Hilbert Spaces I: Basic Properties

Notes from the Functional Analysis Course (Fall 07 - Spring 08)

In this section we introduce an important class of Banach spaces, which carry some additional geometric structure, that enables us to use our two- or three-dimensional intuition.

Convention. Throughout this note all vector spaces are over $\mathbb{C}$.

A. Algebraic Preliminaries: Sesquilinear Forms

Definition. Given two vector spaces $X$ and $Y$, a map $\phi : X \times Y \rightarrow \mathbb{C}$ is said to be a 

The Parallelogram Law. If $X$ is a vector space and $\phi : X \times X \rightarrow \mathbb{C}$ is a sesquilinear form on $X \times X$, then

\[
\phi(x + y, x + y) + \phi(x - y, x - y) = 2[\phi(x, x) + \phi(y, y)], \; \forall x, y \in X. \tag{1}
\]

Proof. This follows directly, using the properties of sesquilinear forms, which yield

\[
\phi(x + y, x + y) = \phi(x, x) + \phi(x, y) + \phi(y, x) + \phi(y, y),
\]

\[
\phi(x - y, x - y) = \phi(x, x) - \phi(x, y) - \phi(y, x) + \phi(y, y),
\]

for all $x, y \in X$.

Lemma 2 (The Polarization Identity). If $X$ is a vector space and $\phi : X \times X \rightarrow \mathbb{C}$ is a sesquilinear form, then

\[
\phi(x, y) = \frac{1}{4} \sum_{k=0}^{3} i^{-k} \phi(x + i^k y, x + i^k y), \; \forall x, y \in X. \tag{2}
\]
Proof. Denote by $\mathbb{T}_4$ the set of roots of unity of order 4, that is, $\mathbb{T}_4 = \{ \pm 1, \pm i \}$, so that the sum in the right-hand side of (2) simply reads $\sum_{\zeta \in \mathbb{T}_4} \overline{\zeta} \phi(x + \zeta y, x + \zeta y)$.

By the properties of sesquilinear forms, we have

$$\phi(x + \zeta y, x + \zeta y) = \phi(x, x) + \phi(\zeta y, x) + \phi(x, \zeta y) + \phi(\zeta y, \zeta y) = $$

$$= \phi(x, x) + \overline{\zeta} \phi(y, x) + \zeta \phi(x, y) + \overline{\zeta} \zeta \phi(y, y), \ \forall \ x, y \in X, \ \zeta \in \mathbb{C}.$$

If we specialize to the case when $\zeta \in \mathbb{T}$, i.e. $|\zeta| = 1 (= \overline{\zeta} \zeta)$, then the above calculation yields

$$\overline{\zeta} \phi(x + \zeta y, x + \zeta y) = \overline{\zeta} \phi(x, x) + \overline{\zeta}^2 \phi(y, x) + \phi(x, y) + \overline{\zeta} \phi(y, y),$$

and then the sum in the right-hand side of (2) is:

$$\sum_{k=0}^{3} i^{-k} \phi(x + i^k y, x + i^k y) = \left( \sum_{\zeta \in \mathbb{T}_4} \overline{\zeta} \right) \cdot \left[ \phi(x, x) + \phi(y, y) \right] + \left( \sum_{\zeta \in \mathbb{T}_4} \overline{\zeta}^2 \right) \cdot \phi(y, x) + 4 \phi(x, y),$$

so by the obvious identities $\sum_{\zeta \in \mathbb{T}_4} \overline{\zeta} = \sum_{\zeta \in \mathbb{T}_4} \overline{\zeta}^2 = 0$, the desired identity (2) follows.

Definitions. Let $X$ be a vector space, and let $\phi : X \times X \to \mathbb{C}$ be a sesquilinear form.

A. We say that $\phi$ is hermitian, if $\phi^* = \phi$, which means:

$$\phi(x, y) = \overline{\phi(y, x)}, \ \forall \ x, y \in X. \quad (3)$$

B. We say that $\phi$ is positive definite, if:

$$\phi(x, x) \geq 0, \ \forall \ x \in X. \quad (4)$$

C. We say that $\phi$ is strictly positive definite, if:

(i) $\phi$ is positive definite, and

(ii) $\phi(x, x) = 0 \Rightarrow x = 0$.

Remarks 1-3. Let $\phi : X \times X \to \mathbb{C}$ be a sesquilinear form.

1. The condition that $\phi$ is hermitian is equivalent to the condition:

$$\phi(x, x) \in \mathbb{R}, \ \forall \ x \in X. \quad (5)$$

The implication “hermitian” $\Rightarrow$ (5) is trivial. Conversely, if $\phi$ satisfies (5), then (using the notations from the proof of Lemma 2), by the Polarization Identity we have

$$\overline{\phi(y, x)} = \frac{1}{4} \sum_{\zeta \in \mathbb{T}_4} \zeta \phi(y + \zeta x, y + \zeta x) = \frac{1}{4} \sum_{\zeta \in \mathbb{T}_4} \zeta \phi(y + \zeta x, y + \zeta x),$$

so using the “change of variable” $\gamma = \zeta$, we also get

$$\overline{\phi(y, x)} = \frac{1}{4} \sum_{\gamma \in \mathbb{T}_4} \overline{\gamma} \phi(y + \gamma x, y + \gamma x), \quad (6)$$

2
and now the desired identity (3) follows from the properties of sesquilinear forms, which imply
\[
\phi(y + \gamma x, y + \gamma x) = \phi(\overline{\gamma}(x + \gamma y), \overline{\gamma}(x + \gamma y)) = \\
\gamma \overline{\gamma} \phi(x + \gamma y, x + \gamma y) = \phi(x + \gamma y, x + \gamma y), \quad \forall \gamma \in \mathbb{T}_4,
\]
so when we go back to (6), again by the Polarization Identity we get
\[
\overline{\phi(y, x)} = \frac{1}{4} \sum_{\gamma \in \mathbb{T}_4} \overline{\gamma}\phi(x + \gamma y, x + \gamma y) = \phi(x, y).
\]

2. If \(\phi\) is hermitian, then it satisfies
\[
\phi(x + y, x + y) = \phi(x, x) + \phi(y, y) + 2\Re \phi(x, y),
\]
\[
\phi(x - y, x - y) = \phi(x, x) + \phi(y, y) - 2\Re \phi(x, y),
\]
for all \(x, y \in \mathcal{X}\). These two identities are referred to as the Law of Cosine.

3. If \(\phi\) is positive definite, then \(\phi\) is hermitian. This follows immediately from Remark 1.

**Lemma 3** (Cauchy-Buniakowski-Schwartz Inequality). If \(\phi : \mathcal{X} \times \mathcal{X} \to \mathbb{C}\), then
\[
|\phi(x, y)|^2 \leq \phi(x, x) \cdot \phi(y, y), \quad \forall x, y \in \mathcal{X}.
\]
Furthermore, if \(\phi(x, x) \neq 0\), then one has equality in (9), if and only if there exists \(\zeta \in \mathbb{C}\), such that
\[
\phi(\zeta x + y, \zeta x + y) = 0.
\]

**Proof.** Fix \(x\) and \(y\), and note that, by the Law of Cosine, one has
\[
0 \leq \phi(\alpha x + \beta y, \alpha x + \beta y) = |\alpha|^2 \phi(x, x) + 2\Re[\overline{\alpha}\beta \phi(x, y)] + |\beta|^2 \phi(y, y), \quad \forall \alpha, \beta \in \mathbb{C}.
\]

Fix now \(\beta \in \mathbb{C}\), with \(|\beta| = 1\), such that \(\beta \phi(x, y) = |\phi(x, y)|\), and define the function \(g : \mathbb{R} \ni t \mapsto \phi(tx + \beta y, tx + \beta y) \in \mathbb{R}\), which by (11) can equivalently be written as
\[
g(t) = t^2 \phi(x, x) + 2t|\phi(x, y)| + \phi(y, y), \quad t \in \mathbb{R}.
\]
By positive definiteness, we also know that
\[
g(t) \geq 0, \quad \forall t \in \mathbb{R}.
\]
We now have two cases to consider: (i) \(\phi(x, x) = 0\); (ii) \(\phi(x, x) > 0\).

In case (i) – when \(g\) is a polynomial of degree 1 – condition (13) forces \(g\) to be constant, so \(\phi(x, y) = 0\), and then (9) is trivial.

In case (ii) it follows that \(g\) is a quadratic polynomial (with positive leading coefficient), so condition (13) forces the discriminant \(\Delta\) to be non-positive, i.e. \(4|\phi(x, y)|^2 - 4\phi(x, x) \cdot \phi(y, y) \leq 0\), from which (9) is trivial. Furthermore, if equality holds in (9), this means that
Δ = 0, so the equation \( g(t) = 0 \) has a unique real solution \( t_0 \), which means precisely that 
\[
\phi(t_0x + \beta y, t_0x + \beta y) = 0,
\]
so we also get
\[
\phi(t_0\overline{\beta}x + y, t\overline{\beta}x + y) = \phi(\overline{\beta}(t_0x + \beta y), \overline{\beta}(t_0x + \beta y)) = \beta\overline{\beta}\phi(t_0x + \beta y, t_0x + \beta y) = 0,
\]
so (10) holds with \( \zeta = t_0\overline{\beta} \).

Conversely, if (10) holds for some \( \zeta \), then by (9) it follows that
\[
0 \leq |\phi(x, \zeta x + y)|^2 \leq \phi(x, x) \cdot \phi(\zeta x + y, \zeta x + y) = 0,
\]
which give \( \phi(x, \zeta x + y) = \phi(y, \zeta x + y) = 0 \), so we get \( \phi(x, y) = -\zeta\phi(x, x) \) and \( \phi(y, y) = -\zeta\overline{\phi}(x, y) = \zeta\overline{\phi}(x, x) \), thus clearly forcing
\[
|\phi(x, y)|^2 = |\zeta|^2\phi(x, x)^2 = \phi(x, x) \cdot \phi(y, y). \quad \square
\]

The final result in this sub-section establishes the characterization of positive definite sesquilinear forms in terms of convexity.

**Theorem 1.** Let \( \mathcal{X} \) be a \( \mathbb{C} \)-vector space. For a map \( q : \mathcal{X} \to [0, \infty) \), the following are equivalent.

(i) There exists a positive definite sesquilinear form \( \phi : \mathcal{X} \times \mathcal{X} \to \mathbb{C} \), such that
\[
q(x) = \sqrt{\phi(x, x)}, \quad \forall x \in \mathcal{X}.
\]

(ii) \( q \) is a seminorm on \( \mathcal{X} \), which satisfies the Parallelogram Law:
\[
q(x + y)^2 + q(x - y)^2 = 2q(x)^2 + 2q(y)^2, \quad \forall x, y \in \mathcal{X}.
\]

Moreover, \( \phi \) is unique and satisfies the Polarization Identity:
\[
\phi(x, y) = \frac{1}{4} \sum_{k=0}^{3} i^{-k} q(x + i^k y)^2,
\]

**Proof.** (i) \( \Rightarrow \) (ii). Assume \( q \) is defined by (14), and let us prove first that \( q \) is a seminorm. Since \( \phi(\zeta x, \zeta x) = \overline{\zeta^2} \phi(x, x) = |\zeta|^2\phi(x, x) \), by taking square roots, it follows that
\[
q(\zeta x) = |\zeta| \cdot q(x), \quad \forall \zeta \in \mathbb{C}, \ x \in \mathcal{X}.
\]

To prove the triangle inequality, we use the Law of Cosine and the Cauchy-Buniakowski-Schwartz Inequality, which gives:
\[
\phi(x + y, x + y) = |\phi(x + y, x + y)| = |\phi(x, x)| + 2\text{Re} \phi(x, y) + \phi(y, y) \leq \\
= |\phi(x, x) + 2\phi(x, y)| + \phi(y, y) \leq \\
\leq \phi(x, x) + 2\sqrt{\phi(x, x) \cdot \phi(y, y)} + \phi(y, y) = [\sqrt{\phi(x, x)} + \sqrt{\phi(y, y)}]^2,
\]
so by taking square roots we get \( q(x + y) \leq q(x) + q(y) \), \( \forall x, y \in \mathcal{X} \).
The fact that \( q \) satisfies the Parallelogram Law (15), as well as the Polarization Identity (16) is immediate from Lemmas 1 and 2.

\((ii) \Rightarrow (i)\). Assume \( q \) be a seminorm that satisfies the Parallelogram Law (15), define \( \phi : \mathcal{X} \times \mathcal{X} \to \mathbb{C} \) by (16), and let us prove that \( \phi \) is positive definite, and it also satisfies (14).

**Claim 1:** \( \phi(y, x) = \phi(x, y), \forall x, y \in \mathcal{X} \).

Fix \( x, y \in \mathcal{X} \), so by definition, we have

\[
\phi(y, x) = \frac{1}{4} \sum_{k=0}^{3} i^{k} q(y + i^{k} x)^{2}. \tag{17}
\]

Since \( q \) is a seminorm, we have \( q(y + i^{k} x) = q(i^{k}(x + i^{-k} y)) = |i^{k}| \cdot q(x + i^{-k} y) = q(x + i^{-k} y), \forall k = 0, \ldots, 3 \), so if we go back to (17) we now have \( \phi(y, x) = \frac{1}{4} \sum_{k=0}^{3} i^{k} q(x + i^{-k} y)^{2} \), and using the notations from Lemma 2 we have

\[
\phi(y, x) = \frac{1}{4} \sum_{\zeta \in \mathcal{T}_{4}} \zeta q(x + \zeta y)^{2} = \frac{1}{4} \sum_{\zeta \in \mathcal{T}_{4}} \zeta q(x + \zeta y)^{2} = \phi(x, y).
\]

(The second equality uses the “change of variable” \( \gamma = \zeta \).

Having proved Claim 1, we now see that, in order to prove that \( \phi \) is sesquilinear, it suffices to prove that, for every \( x \in \mathcal{X} \), the map \( \mathcal{X} \ni y \mapsto \phi(x, y) \in \mathbb{C} \) is \( \mathbb{C} \)-linear.

**Claim 2:** \( \phi(x, 2y) = 2\phi(x, y), \forall x, y \in \mathcal{X} \).

Using the Parallelogram Law (with \( x + i^{k} y \) in place of \( x \) and \( i^{k} y \) in place of \( y \)), we have the equality \( q(x + 2i^{k} y)^{2} + q(i^{k} y)^{2} = 2q(x + i^{k} y)^{2} + 2q(i^{k} y) \), which combined with the equality \( q(i^{k} y) = q(y) \), gives

\[
q(x + 2i^{k} y)^{2} = 2q(x + i^{k} y)^{2} + q(y)^{2},
\]

so by the definition of \( \phi \) we get (the third sum is 0):

\[
\phi(x, 2y) = \frac{1}{4} \sum_{k=0}^{3} i^{-k} q(x + 2i^{k} y)^{2} = \frac{1}{4} \left[ \sum_{k=0}^{3} 2i^{-k} q(x + i^{k} y)^{2} \right] + \frac{1}{4} \left[ \sum_{k=0}^{3} i^{-k} \right] \cdot q(y) = 2\phi(x, y).
\]

**Claim 3:** \( \phi(x, y_{1} + y_{2}) = \phi(x, y_{1} + y_{2}), \forall x, y_{1}, y_{2} \in \mathcal{X} \).

If we apply the Parallelogram Law (with \( \frac{1}{2} x + i^{k} y_{1} \) in place of \( x \) and \( \frac{1}{2} x + i^{k} y_{2} \) in place of \( y \), we get

\[
q(x + i^{k} y_{1} + i^{k} y_{2})^{2} = 2q\left(\frac{1}{2} x + i^{k} y_{1}\right)^{2} + 2q\left(\frac{1}{2} x + i^{k} y_{2}\right)^{2} - q(i^{k} y_{1} - i^{k} y_{2})^{2},
\]

so if we multiply by \( i^{-k} \), sum up, and divide by 4, then using the identity \( q(i^{k} y_{1} - i^{k} y_{2}) = q(y_{1} - y_{2}) \), we get (the third sum is 0):

\[
\phi(x, y_{1} + y_{2}) = \frac{1}{4} \sum_{k=0}^{3} 2i^{-k} q\left(\frac{1}{2} x + i^{k} y_{1}\right)^{2} + \frac{1}{4} \sum_{k=0}^{3} 2i^{-k} q\left(\frac{1}{2} x + i^{k} y_{1}\right)^{2} +
\]

\[
+ \frac{1}{4} \left[ \sum_{k=0}^{3} i^{-k} \right] \cdot q(y_{1} - y_{2})^{2} = 2\phi\left(\frac{1}{2} x, y_{1}\right) + 2\phi\left(\frac{1}{2} x, y_{2}\right). \tag{18}
\]
By Claims 1 and 2, however, we have

\[ 2\phi(\frac{1}{2}x, y_j) = 2\phi(y_j, \frac{1}{2}x) = \bar{\phi}(y_j, x) = \phi(x, y_j), \quad j = 1, 2, \]

so going back to (18) yields: \( \phi(x, y_1 + y_2) = \phi(x, y_1) + \phi(x, y_2). \)

**Claim 4:** \( \phi(x, \zeta y) = \zeta \phi(x, y), \quad \forall \, x, y \in \mathcal{X}, \, \zeta \in \mathbb{C}. \)

First of all, the equality is obvious if \( \zeta = 0, \) since \( \phi(x, 0) = \frac{1}{4} \cdot [\sum_{k=0}^{3} i^{-k}] \cdot q(x)^2 = 0. \)

Second, since the map \( \mathbb{C} \ni \alpha \mapsto q(x + \alpha y) \in [0, \infty) \) is continuous, since it satisfies the inequality \( |q(x + \alpha y) - q(x + \beta y)| \leq q(\alpha y - \beta y) = |\alpha - \beta|q(y), \forall \alpha, \beta \in \mathbb{C}, \) it follows that (for fixed \( x, y \in \mathcal{X} \)), the map \( \zeta \mapsto \phi(x, \zeta y) \) continuous in \( \zeta, \) so it suffices to prove the Claim for \( \zeta \) in a dense subset of \( \mathbb{C}, \) and we choose such a dense subset to be

\[ S = \{(m + ni)2^{-k} : m, n \in \mathbb{Z}, \, k \in \mathbb{N}\}. \]

Since by Claim 2 we clearly have \( \phi(x, 2^{-k}y) = 2^{-k}\phi(x, y), \) it suffices to consider \( \zeta \) in the subset \( S_0 = \{m + ni : m, n \in \mathbb{Z}\}. \) Moreover since by Claim 3 (and the case \( \zeta = 0 \)) we know that

\[ \phi(x, (m + ni)y) = \phi(x, my) + \phi(x, niy) = m\phi(x, y) + n\phi(x, iy), \]

it suffices to prove the Claim only for \( \zeta = i. \) But this is pretty clear since by definition (using again the notations from Lemma 2, and the “change of variable” \( \gamma = i\zeta \)) we have

\[ \phi(x, iy) = \frac{1}{4} \sum_{\zeta \in \mathbb{Z}_4} \bar{\zeta}q(x + i\zeta y)^2 = \frac{i}{4} \sum_{\zeta \in \mathbb{Z}_4} i\bar{\zeta}q(x + i\zeta y)^2 = \frac{i}{4} \sum_{\gamma \in \mathbb{Z}_4} \bar{\gamma}q(x + \gamma y)^2 = i\phi(x, y). \]

Having proved that \( \phi \) is a (hermitian) sesquilinear form, all that remains to prove is the fact that it satisfies the identity

\[ \phi(x, x) = q(x)^2, \quad \forall \, x \in \mathcal{X}. \quad (19) \]

At this point, all we have to do is to notice that

\[ q(x + i^k x) = \begin{cases} q(x + x) = 2q(x), & \text{if } k = 0 \\ q(x - x) = 0, & \text{if } k = 2 \\ q(x \pm i^k x) = |1 \pm i| \cdot q(x) = \sqrt{2}q(x), & \text{if } k = 1, 3 \end{cases} \]

so by the definition of \( \phi \) we have

\[ \phi(x, x) = \frac{1}{4} \sum_{k=0}^{3} i^{-k}q(x + i^k x)^2 = \frac{1}{4} \left[ 4q(x)^2 - 2iq(x)^2 + 2iq(x)^2 \right] = q(x)^2. \]

**B. (Pre-)Hilbert Spaces**

The spaces we introduce in this sub-section are those whose geometry is “governed” by the Parallelogram Law.

**Definitions.** Let \( \mathcal{H} \) be a normed vector space.
A. We say that $\mathcal{H}$ is a pre-Hilbert space, if its norm satisfies the Parallelogram Law:

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2, \ \forall \ x, y \in \mathcal{H}. \quad (20)$$

B. If $\mathcal{H}$ is a pre-Hilbert space, then according to Theorem 1, there exists a unique strictly positive sesquilinear map, denoted from now on by

$$(\cdot | \cdot) : \mathcal{H} \times \mathcal{H} \ni (x, y) \mapsto (x|y) \in \mathbb{C},$$

which is referred to as the inner product, satisfying

$$(x|x) = \|x\|^2, \ \forall \ x \in \mathcal{H}.$$

Moreover, the inner product can be reconstructed out of the norm, by the Polarization Identity:

$$(x|y) = \frac{1}{4} \sum_{k=0}^{3} i^{-k}\|x + i^{k}y\|^2, \ \forall \ x, y \in \mathcal{H}. \quad (21)$$

C. We say that $\mathcal{H}$ is a Hilbert space, if $\mathcal{H}$ is simultaneously a pre-Hilbert space and a Banach space.

Remarks 4-5. Assume $\mathcal{H}$ is a pre-Hilbert space. For future reference, we record here two of the results from the previous sub-section, which in this current setting have the following specific forms.

4. By Remark 2, the Law of Cosine, reads:

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\text{Re}(x|y), \quad (22)$$
$$\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2\text{Re}(x|y), \quad (23)$$

for all $x, y \in \mathcal{H}$.

5. By Lemma 3, the Cauchy-Buniakowksi-Schwartz inequality reads:

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\|, \ \forall \ x, y \in \mathcal{H}. \quad (24)$$

Moreover, if $x \neq 0$, then one has equality in (24), if and only if there exists $\zeta \in \mathbb{C}$, such that $y = \zeta x$.

Remark 6. If $q$ is a seminorm on a $\mathbb{C}$-vector space $\mathcal{X}$, that satisfies the Parallelogram Law (15), then by Theorem 1 there exists a (unique) positive definite sesquilinear form $\phi : \mathcal{X} \times \mathcal{X} \to \mathbb{C}$, such that $\phi(x, x) = q(x)^2, \ \forall \ x \in \mathcal{X}$. Consider the linear subspace $\mathcal{N} = \{x \in \mathcal{X} : q(x) = 0\}$. If we equip the quotient space $\mathcal{X}/\mathcal{N}$ with the induced norm, defined (correctly) by $\|\hat{x}\| = q(x)$, where $\mathcal{X} \ni x \mapsto \hat{x} \in \mathcal{X}/\mathcal{N}$ denotes the quotient map, then it is obvious that $\mathcal{X}/\mathcal{N}$ becomes a pre-Hilbert space. The inner product on the quotient space is defined (correctly) by $\langle \hat{x}|\hat{y} \rangle = \phi(x, y)$.

Remark 7. The completion of a pre-Hilbert space is obviously a Hilbert space, since the norm obviously satisfies the Parallelogram Law.
Remark 8. If $\mathcal{H}$ is a pre-Hilbert space, and we equip the product space $\mathcal{H} \times \mathcal{H}$ with the product topology, then the inner product map $\mathcal{H} \times \mathcal{H} \ni (x,y) \mapsto (x|y) \in \mathbb{C}$ is continuous. Indeed, given two norm-convergent sequences $x_n \to x$ and $y_n \to y$ in $\mathcal{H}$, by the Cauchy-Buniakowski-Schwartz Inequality we have

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

so using the fact that $\|y_n\| \to \|y\|$ it follows immediately that $|(x_n|y_n) - (x|y)| \to 0$.

Example 1. Given a positive integer $n$, we equip $\mathbb{C}^n$ with the norm

$$\|x\|_2 = \sqrt{|x_1|^2 + \cdots + |x_n|^2}, \quad x = (x_1, \ldots, x_n) \in \mathbb{C}^n.$$

It is pretty clear that $(\mathbb{C}^n, \|\cdot\|_2)$ is a Hilbert space, hereafter referred to as the standard $n$-dimensional Hilbert space. The standard inner product is defined by

$$(x|y) = \bar{x}_1y_1 + \cdots + \bar{x}_ny_n + n, \quad x = (x_1, \ldots, x_n), \quad y = (y_1, \ldots, y_n) \in \mathbb{C}^n.$$

Example 2. Given a measure space $(X, \mathcal{A}, \mu)$, the Banach space $L^2(X, \mathcal{A}, \mu)$ is a Hilbert space. The corresponding inner product is

$$(f|g) = \int_X \bar{f}g \, d\mu, \quad f, g \in L^2(X, \mathcal{A}, \mu).$$

Example 3. As a special case of Example 2, if we start with an arbitrary set $S$, and we consider the measure space $(S, \mathcal{P}(S), \nu)$, where $\nu$ is the counting measure, we obtain the Hilbert space $\ell^2(S)$, whose inner product is

$$(x|y) = \sum_{s \in S} \bar{x}_sy_s, \quad x = (x_s)_{s \in S}, \quad y = (y_s)_{s \in S} \in \ell^2(S).$$

The Hilbert space $\ell^2(S)$ is the completion of the pre-Hilbert space $(\bigoplus_{s \in S} \mathbb{C}, \|\cdot\|_2)$.

In the case when $S = \mathbb{N}$, the Hilbert space $\ell^2(\mathbb{N})$ is simply denoted by $\ell^2$.

Example 4. If $\Omega$ is a locally compact Hausdorff space, and $\mu$ is an “honest” Radon measure on $\Omega$, such that

(*) $\mu(D) > 0$, for all non-empty open sets $D \subset \Omega$,

then the space $C_c(\Omega)$ of $\mathbb{C}$-valued continuous functions on $\Omega$, with compact support, is injectively embedded in $L^2(\Omega,\text{Bor}(\Omega),\mu) = L^2(\Omega,\mathcal{M}(\mu),\mu)$, so $(C_c(\Omega), \|\cdot\|_2)$ is a pre-Hilbert space.

In the absence of condition (*), the natural map $C_c(\Omega) \to L^2(\Omega,\text{Bor}(\Omega),\mu)$ is not injective, so we only obtain a seminorm

$$q(f) = \int_X |f|^2 \, d\mu, \quad f \in C_c(\Omega),$$

which still satisfies the Parallelogram Law (15). Passing to quotients and taking completion (as indicated in Remarks 6 and 7, we obtain again the Hilbert space $L^2(\Omega,\text{Bor}(\Omega),\mu) = L^2(\Omega,\mathcal{M}(\mu),\mu)$.)
In what follows we are going to derive an important result concerning the geometry of Hilbert spaces, in part based on the Parallelogram Law.

**Theorem 2.** Assume $H$ is a Hilbert space, and $C \subset H$ is a non-empty, norm-closed, convex subset. For every $x \in H$, there exists a unique $x_0 \in C$, such that

$$\|x - x_0\| \leq \|x - y\|, \quad \forall y \in C,$$

so in effect $\|x - x_0\| = \text{dist}(x, C)$. Moreover, the vector $x_0 \in C$ is the unique one satisfying the inequality

$$\Re(x - x_0|y - x_0) \leq 0, \quad \forall y \in C. \quad (26)$$

**Proof.** Consider the quantity $\delta = \text{dist}(x, C) = \inf_{y \in C} \|x - y\|$. Since $C$ is norm-closed, one has the equivalence $\delta = 0 \leftrightarrow x \in C$, in which case the Theorem is trivially satisfied with $x_0 = x$. (The fact that $x_0 = x$ is the only vector in $C$ that satisfies (26) is quite obvious, because in (26) we can let $y = x$.) Therefore, for the rest of the proof we may assume $\delta > 0$, thus $x \notin C$. By the definition of $\delta$, there exists a sequence $(y_n)_{n=1}^{\infty} \subset C$, such that $\lim_{n \to \infty} \|x - y_n\| = \delta$.

Let us apply now the Parallelogram Law (20) – with $x - y_n$ in place of $x$ and $x - y_n$ in place of $y$ – to obtain the equality

$$\|y_m - y_n\|^2 = 2\|x - y_m\|^2 + 2\|x - y_n\|^2 - 2\|x - y_m - y_n\|^2, \quad \forall m, n \in \mathbb{N}. \quad (27)$$

Since $C$ is convex, the vector $\frac{1}{2}(y_m + y_n)$ belongs to $C$, so (before taking squares) the last term in the right-hand side of (27) satisfies the inequality

$$\|2x - y_m - y_n\| = 2\|x - \frac{1}{2}(y_m + y_n)\| \geq 2\delta,$$

so if we go back to (27) we obtain

$$\|y_m - y_n\|^2 \leq 2\|x - y_m\|^2 + 2\|x - y_n\|^2 - 4\delta^2, \quad \forall m, n \in \mathbb{N}. \quad (28)$$

Since $\lim_{n \to \infty} \|x - y_n\| = \delta$, by (28) it follows immediately that $(y_n)_{n=1}^{\infty}$ is a Cauchy sequence, this convergent to some point $x_0$ (that necessarily belongs to $C$, which is closed), and this point will satisfy $\|x - x_0\| = \lim_{n \to \infty} \|x - y_n\| = \delta$, thus having the desired property (25).

To prove the uniqueness of $x_0$, among all $x_0 \in C$ that satisfy (25), we first prove the following

**Claim 1:** If $x_0 \in C$ is any vector with property (25), then $x_0$ also satisfies (26).

To prove the Claim, we argue by contradiction, assuming there exist $x_0, y \in C$, such that $\|x - x_0\| = \delta$, and $\Re(x - x_0|y - x_0) > 0$. Define, for every $t \in \mathbb{R}$, the vector

$$y_t = x_0 + t(y - x_0) = (1 - t)x_0 + ty, \quad (29)$$

and consider the function $f : \mathbb{R} \to \mathbb{R}$, given by $f(t) = \|x - y_t\|^2 = \|(x - x_0) - t(y - x_0)\|^2$.

Using the Law of Cosine (23) – with $x - x_0$ in place of $x$ and $t(y - x_0)$ in place of $y$ – we can write

$$f(t) = \|x - x_0\|^2 + t^2\|y - x_0\|^2 - 2t\Re(x - x_0)y - x_0),$$

so $f$ is a quadratic function, with $f(0) = \|x - x_0\|^2 = \delta^2$, and $f'(0) = -2\Re(x - x_0)y - x_0).$

Since by assumption we have $f'(0) < 0$, it follows that $f$ is decreasing on a small open
interval around 0, so in particular there exists \( s \in (0, 1) \), such that \( f(s) < f(0) = \delta^2 \). But now we reached a contradiction, because for such an \( s \), by convexity and by (29) it follows that \( y_s \) belongs to \( \mathcal{C} \), so by the definition of \( \delta \) we also have \( f(s) = \|x - y_s\|^2 \geq \delta^2 \).

Having proved Claim 1, we now see that in order to finish the proof of the Theorem, all we need to do is to prove the second uniqueness statement, that is, the following

**Claim 2:** Whenever \( x_0, x_1 \) are such that

\[
\text{Re}(x - x_k|y - x_k) \leq 0, \quad \forall y \in \mathcal{C}, \ k = 0, 1, \tag{30}
\]

it follows that \( x_0 = x_1 \).

Using (30) it follows immediately that

\[
\begin{align*}
\text{Re}(x - x_0|x_1 - x_0) & \leq 0; \\
\text{Re}(x_1 - x|x_1 - x_0) & = \text{Re}(x - x_1|x_0 - x_1) \leq 0,
\end{align*}
\]

so if we add these two inequalities we get \( \text{Re}(x_1 - x_0|x_1 - x_0) = 0 \), which means that \( \|x_1 - x_0\|^2 = 0 \), i.e. \( x_1 = x_0 \).

**Orthogonal Projections and Riesz’ Theorem**

As an important application of Theorem 2 from the preceding sub-section, we have the following.

**Proposition-Definition 1.** Let \( \mathcal{H} \) be a Hilbert space, and let \( \mathcal{Y} \subset \mathcal{H} \) be a closed linear subspace. Denote, for every \( x \in \mathcal{H} \), by \( P_\mathcal{Y}x \) the unique vector \( x_0 \in \mathcal{Y} \), such that

\[
\|x - x_0\| \leq \|x - y\|, \quad \forall y \in \mathcal{Y}. \tag{31}
\]

(See Theorem 2 for existence, uniqueness, and additional properties of \( x_0 \).)

The map \( x \mapsto P_\mathcal{Y}x \) has the following properties.

(i) For every vector \( x \in \mathcal{H} \), the vector \( P_\mathcal{Y}x \) is the unique vector \( x_0 \in \mathcal{Y} \), such that

\[
(x - x_0|y) = 0, \quad \forall y \in \mathcal{Y}. \tag{32}
\]

(ii) \( P_\mathcal{Y} : \mathcal{H} \to \mathcal{H} \) is linear, and satisfies the identity

\[
\|P_\mathcal{X}x\|^2 + \|x - P_\mathcal{Y}x\|^2 = \|x\|^2, \quad \forall x \in \mathcal{H}, \tag{33}
\]

so in particular \( P_\mathcal{Y} \) is also continuous, with \( \|P_\mathcal{Y}\| \leq 1 \).

(iii) \( P_\mathcal{Y}^2 = P_\mathcal{Y} \).

(iv) Range \( P_\mathcal{Y} = \mathcal{Y} \) and Ker \( P_\mathcal{Y} = \text{Range } (I_\mathcal{H} - P_\mathcal{Y}) \), where \( I_\mathcal{H} : \mathcal{H} \to \mathcal{H} \) denotes the identity operator.

The linear continuous operator \( P_\mathcal{Y} : \mathcal{H} \to \mathcal{H} \) is called the orthogonal projection onto \( \mathcal{Y} \).
Proof. (i). Fix $x \in \mathcal{H}$, and let $x_0 \in \mathcal{Y}$ be the unique vector satisfying (31). By Theorem 2 we also know that $x_0$ is the unique vector in $\mathcal{Y}$ satisfying

$$\text{Re}(x - x_0|y - x_0) \leq 0, \quad \forall y \in \mathcal{Y}. \tag{34}$$

Since $x_0 \in \mathcal{Y}$, it is obvious that $\{y - x_0 : y \in \mathcal{Y}\} = \mathcal{Y}$, so condition (34) is equivalent to:

$$\text{Re}(x - x_0|y) \leq 0, \quad \forall y \in \mathcal{Y}. \tag{35}$$

It is obvious that (32)$\Rightarrow$(35), so in order to finish the proof of (i), we only need to prove the implication (35)$\Rightarrow$(32). But this is clear, because if we start with some $y \in \mathcal{Y}$, we can choose some $\zeta \in \mathbb{T}$, such that $|(x - x_0|y)| = \zeta(x - x_0|y) = (x - x_0|\zeta y)$, and $\zeta y$ also belongs to $\mathcal{Y}$.

(ii). By the uniqueness condition (i), in order to prove the linearity of $P_{\mathcal{Y}}$ it suffices to prove the identities

$$((x_1 + x_2) - (P_{\mathcal{Y}}x_1 + P_{\mathcal{Y}}x_2)|y) = 0, \quad \forall x_1, x_2 \in \mathcal{H}, \; y \in \mathcal{Y}; \tag{36}$$

$$((\zeta x - \zeta P_{\mathcal{Y}}x)|y) = 0, \quad \forall x \in \mathcal{H}, \; \zeta \in \mathbb{C}, \; y \in \mathcal{Y}; \tag{37}$$

(The equality (36) yields $P_{\mathcal{Y}}(x_1 + x_2) = P_{\mathcal{Y}}x_1 + P_{\mathcal{Y}}x_2$, while the equality (37) yields $P_{\mathcal{Y}}(\zeta x) = \zeta P_{\mathcal{Y}}x$.)

Both identities (36) and (37) are trivial, since by the sesquilinearity of the inner product we have:

$$((x_1 + x_2) - (P_{\mathcal{Y}}x_1 + P_{\mathcal{Y}}x_2)|y) = (x_1 - P_{\mathcal{Y}}x_1|y) + (x_2 - P_{\mathcal{Y}}x_2|y);$$

$$((\zeta x - \zeta P_{\mathcal{Y}}x)|y) = \zeta(x - P_{\mathcal{Y}}x|y).$$

To prove the equality (33) we use the Law of Cosine, which yields

$$\|x\|^2 = \|x - P_{\mathcal{Y}}x\|^2 + \|P_{\mathcal{Y}}\|^2 + 2\text{Re}(x - P_{\mathcal{Y}}x|P_{\mathcal{Y}}x),$$

and we use property (i) with $y = P_{\mathcal{Y}}x$, which yields $(x - P_{\mathcal{Y}}x|P_{\mathcal{Y}}x) = 0$.

(iii)-(iv). First of all, it is trivial that $\text{Range}(P_{\mathcal{Y}}) \subset \mathcal{Y}$, by construction. Secondly, right from the definition it follows, that

$$P_{\mathcal{Y}}x = x, \quad \forall x \in \mathcal{Y}, \tag{38}$$

so we also have the inclusion $\text{Range}(P_{\mathcal{Y}}) \supset \mathcal{Y}$, so in fact we have the equality $\text{Range}(P_{\mathcal{Y}}) = \mathcal{Y}$. Thirdly, if we start with an arbitrary $x \in \mathcal{H}$, then $P_{\mathcal{Y}}x \in \mathcal{Y}$, and then using (38) we get $P_{\mathcal{Y}}(P_{\mathcal{Y}}x) = P_{\mathcal{Y}}x$, which means that we have the identity $P_{\mathcal{Y}}^2 = P_{\mathcal{Y}}$.

Finally, in order to prove the equality

$$\text{Ker} P_{\mathcal{Y}} = \text{Range}(I_{\mathcal{H}} - P_{\mathcal{Y}}), \tag{39}$$

we first notice that, by the obvious identity $P_{\mathcal{Y}}(I_{\mathcal{H}} - P_{\mathcal{Y}}) = P_{\mathcal{Y}} - P_{\mathcal{Y}}^2 = 0$, we clearly have the inclusion $\text{Range}(I_{\mathcal{H}} - P_{\mathcal{Y}}) \subset \text{Ker} P_{\mathcal{Y}}$. Conversely, if $x \in \text{Ker} P_{\mathcal{Y}}$, then $x = x - P_{\mathcal{Y}}x = (I_{\mathcal{H}} - P_{\mathcal{Y}})x$, so obviously $x$ belongs to $\text{Range}(I_{\mathcal{H}} - P_{\mathcal{Y}})$.

\footnote{There is nothing special about the inclusion $\text{Ker} P \subset \text{Range}(I_{\mathcal{H}} - P)$. It holds for arbitrary linear maps $P : \mathcal{H} \to \mathcal{H}$, on arbitrary vector spaces $\mathcal{H}$.}
Theorem 3 (Riesz). Let $\mathcal{H}$ be a Hilbert space. For a linear functional, $\phi : \mathcal{H} \to \mathbb{C}$, the following are equivalent:

(i) $\phi$ is norm-continuous;

(ii) there exists a vector $x \in \mathcal{H}$, such that

$$\phi(y) = (x|y), \quad \forall y \in \mathcal{H}. \quad (40)$$

Moreover, in this case the vector $x$ is unique and has $\|x\| = \|\phi\|$.

Proof. (i) $\Rightarrow$ (ii). Assume $\phi$ is norm continuous, and let us prove the existence and uniqueness of $x$ satisfying (40). The case when $\phi$ is identically 0 is trivial, by taking $x = 0$. (The uniqueness is obvious, because in the case $\phi = 0$, any vector $x$ satisfying (40) has $\|x\|^2 = (x|x) = \phi(x) = 0$.) For the remainder of the proof we assume that $\phi$ is not identically 0, which implies the existence of some $x_0$ with $\phi(x_0) \neq 0$. Consider then the (proper) linear subspace $Y = \text{Ker } \phi$, and the vector $x_1 = x_0 - P_Y x_0$. By linearity, combined with the fact that $P_Y x_0 \in Y = \text{Ker } \phi$ we have $\phi(x_1) = \phi(x_0) - \phi(P_Y x_0) = \phi(x_0) \neq 0$. Consider then the linear functional

$$\phi_1 : \mathcal{H} \ni y \mapsto (x_1|y) \in \mathbb{C},$$

and let us remark that

$$\text{Ker } \phi = \text{Ker } \phi_1. \quad (41)$$

Indeed, if we start with some $y \in Y$, then by Proposition 1 (i), we have

$$\phi_1(y) = (x_1|y) = (x_0 - P_Y x_0|y) = 0,$$

thus proving the inclusion $\text{Ker } \phi \subset \text{Ker } \phi_1$. Since $\phi_1$ is not identically 0, and $\mathcal{H}/\text{Ker } \phi$ is one-dimensional, this inclusion forces the equality (41). Furthermore, using the same one-dimensional argument, it follows that there exists some $\zeta \in \mathbb{C} \setminus \{0\}$, such that $\phi = \zeta \phi_1$, which means that

$$\phi(y) = \zeta (x_1|y) = (\bar{\zeta} x_1|y), \quad \forall y \in Y,$$

thus proving (40) with $x = \bar{\zeta} x_1$.

For the uniqueness, assume $x' \in \mathcal{H}$ is another vector satisfying (40), and let us prove that $x' = x$. This is, however, trivial, since the equalities $(x|y) = \phi(y) = (x'|y)$ force

$$(x' - x|y) = 0, \quad \forall y \in \mathcal{H},$$

and this clearly implies $\|x' - x\|^2 = (x' - x|x' - x) = 0$.

(ii) $\Rightarrow$ (i). Assume $\phi$ is represented by (40), and let us prove that $\phi$ is continuous, and has $\|\phi\| = \|x\|$. The continuity is obvious by the Cauchy-Buniakowski-Schwartz Inequality, which yields:

$$|\phi(y)| = |(x|y)| \leq \|x\| \cdot \|y\|, \quad \forall y \in \mathcal{H},$$

thus also proving the inequality $\|\phi\| \leq \|x\|$. By the norm inequality, we also know that

$$|\phi(y)| \leq \|\phi\| \cdot \|y\|, \quad \forall y \in \mathcal{H},$$
so if we plug in \( y = x \) we get
\[
\|x\|^2 = (x|x) = |(x|x)| = |\phi(x)| \leq \|\phi\| \cdot \|x\|,
\]
which forces (at least in the case \( x \neq 0 \)) the inequality \( \|x\| \leq \|\phi\| \).

**Comment.** Riesz’ Theorem has many important applications, which will discussed throughout the entire chapter. Besides Exercise 2 below, a whole group of applications will appear in sub-section D.

**Exercise 2.** Prove that all Hilbert spaces are reflexive. This means that, if one considers the dual Banach space \( \mathcal{H}^* \), and if one defines, for every \( x \in \mathcal{H} \), the linear continuous functional \( x^\# : \mathcal{H}^* \ni \phi \mapsto \phi(x) \in \mathbb{C} \), then the map
\[
# : \mathcal{H} \ni x \mapsto x^\# \in (\mathcal{H}^*)^*
\]
is an isometric linear isomorphism.

**D. Orthogonality**

Riesz’ Theorem shows that, the inner products on Hilbert spaces enjoy properties similar to dual pairings (see DT I). To be a bit more precise, we can introduce the following “sesquilinear (hilbertian) analogues” of the polars and absolute polars, as indicated in the following Exercises.

**Exercises 3-4.** Let \( \mathcal{H} \) be a Hilbert space.

3. Given a non-empty set \( \mathcal{A} \subset \mathcal{H} \), define its **hilbertian polar** to be the set
\[
\mathcal{A}^{\circ,\text{H}} = \{ x \in \mathcal{A} : \text{Re}(x|a) \leq 1, \ \forall \ a \in \mathcal{A} \}.
\]

Prove the **hilbertian Bi-polar Theorem:**
\[
(\mathcal{A}^{\circ,\text{H}})^{\circ,\text{H}} = \text{conv}(\mathcal{A} \cup \{0\}).
\]

4. Given a non-empty set \( \mathcal{A} \subset \mathcal{H} \), define its **hilbertian absolute polar** to be the set
\[
\mathcal{A}^{\sqcap,\text{H}} = \{ x \in \mathcal{A} : \text{Re}(x|a) \leq 1, \ \forall \ a \in \mathcal{A} \}.
\]

Prove the **hilbertian Absolute Bi-polar Theorem:**
\[
(\mathcal{A}^{\sqcap,\text{H}})^{\sqcap,\text{H}} = \text{conv} (\text{bal} \mathcal{A}).
\]

Another construction from DT I was the **annihilator**. The hilbertian annihilator has a privileged role in the theory of Hilbert spaces, and therefore the above terminology and notation are slightly modified as follows.

**Definition.** Given a Hilbert space \( \mathcal{H} \), and some non-empty subset \( \mathcal{A} \subset \mathcal{H} \), we define the **orthogonal set to \( \mathcal{A} \)** to be:
\[
\mathcal{A}^\perp = \{ x \in \mathcal{H} : (x|a) = 0, \ \forall \ a \in \mathcal{A} \}.
\]
As a special case (when \(A\) is a singleton), given two vectors \(x, y \in \mathcal{H}\), we shall say that \(x\) is **orthogonal to** \(y\), if \((x|y) = 0\), in which case we write \(x \perp y\). Note that the relation \(\perp\) is symmetric.

The main properties of orthogonality are summarized below.

**Proposition 2.** Suppose \(\mathcal{H}\) is a Hilbert space.

(i) For any non-empty subset \(A \subset \mathcal{H}\), the set \(A^\perp\) is a closed linear subspace.

(ii) If \(A \subset B \subset \mathcal{H}\) are non-empty sets, then \(A^\perp \supset B^\perp\).

(iii) For any non-empty subset \(A\), one has the equality:
\[
\overline{\text{span}} A = (A^\perp)^\perp.
\]

(iv) For any closed linear subspace \(\mathcal{X}\), the linear subspace \(\mathcal{X}^\perp\) satisfies the equalities:

- \(\mathcal{X} \cap \mathcal{X}^\perp = \{0\}\);
- \(\mathcal{X} + \mathcal{X}^\perp = \mathcal{H}\).

**Proof.** (i). Since by definition we clearly have
\[
A^\perp = \bigcap_{x \in A} \{x\}^\perp,
\]
it suffices to show that \(\{x\}^\perp\) is a closed linear subspace, which is trivial by Riesz’ Theorem, since \(\{x\}^\perp = \text{Ker } \phi_x\), where \(\phi_x : \mathcal{H} \ni y \longmapsto (x|y) \in \mathbb{C}\) is known to be linear and continuous.

(ii). This follows immediately from (43).

(iii). From DT I we know that
\[
\overline{\text{span}} A = (A^{\text{ann}})^{\text{ann}},
\]
where the annihilator spaces are defined using the dual pairing between \(\mathcal{H}\) and its topological dual \(\mathcal{H}^*\) as:
\[
A^{\text{ann}} = \{\phi \in \mathcal{H}^* : \phi(a) = 0, \forall a \in A\},
\]
\[
(A^{\text{ann}})^{\text{ann}} = \{y \in \mathcal{H} : \phi(y) = 0, \forall \phi \in A^{\text{ann}}\}.
\]
Using Riesz’ Theorem, we know that the correspondence \(\Phi : \mathcal{H} \ni x \longmapsto \phi_x \in \mathcal{H}^*\) is a bijection. Using this bijection, it follows immediately that \(A^{\text{ann}} = \{\phi_x : x \in A^\perp\} = \Phi(A^\perp)\), which in turn means that
\[
(A^{\text{ann}})^{\text{ann}} = \{y \in \mathcal{H} : \phi_x(y) = 0, \forall x \in A^\perp\} = \{y \in \mathcal{H} : (x|y) = 0, \forall x \in A^\perp\} = (A^\perp)^\perp,
\]
and the desired equality (42) follows immediately from (44).

(iv). The equality \(\mathcal{X} \cap \mathcal{X}^\perp = \{0\}\) is trivial, since any vector \(x \in \mathcal{X} \cap \mathcal{X}^\perp\) is orthogonal to itself, i.e. \(\|x\|^2 = (x|x) = 0\).

To prove the second equality, we start with an arbitrary vector \(y \in \mathcal{H}\), and we consider the vectors \(y_1 = P_Xy\) and \(y_2 = y - y_1\). On the one hand, \(y_1\) clearly belongs to \(\mathcal{X}\). On the other hand, by the properties of the orthogonal projection \(P_X\), we also know that that \((y_2|x) = 0, \forall x \in \mathcal{X}\), which means that \(y_2 \in \mathcal{X}^\perp\), and we are done, since by construction \(y = y_1 + y_2\).  

\[\square\]
Corollary 1. If $\mathcal{H}$ is a Hilbert space, and $\mathcal{A}$ is a non-empty subset, then:
\[ \mathcal{A}^\perp = \left[ \overline{\text{span} \mathcal{A}} \right]^\perp. \]

Proof. Since the left-hand side is closed, we have: $\mathcal{A}^\perp = \left[ (\mathcal{A}^\perp)^\perp \right] = \left[ \overline{\text{span} \mathcal{A}} \right]^\perp$. \( \square \)

Corollary 2 (of the proof). Given a Hilbert space $\mathcal{H}$ and a closed linear subspace $\mathcal{X} \subset \mathcal{H}$, the orthogonal projections $P_\mathcal{X}$ and $P_{\mathcal{X}^\perp}$, satisfy the identity $P_\mathcal{X} + P_{\mathcal{X}^\perp} = I_\mathcal{H}$. In particular, Ker $P_\mathcal{X} = \mathcal{X}^\perp$.

Proof. Start with an arbitrary $y \in \mathcal{H}$, and consider the vectors $y_1 = P_\mathcal{X}y \in \mathcal{X}$ and $y_2 = y - y_1 \in \mathcal{X}^\perp$, constructed in the proof of (iv). Since $(\mathcal{X}^\perp)^\perp = \mathcal{X}$, by the same argument (with $\mathcal{X}^\perp$ in place of $\mathcal{X}$), we can consider the vector $z_1 = P_{\mathcal{X}^\perp}y$, so that the vector $z_2 = y - z_1$ belongs to $\mathcal{X}$. Now we have $y_1 + y_2 = y = z_2 + z_1$, with $y_1, z_2 \in \mathcal{X}$ and $y_2, z_1 \in \mathcal{X}^\perp$, so by the equality $\mathcal{X} \cap \mathcal{X}^\perp = \{0\}$ we must have $y_1 = z_2$ and $y_2 = z_1$. In particular, we have $(I_\mathcal{H} - P_\mathcal{X})y = y_2 = z_1 = P_{\mathcal{X}^\perp}y$, thus proving the desired identity $(I_\mathcal{H} - P_\mathcal{X}) = P_{\mathcal{X}^\perp}$.

The equality Ker $P_\mathcal{X} = \mathcal{X}^\perp$ follows now immediately from Proposition-Definition 1. \( \square \)

E. Geometry in Coordinates.

In this section we introduce the suitable objects that allow one to use coordinates, for geometric calculations in a Hilbert space.

Definition. Assume $\mathcal{H}$ is a pre-Hilbert space. A non-empty subset $\mathcal{S} \subset \mathcal{H}$ is said to be an orthogonal subset, if $x \neq 0$, $\forall x \in \mathcal{S}$, and furthermore,

(0) whenever $x, y \in \mathcal{S}$, $x \neq y$, it follows that $x \perp y$.

Remark 9. If $\mathcal{S}$ is orthogonal, then it is linearly independent. Indeed, if $x_1, \ldots, x_n \in \mathcal{S}$ are $n$ distinct vectors, and $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$ are such that $\sum_{k=1}^{n} \alpha_k x_k = 0$, then

\[ 0 = (x_j | \sum_{k=1}^{n} \alpha_k x_k) = \sum_{k=1}^{n} (x_j | \alpha_k x_k) = \alpha_j \|x_j\|^2, \]

which (by the condition $x_j \neq 0$) forces $\alpha_j = 0$, $\forall j = 1, \ldots, n$.

Proposition-Definition 3. If $\mathcal{H}$ is a Hilbert space and $\mathcal{S} \subset \mathcal{H}$ is an orthogonal subset, then the following are equivalent:

(i) $\mathcal{S}$ is maximally orthogonal, in the sense that there is no orthogonal subset $\mathcal{S}' \supsetneq \mathcal{S}$;

(ii) $\overline{\text{span} \mathcal{S}} = \mathcal{H}$.

Such a set $\mathcal{S}$ is called an orthogonal basis for $\mathcal{H}$.

Proof. (i) $\Rightarrow$ (ii). Assume $\mathcal{S}$ is maximally orthogonal. and let us prove the equality (ii). Argue by contradiction, assuming the closed linear subspace $\mathcal{X} = \overline{\text{span} \mathcal{S}}$ is proper, i.e. $\mathcal{X} \subsetneq \mathcal{H}$. By Proposition 2, it follows that $\mathcal{X}^\perp \neq \{0\}$, so if we take some vector $x \in \mathcal{X}^\perp \setminus \{0\}$,
then \( x \perp y, \forall y \in S \), so the set \( S' = S \cup \{x\} \supsetneq S \) is still orthogonal, thus contradicting the maximality of \( S \).

(ii) \( \Rightarrow \) (i). Assume condition (ii). In order to prove the maximality of \( S \), it suffices to prove the equality

\[
S^\perp = \{0\}. \tag{45}
\]

(Indeed, if (45) holds, then no other set \( S' \supsetneq S \) can be orthogonal, since we would have the inclusion \( S' \setminus S \subseteq S^\perp \) as well as the fact that \( S' \setminus S \) consists of non-zero vectors.) But the equality (45) is trivial, since by Corollary 1 we have \( S^\perp = [\text{span } S]^\perp = H^\perp = \{0\} \).

**Definitions.** Given a pre-Hilbert space \( H \), a non-empty subset \( S \subset H \) is said to be orthonormal, if

- \( S \) is orthogonal;
- \( \|x\| = 1, \forall x \in S \).

If \( H \) is a Hilbert space, we say that \( S \) is an orthonormal basis for \( H \), if \( S \) is both an orthonormal subset and an orthogonal basis for \( H \).

**Proposition 4.** Assume \( H \) is a Hilbert space, which is not zero-dimensional.

(i) For every orthonormal set \( S \subset H \), there exists at least one orthonormal basis \( B \) for \( H \), which contains \( S \).

(ii) Any two orthonormal bases for \( H \) have the same cardinality.

(iii) If \( H \) is finite dimensional, then all orthonormal bases \( B \) have \( \text{card } B = \dim \mathbb{C}H \).

**Proof.** (i). Fix \( S \). Consider the collection \( B \) of all orthonormal subsets \( B \supseteq S \). Remark now that, if \( B_0 \subset B \) is a sub-collection which is totally ordered by \( \supseteq \), then the set \( B_0 = \bigcup_{B \in B_0} B \) clearly belongs to \( B \), and is an upper bound for \( B_0 \). Therefore, by Zorn's Lemma, the collection \( B \) possesses a maximal element \( B \). By construction, \( B \) is orthonormal, and contains \( S \). By Proposition 3, in order to prove that \( B \) is an orthonormal basis, it suffices to show that \( B \) is maximally orthogonal, but this is pretty obvious, for if there exists \( x \neq 0 \), such that \( x \perp y, \forall y \in B \), then the vector \( x_0 = \|x\|^{-1}x \) will have the same property, so the orthonormal set \( B \cup \{x_0\} \supsetneq B \) would contradict the maximality of \( B \).

(iii). This statement is clear, because in the finite dimensional case all linear subspaces are automatically norm-closed, so any orthonormal basis \( B \) for \( H \) will satisfy \( \text{span } B = H \).

(ii). Fix two orthonormal bases \( A \) and \( B \), and let us show that \( \text{card } A \leq \text{card } B \). (By symmetry, the other inequality will also hold, so we will have in fact equality.) The case when \( B \) is finite is trivial by (iii). Assuming \( B \) is infinite, we consider the “rational” complex sub-field \( \mathbb{Q} + i\mathbb{Q} \subset \mathbb{C} \), and the linear span \( X = \text{span}_{\mathbb{Q} + i\mathbb{Q}} B \), consisting of linear combinations of vectors in \( B \), with coefficients in \( B \). On the one hand, since \( B \) is infinite and \( \mathbb{Q} + i\mathbb{Q} \) is countable, it follows immediately that \( \text{card } A = \text{card } B \). On the other hand, since \( \mathbb{Q} + i\mathbb{Q} \) is dense in \( \mathbb{C} \), and \( \text{span}_{\mathbb{C}} B = H \), it follows that \( X \) is dense in \( H \), so in particular, for every \( a \in A \), the set \( X_a = \{x \in X : \|x - a\| < \frac{1}{2}\} \) is non-empty. Let us remark now that, if \( a_1, a_2 \in A \) are distinct, then \( X_{a_1} \cap X_{a_2} = \emptyset \). Indeed, if there existed a vector \( x \in X_{a_1} \cap X_{a_2} \),
then the inequalities \( \|x - a_j\| < \frac{1}{2}, \ j = 1, 2 \) force \( \|a_1 - a_2\| < 1 \), which is impossible, since by the Law of Cosine we have
\[
\|a_1 - a_2\|^2 = \|a_1\|^2 + \|a_2\|^2 - 2\text{Re}(a_1 \bar{a}_2) = 2.
\]
Since all the sets \( \mathcal{X}_a, a \in \mathcal{A} \), are disjoint and nonempty, it follows that \( \text{CARD } \mathcal{A} \leq \text{CARD} \left( \bigcup_{a \in \mathcal{A}} \mathcal{X}_a \right) \leq \text{CARD } \mathcal{X} = \text{CARD } \mathcal{B} \), and we are done.

**Definition.** Given a Hilbert space \( \mathcal{H} \), the cardinality of one (thus of any) orthonormal basis of \( \mathcal{H} \) is called the **Hilbertian dimension of** \( \mathcal{H} \), and is denoted by \( \mathbb{H}\text{-dim } \mathcal{H} \).

**Theorem 4.** Let \( \mathcal{H} \) be a Hilbert space, and let \( \mathcal{B} \subset \mathcal{H} \) be an orthonormal subset, listed\(^2\) as \( \mathcal{B} = \{ e_i \}_{i \in I} \). Let \( \mathcal{X} = \text{SPAN } \mathcal{B} \).

(i) For an \( I \)-tuple \( \alpha = (\alpha_i)_{i \in I} \in \prod_I \mathbb{C} \), the following are equivalent:

- the \( I \)-tuple of vectors \( (\alpha_i e_i)_{i \in I} \) is summable in \( \mathcal{H} \);
- \( \alpha \) belongs to \( \ell^2(I) \), i.e. \( \sum_{i \in I} |\alpha_i|^2 < \infty \).

Moreover, in this case the vector \( x = \sum_{i \in I} \alpha_i e_i \) belongs to \( \mathcal{X} \), its its norm given by the equality
\[
\|x\|^2 = \sum_{i \in I} |\alpha_i|^2,
\]
and the coefficients \( \alpha_i \) are given by
\[
\alpha_i = (e_i | x), \ \forall i \in I.
\]

(ii) For any vector \( v \in \mathcal{H} \), the \( I \)-tuple \( (e_i | v)_{i \in I} \) belongs to \( \ell^2(I) \), and furthermore, the vector \( \sum_{i \in I} (e_i | v)e_i \in \mathcal{X} \) coincides with the orthogonal projection of \( v \) onto \( \mathcal{X} \), that is, one has the equality
\[
P_{\mathcal{X}}v = \sum_{i \in I} (e_i | v)e_i.
\]

**Proof.** Before we proceed with the proof of (i), let us introduce a few notations. Define, for every finite set of indices \( F \in \mathcal{P}_{\text{fin}}(I) \), the finite sum \( x_F = \sum_{i \in F} \alpha_i e_i \), and let us first remark that, since \( i \neq j \Rightarrow (e_i | e_j) = 0 \), we have
\[
\|x_F\|^2 = (x_F | x_F) = \sum_{(i, j) \in F \times F} (\alpha_i e_i | \alpha_j e_j) = \sum_{(i, j) \in F \times F} \bar{\alpha}_i \alpha_j (e_i | e_j) = \sum_{i \in F} |\alpha_i|^2, \ \forall F \in \mathcal{P}_{\text{fin}}(I).
\]

With this notation, the condition that the \( I \)-tuple \( (\alpha_i e_i)_{i \in I} \) is **summable** means that the net \( (x_F)_{F \in \mathcal{P}_{\text{fin}}(I)} \) is norm-convergent in \( \mathcal{H} \).

(i). Assume first that \( (\alpha_i e_i)_{i \in I} \) is summable, and let us prove that \( \sum_{i \in I} |\alpha_i|^2 < \infty \). This is, however, pretty obvious, since the condition \( x_F \rightarrow x \) forces \( \|x_F\| \rightarrow \|x\| \), which by (49) gives
\[
\|x\|^2 = \lim_{F \in \mathcal{P}_{\text{fin}}(I)} \sum_{i \in F} |\alpha_i|^2,
\]
\(^2\) This means that \( i \neq j \Rightarrow e_i \neq e_j \).
thus proving that $\alpha = (\alpha)_{i \in I}$ indeed belongs to $\ell^2(I)$, as well as the equality (46).

Conversely, let us assume now that $(\alpha)_{i \in I}$ belongs to $\ell^2(I)$, and let us show that the net $(x_F)_{F \in \mathcal{P}_{\text{fin}}(I)}$ is convergent. Let $s = \sum_{i \in I} |\alpha_i|^2$, which is equal to the limit of the net $(s_F)_{F \in \mathcal{P}_{\text{fin}}(I)}$ of partial sums, defined as $s_F = \sum_{i \in F} |\alpha_i|^2$. Since we work in a Banach space, it suffices to show that the net $(x_F)_{F \in \mathcal{P}_{\text{fin}}(I)}$ is Cauchy. Since the net $(s_F)_{F \in \mathcal{P}_{\text{fin}}(I)} \subset [0, \infty)$ is convergent, hence Cauchy, it follows that, for every $\varepsilon > 0$, there exists $G_\varepsilon \in \mathcal{P}_{\text{fin}}(I)$, such that $|s_{F_1} - s_{F_2}| < \varepsilon^2/4$, for all $F_1, F_2 \in \mathcal{P}_{\text{fin}}(I)$, with $F_1, F_2 \supseteq G_\varepsilon$. Now, if we fix two such sets $F_1$ and $F_2$, using (49) it follows that

$$
\|x_{F_j} - x_{G_\varepsilon}\|^2 = \|x_{F_j \setminus G_\varepsilon}\|^2 = s_{F_j \setminus G_\varepsilon} = s_{F_j} - s_{G_\varepsilon} < \frac{\varepsilon^2}{4}, \quad j = 1, 2,
$$

which immediately yields

$$
\|x_{F_1} - x_{F_2}\| \leq \|x_{F_1} - x_{G_\varepsilon}\| + \|x_{F_2} - x_{G_\varepsilon}\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
$$

The fact that the vector $\sum_{i \in I} \alpha_i e_i = \lim_F x_F$ belongs to $\mathcal{X}$ is trivial, since $x_F \in \text{SPAN}\{e_i\}_{i \in F} \subset \mathcal{X}$, $\forall F \in \mathcal{P}_{\text{fin}}(I)$.

To prove (47) we use Remark 8, by which it follows that:

$$
(e_i|x) = \lim_{F \in \mathcal{P}_{\text{fin}}(I)} (e_i|x_F).
$$

Since for fixed $i \in I$ and $F \in \mathcal{P}_{\text{fin}}(I)$, we have

$$(e_i|x_F) = \sum_{j \in F} (e_i|\alpha_j e_j) = \sum_{j \in F} \alpha_j (e_i|e_j) = \left\{ \begin{array}{ll} 0 & \text{if } F \neq i \\ \alpha_i & \text{ otherwise if } F \ni i \end{array} \right\}$$

and the desired equality (47) now follows immediately from (50).

(ii). Fix $v \in \mathcal{H}$, and let us prove first that the $I$-tuple $((e_i|v))_{i \in I}$ belongs to $\ell^2(I)$. Denote the coefficients $(e_i|v)$ simply by $\alpha_i$. Fix for the moment some finite set of indices $F \in \mathcal{P}_{\text{fin}}(I)$, and let us consider the vectors $w_F = \sum_{i \in F} \alpha_i e_i$ and $y_F = v - w_F$. Notice now that by construction, for every $i \in F$, we have

$$(e_i|w_F) = \sum_{j \in F} (e_i|\alpha_j e_j) = \sum_{j \in F} \alpha_j (e_i|e_j) = \alpha_i,$$

which means that $(e_i|v - w_F) = 0$, $\forall i \in F$, so in particular we also get

$$(w_F|y_F) = (w_F|v - w_F) = \sum_{i \in F} (\alpha_i e_i|v - w_F) = \sum_{i \in F} \alpha_i (e_i|v - w_F) = 0,$$

so by the Law of Cosine we get

$$
\|v\|^2 = \|w_F + y_F\|^2 = \|w_F\|^2 + \|y_F\|^2.
$$

Since by (i) we already know that $\|w_F\|^2 = \sum_{i \in F} |\alpha_i|^2$, the above calculation yields $\|v\|^2 = \sum_{i \in F} |\alpha_i|^2 + \|y_F\|^2$, so in particular it follows that

$$
\sum_{i \in F} |\alpha_i|^2 \leq \|v\|^2, \quad \forall F \in \mathcal{P}_{\text{fin}}(I),
$$

18
Consider now, using (i), the vector \( w = \sum_{i \in I} \alpha_i e_i \in X \). By (47), we know that \( (e_i|w) = \alpha_i = (e_i|v) \), which means that

\[
(e_i|v - w) = 0, \quad \forall i \in I.
\]

This means, of course that \( v - w \) belongs to \( B^\perp = X^\perp \). Now we have the decomposition \( v = w + (v - w) \) with \( w \in X \) and \( v - w \in X^\perp \), which implies the equality \( w = P_X v \).

**Corollary 3.** Using the same notations as in Theorem 4, assume \( B = \{e_i\}_{i \in I} \) is an orthonormal basis for the Hilbert space \( H \) with the entire space \( B \). Let

\[
\alpha \in H
\]

so the \( I \)-tuple \( (\alpha_i)_{i \in I} \) indeed belongs to \( \ell^2(I) \).

Using the notations from Theorem 4, in this case the space \( \alpha \) establishes an isometric linear isomorphism.

\[
\alpha \in H \ni x \mapsto \left( (e_i|x) \right)_{i \in I} \in \ell^2(I)
\]

establishes an isometric linear isomorphism.

**Proof.** Using the notations from Theorem 4, in this case the space \( X = \text{span} B \) coincides with the entire space \( H \), so the equality (48) simply reads:

\[
x = \sum_{i \in I} (e_i|x) e_i, \quad \forall x \in H.
\]

By Theorem 4 we also know that the map \( T: \ell^2(I) \ni (\alpha_i)_{i \in I} \mapsto \sum_{i \in I} \alpha_i e_i \in H \) is isometric (and trivially linear), and then everything follows from (47) and (51) which imply that \( T \) and \( U \) are inverses of each other.

**Exercise 5** Prove that, if \( H \) and \( H' \) are pre-Hilbert spaces, and \( U: H \to H' \) is a linear map. Prove that \( U \) is isometric, if and only if \( U \) preserves the inner products, i.e. \( (Ux|Uy) = (x|y), \forall x, y \in H \).

**Exercise 6.** Use the notations and assumptions from Theorem 4. Prove that, if \( \alpha = (\alpha_i)_{i \in I} \) and \( \beta = (\beta_i)_{i \in I} \) both belong to \( \ell^2(I) \), then the vectors \( x = \sum_{i \in I} \alpha_i e_i \) and \( y = \sum_{i \in I} \beta_i e_i \) satisfy the equality

\[
(x|y) = \sum_{i \in I} \overline{\alpha_i} \beta_i.
\]

**Conclusion.** Using the notations as above, and assuming that \( B = \{e_i\}_{i \in I} \) is an orthonormal basis for the Hilbert space \( H \), the equality (52) now reads:

\[
(x|y) = \sum_{i \in I} (\overline{e_i|x}) \cdot (e_i|y) = \sum_{i \in I} (x|e_i) \cdot (e_i|y), \quad \forall x, y \in H.
\]

As a special case, one obtains: \( \| x \|^2 = \sum_{i \in I} |(e_i|x)|^2, \forall x \in H \). The formulas (53) and (51) are known as Parseval Identities.

**Exercise 7** Let \( H \) be a Hilbert space, let \( B = \{e_i\}_{i \in I} \) be an orthonormal subset, and let \( X = \text{span} B \). Prove the generalized Parseval Identity:

\[
(P_X x|y) = (x|P_X y) = \sum_{i \in I} (e_i|x) \cdot (e_i|y) = \sum_{i \in I} (x|e_i) \cdot (e_i|y), \quad \forall x, y \in H.
\]