Our goal is a representation of a Hilbert space as a direct sum which is particularly simple and suitable because it makes use of orthogonality. To understand the situation and the problem, let us first introduce the concept of a direct sum. This concept makes sense for any vector space and is defined as follows.

3.3-3 Definition (Direct sum). A vector space X is said to be the *direct sum* of two subspaces Y and Z of X, written

$$X = Y \oplus Z,$$

if each $x \in X$ has a unique representation

$$x = y + z \qquad y \in Y,\ z \in Z.$$

Then Z is called an *algebraic complement* of Y in X and vice versa, and Y, Z is called a *complementary pair* of subspaces in X. ∎

For example, $Y = \mathbf{R}$ is a subspace of the Euclidean plane $\mathbf{R}^2$. Clearly, Y has infinitely many algebraic complements in $\mathbf{R}^2$, each of which is a real line. But most convenient is a complement that is perpendicular. We make use of this fact when we choose a Cartesian coordinate system. In $\mathbf{R}^3$ the situation is the same in principle.

Similarly, in the case of a general Hilbert space H, the main interest concerns representations of H as a direct sum of a closed subspace Y and its **orthogonal complement**

$$Y^\perp = \{z \in H \mid z \perp Y\},$$

which is the set of all vectors orthogonal to Y. This gives our main result in this section, which is sometimes called the *projection theorem*, for reasons to be explained after the proof.

3.3-4 Theorem (Direct sum). *Let Y be any closed subspace of a Hilbert space H. Then*

$$H = Y \oplus Z \qquad Z = Y^\perp. \tag{6}$$

Proof. Since H is complete and Y is closed, Y is complete by Theorem 1.4-7. Since Y is convex, Theorem 3.3-1 and Lemma 3.3-2 imply that for every $x \in H$ there is a $y \in Y$ such that

$$x = y + z \qquad z \in Z = Y^{\perp}. \tag{7}$$

To prove uniqueness, we assume that

$$x = y + z = y_1 + z_1$$

where $y, y_1 \in Y$ and $z, z_1 \in Z$. Then $y - y_1 = z_1 - z$. Since $y - y_1 \in Y$ whereas $z_1 - z \in Z = Y^{\perp}$, we see that $y - y_1 \in Y \cap Y^{\perp} = \{0\}$. This implies $y = y_1$. Hence also $z = z_1$. ∎

y in (7) is called the **orthogonal projection** of x on Y. (or, briefly, the *projection* of x on Y). This term is motivated by elementary geometry. [For instance, we can take $H = \mathbf{R}^2$ and project any point $x = (\xi_1, \xi_2)$ on the ξ_1-axis, which then plays the role of Y; the projection is $y = (\xi_1, 0)$.]

Equation (7) defines a mapping

$$\begin{aligned} P\colon\ H &\longrightarrow Y \\ x &\longmapsto y = Px. \end{aligned}$$

P is called the (orthogonal) **projection** (or *projection operator*) of H onto Y. See Fig. 29. Obviously, P is a bounded linear operator. P

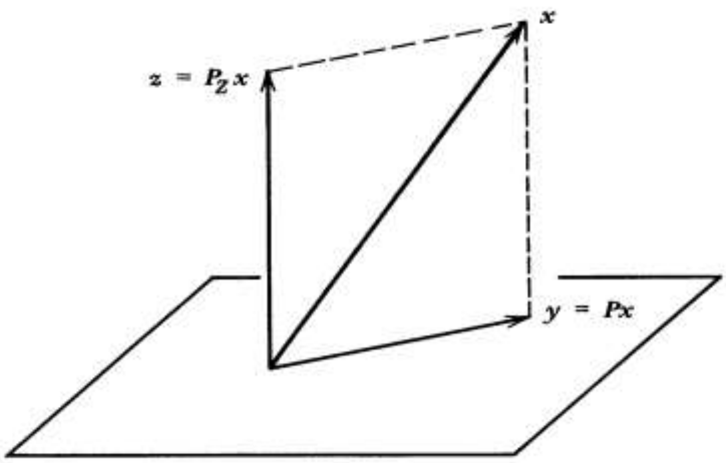

Fig. 29. Notation in connection with Theorem 3.3-4 and formula (9)

maps

$$\begin{aligned}&H \text{ onto } Y,\\ &Y \text{ onto itself},\\ &Z = Y^{\perp} \text{ onto } \{0\},\end{aligned}$$

and is **idempotent,** that is,

$$P^2 = P;$$

thus, for every $x \in H$,

$$P^2 x = P(Px) = Px.$$

Hence $P|_Y$ is the identity operator on Y. And for $Z = Y^{\perp}$ our discussion yields

3.3-5 Lemma (Null space). *The orthogonal complement $Y^{\perp}$ of a closed subspace Y of a Hilbert space H is the null space $\mathcal{N}(P)$ of the orthogonal projection P of H onto Y.*

An orthogonal complement is a special annihilator, where, by definition, the *annihilator* $M^{\perp}$ of a set $M \neq \varnothing$ in an inner product space X is the set[3]

$$M^{\perp} = \{x \in X \mid x \perp M\}.$$

Thus, $x \in M^{\perp}$ if and only if $\langle x, v\rangle = 0$ for all $v \in M$. This explains the name.

Note that $M^{\perp}$ is a vector space since $x, y \in M^{\perp}$ implies for all $v \in M$ and all scalars α, β

$$\langle \alpha x + \beta y, v\rangle = \alpha\langle x, v\rangle + \beta\langle y, v\rangle = 0,$$

hence $\alpha x + \beta y \in M^{\perp}$.

$M^{\perp}$ is closed, as the reader may prove (Prob. 8).

$(M^{\perp})^{\perp}$ is written $M^{\perp\perp}$, etc. In general we have

$$M \subset M^{\perp\perp} \tag{8*}$$

[3] This causes no conflict with Prob. 13, Sec. 2.10, as we shall see later (in Sec. 3.8).

because

$$x \in M \implies x \perp M^{\perp} \implies x \in (M^{\perp})^{\perp}.$$

But for closed subspaces we even have

3.3-6 Lemma (Closed subspace). *If Y is a closed subspace of a Hilbert space H, then*

$$Y = Y^{\perp\perp}. \tag{8}$$

Proof. $Y \subset Y^{\perp\perp}$ by (8*). We show $Y \supset Y^{\perp\perp}$. Let $x \in Y^{\perp\perp}$. Then $x = y + z$ by 3.3-4, where $y \in Y \subset Y^{\perp\perp}$ by (8*). Since $Y^{\perp\perp}$ is a vector space and $x \in Y^{\perp\perp}$ by assumption, we also have $z = x - y \in Y^{\perp\perp}$, hence $z \perp Y^{\perp}$. But $z \in Y^{\perp}$ by 3.3-4. Together $z \perp z$, hence $z = 0$, so that $x = y$, that is, $x \in Y$. Since $x \in Y^{\perp\perp}$ was arbitrary, this proves $Y \supset Y^{\perp\perp}$. ∎

(8) is the main reason for the use of *closed* subspaces in the present context. Since $Z^{\perp} = Y^{\perp\perp} = Y$, formula (6) can also be written

$$H = Z \oplus Z^{\perp}.$$

It follows that $x \longmapsto z$ defines a projection (Fig. 29)

$$P_Z\colon\ H \longrightarrow Z \tag{9}$$

of H onto Z, whose properties are quite similar to those of the projection P considered before.

Theorem 3.3-4 readily implies a characterization of sets in Hilbert spaces whose span is dense, as follows.

3.3-7 Lemma (Dense set). *For any subset $M \neq \varnothing$ of a Hilbert space H, the span of M is dense in H if and only if $M^{\perp} = \{0\}$.*

Proof. (*a*) Let $x \in M^{\perp}$ and assume $V = \operatorname{span} M$ to be dense in H. Then $x \in \bar{V} = H$. By Theorem 1.4-6(*a*) there is a sequence (x_n) in V such that $x_n \longrightarrow x$. Since $x \in M^{\perp}$ and $M^{\perp} \perp V$, we have $\langle x_n, x\rangle = 0$. The continuity of the inner product (cf. Lemma 3.2-2) implies that $\langle x_n, x\rangle \longrightarrow \langle x, x\rangle$. Together, $\langle x, x\rangle = \|x\|^2 = 0$, so that $x = 0$. Since $x \in M^{\perp}$ was arbitrary, this shows that $M^{\perp} = \{0\}$.

(*b*) Conversely, suppose that $M^{\perp}=\{0\}$. If $x \perp V$, then $x \perp M$, so that $x \in M^{\perp}$ and $x=0$. Hence $V^{\perp}=\{0\}$. Noting that V is a subspace of H, we thus obtain $\bar{V}=H$ from 3.3-4 with $Y=\bar{V}$. ∎

Problems

1. Let H be a Hilbert space, $M \subset H$ a convex subset, and (x_n) a sequence in M such that $\|x_n\| \longrightarrow d$, where $d=\inf_{x \in M} \|x\|$. Show that (x_n) converges in H. Give an illustrative example in $\mathbf{R}^2$ or $\mathbf{R}^3$.

2. Show that the subset $M=\{y=(\eta_j) \mid \sum \eta_j=1\}$ of complex space $\mathbf{C}^n$ (cf. 3.1-4) is complete and convex. Find the vector of minimum norm in M.

3. (*a*) Show that the vector space X of all real-valued continuous functions on $[-1,1]$ is the direct sum of the set of all even continuous functions and the set of all odd continuous functions on $[-1,1]$. (*b*) Give examples of representations of $\mathbf{R}^3$ as a direct sum (*i*) of a subspace and its orthogonal complement, (*ii*) of any complementary pair of subspaces.

4. (*a*) Show that the conclusion of Theorem 3.3-1 also holds if X is a Hilbert space and $M \subset X$ is a closed subspace. (*b*) How could we use Appolonius' identity (Sec. 3.1, Prob. 5) in the proof of Theorem 3.3-1?

5. Let $X=\mathbf{R}^2$. Find $M^{\perp}$ if M is (*a*) $\{x\}$, where $x=(\xi_1, \xi_2) \neq 0$, (*b*) a linearly independent set $\{x_1, x_2\} \subset X$.

6. Show that $Y=\{x \mid x=(\xi_j) \in l^2,\ \xi_{2n}=0,\ n \in \mathbf{N}\}$ is a closed subspace of l^2 and find $Y^{\perp}$. What is $Y^{\perp}$ if $Y=\operatorname{span}\{e_1, \cdots, e_n\} \subset l^2$, where $e_j=(\delta_{jk})$?

7. Let A and $B \supset A$ be nonempty subsets of an inner product space X. Show that

(*a*) $A \subset A^{\perp\perp}$, (*b*) $B^{\perp} \subset A^{\perp}$, (*c*) $A^{\perp\perp\perp}=A^{\perp}$.

8. Show that the annihilator $M^{\perp}$ of a set $M \neq \varnothing$ in an inner product space X is a closed subspace of X.

9. Show that a subspace Y of a Hilbert space H is closed in H if and only if $Y=Y^{\perp\perp}$.

10. If $M \neq \varnothing$ is any subset of a Hilbert space H, show that $M^{\perp\perp}$ is the smallest closed subspace of H which contains M, that is, $M^{\perp\perp}$ is contained in any closed subspace $Y \subset H$ such that $Y \supset M$.

3.4 Orthonormal Sets and Sequences

Orthogonality of elements as defined in Sec. 3.1 plays a basic role in inner product and Hilbert spaces. A first impression of this fact was given in the preceding section. Of particular interest are sets whose elements are orthogonal in pairs. To understand this, let us remember a familiar situation in Euclidean space $\mathbf{R}^3$. In the space $\mathbf{R}^3$, a set of that kind is the set of the three unit vectors in the positive directions of the axes of a rectangular coordinate system; call these vectors e_1, e_2, e_3. These vectors form a basis for $\mathbf{R}^3$, so that every $x \in \mathbf{R}^3$ has a unique representation (Fig. 30)

$$x = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3.$$

Now we see a great advantage of the orthogonality. Given x, we can readily determine the unknown coefficients α_1, α_2, α_3 by taking inner products (dot products). In fact, to obtain α_1, we must multiply that representation of x by e_1, that is,

$$\langle x, e_1 \rangle = \alpha_1 \langle e_1, e_1 \rangle + \alpha_2 \langle e_2, e_1 \rangle + \alpha_3 \langle e_3, e_1 \rangle = \alpha_1,$$

and so on. In more general inner product spaces there are similar and other possibilities for the use of orthogonal and orthonormal sets and

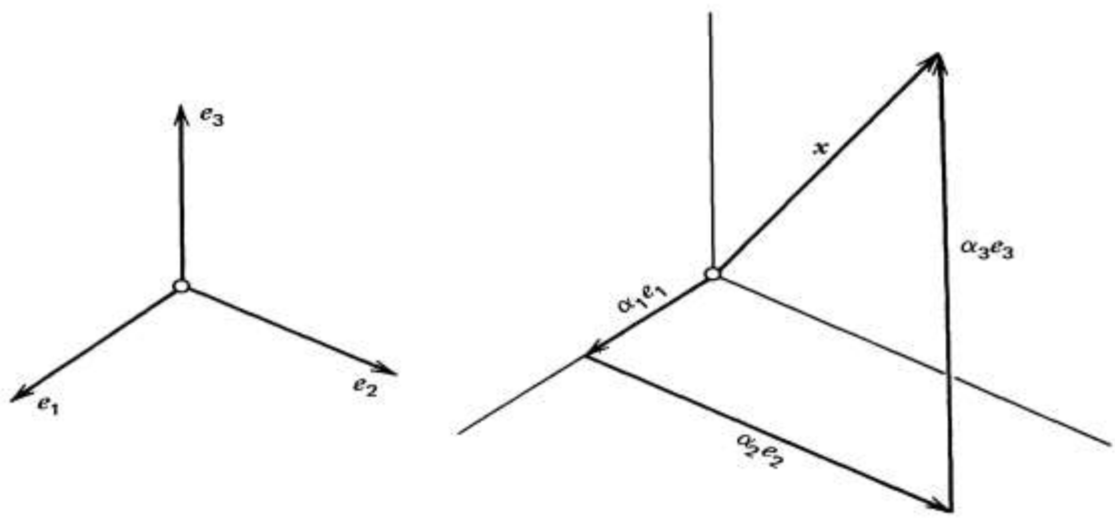

Fig. 30. Orthonormal set $\{e_1, e_2, e_3\}$ in $\mathbf{R}^3$ and representation $x = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3$

sequences, as we shall explain. In fact, the application of such sets and sequences makes up quite a substantial part of the whole theory of inner product and Hilbert spaces. Let us begin our study of this situation by introducing the necessary concepts.

3.4-1 Definition (Orthonormal sets and sequences). *An orthogonal set M in an inner product space X is a subset $M \subset X$ whose elements are pairwise orthogonal. An orthonormal set $M \subset X$ is an orthogonal set in X whose elements have norm 1, that is, for all $x, y \in M$,*

$$\langle x, y\rangle = \begin{cases} 0 & \text{if } x \neq y \\ 1 & \text{if } x = y. \end{cases} \tag{1}$$

If an orthogonal or orthonormal set M is countable, we can arrange it in a sequence (x_n) and call it an *orthogonal* or *orthonormal sequence*, respectively.

More generally, an indexed set, or *family*, (x_α), $\alpha \in I$, is called *orthogonal* if $x_\alpha \perp x_\beta$ for all $\alpha, \beta \in I$, $\alpha \neq \beta$. The family is called *orthonormal* if it is orthogonal and all x_α have norm 1, so that for all $\alpha, \beta \in I$ we have

$$\langle x_\alpha, x_\beta\rangle = \delta_{\alpha\beta} = \begin{cases} 0 & \text{if } \alpha \neq \beta, \\ 1 & \text{if } \alpha = \beta. \end{cases} \tag{2}$$

Here, $\delta_{\alpha\beta}$ is the Kronecker delta, as in Sec. 2.9. ∎

If the reader needs help with families and related concepts, he should look up A1.3 in Appendix 1. He will note that the concepts in our present definition are closely related. The reason is that to any subset M of X we can always find a family of elements of X such that the set of the elements of the family is M. In particular, we may take the family defined by the *natural injection* of M into X, that is, the restriction to M of the identity mapping $x \longmapsto x$ on X.

We shall now consider some simple properties and examples of orthogonal and orthonormal sets.

For orthogonal elements x, y we have $\langle x, y\rangle = 0$, so that we readily obtain the **Pythagorean relation**

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2. \tag{3}$$

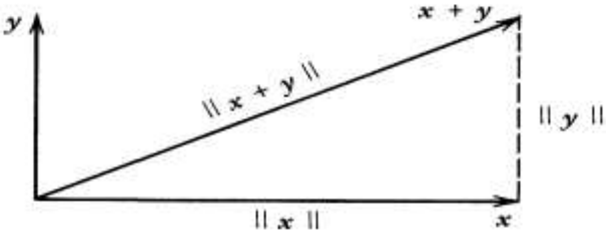

Fig. 31. Pythagorean relation (3) in $\mathbf{R}^2$

Figure 31 shows a familiar example.—More generally, if $\{x_1, \cdots, x_n\}$ is an orthogonal set, then

$$\|x_1 + \cdots + x_n\|^2 = \|x_1\|^2 + \cdots + \|x_n\|^2. \tag{4}$$

In fact, $\langle x_j, x_k \rangle = 0$ if $j \neq k$; consequently,

$$\Big\|\sum_j x_j\Big\|^2 = \Big\langle \sum_j x_j, \sum_k x_k \Big\rangle = \sum_j \sum_k \langle x_j, x_k \rangle = \sum_j \langle x_j, x_j \rangle = \sum_j \|x_j\|^2$$

(summations from 1 to n). We also note

3.4-2 Lemma (Linear independence). *An orthonormal set is linearly independent.*

Proof. Let $\{e_1, \cdots, e_n\}$ be orthonormal and consider the equation

$$\alpha_1 e_1 + \cdots + \alpha_n e_n = 0.$$

Multiplication by a fixed e_j gives

$$\Big\langle \sum_k \alpha_k e_k, e_j \Big\rangle = \sum_k \alpha_k \langle e_k, e_j \rangle = \alpha_j \langle e_j, e_j \rangle = \alpha_j = 0$$

and proves linear independence for any finite orthonormal set. This also implies linear independence if the given orthonormal set is infinite, by the definition of linear independence in Sec. 2.1. ∎

Examples

3.4-3 Euclidean space $\mathbf{R}^3$. In the space $\mathbf{R}^3$, the three unit vectors $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$ in the direction of the three axes of a rectangular coordinate system form an orthonormal set. See Fig. 30.