(a) *We have*

$$|\langle x, y\rangle| \leqq \|x\|\,\|y\| \qquad (1) \qquad \textbf{(Schwarz inequality)}$$

where the equality sign holds if and only if $\{x, y\}$ is a linearly dependent set.

(b) *That norm also satisfies*

$$\|x+y\| \leqq \|x\| + \|y\| \qquad (2) \qquad \textbf{(Triangle inequality)}$$

where the equality sign holds if and only if[2] *$y=0$ or $x=cy$ (c real and $\geqq 0$).*

Proof. **(a)** If $y=0$, then (1) holds since $\langle x, 0\rangle = 0$. Let $y \neq 0$. For every scalar α we have

$$0 \leqq \|x-\alpha y\|^2 = \langle x-\alpha y, x-\alpha y\rangle$$
$$= \langle x, x\rangle - \bar{\alpha}\langle x, y\rangle - \alpha[\langle y, x\rangle - \bar{\alpha}\langle y, y\rangle].$$

We see that the expression in the brackets $[\cdots]$ is zero if we choose $\bar{\alpha} = \langle y, x\rangle / \langle y, y\rangle$. The remaining inequality is

$$0 \leqq \langle x, x\rangle - \frac{\langle y, x\rangle}{\langle y, y\rangle}\langle x, y\rangle = \|x\|^2 - \frac{|\langle x, y\rangle|^2}{\|y\|^2};$$

here we used $\langle y, x\rangle = \overline{\langle x, y\rangle}$. Multiplying by $\|y\|^2$, transferring the last term to the left and taking square roots, we obtain (1).

Equality holds in this derivation if and only if $y=0$ or $0 = \|x-\alpha y\|^2$, hence $x - \alpha y = 0$, so that $x = \alpha y$, which shows linear dependence.

(b) We prove (2). We have

$$\|x+y\|^2 = \langle x+y, x+y\rangle = \|x\|^2 + \langle x, y\rangle + \langle y, x\rangle + \|y\|^2.$$

By the Schwarz inequality,

$$|\langle x, y\rangle| = |\langle y, x\rangle| \leqq \|x\|\,\|y\|.$$

[2] Note that this condition for equality is perfectly "symmetric" in x and y since $x=0$ is included in $x=cy$ (for $c=0$) and so is $y=kx$, $k=1/c$ (for $c>0$).

By the triangle inequality for numbers we thus obtain

$$\begin{aligned}\|x+y\|^2 &\leqq \|x\|^2 + 2\,|\langle x, y\rangle| + \|y\|^2 \\ &\leqq \|x\|^2 + 2\,\|x\|\,\|y\| + \|y\|^2 \\ &= (\|x\| + \|y\|)^2.\end{aligned}$$

Taking square roots on both sides, we have (2).

Equality holds in this derivation if and only if

$$\langle x, y\rangle + \langle y, x\rangle = 2\,\|x\|\,\|y\|.$$

The left-hand side is $2 \operatorname{Re}\langle x, y\rangle$, where Re denotes the real part. From this and (1),

$$\operatorname{Re}\langle x, y\rangle = \|x\|\,\|y\| \geqq |\langle x, y\rangle|. \tag{3}$$

Since the real part of a complex number cannot exceed the absolute value, we must have equality, which implies linear dependence by part (a), say, $y = 0$ or $x = cy$. We show that c is real and $\geqq 0$. From (3) with the equality sign we have $\operatorname{Re}\langle x, y\rangle = |\langle x, y\rangle|$. But if the real part of a complex number equals the absolute value, the imaginary part must be zero. Hence $\langle x, y\rangle = \operatorname{Re}\langle x, y\rangle \geqq 0$ by (3), and $c \geqq 0$ follows from

$$0 \leqq \langle x, y\rangle = \langle cy, y\rangle = c\,\|y\|^2. \qquad \blacksquare$$

The Schwarz inequality (1) is quite important and will be used in proofs over and over again. Another frequently used property is the continuity of the inner product:

3.2-2 Lemma (Continuity of inner product). *If in an inner product space,* $x_n \longrightarrow x$ *and* $y_n \longrightarrow y$, *then* $\langle x_n, y_n\rangle \longrightarrow \langle x, y\rangle$.

Proof. Subtracting and adding a term, using the triangle inequality for numbers and, finally, the Schwarz inequality, we obtain

$$\begin{aligned}|\langle x_n, y_n\rangle - \langle x, y\rangle| &= |\langle x_n, y_n\rangle - \langle x_n, y\rangle + \langle x_n, y\rangle - \langle x, y\rangle| \\ &\leqq |\langle x_n, y_n - y\rangle| + |\langle x_n - x, y\rangle| \\ &\leqq \|x_n\|\,\|y_n - y\| + \|x_n - x\|\,\|y\| \qquad \longrightarrow \qquad 0\end{aligned}$$

since $y_n - y \longrightarrow 0$ and $x_n - x \longrightarrow 0$ as $n \longrightarrow \infty$. $\blacksquare$

As a first application of this lemma, let us prove that every inner product space can be completed. The completion is a Hilbert space and is unique except for isomorphisms. Here the definition of an isomorphism is as follows (as suggested by our discussion in Sec. 2.8).

An **isomorphism** T of an inner product space X onto an inner product space $\tilde{X}$ over the same field is a bijective linear operator $T: X \longrightarrow \tilde{X}$ which preserves the inner product, that is, for all $x, y \in X$,

$$\langle Tx, Ty\rangle = \langle x, y\rangle,$$

where we denoted inner products on X and $\tilde{X}$ by the same symbol, for simplicity. $\tilde{X}$ is then called *isomorphic* with X, and X and $\tilde{X}$ are called *isomorphic inner product spaces.* Note that the bijectivity and linearity guarantees that T is a vector space isomorphism of X onto $\tilde{X}$, so that T preserves the whole structure of inner product space. T is also an isometry of X onto $\tilde{X}$ because distances in X and $\tilde{X}$ are determined by the norms defined by the inner products on X and $\tilde{X}$.

The theorem about the completion of an inner product space can now be stated as follows.

3.2-3 Theorem (Completion). *For any inner product space X there exists a Hilbert space H and an isomorphism A from X onto a dense subspace $W \subset H$. The space H is unique except for isomorphisms.*

Proof. By Theorem 2.3-2 there exists a Banach space H and an isometry A from X onto a subspace W of H which is dense in H. For reasons of continuity, under such an isometry, sums and scalar multiples of elements in X and W correspond to each other, so that A is even an isomorphism of X onto W, both regarded as normed spaces. Lemma 3.2-2 shows that we can define an inner product on H by setting

$$\langle \hat{x}, \hat{y}\rangle = \lim_{n\to\infty} \langle x_n, y_n\rangle,$$

the notations being as in Theorem 2.3-2 (and 1.6-2), that is, (x_n) and (y_n) are representatives of $\hat{x} \in H$ and $\hat{y} \in H$, respectively. Taking (9) and (10), Sec. 3.1, into account, we see that A is an isomorphism of X onto W, both regarded as inner product spaces.

Theorem 2.3-2 also guarantees that H is unique except for isometries, that is, two completions H and $\tilde{H}$ of X are related by an

isometry $T: H \longrightarrow \tilde{H}$. Reasoning as in the case of A, we conclude that T must be an isomorphism of the Hilbert space H onto the Hilbert space $\tilde{H}$. ∎

A **subspace** Y of an inner product space X is defined to be a vector subspace of X (cf. Sec. 2.1) taken with the inner product on X restricted to $Y \times Y$.

Similarly, a **subspace** Y of a Hilbert space H is defined to be a subspace of H, regarded as an inner product space. Note that Y need not be a Hilbert space because Y may not be complete. In fact, from Theorems 2.3-1 and 2.4-2 we immediately have the statements (a) and (b) in the following theorem.

3.2-4 Theorem (Subspace). *Let Y be a subspace of a Hilbert space H. Then:*

(a) *Y is complete if and only if Y is closed in H.*

(b) *If Y is finite dimensional, then Y is complete.*

(c) *If H is separable, so is Y. More generally, every subset of a separable inner product space is separable.*

The simple proof of (c) is left to the reader.

Problems

1. What is the Schwarz inequality in $\mathbf{R}^2$ or $\mathbf{R}^3$? Give another proof of it in these cases.

2. Give examples of subspaces of l^2.

3. Let X be the inner product space consisting of the polynomial $x = 0$ (cf. the remark in Prob. 9, Sec. 2.9) and all real polynomials in t, of degree not exceeding 2, considered for real $t \in [a, b]$, with inner product defined by (7), Sec. 3.1. Show that X is complete. Let Y consist of all $x \in X$ such that $x(a) = 0$. Is Y a subspace of X? Do all $x \in X$ of degree 2 form a subspace of X?

4. Show that $y \perp x_n$ and $x_n \longrightarrow x$ together imply $x \perp y$.

5. Show that for a sequence (x_n) in an inner product space the conditions $\|x_n\| \longrightarrow \|x\|$ and $\langle x_n, x\rangle \longrightarrow \langle x, x\rangle$ imply convergence $x_n \longrightarrow x$.

6. Prove the statement in Prob. 5 for the special case of the complex plane.

7. Show that in an inner product space, $x \perp y$ if and only if we have $\|x+\alpha y\| = \|x-\alpha y\|$ for all scalars α. (See Fig. 25.)

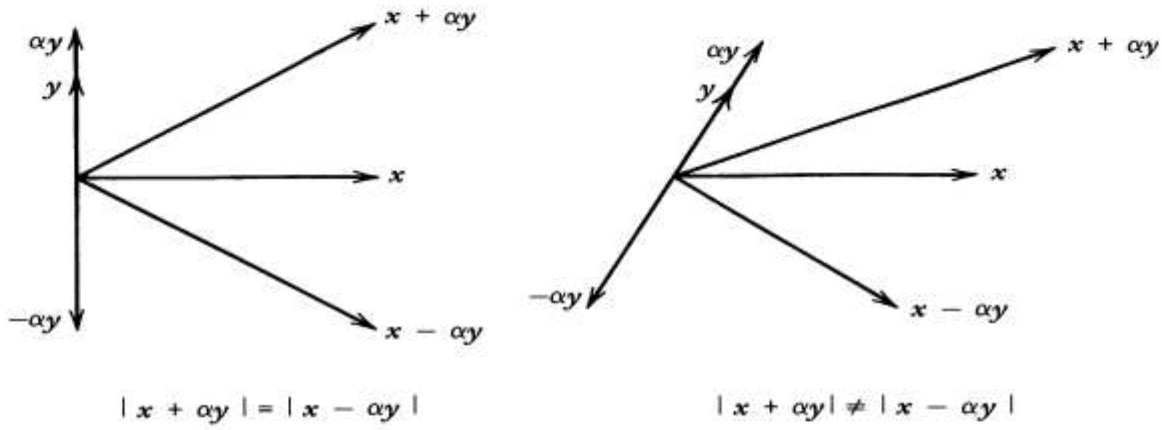

Fig. 25. Illustration of Prob. 7 in the Euclidean plane $\mathbf{R}^2$

8. Show that in an inner product space, $x \perp y$ if and only if $\|x+\alpha y\| \geqq \|x\|$ for all scalars α.

9. Let V be the vector space of all continuous complex-valued functions on $J = [a, b]$. Let $X_1 = (V, \|\cdot\|_\infty)$, where $\|x\|_\infty = \max_{t \in J} |x(t)|$; and let $X_2 = (V, \|\cdot\|_2)$, where

$$\|x\|_2 = \langle x, x\rangle^{1/2}, \qquad \langle x, y\rangle = \int_a^b x(t)\overline{y(t)}\, dt.$$

Show that the identity mapping $x \longmapsto x$ of X_1 onto X_2 is continuous. (It is not a homeomorphism. X_2 is not complete.)

10. **(Zero operator)** Let $T\colon X \longrightarrow X$ be a bounded linear operator on a complex inner product space X. *If* $\langle Tx, x\rangle = 0$ for all $x \in X$, show that $T = 0$.

Show that this does not hold in the case of a *real* inner product space. *Hint.* Consider a rotation of the Euclidean plane.

3.3 Orthogonal Complements and Direct Sums

In a metric space X, the *distance* δ from an element $x \in X$ to a nonempty subset $M \subset X$ is defined to be

$$\delta = \inf_{\tilde{y} \in M} d(x, \tilde{y}) \qquad (M \neq \varnothing).$$

In a normed space this becomes

$$\delta = \inf_{\tilde{y} \in M} \|x - \tilde{y}\| \qquad (M \neq \varnothing). \tag{1}$$

A simple illustrative example is shown in Fig. 26.

We shall see that it is important to know whether there is a $y \in M$ such that

$$\delta = \|x - y\|, \tag{2}$$

that is, intuitively speaking, a point $y \in M$ which is closest to the given x, and if such an element exists, whether it is unique. This is an *existence and uniqueness problem.* It is of fundamental importance, theoretically as well as in applications, for instance, in connection with approximations of functions.

Figure 27 illustrates that even in a very simple space such as the Euclidean plane $\mathbf{R}^2$, there may be no y satisfying (2), or precisely one such y, or more than one y. And we may expect that other spaces, in particular infinite dimensional ones, will be much more complicated in that respect. For general normed spaces this is the case (as we shall see in Chap. 6), but for Hilbert spaces the situation remains relatively

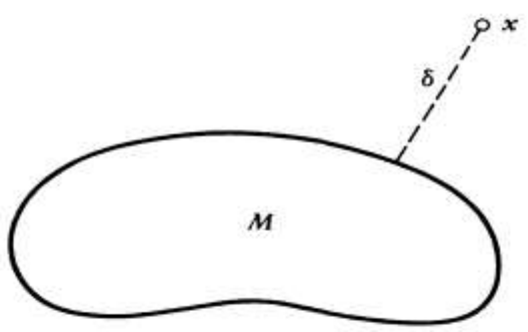

Fig. 26. Illustration of (1) in the case of the plane $\mathbf{R}^2$

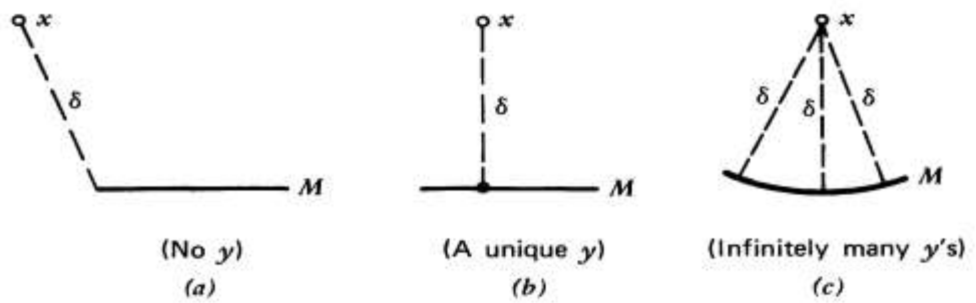

Fig. 27. Existence and uniqueness of points $y \in M$ satisfying (2), where the given $M \subset \mathbf{R}^2$ is an open segment [in (a) and (b)] and a circular arc [in (c)]

simple. This fact is surprising and has various theoretical and practical consequences. It is one of the main reasons why the theory of Hilbert spaces is simpler than that of general Banach spaces.

To consider that existence and uniqueness problem for Hilbert spaces and to formulate the key theorem (3.3-1, below), we need two related concepts, which are of general interest, as follows.

The **segment** joining two given elements x and y of a vector space X is defined to be the set of all $z \in X$ of the form

$$z = \alpha x + (1-\alpha)y \qquad (\alpha \in \mathbf{R},\ 0 \leqq \alpha \leqq 1).$$

A subset M of X is said to be **convex** if for every x, $y \in M$ the segment joining x and y is contained in M. Figure 28 shows a simple example.

For instance, every subspace Y of X is convex, and the intersection of convex sets is a convex set.

We can now provide the main tool in this section:

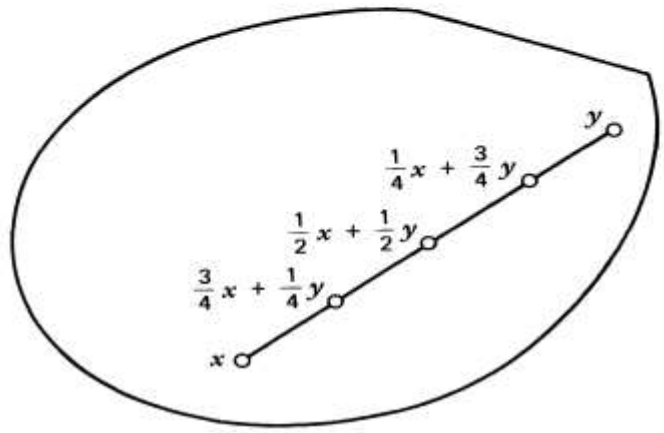

Fig. 28. Illustrative example of a segment in a convex set

3.3-1 Theorem (Minimizing vector). *Let X be an inner product space and $M \neq \varnothing$ a convex subset which is complete (in the metric induced by the inner product). Then for every given $x \in X$ there exists a unique $y \in M$ such that*

$$\delta = \inf_{\tilde{y} \in M} \|x - \tilde{y}\| = \|x - y\|. \tag{3}$$

Proof. **(a)** *Existence.* By the definition of an infimum there is a sequence (y_n) in M such that

$$\delta_n \longrightarrow \delta \qquad \text{where} \qquad \delta_n = \|x - y_n\|. \tag{4}$$

We show that (y_n) is Cauchy. Writing $y_n - x = v_n$, we have $\|v_n\| = \delta_n$ and

$$\|v_n + v_m\| = \|y_n + y_m - 2x\| = 2\,\|\tfrac{1}{2}(y_n + y_m) - x\| \geqq 2\delta$$

because M is convex, so that $\frac{1}{2}(y_n + y_m) \in M$. Furthermore, we have $y_n - y_m = v_n - v_m$. Hence by the parallelogram equality,

$$\begin{aligned}\|y_n - y_m\|^2 = \|v_n - v_m\|^2 &= -\|v_n + v_m\|^2 + 2(\|v_n\|^2 + \|v_m\|^2)\\ &\leqq -(2\delta)^2 + 2(\delta_n{}^2 + \delta_m{}^2),\end{aligned}$$

and (4) implies that (y_n) is Cauchy. Since M is complete, (y_n) converges, say, $y_n \longrightarrow y \in M$. Since $y \in M$, we have $\|x - y\| \geqq \delta$. Also, by (4),

$$\|x - y\| \leqq \|x - y_n\| + \|y_n - y\| = \delta_n + \|y_n - y\| \longrightarrow \delta.$$

This shows that $\|x - y\| = \delta$.

(b) *Uniqueness.* We assume that $y \in M$ and $y_0 \in M$ both satisfy

$$\|x - y\| = \delta \qquad \text{and} \qquad \|x - y_0\| = \delta$$

and show that then $y_0 = y$. By the parallelogram equality,

$$\begin{aligned}\|y - y_0\|^2 &= \|(y - x) - (y_0 - x)\|^2\\ &= 2\,\|y - x\|^2 + 2\,\|y_0 - x\|^2 - \|(y - x) + (y_0 - x)\|^2\\ &= 2\delta^2 + 2\delta^2 - 2^2\,\|\tfrac{1}{2}(y + y_0) - x\|^2.\end{aligned}$$

On the right, $\frac{1}{2}(y+y_0)\in M$, so that

$$\|\tfrac{1}{2}(y+y_0)-x\|\geqq\delta.$$

This implies that the right-hand side is less than or equal to $2\delta^2+2\delta^2-4\delta^2=0$. Hence we have the inequality $\|y-y_0\|\leqq 0$. Clearly, $\|y-y_0\|\geqq 0$, so that we must have equality, and $y_0=y$. ∎

Turning from arbitrary convex sets to subspaces, we obtain a lemma which generalizes the familiar idea of elementary geometry that the unique point y in a given subspace Y closest to a given x is found by "dropping a perpendicular from x to Y."

3.3-2 Lemma (Orthogonality). *In Theorem* 3.3-1, *let M be a complete subspace Y and $x\in X$ fixed. Then $z=x-y$ is orthogonal to Y.*

Proof. If $z\perp Y$ were false, there would be a $y_1\in Y$ such that

$$\langle z, y_1\rangle=\beta\neq 0. \tag{5}$$

Clearly, $y_1\neq 0$ since otherwise $\langle z, y_1\rangle=0$. Furthermore, for any scalar α,

$$\begin{aligned}\|z-\alpha y_1\|^2&=\langle z-\alpha y_1, z-\alpha y_1\rangle\\&=\langle z, z\rangle-\bar\alpha\langle z, y_1\rangle-\alpha[\langle y_1, z\rangle-\bar\alpha\langle y_1, y_1\rangle]\\&=\langle z, z\rangle-\bar\alpha\beta-\alpha[\bar\beta-\bar\alpha\langle y_1, y_1\rangle].\end{aligned}$$

The expression in the brackets $[\cdots]$ is zero if we choose

$$\bar\alpha=\frac{\bar\beta}{\langle y_1, y_1\rangle}.$$

From (3) we have $\|z\|=\|x-y\|=\delta$, so that our equation now yields

$$\|z-\alpha y_1\|^2=\|z\|^2-\frac{|\beta|^2}{\langle y_1, y_1\rangle}<\delta^2.$$

But this is impossible because we have

$$z-\alpha y_1=x-y_2 \qquad \text{where} \qquad y_2=y+\alpha y_1\in Y,$$

so that $\|z-\alpha y_1\|\geqq\delta$ by the definition of δ. Hence (5) cannot hold, and the lemma is proved. ∎