Chapter 7

Hilbert Spaces

7.1 Basic Properties

Hilbert spaces form a special class of Banach spaces with the geometric notion of orthogonality of vectors, or more generally, the notion of an angle between vectors, built into them.

Consider the space $\mathbb{R}^2$. If $x = (x_1, x_2)$ and $y = (y_1, y_2)$ are vectors in $\mathbb{R}^2$, then we define the scalar product of these vectors by

$$x.y \;=\; x_1y_1 + x_2y_2 \;=\; |x|.|y|\cos\theta$$

where $|x| = \|x\|_2,\ |y| = \|y\|_2$ and θ is the angle between the two vectors. The scalar product is linear in each of the two variables. It is symmetric in these variables and $x.x = \|x\|_2^2$. It turns out that these properties are crucial and we generalize these to other vector spaces.

Definition 7.1.1 *Let V be a real normed linear space. An* **inner product** *on V is a form $(.,.) : V \times V \to \mathbb{R}$ such that*
(i) *it is symmetric, i.e. for all x and $y \in V$,*

$$(x, y) \;=\; (y, x);$$

(ii) *it is bilinear: in particular, if x, y and $z \in V$ and if α and $\beta \in \mathbb{R}$, then*

$$(\alpha x + \beta y, z) \;=\; \alpha(x, z) + \beta(y, z);$$

(iii) *for all $x \in V$,*

$$(x, x) \;=\; \|x\|^2. \blacksquare$$

Remark 7.1.1 The linearity with respect to the second variable is, clearly, a consequence of conditions (i) and (ii) above. ■

Remark 7.1.2 In case the base field is $\mathbb{C}$, then the inner product is a *sesquilinear form*. If x and $y \in V$, we have

$$(y,x) = \overline{(x,y)}.$$

Thus, we have conjugate linearity with respect to the second variable, *i.e.* if x, y and $z \in V$ and if α and $\beta \in \mathbb{C}$, then

$$(x, \alpha y + \beta z) = \overline{\alpha}(x,y) + \overline{\beta}(x,z). \blacksquare$$

In view of condition (iii), we say that the norm comes from the inner product.

Definition 7.1.2 *A* **Hilbert space** *is a Banach space whose norm comes from an inner product.* ■

Example 7.1.1 Consider the space $\mathbb{R}^n$. For $x = (x_1, \cdots, x_n)$ and $y = (y_1, \cdots, y_n)$, we define

$$(x,y) = \sum_{i=1}^{n} x_i y_i.$$

This defines an inner product and the norm associated to it is the norm $\|.\|_2$. Thus ℓ_2^n is a Hilbert space. In the case of $\mathbb{C}^n$, the inner product is given by

$$(x,y) = \sum_{i=1}^{n} x_i \overline{y_i}.$$

Again the norm is $\|.\|_2$. ■

Example 7.1.2 Consider the space ℓ_2. For x and $y \in \ell_2$, define

$$(x,y) = \sum_{i=1}^{\infty} x_i y_i$$

where $x = (x_i)$ and $y = (y_i)$ are real sequences. Again, if the base field is $\mathbb{C}$, then we define

$$(x,y) = \sum_{i=1}^{\infty} x_i \overline{y_i}.$$

This makes ℓ_2 into a Hilbert space. ■

As we have seen in the previous chapter, these are particular cases of the Lebesgue spaces L^2.

Example 7.1.3 Let $(X, \mathcal{S}, \mu)$ be a measure space. If f and $\mathsf{g} \in L^2(\mu)$, and if f and g represent these classes respectively, then

$$(\mathsf{f}, \mathsf{g}) = \int_X fg \, d\mu$$

defines an inner product which makes $L^2(\mu)$ as a Hilbert space. ■

Example 7.1.4 Let $(a, b) \subset \mathbb{R}$ be a finite interval. We denote by $H^1(a, b)$ the space $W^{1,2}(a, b)$ and by $H_0^1(a, b)$ the space $W_0^{1,2}(a, b)$. Then both these spaces are Hilbert spaces with the inner product given by

$$(\mathsf{f}, \mathsf{g}) = \int_a^b (fg + f'g') \, dx.$$

By virtue of the Poincaré inequality (cf. Theorem 6.4.6), the space $H_0^1(a, b)$ is also a Hilbert space with the inner product

$$(\mathsf{f}, \mathsf{g})_1 = \int_a^b f'g' \, dx. \blacksquare$$

Let H be a Hilbert space and let x and $y \in H$. Then

$$\|x + y\|^2 = (x + y, x + y) = \|x\|^2 + 2(x, y) + \|y\|^2 \tag{7.1.1}$$

in the real case; if the field is $\mathbb{C}$ then the middle term on the right will be replaced by $2\text{Re}(x, y)$, where Re z denotes the real part of a complex number z. Writing a similar expression for $\|x - y\|^2$ and adding the two, we get

$$\left\|\frac{1}{2}(x + y)\right\|^2 + \left\|\frac{1}{2}(x - y)\right\|^2 = \frac{1}{2}(\|x\|^2 + \|y\|^2). \tag{7.1.2}$$

This is known as the *parallelogram identity.* In case of $\mathbb{R}^2 = \ell_2^2$, this is the familiar result from plane geometry which relates the sum of the squares of the lengths of the diagonals of a parallellogram to that of the sides. It is also known as *Apollonius' theorem.*

Remark 7.1.3 A theorem of Fréchet, Jordan and von Neumann states that a Banach space whose norm satisfies the parallelogram identity

(7.1.2) is a Hilbert space, *i.e.* the norm comes from an inner product. ■

Example 7.1.5 The space $\mathcal{C}[-1,1]$ cannot be made into a Hilbert space. To see this, consider the functions

$$u(x) = \min\{x, 0\}, \text{ and } v(x) = x$$

defined on $[-1,1]$. Then $\|u\| = \|v\| = 1$ while we have

$$\left\|\frac{1}{2}(u+v)\right\| = 1 \text{ and } \left\|\frac{1}{2}(u-v)\right\| = \frac{1}{2}$$

and the parallelogram identity is not satisfied by this pair of functions. ■

Proposition 7.1.1 *Every Hilbert space is uniformly convex and hence , is reflexive.*

Proof: The proof is of the uniform convexity follows from the parallelogram identity (7.1.2) exactly as described in Example 5.5.2; the reflexivity now follows from Theorem 5.5.1. ■

We now prove a fundamental ineqaulity for Hilbert spaces.

Theorem 7.1.1 (**Cauchy-Schwarz Inequality**) *Let H be a Hilbert space and let x and $y \in H$. Then*

$$|(x,y)| \leq \|x\| \; \|y\|. \tag{7.1.3}$$

Equality occurs in this inequality if, and only if, x and y are scalar multiples of each other.

Proof: Let θ be a complex number such that $|\theta| = 1$ and $\theta(x,y) = |(x,y)|$. Let $t \in \mathbb{R}$. We have

$$\begin{aligned} 0 &\leq \|\theta x - ty\|^2 \\ &= \|x\|^2 - 2t\mathrm{Re}(\theta x, y) + t^2\|y\|^2 \\ &= \|x\|^2 - 2t|(x,y)| + t^2\|y\|^2. \end{aligned}$$

Since we have a quadratic polynomial which is always of constant sign, the roots of this polynomial must be complex. Thus, we deduce that

$$4|(x,y)|^2 \leq 4\|x\|^2\|y\|^2$$

which yields (7.1.3).

Equality occurs in (7.1.3) if, and only if, the polynomial has two coincident roots. Thus, there exists t_0 such that $\theta x = t_0 y$, or, in other words $x = \alpha y$ where $\alpha = \theta^{-1} t_0$. ■

Corollary 7.1.1 *Let H be a Hilbert space. Let $y \in H$. Define*

$$f_y(x) = (x, y)$$

for all $x \in H$. Then $f_y \in H^*$ and $\|f_y\| = \|y\|$.

Proof: Clearly f_y is a linear functional. By the Cauchy-Schwarz inequality, we have

$$|f_y(x)| \leq \|x\|.\|y\|$$

which shows that $f_y \in H^*$ and that $\|f_y\| \leq \|y\|$. If $y \neq \mathbf{0}$, then set $x = y/\|y\|$. Then $f_y(x) = \|y\|$, which shows that $\|f_y\| = \|y\|$. ■

Remark 7.1.4 We will see in the next section that all continuous linear functionals on a Hilbert space occur in this manner. ■

Corollary 7.1.2 *Let H be a Hilbert space and let $x_n \rightharpoonup x$ and $y_n \to y$ in H. Then*

$$(x_n, y_n) \to (x, y).$$

Proof: Observe that

$$\begin{aligned} |(x_n, y_n) - (x, y)| &\leq |(x_n, y_n - y)| + |(x_n - x, y)| \\ &\leq \|x_n\|.\|y_n - y\| + |f_y(x_n - x)| \end{aligned}$$

by the Cauchy-Schwarz inequality and the preceding corollary. Now, since any weakly converging sequence is bounded and since $y_n \to y$, the first term on the right-hand side tends to zero. The second term also tends to zero by virtue of the preceding corollary, since $x_n \rightharpoonup x$ in H. ■

Remark 7.1.5 Since norm convergence implies weak convergence, it follows *a fortiori* that if $x_n \to x$ and $y_n \to y$ in H, then $(x_n, y_n) \to (x, y)$. ■

Theorem 7.1.2 *Let H be a Hilbert space and let $K \subset H$ be a closed and convex subset of H. Then, for every $x \in H$, there exists a unique element $P_K(x) \in K$ such that*

$$\|x - P_K(x)\| = \min_{y \in K} \|x - y\|. \tag{7.1.4}$$

Further, if H is a real Hilbert space, then $P_K(x) \in K$ is characterized by the following relations:

$$(x - P_K(x), y - P_K(x)) \ \leq \ 0 \qquad (7.1.5)$$

for every $y \in K$.

Proof: Since H is uniformly convex, the existence and uniqueness of $P_K(x)$ has been proved in Theorem 5.6.1. Let $y \in K$. For any $0 < t < 1$, set $z = (1-t)P_K(x) + ty$ which belongs to K by convexity. Then, by virtue of (7.1.4),

$$\|x - P_K(x)\| \ \leq \ \|x - z\| \ = \ \|(x - P_K(x)) - t(y - P_K(x))\|.$$

Squaring both sides, we get

$$\|x-P_K(x)\|^2 \ \leq \ \|x-P_K(x)\|^2 - 2t(x-P_K(x), y-P_K(x)) + t^2\|y-P_K(x)\|^2.$$

Cancelling the common term *viz.* $\|x - P_K(x)\|^2$, dividing throughout by t and letting $t \to 0$, we get (7.1.5).

Conversely, if $P_K(x) \in K$ is an element satisfying (7.1.5), then, for any $y \in K$, we have

$$\begin{array}{rcl} \|x - P_K(x)\|^2 & = & \|(x-y) + (y - P_K(x)\|^2 \\ & = & \|x-y\|^2 + 2(x-y, y-P_K(x)) + \|y - P_K(x)\|^2 \\ & = & \|x-y\|^2 + 2(x - P_K(x), y - P_K(x)) - \|y - P_K(x)\|^2 \\ & \leq & \|x-y\|^2. \end{array}$$

Thus $P_K(x)$ also satisfies (7.1.4). ■

Remark 7.1.6 If H is a complex Hilbert space, then $(x - P_K(x), y - P_K(x))$ is replaced by its real part in (7.1.5). ■

Remark 7.1.7 The element $P_K(x)$, which is closest to x in K, is called the *projection of x onto K*. In general, the mapping $x \mapsto P_K(x)$ is *not* linear. In $\mathbb{R}^2$, the condition (7.1.5) means that, for all $y \in K$, the lines joining x to $P_K(x)$ and y to $P_K(x)$ will always make an obtuse angle. ■

We now study some properties of the mapping $P_K : H \to K$.

Proposition 7.1.2 *Let H be a Hilbert space and let K be a closed and convex subset of H. Let $P_K : H \to K$ be as defined by the preceding theorem. Then, for all x and $y \in H$, we have*

$$\|P_K(x) - P_K(y)\| \ \leq \ \|x - y\|.$$

Proof: Assume, for simplicity, that H is a real Hilbert space. By virtue of (7.1.5), we have

$$(x - P_K(x), P_K(y) - P_K(x)) \leq 0$$

and

$$(y - P_K(y), P_K(x) - P_K(y)) \leq 0.$$

Adding these two inequalities, we get

$$(x - y, P_K(y) - P_K(x)) + \|P_K(y) - P_K(x)\|^2 \leq 0.$$

Thus,

$$\|P_K(y) - P_K(x)\|^2 \leq (y - x, P_K(y) - P_K(x))$$

and the result now follows from the Cauchy-Schwarz inequality being applied to the term on the right-hand side. ■

Corollary 7.1.3 *Let M be a closed subspace of a Hilbert space H. Then the projection P_M is a continuous linear mapping. Further, for $x \in H$, the element $P_M(x) \in M$ is characterized by*

$$(P_M(x), y) = (x, y) \tag{7.1.6}$$

for every $y \in M$.

Proof: If (7.1.6) holds, then (7.1.5) holds trivially. Conversely, if $P_M(x) \in M$ is the projection of x onto M, then $P_M(x)$ satisfies (7.1.5). Let $y \in M$. Set $z = y + P_M(x) \in M$, since M is a subspace. Then (7.1.5) yields

$$(x - P_M(x), y) \leq 0$$

for all $y \in M$. Since we also have $-y \in M$, we get the reverse inequality as well and this proves (7.1.6). It now follows from (7.1.6) that P_M is a linear map and it is continuous by the preceding proposition. This completes the proof. ■

Remark 7.1.8 If M is a closed subspace of a Hilbert space H, the vector $x - P_M(x)$ is orthogonal to every vector in M. Thus, P_M is called the **orthogonal projection** of H onto M. ■

Theorem 7.1.3 *Let H be a Hilbert space and let M be a closed subspace. then M is complemented in H.*

Proof: Set

$$M^\perp = \{y \in H \mid (x,y) = 0 \ \text{ for all } x \in M\}.$$

It is immediate to check that $M^\perp$ is a subspace. It is also closed. For, let $\{y_n\}$ be a sequence in $M^\perp$ and let $y_n \to y$ in H. If $x \in M$ is arbitrary, then since $(x, y_n) = 0$ for all n, we get that $(x, y) = 0$ as well and so $y \in M^\perp$ which establishes our claim. If $x \in H$, then $P_M(x) \in M$ and $x - P_M(x) \in M^\perp$ by the preceding corollary. Thus $H = M + M^\perp$. Further, if $x \in M \cap M^\perp$, we then have that $\|x\|^2 = (x,x) = 0$ and so $M \cap M^\perp = \{\mathbf{0}\}$. Thus $H = M \oplus M^\perp$ and the proof is complete. ■

Remark 7.1.9 The subspace $M^\perp$ consisting of all vectors orthogonal to all elements of M is called the **orthogonal complement** of M. The notation is not accidental. We will see in the next section that (at least in the case of real Hilbert spaces) the orthogonal complement can be identified with the annihilator of M. ■

7.2 The Dual of a Hilbert Space

Earlier, we saw that every vector in a Hilbert space gave rise to a continuous linear functional. The main result of this section is to show that all continuous linear functionals arise in this way.

Theorem 7.2.1 (**Riesz Representation Theorem**) *Let H be a Hilbert space. Let $\varphi \in H^*$. Then, there exists a unique vector $y \in H$ such that*

$$\varphi(x) = (x,y) \tag{7.2.1}$$

for all $x \in H$. Further, $\|\varphi\| = \|y\|$.

Proof: We saw that (cf. Corollary 7.1.1), given $y \in H$, the functional f_y defined by

$$f_y(x) = (x,y)$$

is in H^* and that $\|f_y\| = \|y\|$. Thus, the mapping $\Phi : H \to H^*$ defined by $\Phi(y) = f_y$ is an isometry of H into H^* and so its image is closed in H^*. If we show that the image is dense in H^*, then it will follow that $H^* = \Phi(H)$, or, in other words, that Φ is onto and this will complete the proof.

Consider a linear functional φ on H^* which vanishes on $\Phi(H)$. Since every Hilbert space is uniformly convex and hence, reflexive, this means that there exists $x \in H$ such that $f_y(x) = 0$ for all $y \in H$. This implies that $(x, y) = 0$ for all $y \in H$. In particular,

$$\|x\|^2 = (x, x) = 0$$

which shows that $x = \mathbf{0}$, *i.e.* φ is zero. This shows that $\Phi(H)$ is dense in H^* and the proof is complete. ■

Remark 7.2.1 It is also possible to directly prove this theorem without using the reflexivity of H. This will be outlined in the exercises at the end of this chapter. ■

Remark 7.2.2 Let H be a Hilbert space. Then every element of the dual, H^*, can be represented as f_x, where $x \in H$. We can then define an inner product on H^* by

$$(f_x, f_y)_* = (y, x).$$

It is easy to see that this defines an inner product which gives rise to the usual norm on H^*. Thus, H^* also becomes a Hilbert space in its own right. In the same way, H^{**} also becomes a Hilbert space. Now, we have two natural mappings from H into H^{**}. The first is the usual canonical imbedding $x \mapsto J(x)$. The second is the mapping $x \mapsto f_{f_x}$, *i.e.* the composition of the Riesz map $H \to H^*$ and that of $H^* \to H^{**}$. We will show that these are the same. The latter map, by the Riesz representation theorem, is onto and so J will be onto, giving another proof of the reflexivity of a Hilbert space, provided we prove the Riesz representation theorem independently, as suggested in Remark 7.2.1. To see that the maps are the same, observe that if $f = f_y \in H^*$, then

$$\begin{array}{rcccccc} f_{f_x}(f) & = & (f, f_x)_* & = & (f_y, f_x)_* & = & (x, y) \\ J(x)(f) & = & f(x) & = & f_y(x) & = & (x, y). \end{array}$$

This establishes the claim. ■

Remark 7.2.3 As a consequence of the Riesz representation theorem, the map $y \mapsto f_y$ is an isometry of H onto H^*. It is linear if H is real and conjugate linear if H is complex. Thus, at least in the real case, we can identify a Hilbert space with its own dual via the Riesz isometry. ■

Remark 7.2.4 In the case of real Hilbert spaces, while we can identify a Hilbert space with its dual, we have to be careful in doing so and we cannot do it to *every* space under consideration at a time. A typical example of such a situation is the following. Let V and H be real Hilbert spaces. Let $V \subset H$ and let V be dense in H. Let us assume further that there exists a constant $C > 0$ such that

$$\|v\|_H \ \leq \ C\|v\|_V$$

for every $v \in V$.

Let us now identify H^* with H via the Riesz representation theorem. Let $f \in H$. Then the map $v \mapsto (v, f)_H$ defines a continuous linear functional on V since

$$|(v, f)_H| \ \leq \ \|v\|_H \|f\|_H \ \leq \ C\|v\|_V \|f\|_H$$

for all $v \in V$. Let us denote this linear functional by $T(f)$. Thus $T \in \mathcal{L}(H, V^*)$ and

$$\|T(f)\|_{\mathcal{L}(H,V^*)} \ \leq \ C\|f\|_H.$$

If $T(f) = \mathbf{0}$, then $(v, f) = 0$ for all $v \in V$ and so, by density, we have $f = \mathbf{0}$. Thus T is one-one as well. Finally, we claim that the image of T is dense in V^*. Indeed, if $\varphi \in V^{**}$ vanishes on $T(H)$, then, by reflexivity, there exists $v \in V$ such that $T(f)(v) = 0$ for all $f \in H$ *i.e.* $(v, f) = 0$ for all $f \in H$. Since $V \subset H$, it follows that $(v, v) = 0$, *i.e.* $v = \mathbf{0}$, which means that φ is idertically zero, which establishes the claim.

Thus we have the following scheme:

$$V \ \subset \ H \ \cong \ H^* \ \subset \ V^*$$

where both the inclusions are dense. It would now be clearly absurd for us to identify V with V^* as well. Thus we cannot simultaneously identify V and H with their respective duals. The space H in this case is called the *pivot space* and is identified with its dual, whereas the other spaces, though they are also Hilbert spaces, will not be identified with their respective duals. This situation typically arises when we have a parametrized family of Hilbert spaces as in the case of the ***Sobolev spaces*** (cf. Kesavan [3]). In particular, we can set $V = H_0^1(a, b)$ (cf. Example 7.1.4) and $H = L^2(a, b)$. We identify $L^2(a, b)$ with its dual while we denote the dual of $H_0^1(a, b)$ by $H^{-1}(a, b)$ and we have the inclusions

$$H_0^1(a, b) \ \subset \ L^2(a, b) \ \cong \ (L^2(a, b))^* \ \subset \ H^{-1}(a, b). \ \blacksquare$$

Let H be a Hilbert space and let $A \in \mathcal{L}(H)$. For a fixed $y \in H$, the map $x \mapsto (A(x), y)$ clearly defines a continuous linear functional on H and so, by the Riesz representation theorem, this functional can be written as the inner product of x with a vector (which depends on y). This leads us to the following definition.

Definition 7.2.1 *Let H be a Hilbert space and let $A \in \mathcal{L}(H)$. We define the* **adjoint** *of A as the mapping $A^* : H \to H$ given by*

$$(x, A^*(y)) \;=\; (A(x), y) \tag{7.2.2}$$

for all x and $y \in H$. ■

Remark 7.2.5 In the case of real Hilbert spaces, since H and H^* can be identified via the Riesz isometry, the map A^* is just the adjoint in the sense of Definition 4.7.2. ■

The following proposition lists the propoerties of the adjoint map.

Proposition 7.2.1 *Let H be a Hilbert space. Let A_i, $i = 1, 2$ and A be continuous linear operators on H. Let α be a scalar. Then*
(i) $\|A\| = \|A^*\|$;
(ii) $\|A^*A\| = \|A\|^2$;
(iii) $A^{**} = A$;
(iv) $(A_1 + A_2)^* = A_1^* + A_2^*$;
(v) $(A_1A_2)^* = A_2^*A_1^*$;
(vi) $(\alpha A)^* = \overline{\alpha}A^*$ *(to be interpreted as αA^* in the real case).*

Proof: By the Cauchy-Schwarz inequality, we have for any $x \in H$,

$$\|x\| \;=\; \sup_{\|y\| \leq 1} |(x, y)|.$$

It then follows that if $A \in \mathcal{L}(H)$, then

$$\|A\| \;=\; \sup_{\|x\| \leq 1} \|A(x)\| \;=\; \sup_{\|x\| \leq 1} \sup_{\|y\| \leq 1} |(A(x), y)|.$$

Then, using (7.2.2) and the Cauchy-Schwarz inequality, we get

$$\|A\| \;=\; \sup_{\|x\| \leq 1} \sup_{\|y\| \leq 1} |(A(x), y)| \;=\; \sup_{\|x\| \leq 1} \sup_{\|y\| \leq 1} |(x, A^*(y))| \;\leq\; \|A^*\|.$$

Similarly,

$$\|A^*\| = \sup_{\|y\|\leq 1}\sup_{\|x\|\leq 1}|(x, A^*(y))| = \sup_{\|y\|\leq 1}\sup_{\|x\|\leq 1}|(A(x), y)| \leq \|A\|.$$

This proves (i). Again, if x and $y \in H$, then

$$|(A^*A(x), y)| = |(A(x), A(y))| \leq \|A\|^2\|x\|.\|y\|$$

from which, we deduce that

$$\|A^*A\| \leq \|A\|^2.$$

On the other hand,

$$\|A(x)\|^2 = (A(x), A(x)) = (A^*A(x), x) \leq \|A^*A\|.\|x\|^2$$

by the Cauchy-Schwarz inequality and we deduce that

$$\|A\|^2 \leq \|A^*A\|.$$

This proves (ii). The other relations follow trivially from (7.2.2). ■

Remark 7.2.6 A Banach algebra, B, is said to be a **-algebra* if there exists a mapping $x \mapsto x^*$ from B into itself satisfying the properties analogous to (iii) - (vi) of the above proposition. Such a mapping is said to be an *involution*. If, in addition, properties (i) and (ii) are also true, it is said to be a *B^*-algebra*. Thus, if H is a Hilbert space, the $\mathcal{L}(H)$ is a B^*-algebra with the involution being given by the adjoint mapping. ■

Definition 7.2.2 *Let H be a Hilbert space and let $A \in \mathcal{L}(H)$. A is said to be* **self-adjoint** *if $A^* = A$. It is said to be* **normal** *if $AA^* = A^*A$. It is said to be* **unitary** *if $AA^* = A^*A = I$, where I is the identity operator on H.* ■

Example 7.2.1 Any orthogonal projection in a Hilbert space is self-adjoint. If $P : H \to M$ is the orthogonal projection of a Hilbert space H onto a closed subspace M, then, for any x and $y \in H$, we have

$$(P^*(x), y) = (x, P(y)) = (P(x), P(y)) = (P(x), y)$$

by repeated application of Corollary 7.1.3. Since x and y are arbitrary elements of H, it follows that $P = P^*$. ■

Example 7.2.2 In ℓ_2^n, the operator defined by a hermetian matrix is self-adjoint, that defined by a normal matrix is normal and that defined by a unitary matrix (orthogonal matrix, if the base field is $\mathbb{R}$) is unitary (cf. Definition 1.1.14). ■

Remark 7.2.7 If $A : D(A) \subset H \to H$ is a densely defined linear transformation in a Hilbert space H, it is easy to see how to define the adjoint $A^* : D(A^*) \subset H \to H$. Again, we have for $u \in D(A)$ and $v \in D(A^*)$,

$$(A(u), v) \;=\; (u, A^*(v)).$$

All the results of Section 4.7, in particular, Proposition 4.7.3 and Theorem 4.7.1, are true. ■

7.3 Application: Variational Inequalities

Let H be a real Hilbert space and let $a(.,.) : H \times H \to \mathbb{R}$ be a continuous bilinear form (cf. Example 4.7.4). Let $M > 0$ such that

$$|a(x,y)| \;\le\; M\|x\|.\|y\| \tag{7.3.1}$$

for all x and $y \in H$. Assume further that $a(.,.)$ is H-elliptic (or, coercive; cf. Exercise 5.16). Let $\alpha > 0$ such that

$$a(x,x) \;\ge\; \alpha\|x\|^2 \tag{7.3.2}$$

for all $x \in H$.

Example 7.3.1 The inner product of a real Hilbert space is a symmetric, continuous and coercive bilinear form. Conversely, if $a(.,.)$ is a symmetric, continuous and coercive bilinear form, then

$$(x,y)_a \;\stackrel{\text{def}}{=}\; a(x,y)$$

defines a new inner product on H. The associated norm is

$$\|x\|_a \;=\; \sqrt{a(x,x)}.$$

Thanks to the continuity and coercivity of the bilinear form, we have

$$\sqrt{\alpha}\|x\| \;\le\; \|x\|_a \;\le\; \sqrt{M}\|x\|.$$

Thus the two norms on H are equivalent. ■

Example 7.3.2 Let $H = \ell_2^n$. Let $\mathbf{A}$ be an $n \times n$ matrix. If $\mathbf{x}$ and $\mathbf{y} \in \mathbb{R}^n = \ell_2^N$ are vectors, define

$$a(x,y) \;=\; \mathbf{y}'\mathbf{A}\mathbf{x}$$

where $\mathbf{y}'$ is the transpose of the column vector $\mathbf{y}$. Then $a(.,.)$ defines a continuous bilinear form on ℓ_2^n. If $\mathbf{A}$ is symmetric, then the bilinear form is symmetric as well. If $\mathbf{A}$ is positive definite, then the bilinear form is coercive. ■

Theorem 7.3.1 *(***Stampacchia's Theorem***) Let H be a real Hilbert space and let $a(.,.) : H \times H \to \mathbb{R}$ be a continuous and coercive bilinear form on H (satisfying (7.3.1) and (7.3.2)). Let K be a closed and convex subset of H. Let $f \in H$. Then, there exists a unique $x \in K$ such that, for all $y \in K$,*

$$a(x, y-x) \;\geq\; (f, y-x). \tag{7.3.3}$$

Proof: Let $u \in H$ be fixed. The map $v \mapsto a(u,v)$ is a continuous linear functional on H, by the continuity of the bilinear form. Thus, by the Riesz representation theorem, there exists $A(u) \in H$ such that

$$(A(u), v) \;=\; a(u,v)$$

for all $v \in H$. Clearly, the map $u \mapsto A(u)$ is linear. Further, by (7.3.1) and (7.3.2) we have

$$\|A(u)\| \;\leq\; M\|u\| \text{ and } (A(u), u) \;\geq\; \alpha\|u\|^2$$

for all $u \in H$. Thus $A \in \mathcal{L}(H)$. Now, (7.3.3) is equivalent to finding $x \in K$ such that

$$(A(x), y-x) \;\geq\; (f, y-x)$$

for all $y \in K$. If $\rho > 0$ is any constant (to be determined suitably), this is equivalent to finding $x \in K$ such that

$$(-\rho A(x) + \rho f + x - x, y - x) \;\leq\; 0$$

for all $y \in K$. In other words (cf. Theorem 7.1.2),

$$x \;=\; P_K(x - \rho A(x) + \rho f) \;\stackrel{\text{def}}{=}\; S(x).$$

Thus, we seek a fixed point of the mapping $S : K \to K$. Let x_1 and $x_2 \in K$. Then, by Proposition 7.1.2, we have

$$\|S(x_1) - S(x_2)\| \leq \|x_1 - x_2 - \rho(A(x_1) - A(x_2))\|.$$

Squaring both sides, we get

$$\begin{aligned}\|S(x_1) - S(x_2)\|^2 &\leq \|x_1 - x_2\|^2 - 2\rho(x_1 - x_2, A(x_1) - A(x_2)) \\ &\quad +\rho^2\|A(x_1) - A(x_2)\|^2 \\ &\leq (1 - 2\rho\alpha + \rho^2 M^2)\|x_1 - x_2\|^2\end{aligned}$$

using (7.3.1) and (7.3.2). Now, choosing ρ such that

$$0 < \rho < \frac{2\alpha}{M^2},$$

we have $1 - 2\rho\alpha + \rho^2 M^2 < 1$ so that $S : K \to K$ is a contraction. Since K is closed, by the contraction mapping theorem (cf. Theorem 2.4.1) we deduce that there exists a unique fixed point $x \in K$ for S which completes the proof. ■

Corollary 7.3.1 (**Lax-Milgram Lemma**) *Let H be a Hilbert space and let $a(.,.) : H \times H \to \mathbb{R}$ be a continuous and coercive bilinear form. Let $f \in H$. Then, there exists a unique $x \in H$ such that*

$$a(x, y) = (f, y)$$

for every $y \in H$.

Proof: Applying the preceding theorem with $K = H$, there exists a unique $x \in H$ satisfying (7.3.3). Replacing y by $y + x$, we get

$$a(x, y) \geq (f, y)$$

for every $y \in H$. Since $-y \in H$ as well, we also get the reverse inequality. Hence the result. ■

Remark 7.3.1 The Lax-Milgram lemma was already proved in Exercise 5.16. If, in addition $a(.,.)$ is symmetric, then the preceding results have been proved via Exercise 5.17. In that case, the solution x has a variational characterization, *viz.* $x \in K$ is the minimizer of the functional

$$J(y) = \frac{1}{2}a(y, y) - (f, y)$$

over K. For this reason, (7.3.3) is called a *variational inequality.* In the terminology of the calculus of variations, (7.3.3) is the equivalent of the *Euler-Lagrange* condition for the minimization of a functional. In the case of *unconstrained minimization i.e.* $K = H$, this becomes an *equation* instead of an inequality, as seen in the Lax-Milgram Lemma, and corresponds to the vanishing of the 'first variation' of J (cf. Kesavan [4]).

Indeed, it is easy to see that J is Fréchet differentiable (cf. Exercise 2.38) and that

$$J'(x)(y) \;=\; a(x,y) - (f,y)$$

for any x and $y \in H$. Thus (7.3.3) and the Lax-Milgram lemma are just the results of Exercise 2.44 when $a(.,.)$ is symmetric.

The Lax-Milgram lemma forms the basis of a wide class of numerical methods, known as *finite element methods*, to solve boundary value problems for elliptic partial differential equations (cf. Kesavan [3]). ■

Remark 7.3.2 In the symmetric case, as explained in Example 7.3.1, $a(.,.)$ defines a new inner product whose norm is equivalent to the usual norm. Thus the dual space remains the same and so the Lax-Milgram lemma is just a restatement of the Riesz representation theorem. ■

7.4 Orthonormal Sets

As mentioned earlier, orthogonality is a very important notion special to Hilbert spaces. In this section, we will take a closer look at this property.

Definition 7.4.1 *Let H be a Hilbert space and let $\mathcal{I}$ be an indexing set. A subset $S = \{u_i \in H \mid i \in \mathcal{I}\}$ is said to be* **orthonormal** *if*

$$\|u_i\| \;=\; 1 \;\; \text{for all} \;\; i \in \mathcal{I}$$

and

$$(u_i, u_j) \;=\; 0 \;\; \text{for all} \;\; i,j \in \mathcal{I},\; i \neq j. \;■$$

Remark 7.4.1 If we use the *Kronecker symbol, viz.* δ_{ij} which equals unity if $i = j$ and equals zero if $i \neq j$, then the above relations can be written as

$$(u_i, u_j) \;=\; \delta_{ij}$$

for all i and $j \in \mathcal{I}$. ■

Remark 7.4.2 An orthonormal set of vectors is automatically linearly independent. For, if we have a linear relation of the form

$$\sum_{k=1}^{n} \alpha_k u_{i_k} \ = \ \mathbf{0},$$

then, taking the inner product with u_{i_j} and using the orthonormality of the vectors, we get $\alpha_j = 0$ for any $1 \leq j \leq n$. ■

Example 7.4.1 The sequence $\{\mathsf{e}_n\}$ in ℓ_2 (cf. Example 2.3.12) forms an orthonormal set. Similarly, the standard basis in ℓ_2^n (cf. Example 1.1.2) forms an orthonormal set. ■

Example 7.4.2 Consider the interval $X = [0,1]$ endowed with the Lebesgue measure. The corresponding space $L^2(\mu)$ is denoted $L^2(0,1)$ (cf. Section 6.3). The sequence $\{\mathsf{f}_n\}$ where f_n is the equivalence class represented by the function $f_n(t) = \sqrt{2}\sin n\pi t$, forms an orthonormal set. ■

Proposition 7.4.1 (**Gram-Schmidt Orthogonalization**) *Let H be a Hilbert space and let $\{x_1, \cdots, x_n\}$ be a set of linearly independent vectors in H. Then there exists an orthonormal set of vectors $\{e_1, \cdots, e_n\}$ in H such that, for each $1 \leq i \leq n$, the vector e_i is a linear combination of the vectors $x_1, \cdots, x_i$.*

Proof: Clearly, none of the x_i can be the null vector. Define

$$e_1 \ = \ \frac{1}{\|x_1\|} x_1.$$

Next, consider the vector $x_2 - (x_2, e_1)e_1$. This vector cannot vanish since x_1 and x_2 are linearly independent and e_1 is a scalar multiple of x_1. Thus, we can define

$$e_2 \ = \ \frac{1}{\|x_2 - (x_2, e_1)e_1\|}[x_2 - (x_2, e_1)e_1].$$

It is now immediate to check that $\|e_1\| = \|e_2\| = 1$ and that $(e_1, e_2) = 0$. Further, e_2 is a linear combination of x_1 and x_2, since e_1 is just a scalar multiple of x_1.

We can now proceed inductively. Assume that we have constructed the vectors $e_1, \cdots, e_k$, for $1 \leq k \leq n-1$. We then define

$$e_{k+1} \ = \ \frac{1}{\|x_{k+1} - \sum_{i=1}^{k}(x_{k+1}, e_i)e_i\|}\left[x_{k+1} - \sum_{i=1}^{k}(x_{k+1}, e_i)e_i\right].$$

It is now easy to verify that the set $\{e_1, \cdots, e_n\}$ verifies the conditions mentioned in the statement of the proposition. ■

In the exercises at the end of this chapter, we will see important examples of the Gram-Schmidt orthogonalization process leading to various well known special functions of mathematical physics.

Remark 7.4.3 Consider the space $\mathbb{R}^n = \ell_2^n$. For any $1 \leq j \leq n$, the sets $\{x_1, \cdots, x_j\}$ and $\{e_1, \cdots, e_j\}$ span the same subspace. Thus, we can write

$$x_j = \sum_{i=1}^{j} r_{ij} e_i.$$

Let $\mathbf{A}$ be the matrix whose columns are the x_j, and $\mathbf{Q}$ the matrix whose columns are the e_j. Let $\mathbf{R}$ be the matrix whose entries are the r_{ij}. For any j, we have that $r_{ij} = 0$ if $i > j$. Thus, $\mathbf{R}$ is an upper triangular matrix. Further, we see that

$$\mathbf{A} = \mathbf{QR}.$$

Since the columns of $\mathbf{A}$ are linearly independent, the matrix $\mathbf{A}$ is invertible. Since the columns of $\mathbf{Q}$ are orthonormal, the matrix $\mathbf{Q}$ is orthogonal. Thus, the Gram-Schmidt orthogonalization process proves the following result from matrix theory: *every invertible matrix can be decomposed into the product of an orthogonal matrix and an upper triangular matrix.* ■

Remark 7.4.4 The process of producing orthonormal vectors from linearly independent ones is quite useful in several contexts. For instance, let us consider a continuous function on the interval $[0, 1]$. We wish to approximate it by a polynomial. Amongst several ways of doing this, one is the *least squares approximation.* We look for a polynomial p of degree at most n such that

$$\int_0^1 |f(t) - \mathsf{p}(t)|^2 \, dt = \min_{\mathsf{q} \in \mathcal{P}_n} \int_0^1 |f(t) - \mathsf{q}(t)|^2 \, dt$$

where $\mathcal{P}_n$ is the space of all polynomials (in one variable) of degree less than or equal to n. In other words, we are looking for the projection of f onto the subspace of polynomials of degree less than , or equal to, n

in the space $L^2(0,1)$. We know that such a p exists uniquely and that it is characterized by (cf. Corollary 7.1.3)

$$(\mathsf{p},\mathsf{q}) \;=\; (f,\mathsf{q})$$

for all $\mathsf{q} \in \mathcal{P}_n$. By linearity, it is sufficient to check the above relation for just the basis elements of $\mathcal{P}_n$. The standard basis of $\mathcal{P}_n$ consists of the functions $\mathsf{p}_0, \mathsf{p}_1, \cdots, \mathsf{p}_n$ where

$$\mathsf{p}_0(t) \;\equiv\; 1 \text{ and } \mathsf{p}_k(t) \;=\; t^k$$

for $t \in [0,1]$ and for all $1 \leq k \leq n$. Writing

$$\mathsf{p}(t) \;=\; \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \cdots + \alpha_n t^n,$$

we then derive the following linear system:

$$\mathbf{Ax} \;=\; \mathbf{f}$$

where, $\mathbf{A} = (a_{ij})$ is the $(n+1) \times (n+1)$ matrix given by

$$a_{ij} \;=\; (\mathsf{p}_j, \mathsf{p}_i) \;=\; \int_0^1 t^{i+j}\, dt \;=\; \frac{1}{i+j+1};$$

$\mathbf{x}$ is the $(n+1) \times 1$ column vector whose components are the unknown coefficients of p, $\alpha_0, \alpha_1, \cdots, \alpha_n$; $\mathbf{f}$ is the $(n+1) \times 1$ column vector whose i-th component is

$$f_i \;=\; (f, \mathsf{p}_i) \;=\; \int_0^1 f(t) t^i\, dt.$$

Solving this linear system yields p. However, especially when n is large, the matrix $\mathbf{A}$ is known to be very difficult to invert numerically; it is an example of what is known as a *highly ill-conditioned matrix, i.e.* even small errors in the data can lead to very large errors in the solution of any linear system involving this matrix.

On the other hand, if we replace the standard basis by a basis consisting of orthonormal polynomials $\mathsf{q}_0, \mathsf{q}_1, \cdots, \mathsf{q}_n$, then we can write

$$\mathsf{p} \;=\; \sum_{j=0}^{n} \beta_j \mathsf{q}_j.$$

Now,

$$(\mathsf{p}, \mathsf{q}_i) \;=\; \sum_{j=0}^{n} \beta_j (\mathsf{q}_j, \mathsf{q}_i) \;=\; \sum_{j=0}^{n} \beta_j \delta_{ji} \;=\; \beta_i$$

and so

$$\beta_i \;=\; \int_0^1 f(t)\mathsf{q}_i(t)\, dt.$$

Thus, without solving any linear system, we can directly compute the least squares approximation. ■

Example 7.4.3 Let us compute some of the elements of the orthonormal set obtained from the standard basis of the space of polynomials of degree at most n in $L^2(-1,1)$. Recall that $\mathsf{p}_i(t) = t^i$ for $0 \le i \le n$.

$$\|\mathsf{p}_0\|_2 \;=\; \left(\int_{-1}^1 dt\right)^{\frac{1}{2}} \;=\; \sqrt{2}.$$

Thus $\mathsf{q}_0(t) \;=\; 1/\sqrt{2}$ for all $t \in [-1,1]$. Now consider the function

$$q_1(t) \;=\; t - \frac{1}{\sqrt{2}}\left(\int_{-1}^1 t\, dt\right)\mathsf{q}_0(t) \;=\; t.$$

Then $\|q_1\|_2 = \sqrt{2}/\sqrt{3}$. Thus,

$$\mathsf{q}_1(t) \;=\; \frac{\sqrt{3}}{\sqrt{2}}t$$

for all $t \in [-1,1]$. Next, we consider

$$q_2(t) \;=\; t^2 - \frac{\sqrt{3}}{\sqrt{2}}\left(\int_{-1}^1 t^3\, dt\right)\frac{\sqrt{3}}{\sqrt{2}}t - \frac{1}{\sqrt{2}}\left(\int_{-1}^1 t^2\, dt\right)\frac{1}{\sqrt{2}} \;=\; t^2 - \frac{1}{3}.$$

Then, a straight forward computation yields that $\|q_2\|_2 \;=\; 2\sqrt{2}/3\sqrt{5}$. Thus, for all $t \in [-1,1]$, we get

$$\mathsf{q}_2(t) \;=\; \frac{\sqrt{5}}{2\sqrt{2}}(3t^2 - 1).$$

Similarly, we can show that

$$\mathsf{q}_3(t) \;=\; \frac{\sqrt{7}}{2\sqrt{2}}(5t^3 - 3t)$$

and so on.

An easier way of computing these polynomials will be seen in the exercises at the end of this chapter. ■

Proposition 7.4.2 *Let $\{e_1, \cdots, e_n\}$ be a finite orthonormal set in a Hilbert space H. Then, for any $x \in H$,*

$$\sum_{i=1}^{n} |(x, e_i)|^2 \ \leq \ \|x\|^2. \tag{7.4.1}$$

Further, $x - \sum_{i=1}^{n}(x, e_i)e_i$ is orthogonal to e_j for all $1 \leq j \leq n$.

Proof: We know that $\|x - \sum_{i=1}^{n}(x, e_i)e_i\|^2 \geq 0$. Expanding this, and using the orthonormality of the set, we get (7.4.1) immediately. Further,

$$\left(x - \sum_{i=1}^{n}(x, e_i)e_i, e_j\right) \ = \ (x, e_j) - \sum_{i=1}^{n}(x, e_i)\delta_{ij} \ = \ (x, e_j) - (x, e_j) \ = \ 0.$$

This completes the proof. ■

Proposition 7.4.3 *Let H be a Hilbert space. Let $\mathcal{I}$ be an indexing set and let $\{e_i \mid i \in \mathcal{I}\}$ be an orthonormal set in H. Let $x \in H$. Define*

$$\mathcal{S} \ = \ \{i \in \mathcal{I} \mid (x, e_i) \neq 0\}. \tag{7.4.2}$$

Then, $\mathcal{S}$ is atmost countable.

Proof: Define

$$\mathcal{S}_n \ = \ \left\{ i \in \mathcal{I} \mid |(x, e_i)|^2 > \frac{\|x\|^2}{n} \right\}.$$

By (7.4.1), it follows that $\mathcal{S}_n$ has at most $n - 1$ elements for any positive integer n. Since

$$\mathcal{S} \ = \ \cup_{n=1}^{\infty} \mathcal{S}_n,$$

it follows that $\mathcal{S}$ is at most countable. ■

The preceding proposition helps us to define (infinite) sums over arbitrary orthonormal sets. Let $\{e_i \mid i \in \mathcal{I}\}$ be an orthonormal set in H for an indexing set $\mathcal{I}$. Let $x \in H$. We wish to define the sum

$$\sum_{i \in \mathcal{I}} |(x, e_i)|^2.$$

Let $\mathcal{S}$ be the set defined by (7.4.2). If $\mathcal{S} = \emptyset$, we define the above sum to be zero. If it is a finite set, then the above sum is just the finite sum of the corresponding non-zero terms. If it is countably infinite, then we choose a numbering $e_1, e_2, \cdots, e_n, \cdots$ for the elements in the orthonormal set whose inner product with x is non-zero. Then we define the above sum to be

$$\sum_{n=1}^{\infty} |(x, e_n)|^2.$$

The sum is independent of the numbering chosen since this is a series of positive terms and so any rearrangement thereof will yield the same sum.

We are now in a position to generalize (7.4.1).

Theorem 7.4.1 (**Bessel's Inequality**) *Let H be a Hilbert space and let $\{e_i \mid i \in \mathcal{I}\}$ be an orthonormal set in H, for some indexing set $\mathcal{I}$. Let $x \in H$. Then*

$$\sum_{i \in \mathcal{I}} |(x, e_i)|^2 \;\leq\; \|x\|^2. \tag{7.4.3}$$

Proof: Let $\mathcal{S}$ be defined by (7.4.2). If $\mathcal{S}$ is empty, there is nothing to prove. If it is finite, the result is the same as (7.4.1), which has already been proved. If $\mathcal{S}$ is countably infinite, then, since (7.4.1) establishes the result for all partial sums, (7.4.3) follows. ■

Let $\{e_i| \ i \in \mathcal{I}\}$ be an orthonormal set in a Hilbert space H. Given a vector $x \in H$, we now try to give a meaning to the sum

$$\sum_{i \in \mathcal{I}} (x, e_i) e_i$$

as a vector in H. Once again, let $\mathcal{S}$ be the set defined by (7.4.2). If it is empty, we define the above sum to be the null vector. If it is finite, then we define it to be the (finite) sum of the corresponding terms. Let us, therefore, assume now that $\mathcal{S}$ is a countably infinite set. Let us number the elements $E = \{e_i \mid i \in \mathcal{S}\}$ as $\{e_1, e_2, \cdots, e_n, \cdots\}$. Define

$$y_n \;=\; \sum_{i=1}^{n} (x, e_i) e_i.$$

If $m > n$, then

$$\|y_m - y_n\|^2 \;=\; \sum_{i=n+1}^{m} |(x, e_i)|^2$$

using the orthonormality of the set. But the sum on the right-hand side can be made arbitrarily small for large n and m since it is part of the tail of a convergent series (cf. (7.4.3)). Thus, the sequence $\{y_n\}$ is Cauchy and hence converges to a limit, say, y in H.

Assume now that the elements of the set E above are rearranged so that

$$E = \{e'_1, e'_2, \cdots, e'_n, \cdots\}$$

where each e'_i is equal to a unique e_j. Once again, we define

$$y'_n = \sum_{i=1}^{n} (x, e'_i)e'_i.$$

As before $\{y'_n\}$ is Cauchy and will converge to an element $y' \in H$.

We claim that $y = y'$ so that, whatever the manner in which we number the elements of E, we get the same limit vector, which we will unambiguously define as the required infinite vector sum.

Let $\varepsilon > 0$. Choose N sufficiently large such that, for all $n \geq N$, we have

$$\|y_n - y\| < \varepsilon, \ \|y'_n - y'\| < \varepsilon \text{ and } \sum_{i=N+1}^{\infty} |(x, e_i)|^2 < \varepsilon^2.$$

Fix $n \geq N$. Then we can find $m \geq N$ such that

$$\{e_1, \cdots, e_n\} \subset \{e'_1, \cdots, e'_m\}.$$

Then, the difference $y'_m - y_n$ will consist of a finite number of terms of the form $(x, e_i)e_i$ where all the i concerned are greater than $n(\geq N)$. Hence, it follows that

$$\|y'_m - y_n\|^2 \leq \sum_{i=N+1}^{\infty} |(x, e_i)|^2 < \varepsilon^2.$$

Thus,

$$\|y - y'\| \leq \|y - y_n\| + \|y_n - y'_m\| + \|y'_m - y'\| < 3\varepsilon$$

which proves that $y = y'$ since $\varepsilon > 0$ can be chosen arbitrarily small.

To sum up, we choose an arbitrary numbering of E and write $E = \{e_1, e_2, \cdots, e_n, \cdots\}$ and define

$$\sum_{i \in \mathcal{I}} (x, e_i)e_i = \lim_{n \to \infty} \sum_{j=1}^{n} (x, e_j)e_j.$$

The following result is now an immediate consequence of this definition and of Proposition 7.4.2.

Proposition 7.4.4 *Let H be a Hilbert space and let $\{e_i \mid i \in \mathcal{I}\}$ be an orthonormal set in H. let $x \in H$. Then*

$$x - \sum_{i \in \mathcal{I}} (x, e_i) e_i$$

is orthogonal to every e_j, $j \in \mathcal{I}$. ■

Definition 7.4.2 *An orthonormal set in a Hilbert space is said to be* **complete** *if it is maximal with respect to the partial ordering on orthonormal sets induced by set inclusion. A complete orthonormal set is also called an* **orthonormal basis**. ■

Proposition 7.4.5 *Every Hilbert space admits an orthonormal basis.*

Proof: Given any chain (with respect to the partial ordering induced by set inclusion on orthonormal sets), the union of its members gives an upper bound. Hence, by Zorn's lemma, there exists a maximal orthonormal set. ■

Theorem 7.4.2 *Let H be a Hilbert space and let $\{e_i \mid i \in \mathcal{I}\}$ be an orthonormal set in H. The following are equivalent:*
(i) *The orthonormal set is complete.*
(ii) *If $x \in H$ is such that $(x, e_i) = 0$ for all $i \in \mathcal{I}$, then $x = \mathbf{0}$.*
(iii) *If $x \in H$, then*

$$x = \sum_{i \in \mathcal{I}} (x, e_i) e_i. \tag{7.4.4}$$

(iv) *If $x \in H$, then*

$$\|x\|^2 = \sum_{i \in \mathcal{I}} |(x, e_i)|^2 \tag{7.4.5}$$

(This is known as **Parseval's identity**.*)*

Proof: (i) ⇒ (ii). Assume that the orthonormal set is complete and that $(x, e_i) = 0$ for all $i \in \mathcal{I}$. If $x \neq \mathbf{0}$, then the set

$$\{e_i \mid i \in \mathcal{I}\} \cup \left\{ \frac{1}{\|x\|} x \right\}$$

is also an orthonormal set strictly larger than the given set which contradicts the maximality of the given set.

(ii) $\Rightarrow$ (iii). We know that $x - \sum_{i\in\mathcal{I}}(x, e_i)e_i$ is orthogonal to every e_j, $j \in \mathcal{I}$. This immediately gives (7.4.4).

(iii) $\Rightarrow$ (iv). Set $y_n = \sum_{i=1}^{n}(x, e_i)e_i$ where $\{e_1, e_2, \cdots, e_n \cdots\}$ is a numbering of the elements of the set E explained earlier when defining the sum $\sum_{i\in\mathcal{I}}(x, e_i)e_i$. Then a straight forward computation yields

$$\|y_n\|^2 \;=\; \sum_{i=1}^{n} |(x, e_i)|^2.$$

This immediately yields (7.4.5) on passing to the limit as $n \to \infty$.

(iv) $\Rightarrow$ (i). If the given set were not complete, then there exists $e \in H$ such that $\|e\| = 1$ and $(e, e_i) = 0$ for every $i \in \mathcal{I}$. But then, this will contradict (7.4.5) (applied to the vector e). ■

Corollary 7.4.1 *Let H be a Hilbert space and let $\{e_i \mid i \in \mathcal{I}\}$ be an orthonormal set. It is complete if, and only if, the subspace of all (finite) linear combinations of the e_i is dense in H.*

Proof: If the orthonormal set is complete, then by the preceding theorem, every element of H is the limit of finite linear combinations of the e_i by (7.4.4) and so the subspace spanned by the e_i is dense in H.

Conversely, if the subspace spanned by the e_i is dense in H, then if $x \in H$ is such that it is orthogonal to all the elements of this subspace, then $x = \mathbf{0}$. In particular, if $(x, e_i) = 0$ for all $i \in \mathcal{I}$, then, clearly, x is orthogonal to the subspace spanned by the e_i and so it must vanish. Thus, statement (ii) of the preceding theorem is satisfied and so the orthonormal set is complete. ■

Corollary 7.4.2 *Let H be a Hilbert space and let $\{e_1, e_2, \cdots, e_n \cdots\}$ be a sequence in H which is also an orthonormal basis for H. Then $e_n \rightharpoonup \mathbf{0}$.*

Proof: Let $x \in H$. Then, by (7.4.5), it follows that $(x, e_n) \to 0$ as $n \to \infty$. Then, by the Riesz representation theorem, it follows that $e_n \rightharpoonup \mathbf{0}$. ■

Remark 7.4.5 Notice that if $\{e_n\}$ is an orthonormal sequence which is also complete in a Hilbert space H, then it weakly converges to the null vector while it does not have a norm convergent subsequence, since

$$\|e_n - e_m\| \;=\; \sqrt{2}$$

for all $n \neq m$. ■

Theorem 7.4.3 *A Hilbert space has a countable orthonormal basis if, and only if, it is separable.*

Proof: For simplicity, assume that the space is a real Hilbert space. If the space has a countable orthonormal basis $\{e_n\}$, then the set of all finite linear combinations of the $\{e_n\}$ is dense in H, by (7.4.4). The set of all finite linear combinations of the $\{e_n\}$ with rational coefficients then forms a countable dense subset.

Conversely, assume that the space is separable. Let $\{x_n\}$ be a countable dense subset. Consider the balls $B_n = B(x_n; \sqrt{2}/4)$. If $\{e_i \mid i \in \mathcal{I}\}$ is an orthonormal set, then, since $\|e_i - e_j\| = \sqrt{2}$ for $i \neq j$, it follows that each ball can contain at most one element from the orthonormal basis. But the set $\{x_n\}$ being dense, each ball $B(e_i; \sqrt{2}/4)$ must contain some x_n, or, in other words, each e_i must lie in one of the balls B_n. Thus, the orthonormal set can be at most countable. Thus an orthonormal basis can also only be at most countable. ■

Example 7.4.4 By virtue of (7.4.5), the orthonormal sets in ℓ_2 and ℓ_2^n described in Example 7.4.1 are orthonormal bases of those spaces. ■

Example 7.4.5 (Fourier series) Consider the space $L^2(-\pi, \pi)$. The set

$$\{\mathsf{f}_0\} \cup \{\mathsf{f}_n, \mathsf{g}_n \mid n \in \mathbb{N}\}$$

where

$$f_0(t) = \frac{1}{\sqrt{2\pi}}, \ f_n(t) = \frac{\cos nt}{\sqrt{\pi}} \text{ and } g_n(t) = \frac{\sin nt}{\sqrt{\pi}}$$

for $t \in (-\pi, \pi)$, forms an orthonormal set. By Theorem 6.3.1, the space of continuous functions with compact support contained in $(-\pi, \pi)$ is dense in $L^2(-\pi, \pi)$. Such functions vanish at the end points of the interval $[-\pi, \pi]$ and so they are 2π-periodic on the interval $[-\pi, \pi]$. Consider the space spanned by the orthonormal set mentioned above. By an application of the Stone-Weierstrass theorem (cf. Rudin [6]), it follows that this space is dense in the space of all 2π-periodic continuous functions with respect to the sup norm, which is nothing but the norm $\|.\|_\infty$. Since the interval $(-\pi, \pi)$ has finite measure, this implies that this space is dense with respect to the norm $\|.\|_2$ as well (cf. Proposition 6.1.3). This shows that the space spanned by this orthonormal set is dense in $L^2(-\pi, \pi)$ as well and so, by Corollary 7.4.1, it follows that it is a complete orthonormal set in $L^2(-\pi, \pi)$.

Thus if $\mathsf{f} \in L^2(-\pi, \pi)$, we have that

$$\begin{aligned} f(t) &= \left(\int_{-\pi}^{\pi} f(t)\tfrac{1}{\sqrt{2\pi}}\, dt\right)\tfrac{1}{\sqrt{2\pi}} + \sum_{n=1}^{\infty}\left(\int_{-\pi}^{\pi} f(t)\tfrac{\cos nt}{\sqrt{\pi}}\, dt\right)\tfrac{\cos nt}{\sqrt{\pi}} \\ &\quad + \sum_{n=1}^{\infty}\left(\int_{-\pi}^{\pi} f(t)\tfrac{\sin nt}{\sqrt{\pi}}\, dt\right)\tfrac{\sin nt}{\sqrt{\pi}} \end{aligned}$$

by virtue of (7.4.4). This can be rewritten as

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos nt + b_n \sin nt)$$

where

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\, dt$$

and

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos nt\, dt,\ b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin nt\, dt.$$

This is nothing but the classical Fourier series of a function f and the a_n, $n \geq 0$ and b_n, $n \geq 1$ are the usual Fourier coefficients of f. The above series expansion means that the partial sums of the Fourier series converge to f in the $\|.\|_2$ norm. In other words if

$$f_N(t) = \frac{a_0}{2} + \sum_{n=1}^{N}(a_n \cos nt + b_n \sin nt),$$

for $t \in (-\pi, \pi)$, then

$$\int_{-\pi}^{\pi} |f_N(t) - f(t)|^2\, dt \to 0$$

as $N \to \infty$. The analogue of the Parseval identity (7.4.5) reads as follows:

$$\frac{1}{\pi}\int_{-\pi}^{\pi} |f(t)|^2\, dt = \frac{a_0^2}{2} + \sum_{n=1}^{\infty}(|a_n|^2 + |b_n|^2). \blacksquare$$

Example 7.4.6 (Fourier sine series) Consider the space $L^2(0, \pi)$. Consider the set

$$\left\{\sqrt{\frac{2}{\pi}}\sin nt \mid n \in \mathbb{N}\right\}.$$

This is an orthonormal set in $L^2(0, \pi)$ as one can easily verify. Let $\mathsf{f} \in L^2(0, \pi)$ be orthogonal to every element of this set. Extend f as an odd function to all the interval $(-\pi,\pi)$. Thus, we set $f(x) = -f(-x)$ if $x \in [-\pi, 0)$. Since f is an odd function, it follows that

$$\int_{-\pi}^{\pi} f(t)\, dt \;=\; \int_{-\pi}^{\pi} f(t)\cos nt\, dt \;=\; 0$$

for all $n \in \mathbb{N}$. Since we also have that

$$\int_{-\pi}^{\pi} f(t)\sin nt\, dt \;=\; 2\int_{0}^{\pi} f(t)\sin nt \;=\; 0$$

for all $n \in \mathbb{N}$, it follows that $\mathsf{f} = \mathbf{0}$ in $L^2(-\pi,\pi)$ and so $\mathsf{f} = \mathbf{0}$ in $L^2(0,\pi)$ as well. Thus, by Theorem 7.4.2 (ii), it follows that the given set is complete in $L^2(0,\pi)$. In particular, if $\mathsf{f} \in L^2(0,\pi)$, we can write the series expansion

$$f(t) \;=\; \sum_{n=1}^{\infty} b_n \sin nt$$

where

$$b_n \;=\; \frac{2}{\pi}\int_0^{\pi} f(t)\sin nt\, dt.$$

This is called the Fourier sine series of the function f. ■

By analogy, if H is a separable Hilbert space with an orthonormal basis $\{e_n \mid n \in \mathbb{N}\}$ and if $x \in H$, we call

$$x \;=\; \sum_{n=1}^{\infty} (x, e_n) e_n$$

as its Fourier expansion and the coefficients (x, e_n) are called its Fourier coefficients.

Remark 7.4.6 Let V be a Banach space and let $\{e_n\}$ be a sequence of vectors in V such that every vector $x \in V$ can be written as

$$x \;=\; \sum_{n=1}^{\infty} \alpha_n(x) e_n \;=\; \lim_{N\to\infty} \sum_{n=1}^{N} \alpha_n(x) e_n.$$

Then $\{e_n\}$ is called a *Schauder basis* for V. Thus, in every separable Hilbert space, an orthonormal basis forms a Schauder basis. The set $\{\mathsf{e}_n\}$ in ℓ_p (cf. Example 3.1.1) is a Schauder basis for ℓ_p for $1 \le p < \infty$.

In the literature, a usual basis of the vector space, *i.e.* a set of linearly independent elements such that every vector is a finite linear combination of vectors from the set, is called a *Hamel basis.* Notice that, by Baire's theorem (cf. Exercise 4.1), a Banach space cannot have a countable Hamel basis, while it may have a (countable) Schauder basis. ■

In the next chapter we will see how orthonormal bases occur very naturally in Hilbert spaces.

7.5 Exercises

7.1 Let V be a real Banach space and assume that the parallelogram identity holds in V. Define

$$(u,v) \;=\; \frac{1}{4}(\|u+v\|^2 - \|u-v\|^2).$$

Show that this defines an inner product which induces the given norm and hence that V is a Hilbert space.

7.2 Let V be a complex Banach space and assume that the parallelogram identity holds in V. Define

$$(u,v) \;=\; \frac{1}{4}[\|u+v\|^2 - \|u-v\|^2 + i\|u+iv\|^2 - i\|u-iv\|^2].$$

Show that this defines an inner product which induces the given norm and hence that V is a Hilbert space.

7.3 Let H be a Hilbert space and let M be a closed subspace of H. Let $P : H \to M$ be the orthogonal projection of H onto M. Show that $\|P\| = 1$.

7.4 Let $H = \ell_2^n$. Let $\mathbf{J}$ be the $n \times n$ matrix all of whose entries are equal to $1/n$. Show that

$$\|\mathbf{J}\|_{2,n} \;=\; \|\mathbf{I} - \mathbf{J}\|_{2,n} \;=\; 1$$

where $\mathbf{I}$ is the $n \times n$ identity matrix.

7.5(a) Let H be a Hilbert space and let φ be a non-zero continuous linear functional on H. Let $M = \mathrm{Ker}(\varphi)$. Show that M has codimension one.
(b) Let $g \in M^{\perp}$ be a unit vector such that any $y \in H$ can be written as

$$y = \lambda g + z$$

where $z \in M$. Define $x = \varphi(g)g$. Show that x is such that

$$\varphi(y) = (y, x)$$

for all $y \in H$. (This gives a direct proof of the Riesz representation theorem.)

7.6 Let H be a Hilbert space and let $U : H \to H$ be a unitary operator. Show that U is an isometry, *i.e.* $\|Ux\| = \|x\|$ for all $x \in H$.

7.7 Let H be a real Hilbert space and let $a(.,.) : H \times H \to \mathbb{R}$ be a continuous and H-elliptic bilinear form (cf. Section 7.3) with constants $M > 0$ (for continuity) and $\alpha > 0$ (for ellipticity). Let $f \in H$.
(a) Let $W \subset H$ be a closed subspace. Show that there exists a unique $w \in W$ such that

$$a(w, v) = (f, v) \tag{7.5.1}$$

for all $v \in W$.
(b) Show that, if $w \in W$ is as above, then

$$\|w\| \leq \frac{1}{\alpha}\|f\|.$$

(c) Let $u \in H$ be the unique vector such that

$$a(u, v) = (f, v)$$

for all $v \in H$. Show that

$$\|u - w\| \leq \frac{M}{\alpha} \inf_{v \in W} \|u - v\|.$$

(d) Let H be separable and let $\{u_n\}_{n=1}^{\infty}$ be an orthonormal basis for H. Let $W_n = \mathrm{span}\{u_1, \cdots, u_n\}$. Let w_n be the solution of (7.5.1) when $W = W_n$. Show that $w_n \to u$ as $n \to \infty$.

7.8 Consider the space $L^2(0,1)$. Define $r_0(t) \equiv 1$ and

$$r_n(t) = \sum_{i=1}^{2^n} (-1)^{i-1} \chi_{[\frac{i-1}{2^n}, \frac{i}{2^n}]}(t)$$

where χ_E denotes the characteristic function of a set E.
(a) Show that

$$r_n(t) = \operatorname{sgn}(\sin 2^n \pi t), \ 0 \leq t \leq 1$$

where $\operatorname{sgn}(t)$ equals 1 when $t \geq 0$ and equals -1 when $t < 0$.
(b) Show that $\{\mathsf{r}_n(t)\}_{n=0}^{\infty}$ is orthonormal in $L^2(0,1)$ but that it is not complete.

7.9 Let $(a,b) \subset \mathbb{R}$ be a finite interval and let $\{\phi_n\}_{n\in\mathbb{N}}$ be an orthonormal basis for $L^2(a,b)$. Define

$$\Phi_{ij}(t,s) = \phi_i(t)\phi_j(s)$$

for $(t,s) \in (a,b) \times (a,b)$. Show that $\{\Phi_{ij}\}_{(i,j)\in\mathbb{N}\times\mathbb{N}}$ forms an orthonormal basis for $L^2((a,b) \times (a,b))$.

7.10 Show that the sets

$$\left\{\frac{1}{\sqrt{\pi}}\right\} \cup \left\{\sqrt{\frac{2}{\pi}} \cos nt \mid n \in \mathbb{N}\right\}$$

is a complete orthonormal set in $L^2(0,\pi)$. (Thus, a function in $L^2(0,\pi)$ can be expanded as a series of cosines and this is called its *Fourier cosine series.*)

7.11 Consider $L^2(-1,1)$ and the linearly independent set of functions p_n where $\mathsf{p}_n(t) = t^n$. Applying the Gram-Schmidt orthogonalization procedure, we obtain an orthonormal sequence $\{\mathsf{q}_n\}$ of polynomials (cf. Example 7.4.3).
(a) Define

$$P_n(t) = \sqrt{\frac{2}{2n+1}} \mathsf{q}_n(t).$$

These are the *Legendre polynomials.* Show that $P_n(t)$ consists only of even powers of t when n is even, and of only odd powers of t, when n is

odd.
(b) Show that $P_0(t) \equiv 1,\ P_1(t) = t$ and that, for $n \geq 1$,

$$(n+1)P_{n+1}(t) = (2n+1)tP_n(t) - nP_{n-1}(t)$$

for $t \in [-1, 1]$, given that $P_n(1) = 1$ for all $n \geq 0$.
(This gives a simple recursive formula to compute the Legendre polynomials.) (c) Prove Rodrigues' Formula:

$$P_n(t) = \frac{1}{2^n n!}\frac{d^n}{dt^n}(t^2 - 1)^n.$$

7.12 (a) Consider the space

$$\widetilde{H} = \left\{ f : \mathbb{R} \to \mathbb{R} \mid \int_{-\infty}^{\infty} e^{-x^2} |f(x)|^2 \, dx < \infty \right\}$$

and let H be the space of all equivalence classes (with respect to equality almost everywhere) of functions in $\widetilde{H}$. Define the inner-product

$$(\mathsf{f}, \mathsf{g}) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-x^2} f(x) g(x) \, dx.$$

Show that H is a Hilbert space.
(b) Show that if $f_n(x) = x^n$, then f_n belongs to H for every $n \in \{0\} \cup \mathbb{N}$.
(c) Apply the Gram-Schmidt process to the linearly independent set $\{\mathsf{f}_n\}$ to obtain an orthonormal set h_n. Define

$$H_n(x) = \sqrt{2^n n!} h_n(x).$$

These are the *Hermite polynomials*. Compute H_0 and H_1.
(d) Prove Rodrigues' Formula:

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}.$$

7.13 Let $\mathsf{f}, \mathsf{g} \in L^2(-\pi, \pi)$ and let their Fourier series be given by

$$\begin{aligned} f(t) &= \tfrac{a_0}{2} + \textstyle\sum_{n=1}^{\infty} (a_n \cos nt + b_n \sin nt) \\ g(t) &= \tfrac{c_0}{2} + \textstyle\sum_{n=1}^{\infty} (c_n \cos nt + d_n \sin nt). \end{aligned}$$

Show that

$$\frac{1}{\pi} \int_{-\pi}^{\pi} f(t) g(t) \, dt = \frac{a_0 c_0}{2} + \sum_{n=1}^{\infty} (a_n c_n + b_n d_n).$$

7.14 Compute the Fourier series of the function:

$$f(t) \;=\; \begin{cases} -1 & -\pi \le t < 0 \\ 1 & 0 < t \le \pi. \end{cases}$$

7.15 Compute the Fourier cosine series of the function $f(t) = \sin t$ on $[0, \pi]$.

7.16 (a) Compute the Fourier sine series and the Fourier cosine series of the function $f(t) = t$ on $[0, \pi]$.
(b) Evaluate:

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \quad \text{and} \quad \sum_{n=1}^{\infty} \frac{1}{n^4}$$

using Parseval's identity.

7.17 (a) Let $\mathsf{f} \in L^2(-\pi, \pi)$. Let its Fourier series be given by

$$f(t) \;=\; \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nt + b_n \sin nt).$$

Extend the function f to all of $\mathbb{R}$ by periodicity, *i.e.* such that $f(t+2\pi) = f(t)$ for all $t \in \mathbb{R}$. Define

$$F(t) = \int_0^t \left(f(s) - \frac{a_0}{2} \right) \, ds.$$

Show that $F : \mathbb{R} \to \mathbb{R}$ is also 2π-periodic.
(b) Show that its Fourier series is given by

$$F(t) \;=\; \sum_{n=1}^{\infty} \frac{b_n}{n} + \sum_{n=1}^{\infty} \left(\frac{a_n}{n} \sin nt - \frac{b_n}{n} \cos nt \right).$$

(c) Show that the above series converges to F uniformly on $\mathbb{R}$.

7.18 Let $\mathsf{f} \in H_0^1(-\pi, \pi)$. Show that if its Fourier series expansion is given by

$$f(t) \;=\; \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nt + b_n \sin nt),$$

then the Fourier series expansion of f' is given by

$$f'(t) \;=\; \sum_{n=1}^{\infty} (n b_n \cos nt - n a_n \sin nt).$$

7.19 Let H be a Hilbert space and let $A : D(A) \subset H \to H$ be a linear operator. We say that it is *dissipative* if $(A(u), u) \leq 0$ for all $u \in D(A)$. We say that it is *maximal dissipative* if, in addition $\mathcal{R}(I - A) = H$, where I denotes the identity operator on H. Let A be the infinitesimal generator of a c_0-semigroup of contractions (cf. Exercises 4.4, 4.13 and 4.19) on H. Show that it is maximal dissipative.

7.20 Let H be a Hilbert space and let $A : D(A) \subset H \to H$ be a maximal dissipative operator. Show that if $B : D(B) \subset H \to H$ is dissipative and is an extension of A, *i.e.* $D(A) \subset D(B)$ and $B|_{D(A)} = A$, then $D(B) = D(A)$. (This justifies the adjective 'maximal').

7.21 Let H be a Hilbert space and let $A : D(A) \subset H \to H$ be a maximal dissipative operator. Show that it is closed and densely defined.

7.22 (a) Let H be a Hilbert space and let $A : D(A) \subset H \to H$ be a dissipative operator. Let $\lambda > 0$. If $\mathcal{R}(\lambda I - A) = H$, show that $(\lambda I - A)^{-1}$ exists in $\mathcal{L}(H)$ and that

$$\|(\lambda I - A)^{-1}\| \leq \frac{1}{\lambda}.$$

(b) If A is a dissipative operator and if $\mathcal{R}(\lambda_0 I - A) = H$ for some $\lambda_0 > 0$, show that $\mathcal{R}(\lambda I - A) = H$ for all $0 < \lambda < 2\lambda_0$.
(c) Deduce that if A is a maximal dissipative operator, then $\mathcal{R}(\lambda I - A) = H$ for all $\lambda > 0$.

Remark 7.5.1 Comparing the results of the above exercises with the comments made in Remark 4.8.1, we deduce that an operator $A : D(A) \subset H \to H$ will be the infinitesimal generator of a c_0-semigroup of contractions if, and only if, it is maximal dissipative. Unlike Banach spaces, where the Hille-Yosida theorem involves verification of infinitely many conditions, one for each $\lambda > 0$, this is much easier to verify in Hilbert spaces. The dissipativity is usually very easy to check. Further, it is enough to verify that the equation $(\lambda I - A)x = f$ has a solution $x \in D(A)$ for every $f \in H$ just for *one* fixed $\lambda > 0$.

7.23 Let H_i, $i = 1, 2, 3$ be Hilbert spaces with norms $\|.\|_i$, $i = 1, 2, 3$ respectively. Let $T : D(T) \subset H_1 \to H_2$ and $S : D(S) \subset H_2 \to H_3$ be closed and densely defined linear transformations. Assume that $\mathcal{R}(T) \subset \mathcal{N}(S)$. Assume further, that there exists a constant $C > 0$ such that, for all $x \in D(S) \cap D(T^*)$, we have

$$\|T^*x\|_1^2 + \|Sx\|_3^2 \geq C\|x\|_2^2.$$

(a) Let $\widetilde{H}_2 = \mathcal{N}(S)$ and let $\widetilde{T}^*$ denote the adjoint of $T : D(T) \subset H_1 \to \widetilde{H}_2$. Show that $\widetilde{T}^*$ has closed range.
(b) If $P : H_2 \to \mathcal{N}(S)$ is the orthogonal projection, show that

$$T^*(x) = \widetilde{T}^*(Px)$$

for all $x \in D(T^*)$.
(c) Deduce that T has closed range.
(d) Show that $\mathcal{R}(T) = \mathcal{N}(S)$.

7.24 Let H be a Hilbert space and let $GL(H)$ be the set of all invertible continuous linear operators on H. Then $GL(H)$ is a group with respect to the binary operation defined via composition of operators. Consider the unit circle $S^1 \subset \mathbb{R}^2$ with its usual topology inherited from $\mathbb{R}^2$. Representing a point $g \in S^1$ as $(\cos\theta, \sin\theta)$ where $\theta \in [0, 2\pi)$, we have that S^1 is a group under the operation defined via $(\theta_1, \theta_2) \mapsto (\theta_1 + \theta_2)\text{mod}(2\pi)$. A *representation* of S^1 is a group homomorphism $\widehat{\pi} : S^1 \to GL(H)$ which is also continuous. For simplicity, we will denote the image of $g = (\cos\theta, \sin\theta)$ by $\widehat{\pi}(\theta)$.
(a) Show that every representation is uniformly bounded, *i.e.* there exists a constant $C > 0$ such that

$$\|\widehat{\pi}(\theta)\| \leq C$$

for all $\theta \in [0, 2\pi)$.
(b) Define

$$< u, v > = \frac{1}{2\pi}\int_0^{2\pi} (\widehat{\pi}(2\pi - \theta)(u), \widehat{\pi}(2\pi - \theta)(v))\, d\theta.$$

Show that $< .,. >$ defines a new inner product on H whose induced norm is equivalent to the original norm on H.
(c) With respect to the inner-product $< .,. >$ on H, show that $\widehat{\pi}(\theta)$ is a unitary operator on H for every $\theta \in [0, 2\pi)$. (We say that every representation of S^1 is equivalent to a *unitary representation*.)

7.25 Let $V = \ell_2^N$ and let $\{\mathbf{A}_n\}$ be a sequence of $N \times N$ matrices such that $\mathbf{A}_n = \mathbf{A}_n^*$ for each n. Assume further that, for each $\mathbf{v} \in V$, we have that the sequence $\{(\mathbf{A}_n\mathbf{v}, \mathbf{v})\}$ decreases to zero as $n \to \infty$. Show that there exists a matrix $\mathbf{A} = \mathbf{A}^*$ such that $(\mathbf{A}\mathbf{v}, \mathbf{v}) \geq 0$ for all $\mathbf{v} \in V$ and such that $\mathbf{A}_n \to \mathbf{A}$ in $\mathcal{L}(V)$.