetry we also have

$$\text{Card}\,\mathcal{B}' \leq \text{Card}\,\mathcal{B},$$

and then using the Cantor-Bernstein Theorem, we finally get

$$\text{Card}\,\mathcal{B}\text{Card}\,\mathcal{B}'.$$

□

Corollary 6.7 (of the proof). *A Hilbert space is separable, in the norm topology, if and only of it has an orthonormal basis which is at most countable.*

Proof. Use Claims 1 and 2 from the proof of the Theorem. □

Definition. Let $\mathcal{H}$ be a Hilbert space, and let $\mathcal{B}$ be an orthonormal basis for $\mathcal{H}$. By the above theorem, the cardinal number $\text{Card}\,\mathcal{B}$ does not depend on the choice of $\mathcal{B}$. This number is called the *hilbertian* (or *orthogonal*) *dimension of* $\mathcal{H}$, and is denoted by $\text{H-dim}\,\mathcal{H}$.

Corollary 6.8. *For two Hilbert spaces $\mathcal{H}$ and $\mathcal{H}'$, the following are equivalent:*

(i) $\text{H-dim}\,\mathcal{H} = \text{H-dim}\,\mathcal{H}'$;

(ii) *There exists an isometric linear isomorphism $U : \mathcal{H} \to \mathcal{H}'$.*

Proof. $(i) \Rightarrow (ii)$. Choose a set I with $\text{H-dim}\,\mathcal{H} = \text{H-dim}\,\mathcal{H}' = \text{Card}\,I$. Apply Theorem ?? to produce isometric linear isomorphisms $T : \mathcal{H} \to \ell^2_{\mathbb{K}}(I)$ and $T' : \mathcal{H}' \to \ell^2_{\mathbb{K}}(I)$. Then define $U = T'^{-1} \circ T$.

$(ii) \Rightarrow (i)$. Assume one has an isometric linear isomorphism $U : \mathcal{H} \to \mathcal{H}'$. Choose an orthonormal basis $\mathcal{B}$ for $\mathcal{H}$. Then $U(\mathcal{B})$ is clearly and orthonormal basis for $\mathcal{H}'$, and since $U : \mathcal{B} \to U(\mathcal{B})$ is bijective, we get

$$\text{H-dim}\,\mathcal{H} = \text{Card}\,\mathcal{B} = \text{Card}\,U(\mathcal{B}) = \text{H-dim}\,\mathcal{H}'.$$

□