3. Hilbert spaces

In this section we examine a special type of Banach spaces. We start with some algebraic preliminaries.

**Definition.** Let $K$ be either $\mathbb{R}$ or $\mathbb{C}$, and let $X$ and $Y$ be vector spaces over $K$. A map $\phi : X \times Y \to K$ is said to be $K$-sesquilinear, if

- for every $x \in X$, then map $\phi_x : Y \ni y \mapsto \phi(x, y) \in K$ is linear;
- for every $y \in Y$, then map $\phi^y : X \ni x \mapsto \phi(x, y) \in K$ is conjugate linear, i.e. the map $X \ni x \mapsto \phi^y(x) \in K$ is linear.

In the case $K = \mathbb{R}$ the above properties are equivalent to the fact that $\phi$ is bilinear.

**Remark 3.1.** Let $X$ be a vector space over $\mathbb{C}$, and let $\phi : X \times X \to \mathbb{C}$ be a $\mathbb{C}$-sesquilinear map. Then $\phi$ is completely determined by the map

$$Q_\phi : X \ni x \mapsto \phi(x, x) \in \mathbb{C}.$$ 

This can be seen by computing, for $k \in \{0, 1, 2, 3\}$ the quantity

$$Q_\phi(x + i^k y) = \phi(x + i^k y, x + i^k y) = \phi(x, x) + \phi(x, i^k y) + \phi(i^k y, x) + \phi(i^k y, i^k y) =$$

$$= \phi(x, x) + i^k \phi(x, y) + i^{-k} \phi(y, x) + \phi(y, y),$$

which then gives

$$\phi(x, y) = \frac{1}{4} \left\{ \sum_{k=0}^{3} i^{-k} Q_\phi(x + i^k y) \right\}, \quad \forall x, y \in X. \quad (1)$$

The map $Q_\phi : X \to \mathbb{C}$ is called the **quadratic form determined by** $\phi$. The identity (1) is referred to as the **Polarization Identity**. Remark that the Polarization Identity can also gives

$$\phi(y, x) = \frac{1}{4} \left\{ \sum_{k=0}^{3} i^k Q_\phi(x + i^k y) \right\}, \quad \forall x, y \in X.$$ 

In particular, we see that the following conditions are equivalent

- (i) $Q_\phi$ is real-valued, i.e. $Q_\phi(x) \in \mathbb{R}, \forall x \in X$;
- (ii) $\phi$ is sesqui-symmetric, i.e.

$$\phi(y, x) = \overline{\phi(x, y)}, \quad \forall x, y \in X.$$ 

**Definition.** Let $K$ be one of the fields $\mathbb{R}$ or $\mathbb{C}$, and let $X$ be a vector space over $K$. A $K$-sesquilinear map $\phi : X \times X \to K$ is said to be **positive definite**, if

- $\phi$ is sesqui-symmetric;
- $\phi(x, x) \geq 0, \forall x \in X$.

As observed before, in the complex case, the second condition actually forces the first one.

With this terminology, we have the following useful technical result.
Proposition 3.1 (Cauchy-Bunyakowski-Schwartz Inequality). Let \( \mathbb{K} \) be a \( \mathbb{K} \)-vector space, and let \( \phi : \mathcal{X} \times \mathcal{X} \to \mathbb{K} \) be a positive definite \( \mathbb{K} \)-sesquilinear map. Then

\[
|\phi(x,y)|^2 \leq \phi(x,x) \cdot \phi(y,y), \quad \forall x, y \in \mathcal{X}.
\]

Moreover, if equality holds then either \( x = y = 0 \), or there exists some \( \alpha \in \mathbb{K} \) with \( \phi(x + \alpha y, x + \alpha y) = 0 \).

Proof. Fix \( x, y \in \mathcal{X} \), and choose a number \( \lambda \in \mathbb{K} \), with \( |\lambda| = 1 \), such that

\[
|\phi(x,y)| = \lambda \phi(x,y) = \phi(x,\lambda y).
\]

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

\[
F(\zeta) = \phi(\zeta y + x, \zeta y + x), \quad \forall \zeta \in \mathbb{K}.
\]

A simple computation gives

\[
F(\zeta) = \zeta \overline{\lambda} \overline{\lambda} \phi(y, y) + \zeta \lambda \phi(x, y) \overline{\lambda} \phi(y, x) + \phi(x, x) = \\
= |\zeta|^2 \phi(y, y) + \zeta \lambda \phi(x, y) + \zeta \overline{\lambda} \phi(y, x) + \phi(x, x) = \\
= |\zeta|^2 \phi(y, y) + \zeta |\phi(x, y)| + \zeta \overline{|\phi(x, y)|} + \phi(x, x), \quad \forall \zeta \in \mathbb{R}.
\]

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

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

where \( a = \phi(y, y) \geq 0, b = 2|\phi(x, y)|, c = \phi(x, x) \). Notice that we have

\[
F(t) \geq 0, \quad \forall t \in \mathbb{R}.
\]

This forces either \( a = b = 0 \), in which case the inequality (2) is trivial, or \( a > 0 \), and \( b^2 - 4ac \leq 0 \), which reads

\[
4|\phi(x, y)|^2 - 4\phi(x, x)\phi(y, y) \leq 0,
\]

and again (2) follows.

Let us examine now when equality holds in (2). We assume that either \( x \neq 0 \) or \( y \neq 0 \). In fact we can also assume that \( \phi(x, x) > 0 \) and \( \phi(y, y) > 0 \), so the fact that we have equality in (2) is equivalent to \( b^2 - 4ac = 0 \), which in terms of quadratic equations says that the equation

\[
F(t) = at^2 + bt + c = 0
\]

has a (unique) solution \( t_0 \). If we take \( \alpha = t_0 \lambda \), then the equality \( F(t_0) = 0 \) is precisely \( \phi(\alpha y + x, \alpha y + x) = 0 \).

\[\square\]

Definition. An inner product on \( \mathcal{X} \) is a positive definite \( \mathbb{K} \)-sesquilinear map

\[
\mathcal{X} \times \mathcal{X} \ni (\xi, \eta) \mapsto (\xi | \eta) \in \mathbb{K},
\]

which is non-degenerate, in the sense that

- if \( \xi \in \mathcal{X} \) satisfies \((\xi | \xi) = 0\), then \( \xi = 0 \).

Proposition 3.2. Let \( (\cdot | \cdot) \) be an inner product on the \( \mathbb{K} \)-vector space \( \mathcal{X} \).

(i) If we define, for every \( \xi \in \mathcal{X} \), the quantity \( \|\xi\| = \sqrt{(\xi | \xi)} \), then the map

\[
\mathcal{X} \ni \xi \mapsto \|\xi\| \in [0, \infty)
\]

defines a norm on \( \mathcal{X} \).
(ii) For every $\xi, \eta \in X$, one has the Cauchy-Bunyakowski-Schwartz inequality:

$$
|\langle \xi \mid \eta \rangle| \leq \|\xi\| \cdot \|\eta\|. \tag{3}
$$

Moreover, if equality holds then $\xi$ and $\eta$ are proportional, in the sense that either $\xi = 0$, or $\eta = 0$, or $\xi = \lambda \eta$, for some $\lambda \in \mathbb{K}$.

**Proof.** First of all, remark that by the non-degeneracy of the inner product, one has the implication

$$
\|\xi\| = 0 \Rightarrow \xi = 0. \tag{4}
$$

The inequality (3) is immediate from Proposition 3.1. Moreover, if we have equality, then by Proposition 3.1, we also know that either $\xi = 0$, or $\langle \xi + \alpha \eta \mid \xi + \alpha \eta \rangle = 0$, for some $\alpha \in \mathbb{K}$. This is equivalent to $\|\xi + \alpha \eta\|^2 = 0$, so by (4), we must have $\xi = -\alpha \eta$.

Finally, in order to check that $\|\cdot\|$ is a norm, we must prove

(a) $\|\lambda \xi\| = |\lambda| \cdot \|\xi\|$, $\forall \lambda \in \mathbb{K}$, $\xi \in X$;

(b) $\|\xi + \eta\| \leq \|\xi\| + \|\eta\|$, $\forall \xi, \eta \in X$.

Property (a) is trivial, by sesquilinearity:

$$
\|\lambda \xi\|^2 = (\lambda \xi \mid \lambda \xi) = \overline{\lambda} \lambda (\xi \mid \xi) = |\lambda|^2 \cdot \|\xi\|^2.
$$

Finally, for $\xi, \eta \in X$, we have

$$
\|\xi + \eta\|^2 = (\xi + \eta \mid \xi + \eta) = (\xi \mid \xi) + (\eta \mid \eta) + (\xi \mid \eta) + (\eta \mid \xi) = \|\xi\|^2 + \|\eta\|^2 + (\xi \mid \eta) + (\eta \mid \xi) = \|\xi\|^2 + \|\eta\|^2 + 2\text{Re}(\xi \mid \eta).
$$

We now use the C-B-S inequality (3), which allows us to continue the above estimate to get

$$
\|\xi + \eta\|^2 = \|\xi\|^2 + \|\eta\|^2 + 2\text{Re}(\xi \mid \eta) \leq \|\xi\|^2 + \|\eta\|^2 + 2(\|\xi\| \cdot \|\eta\|) \leq \|\xi\|^2 + \|\eta\|^2 + 2\|\xi\| \cdot \|\eta\| = (\|\xi\| + \|\eta\|)^2,
$$

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

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

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

(i) [Polarization Identities]

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

$$
(\xi \mid \eta) = \frac{1}{4} (\|\xi + \eta\|^2 - \|\xi - \eta\|^2), \quad \forall \xi, \eta \in X.
$$

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

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

**Proof.** (i). This is obvious, by the computations from the proof of Proposition 3.2, which give

$$
\|\xi \pm \eta\|^2 = \|\xi\|^2 + \|\eta\|^2 \pm 2\text{Re}(\xi \mid \eta).
$$
(ii).(a). In the real case, the above identity gives
\[ \|\xi \pm \eta\|^2 = \|\xi\|^2 + \|\eta\|^2 \pm 2(\xi \mid \eta), \]
so we immediately get
\[ \|\xi + \eta\|^2 - \|\xi - \eta\|^2 = 4(\xi \mid \eta). \]

(b). This case is a particular instance of the Polarization Identity discussed in Remark 3.1.

Corollary 3.1. Let \(X\) be a \(K\)-vector space equipped with an inner product \((\cdot \mid \cdot)\). Then the map
\[ X \times X \ni (\xi, \eta) \mapsto (\xi \mid \eta) \in K \]
is continuous, with respect to the product topologies.

Proof. Immediate from the polarization identities.

Corollary 3.2. Let \(X\) and \(Y\) be two \(K\)-vector spaces equipped with inner products \((\cdot \mid \cdot)_X\) and \((\cdot \mid \cdot)_Y\). If \(T : X \to K\) is an isometric linear map, then
\[ (T\xi \mid T\eta)_Y = (\xi \mid \eta)_X, \quad \forall \xi, \eta \in X. \]

Proof. Immediate from the polarization identities.

Exercise 2. Let \(X\) be a normed \(K\)-vector space. Assume the norm satisfies the Parallelogram Law. Prove that there exists an inner product \((\cdot \mid \cdot)\) on \(X\), such that
\[ \|\xi\| = \sqrt{(\xi \mid \xi)}, \quad \forall \xi \in X. \]

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

Proposition 3.3. Let \(X\) be a \(K\)-vector space, equipped with an inner product \((\cdot \mid \cdot)_X\). Let \(Z\) be the completion of \(X\) with respect to the norm defined by the inner product. Then \(Z\) carries a unique inner product \((\cdot \mid \cdot)_Z\), so that the norm on \(Z\) is defined by \((\cdot \mid \cdot)_Z\). Moreover, this inner product extends \((\cdot \mid \cdot)_X\), in the sense that
\[ (\langle \xi \rangle \mid \langle \eta \rangle)_Z = (\xi \mid \eta)_X, \quad \forall \xi, \eta \in X. \]

Proof. It is obvious that the norm on \(Z\) satisfies the Parallelogram Law. We then apply Exercise 2.

Definitions. Let \(K\) be one of the fields \(\mathbb{R}\) or \(\mathbb{C}\). A Hilbert space over \(K\) is a \(K\)-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 3.1. For \(J\) a non-empty set, the space \(\ell^2_K(J)\) is a Hilbert space. We know that this is a Banach space. The inner product defining the norm is
\[ (\alpha \mid \beta) = \sum_{j \in J} \alpha(j)\beta(j), \quad \forall \alpha, \beta \in \ell^2_K(J). \]
§3. Hilbert spaces

(The fact that function \( \pi \beta : J \to \mathbb{K} \) belongs to \( \ell^1_\mathbb{K}(J) \) is discussed in Section 1, Exercise 5.)

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

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

\[ \xi \perp \eta, \quad \forall \eta \in M. \]

Finally, two subsets \( M, N \subset X \) are said to be orthogonal, and we write \( M \perp N \), if

\[ \xi \perp \eta, \quad \forall \xi \in M, \eta \in N. \]

**Notation.** Let \( X \) be a vector space equipped with an inner product. For a subset \( M \subset X \), we define the set

\[ M^\perp = \{ \xi \in X : \xi \perp M \}. \]

**Remarks 3.2.** Let \( X \) be a \( \mathbb{K} \)-vector space equipped with an inner product.

A. The relation \( \perp \) is symmetric.

B. If \( \xi, \eta \in X \) satisfy \( \xi \perp \eta \), then one has the Pythagorean Theorem:

\[ \| \xi + \eta \|^2 = \| \xi \|^2 + \| \eta \|^2. \]

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

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

D. For sets \( M \subset N \subset X \), one has

\[ M^\perp \supset N^\perp. \]

E. For any set \( M \subset X \), one has

\[ M^\perp = (\text{Span } M)^\perp, \]

where \( \text{Span } M \) denotes the norm closure of the linear span of \( M \). The inclusion

\[ M^\perp \supset (\text{Span } M)^\perp \]

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

\[ \text{Span } M \subset \{ \xi \}^\perp, \]

i.e. \( \xi \in (\text{Span } M)^\perp \).

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

**Proposition 3.4.** 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
\[ \|\eta_m - \eta_n\|^2 \leq 2\|\xi - \eta_m\|^2 + 2\|\xi - \eta_n\|^2 - 4\delta^2, \quad \forall m, n \geq 1. \]

Use the Parallelogram Law
\[ 2\|\xi - \eta_m\|^2 + 2\|\xi - \eta_n\|^2 = \|2\xi - \eta_m - \eta_n\|^2 + \|\eta_m - \eta_n\|^2. \]  

We notice that, since \( \frac{1}{2}(\eta_m + \eta_n) \in \mathcal{C} \), we have
\[ \|\xi - \frac{1}{2}(\eta_m + \eta_n)\| \geq \delta, \]
so we get
\[ \|2\xi - \eta_m - \eta_n\|^2 = 4\|\xi - \frac{1}{2}(\eta_m + \eta_n)\|^2 \geq 4\delta^2, \]
so if we go back to (5) we get
\[ 2\|\xi - \eta_m\|^2 + 2\|\xi - \eta_n\|^2 \geq 4\delta^2 + \|\eta_m - \eta_n\|^2, \]
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 \( \xi' \). 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
\[ 4\delta^2 > \|2\xi - \xi' - \xi''\|^2 = 4\|\xi - \frac{1}{2}(\xi' + \xi'')\|^2, \]
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 \( \delta \).

**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 on** \( \mathcal{X} \).

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

**Proposition 3.5.** 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
\[ \zeta = P_{\mathcal{X}}\xi \iff (\xi - \zeta) \perp \mathcal{X}. \]

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

(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). “\( \Rightarrow \).” Assume \( \zeta = P_\mathcal{X} \xi \). Fix an arbitrary vector \( \eta \in \mathcal{X} \setminus \{0\} \), and choose a number \( \lambda \in \mathbb{K} \), with \( |\lambda| = 1 \), such that \[
\lambda \langle \xi - \zeta \mid \eta \rangle = |\langle \xi - \zeta \mid \eta \rangle|.
\]
In particular, we have
\[
|\langle \xi - \zeta \mid \eta \rangle| = \text{Re} \langle \xi - \zeta \mid \lambda \eta \rangle.
\]
Define the map \( F : \mathbb{R} \rightarrow \mathbb{R} \) by
\[
F(t) = \|\xi - \zeta - t\lambda \eta\|^2 - \|\xi - \zeta\|^2.
\]
By the definition of \( \zeta = P_\mathcal{X} \xi \), we have
\[
(6) \quad F(t) > 0, \quad \forall \ t \in \mathbb{R} \setminus \{0\}.
\]
Notice that \( F(t) = at^2 + bt \), \( \forall \ t \in \mathbb{R} \), where \( a = \langle \lambda \eta \mid \lambda \eta \rangle = \|\eta\|^2 \), and \( b = 2\text{Re} \langle \xi - \zeta \mid \lambda \eta \rangle = 2|\langle \xi - \zeta \mid \eta \rangle| \). Of course, the property
\[
at^2 + bt > 0, \quad \forall \ t \in \mathbb{R} \setminus \{0\}
\]
forces \( b = 0 \), so we indeed get \( \langle \xi - \zeta \mid \eta \rangle = 0 \).

“\( \Leftarrow \).” Assume \( \langle \xi - \zeta \mid \eta \rangle \perp \mathcal{X} \). For any \( \eta \in \mathcal{X} \), we have \( \langle \xi - \zeta \mid (\zeta - \eta) \rangle \), so using the Pythagorean Theorem, we get
\[
\|\xi - \eta\|^2 = \|\xi - \zeta\|^2 + \|\zeta - \eta\|^2,
\]
which forces
\[
\|\xi - \eta\| \geq |\xi - \zeta|, \quad \forall \eta \in \mathcal{X}.
\]
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
\[
(\lambda \zeta_1 + \xi_2 - \zeta \mid \eta) = (\lambda \zeta_1 \mid \eta) + (\zeta_2 \mid \eta) = \lambda (\zeta_1 \mid \eta) + (\xi_2 - \zeta \mid \eta) = 0.
\]
By (i) we have \( \langle \xi_1 - \zeta_1 \mid \perp \mathcal{X} \) and \( \langle \xi_2 - \zeta_2 \mid \perp \mathcal{X} \), so the above computation proves that
\[
(\lambda \zeta_1 + \xi_2 - \zeta) \perp \mathcal{X},
\]
so using (i) we get
\[
P_\mathcal{X}(\lambda \zeta_1 + \xi_2) = \zeta = \lambda \zeta_1 + \zeta_2 = \lambda P_\mathcal{X} \xi_1 + P_\mathcal{X} \xi_2,
\]
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 \( \langle \xi - P_\mathcal{X} \xi \mid \perp P_\mathcal{X} \xi \). By the Pythagorean Theorem we then have
\[
\|\xi\|^2 = \|\xi - P_\mathcal{X} \xi\|^2 + \|P_\mathcal{X} \xi\|^2 = \|\xi - P_\mathcal{X} \xi\|^2 + \|P_\mathcal{X} \xi\|^2 \geq \|P_\mathcal{X} \xi\|^2.
\]
In other words, we have
\[
\|P_\mathcal{X} \xi\| \leq \|\xi\|, \quad \forall \xi \in \mathcal{H},
\]
so \( P_X \) is indeed continuous, and we have \( \|P_X\| \leq 1 \). Using (ii) we immediately get that, when \( X \neq \{0\} \), we have \( \|P_X\| = 1 \).

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

**Corollary 3.3.** If \( \mathcal{H} \) is a Hilbert space, and \( X \subset \mathcal{H} \) is a closed linear subspace, then

\[ X + X^\perp = \mathcal{H} \text{ and } X \cap X^\perp. \]

In other words the map

\[ X \times X^\perp \ni (\eta, \zeta) \mapsto \eta + \zeta \in \mathcal{H} \]

is a linear isomorphism.

**Proof.** If \( \xi \in \mathcal{H} \) then \( P_X \xi \in X \), and \( \xi - P_X \xi \in X^\perp \), and then the equality

\[ \xi = P_X \xi + (\xi - P_X \xi) \]

proves that \( \xi \in X + X^\perp \). The equality \( X \cap X^\perp = \{0\} \) is trivial, since for \( \zeta \in X \cap X^\perp \), we must have \( \zeta \perp \zeta \), which forces \( \zeta = 0 \). \( \square \)

**Exercise 3.** Let \( \mathcal{H} \) be a Hilbert space.

(i) Prove that, for any closed subspace \( X \subset \mathcal{H} \), one has the equality

\[ P_{X^\perp} = I - P_X. \]

(ii) Prove that two closed subspaces \( X, Y \subset \mathcal{H} \), the following are equivalent:

- \( X \perp Y \);
- \( P_X P_Y = 0 \);
- \( P_Y P_X = 0 \).

(iii) Prove that two closed subspaces \( X, Y \subset \mathcal{H} \), the following are equivalent:

- \( X \subset Y \);
- \( P_X P_Y = P_X \);
- \( P_Y P_X = P_X \).

(iv) Let \( X, Y \subset \mathcal{H} \) are closed subspaces, such that \( X \perp Y \), then

- \( X + Y \) is c closed linear subspace of \( \mathcal{H} \);
- \( P_{X+Y} = P_X + P_Y \).

**Corollary 3.4.** Let \( \mathcal{H} \) be a Hilbert space, and let \( X \subset \mathcal{H} \) be a linear (not necessarily closed) subspace. Then on has the equality

\[ X = (X^\perp)^\perp. \]

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

\[ P_Z = I - P_{X^\perp} = I - P_{\overline{X^\perp}} = I - (I - P_X) = P_X, \]

which forces

\[ Z = \text{Ran} \ P_Z = \text{Ran} \ P_X = \overline{X}. \] \( \square \)
Theorem 3.1 (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

$$\phi(\eta) = (\xi \mid \eta), \quad \forall \eta \in \mathcal{H}.$$ 

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 $X = \text{Ker} \phi$. Notice that $X$ is closed. Using the linear isomorphism (7) we see that the composition

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

is a linear isomorphism. Since $\mathcal{H}/X = \mathcal{H}/\text{Ker} \phi \cong \text{Ran} \phi = \mathbb{K}$, it follows that $\dim(X^\perp) = 1$. In other words, there exists $\xi_0 \in X^\perp$, $\xi_0 \neq 0$, such that

$$X^\perp = \mathbb{K} \xi_0.$$

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

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

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

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

On the other hand, we have

$$\langle \xi_0 \mid \eta \rangle = \langle \xi_0 \mid \lambda \xi_0 \rangle + \langle \xi_0 \mid \zeta \rangle = \lambda \|\xi_0\|^2,$$

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

$$\langle \xi \mid \eta \rangle = \langle \frac{\phi(\xi_0)}{\|\xi_0\|^2} \xi_0^{-2} \xi_0 \mid \eta \rangle = \phi(\xi_0)\|\xi_0\|^{-2} \langle \xi_0 \mid \eta \rangle = \lambda \phi(\xi_0) = \phi(\eta).$$

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

$$\phi(\eta) = (\xi' \mid \eta), \quad \forall \eta \in \mathcal{H}.$$

In particular, we have

$$\|\xi - \xi'\|^2 = \langle \xi - \xi' \mid \xi - \xi' \rangle = \langle \xi \mid \xi - \xi' \rangle - \langle \xi' \mid \xi - \xi' \rangle = \phi(\xi - \xi') - \phi(\xi - \xi') = 0,$$

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

$$|\phi(\eta)| = |(\xi \mid \eta)| \leq \|\xi\| \cdot \|\eta\|, \quad \forall \eta \in \mathcal{H},$$

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

$$\phi(\zeta) = (\xi \mid \|\xi\|^{-1} \xi) = \|\xi\|,$$

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 \( X \) be a \( \mathbb{K} \)-vector space, equipped with an inner product. A set \( F \subset X \) is said to be **orthogonal**, if \( 0 \not\in F \), and

\[
\xi \perp \eta, \quad \forall \xi, \eta \in F, \text{ with } \xi \neq \eta.
\]

A set \( F \subset X \) is said to be **orthonormal**, if it is orthogonal, but it also satisfies:

\[
\|\xi\| = 1, \quad \forall \xi \in F.
\]

Remark that, if one starts with an orthogonal set \( F \subset X \), then the set

\[
F^{(1)} = \{ \|\xi\|^{-1} \xi : \xi \in F \}
\]

is orthonormal.

**Proposition 3.6.** Let \( X \) be a \( \mathbb{K} \)-vector space equipped with an inner product. Any orthogonal set \( F \subset X \) is linearly independent.

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

\[
\lambda_1 \xi_1 + \cdots + \lambda_n \xi_n = 0,
\]

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

\[
\lambda_k \|\xi_k\|^2 = (\xi_k | \lambda - 1\xi_1 + \cdots + \lambda_n \xi_n) = 0,
\]

and since \( \xi_k \neq 0 \), we get \( \lambda_k = 0 \). \( \Box \)

**Lemma 3.2.** Let \( X \) be a \( \mathbb{K} \)-vector space equipped with an inner product, and let \( F \subset X \) be an orthogonal set. Then there exists a maximal (with respect to inclusion) orthogonal set \( G \subset X \) with \( F \subset G \).

**Proof.** Consider the sets

\[
\mathcal{A} = \{ G : G \text{ orthogonal subset of } X \},
\]

\[
\mathcal{A}_F = \{ G \in \mathcal{A} : G \supset F \},
\]

ordered with the inclusion. We are going to apply Zorn’s Lemma to \( \mathcal{A}_F \). Let \( F \subset \mathcal{A}_F \) be a subcollection, which is totally ordered, i.e. for any \( G_1, G_2 \in F \) one has \( G_1 \subset G_2 \) or \( G_1 \supset G_2 \). Define the set

\[
M = \bigcup_{G \in F} G.
\]

Since \( G \subset X \setminus \{0\} \), for all \( G \in F \), it is clear that \( M \subset X \setminus \{0\} \). If \( \xi_1, \xi_2 \in M \) are vectors with \( \xi_1 \neq \xi_2 \), then we can find \( G_1, G_2 \in F \) with \( \xi_1 \in G_1 \) and \( \xi_2 \in G_2 \). Using the fact that \( F \) is totally ordered, it follows that there is \( k \in \{1, 2\} \) such that \( \xi_1, \xi_2 \in G_k \), so we indeed get \( \xi_1 \perp \xi_2 \). It is now clear that \( M \in \mathcal{A}_F \), and \( M \supset G \), for all \( G \in F \). In other words, we have shown that **every totally ordered subset of** \( \mathcal{A}_F \) **has an upper bound, in** \( \mathcal{A}_F \). By Zorn’s Lemma, \( \mathcal{A}_F \) has a maximal element. Finally, it is clear that any maximal element for \( \mathcal{A}_F \) is also a maximal element in \( \mathcal{A} \). \( \Box \)

**Remark 3.3.** Using the notations from the proof above, given an orthonormal set \( M \subset X \), the following are equivalent:

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

(ii) \( M \) is maximal in

\[
\mathcal{A}^{(1)} = \{ G : G \text{ orthonormal subset of } X \}.
\]
The implication \((i) \Rightarrow (ii)\) is trivial. Conversely, if \(M\) is maximal in \(\mathcal{A}^{(1)}\), we use the Lemma to find a maximal \(N \in \mathcal{A}\) with \(N \supseteq M\). But then \(N^{(1)}\) is orthonormal, and \(N^{(1)} \supseteq M\), which by the maximality of \(M\) in \(\mathcal{A}^{(1)}\) will force \(N^{(1)} = M\). Since \(N\) is linearly independent, the relations

\[ N^{(1)} = M \subset N, \]

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

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

1. \(\mathcal{F}\) is maximal among all orthogonal subsets of \(\mathcal{H}\);
2. \(\text{Span}\ \mathcal{F}\) is dense in \(\mathcal{H}\) in the norm topology.

**Proof.** \((i) \Rightarrow (ii)\). Assume \(\mathcal{F}\) is maximal. We are going to show that \(\text{Span}\ \mathcal{F}\) 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 \(\mathcal{X} = \mathcal{X}^\perp\), we see that, the strict inclusion \(\mathcal{X} \subsetneq \mathcal{H}\) forces \(\mathcal{X} \neq \{0\}\). 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 contradiction. If \(\mathcal{F}\) is not maximal, then there exists \(\xi \in \mathcal{H} \setminus \mathcal{F}\), 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 orthogonal. (Recall that all elements of an orthogonal 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 3.2.** Let \(J\) be a non-empty set. Consider the Hilbert space \(\ell^2_\mathbb{K}(I)\). For every \(j \in J\), let \(\delta_j : J \to \mathbb{K}\) be the function defined by

\[ \delta_j(k) = \begin{cases} 1 & \text{if } k = j \\ 0 & \text{if } k \neq j \end{cases} \]

If we consider the set \(\mathcal{B} = \{\delta_j : j \in J\}\), then

\[ \text{Span}\ \mathcal{B} = \text{fin}_\mathbb{K}(J), \]

which is dense in \(\ell^2_\mathbb{K}(J)\). The above result then says that \(\mathcal{B}\) is an orthonormal basis for \(\ell^2_\mathbb{K}(J)\).
Lemma 3.3. Let $\mathcal{B}$ be an orthonormal basis for the Hilbert space $\mathcal{H}$, and let $\mathcal{F} \subset \mathcal{B}$ be an arbitrary non-empty subset.

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

Proof. (i). This is clear, since $\mathcal{F}$ is orthonormal and has dense span.
(ii). Denote for simplicity $\operatorname{Span}\mathcal{F} = X$ and $\operatorname{Span}(\mathcal{B} \setminus \mathcal{F}) = Y$. Since $\xi \perp \eta$, $\forall \xi \in \mathcal{F}, \eta \in \mathcal{B} \setminus \mathcal{F}$, it is pretty obvious that $X \perp Y$. Since $X + Y$ clearly contains $\operatorname{Span} \mathcal{B}$, it follows that $X + Y$ is dense in $\mathcal{H}$. We know however that $X + Y$ is closed, so we have in fact the equality $X + Y = \mathcal{H}$.

This will then give $I = P_X = P_X + P_Y$, so we get $P_Y = I - P_X = P_X^\perp$, so

$X^\perp = \operatorname{Ran} P_X^\perp = \operatorname{Ran} P_Y = Y$.

\[\square\]

Theorem 3.3. Let $\mathcal{H}$ be a Hilbert space, and let $\mathcal{B}$ be an orthonormal basis for $\mathcal{H}$, labelled\(^1\) as $\mathcal{B} = \{\xi_j : j \in J\}$. For every vector $\eta \in \mathcal{H}$, let $\alpha^\eta : J \rightarrow \mathbb{K}$ be the map defined by $\hat{\eta}(j) = (\xi_j | \eta)$, $\forall j \in J$.

(i) For every $\eta \in \mathcal{H}$, the map $\hat{\eta}$ belongs to $\ell^2_{\mathbb{K}}(J)$.
(ii) The map $T : \mathcal{H} \ni \eta \mapsto \hat{\eta} \in \ell^2_{\mathbb{K}}(J)$ is an isometric linear isomorphism.

Proof. (i). Fix for the moment $\eta \in \mathcal{H}$. One way to prove that $\hat{\eta}$ belongs to $\ell^2_{\mathbb{K}}(J)$ is by showing that

$$\sup \{ \sum_{j \in F} |\hat{\eta}(j)|^2 : F \in \mathcal{P}_{\text{fin}}(J) \} < \infty.$$ 

For any non-empty finite subset $F \subset J$, we define the subspace $\mathcal{H}_F = \operatorname{Span}\{\xi_j : j \in F\}$, and define the vector

$$\eta_F \equiv \sum_{j \in F} (\xi_j | \eta) \cdot \xi_j.$$ 

Claim: For every finite set $F \subset J$, one has the equality $\eta_F = P_{\mathcal{H}_F} \eta$.

\(^1\)This notation implicitly assumes that $\xi_j \neq \xi_k$, for all $j, k \in J$ with $j \neq k$. 
§3. Hilbert spaces

It suffices to prove that
\[(\eta - \eta_F) \perp H_F.\]
But this is obvious, since if we start with some \(k \in F\), then using the fact that \((\xi_k | \eta) = 0\), for all \(j \in F \setminus \{k\}\), together with the equality \(|\xi_k| = 1\), we get
\[(\xi_k | \eta - \eta_F) = (\xi_k | \eta) - \sum_{j \in F} (\xi_j | \eta) \cdot (\xi_k | \xi_j) = (\xi_k | \eta) - (\xi_k | \eta) \cdot (\xi_k | \xi_k) = 0.\]

Having proven the Claim, let us observe that, since the terms in the sum that defines \(\eta_F\) are all orthogonal, we get
\[\|\eta_F\|^2 = \sum_{j \in F} |(\xi_j | \eta) \cdot \xi_j|^2 = \sum_{j \in F} |(\xi_j | \eta)|^2 \cdot \|\xi_j\|^2 = \sum_{j \in F} |\hat{\eta}(j)|^2.\]
Combining this computation with the Claim, we now have
\[\sum_{j \in F} |\hat{\eta}(j)|^2 = \|\eta_F\|^2 = \|P_{H_F} \eta\|^2 \leq \|\eta\|^2,\]
which proves that
\[\sup \{ \sum_{j \in F} |\hat{\eta}(i)|^2 : F \in \mathcal{P}_i \} < \|\eta\|^2.\]

(ii). The linearity of \(T\) is obvious. The above inequality actually proves that
\[\|T \eta\| \leq \|\eta\|, \ \forall \eta \in H.\]
We now prove that in fact \(T\) is isometric. Since \(T\) is linear and continuous, it suffices to prove that \(T|_{\text{Span } B}\) is isometric. Start with some vector \(\eta \in \text{Span } B\), which means that there exists some finite set \(F \subset J\), and scalars \((\lambda_k)_{k \in F} \subset H\), such that \(\eta = \sum_{k \in F} \lambda_k \xi_k\). Remark that
\[(\xi_j | \eta) = \sum_{k \in F} \lambda_k (\xi_j | \xi_k) = \left\{ \begin{array}{ll} \lambda_k & \text{if } k \in F \\ 0 & \text{if } k \in F \end{array} \right.\]
so the element \(\hat{\eta} = T \eta \in \ell_2^J(J)\) is defined by
\[\hat{\eta}(k) = \left\{ \begin{array}{ll} \lambda_k & \text{if } k \in F \\ 0 & \text{if } k \in F \end{array} \right.\]
This gives
\[\|\eta\|^2 = \sum_{j,k \in F} \lambda_j \bar{\lambda}_k (\xi_j | \xi_k) = \sum_{k \in F} |\lambda_k|^2 = \sum_{k \in F} |\hat{\eta}(k)|^2 = \|\hat{\eta}\|^2,\]
so we indeed get
\[\|\eta\| = \|T \eta\|, \ \forall \eta \in \text{Span } B.\]
Let us prove that \(T\) is surjective. Notice that, the above computation, applied to singleton sets \(F = \{k\}\), \(k \in J\), proves that
\[T \xi_k = \delta_k, \ \forall k \in J.\]
In particular, we have
\[\text{Ran } T \supset T(\text{Span } B) = \text{Span } T(B) = \text{Span } \{ T \xi_k : k \in J \} = \text{Span } \{ \delta_k : k \in J \} = \text{fin}_k(J),\]
which proves that \(\text{Ran } T\) is dense in \(\ell_2^J(J)\). We know however that \(T\) is isometric, so \(\text{Ran } T \subset \ell_2^J(J)\) is closed. This forces \(\text{Ran } T = \ell_2^J(J)\).
\[\Box\]
COROLLARY 3.5 (Parseval Identity). Let \( \mathcal{H} \) be a Hilbert space, and let \( \mathcal{B} = \{ \xi_j : j \in J \} \) be an orthonormal basis for \( \mathcal{H} \). One has:

\[
(\zeta \mid \eta) = \sum_{j \in I} (\xi_j \mid \eta) \cdot (\xi_j \mid \eta), \quad \forall \zeta, \eta \in \mathcal{H}.
\]

**Proof.** If we define \( \alpha(j) = (\xi_j \mid \zeta) \) and \( (\xi_j \mid \eta), \forall j \in I \), 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

\[
\sum_{j \in J} \overline{\alpha(j)}\beta(j) = (\alpha \mid \beta) = (T\zeta \mid T\eta) = (\zeta \mid \eta).
\]

\( \square \)

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

\[
\sum_{j \in J} \alpha(j)\xi_j.
\]

The summation notation is justified by the following fact.

**Proposition 3.7.** With the above notations, the net \( (\eta_F)_{F \in \mathcal{P}_{\text{fin}}(J)} \subset \mathcal{H} \) of partial sums, defined by

\[
\eta_F = \sum_{k \in F} \alpha(k)\xi_k, \quad \forall F \in \mathcal{P}_{\text{fin}}(J),
\]

is convergent in the norm topology to the vector \( \sum_{j \in J} \alpha(j)\xi_j \).

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

\[
\alpha_F(k) = \begin{cases} 
\alpha(k) & \text{if } k \in F \\
0 & \text{if } k \in J \setminus F
\end{cases}
\]

then \( T^{-1}\alpha_F = \eta_F \). Using the fact that \( T \) is an isometry, we have

\[
\|\eta - \eta_F\| = \|T\eta - T\eta_F\| = \|\alpha - \alpha_F\|
\]

and the desired property follows from the well-known properties of \( \ell^2_{\mathbb{R}}(J) \). \( \square \)

**Exercise 4.** Let \( \mathcal{H} \) be a Hilbert space, let \( \mathcal{F} = \{ \xi_j : j \in J \} \) be an orthonormal set. Define the closed linear subspace \( \mathcal{H}_F = \overline{\text{Span} \mathcal{F}} \). Prove that the orthogonal projection \( P_{\mathcal{H}_F} \) is defined by

\[
P_{\mathcal{H}_F} \eta = \sum_{j \in J} (\xi_j \mid \eta)\xi_j, \quad \forall \eta \in \mathcal{H}.
\]

**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 \( \beta^j = T\eta \mid_{J} \) belongs to \( \ell^2_{\mathbb{R}}(J) \). In particular, the sum

\[
\eta_F = \sum_{j \in J} (\xi_j \mid \eta)\xi_j
\]

is “legitimate” and defines an element in \( \mathcal{H}_F \) (use the fact that \( \mathcal{F} \) is an orthonormal basis for \( \mathcal{H}_F \)). Finally, prove that \( (\eta - \eta_F) \perp \mathcal{F} \), using Parseval Identity.
Theorem 3.4. Let \( \mathcal{H} \) be a Hilbert space. Then any two orthonormal bases of \( \mathcal{H} \) have the same cardinality.

**Proof.** Fix two orthonormal bases \( B \) and \( B' \). There are two possible cases.

**Case I:** One of the sets \( B \) or \( B' \) is finite.

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

\[
\text{Span} \ B = \text{Span} \ B' = \mathcal{H},
\]

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

\[
\text{Card} \ B = \text{Card} \ B' = \dim \mathcal{H}.
\]

**Case II:** Both \( B \) and \( B' \) are infinite.

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

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

\[
\text{Card} \ Z = \text{Card} \ B'.
\]

To prove this fact, we define the set

\[
X = \text{Span}_\mathbb{Q} B'.
\]

It is clear that

\[
\text{Card} \ X = \text{Card} \ B'.
\]

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

\[
Z = X + iX,
\]

and we will still have

\[
\text{Card} \ Z = \text{Card} \ X = \text{Card} \ B'.
\]

Now we are done, since clearly \( Z \) is dense in \( \text{Span}_\mathbb{C} B' \).

Choose \( Z \) as in Claim 1. For every \( \xi \in B \) we choose a vector \( \zeta_\xi \in Z \), such that

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

**Claim 2:** The map \( B \ni \xi \mapsto \zeta_\xi \in Z \) is injective.

Start with two vectors \( \xi_1, \xi_2 \in B \), 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 Pythagorean 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

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

\[\square\]

**Corollary 3.6** (of the proof). A Hilbert space is separable, in the norm topology, if and only if it has an orthonormal basis which is at most countable.

**Proof.** Use Claims 1 and 2 from the proof of the Theorem. \[\square\]

**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 3.7.** 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 $T : \mathcal{H} \to \ell_2(I)$ and $T' : \mathcal{H}' \to \ell_2(I)$. 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

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

\[\square\]