3.4-4 Space l^2. In the space l^2, an orthonormal sequence is (e_n), where $e_n = (\delta_{nj})$ has the nth element 1 and all others zero. (Cf. 3.1-6.)

3.4-5 Continuous functions. Let X be the inner product space of all real-valued continuous functions on $[0, 2\pi]$ with inner product defined by

$$\langle x, y \rangle = \int_0^{2\pi} x(t)y(t)\,dt$$

(cf. 3.1-5). An orthogonal sequence in X is (u_n), where

$$u_n(t) = \cos nt \qquad n = 0, 1, \cdots.$$

Another orthogonal sequence in X is (v_n), where

$$v_n(t) = \sin nt \qquad n = 1, 2, \cdots.$$

In fact, by integration we obtain

$$\text{(5)} \qquad \langle u_m, u_n \rangle = \int_0^{2\pi} \cos mt \cos nt\,dt = \begin{cases} 0 & \text{if } m \neq n \\ \pi & \text{if } m = n = 1, 2, \cdots \\ 2\pi & \text{if } m = n = 0 \end{cases}$$

and similarly for (v_n). Hence an orthonormal sequence is (e_n), where

$$e_0(t) = \frac{1}{\sqrt{2\pi}}, \qquad e_n(t) = \frac{u_n(t)}{\|u_n\|} = \frac{\cos nt}{\sqrt{\pi}} \qquad (n = 1, 2, \cdots).$$

From (v_n) we obtain the orthonormal sequence $(\tilde{e}_n)$, where

$$\tilde{e}_n(t) = \frac{v_n(t)}{\|v_n\|} = \frac{\sin nt}{\sqrt{\pi}} \qquad (n = 1, 2, \cdots).$$

Note that we even have $u_m \perp v_n$ for all m and n. (Proof?) These sequences appear in *Fourier series*, as we shall discuss in the next section. Our examples are sufficient to give us a first impression of what is going on. Further orthonormal sequences of practical importance are included in a later section (Sec. 3.7). ∎

A great advantage of orthonormal sequences over arbitrary linearly independent sequences is the following. If we know that a given x can be represented as a linear combination of some elements of an orthonormal sequence, then the orthonormality makes the actual determination of the coefficients very easy. In fact, if $(e_1, e_2, \cdots)$ is an orthonormal sequence in an inner product space X and we have $x \in \operatorname{span}\{e_1, \cdots, e_n\}$, where n is fixed, then by the definition of the span (Sec. 2.1),

$$x = \sum_{k=1}^{n} \alpha_k e_k, \tag{6}$$

and if we take the inner product by a fixed e_j, we obtain

$$\langle x, e_j\rangle = \left\langle \sum \alpha_k e_k, e_j \right\rangle = \sum \alpha_k \langle e_k, e_j\rangle = \alpha_j.$$

With these coefficients, (6) becomes

$$x = \sum_{k=1}^{n} \langle x, e_k\rangle e_k. \tag{7}$$

This shows that the determination of the unknown coefficients in (6) is simple. Another advantage of orthonormality becomes apparent if in (6) and (7) we want to add another term $\alpha_{n+1}e_{n+1}$, to take care of an

$$\tilde{x} = x + \alpha_{n+1}e_{n+1} \in \operatorname{span}\{e_1, \cdots, e_{n+1}\};$$

then we need to calculate only one more coefficient since the other coefficients remain unchanged.

More generally, if we consider any $x \in X$, not necessarily in $Y_n = \operatorname{span}\{e_1, \cdots, e_n\}$, we can define $y \in Y_n$ by setting

$$y = \sum_{k=1}^{n} \langle x, e_k\rangle e_k, \tag{8a}$$

where n is fixed, as before, and then define z by setting

$$x = y + z, \tag{8b}$$

that is, $z = x - y$. We want to show that $z \perp y$. To really understand what is going on, note the following. *Every* $y \in Y_n$ is a linear combination

$$y = \sum_{k=1}^{n} \alpha_k e_k.$$

Here $\alpha_k = \langle y, e_k \rangle$, as follows from what we discussed right before. Our claim is that for the particular choice $\alpha_k = \langle x, e_k \rangle$, $k = 1, \cdots, n$, we shall obtain a y such that $z = x - y \perp y$.

To prove this, we first note that, by the orthonormality,

$$\|y\|^2 = \left\langle \sum \langle x, e_k \rangle e_k, \sum \langle x, e_m \rangle e_m \right\rangle = \sum |\langle x, e_k \rangle|^2. \tag{9}$$

Using this, we can now show that $z \perp y$:

$$\begin{aligned} \langle z, y \rangle = \langle x - y, y \rangle &= \langle x, y \rangle - \langle y, y \rangle \\ &= \left\langle x, \sum \langle x, e_k \rangle e_k \right\rangle - \|y\|^2 \\ &= \sum \langle x, e_k \rangle \overline{\langle x, e_k \rangle} - \sum |\langle x, e_k \rangle|^2 \\ &= 0. \end{aligned}$$

Hence the Pythagorean relation (3) gives

$$\|x\|^2 = \|y\|^2 + \|z\|^2. \tag{10}$$

By (9) it follows that

$$\|z\|^2 = \|x\|^2 - \|y\|^2 = \|x\|^2 - \sum |\langle x, e_k \rangle|^2. \tag{11}$$

Since $\|z\| \geqq 0$, we have for every $n = 1, 2, \cdots$

$$\sum_{k=1}^{n} |\langle x, e_k \rangle|^2 \leqq \|x\|^2. \tag{12*}$$

These sums have nonnegative terms, so that they form a monotone increasing sequence. This sequence converges because it is bounded by $\|x\|^2$. This is the sequence of the partial sums of an infinite series, which thus converges. Hence (12*) implies

3.4-6 Theorem (Bessel inequality). *Let (e_k) be an orthonormal sequence in an inner product space X. Then for every $x \in X$*

$$\sum_{k=1}^{\infty} |\langle x, e_k\rangle|^2 \leqq \|x\|^2 \qquad \textbf{(Bessel inequality).} \tag{12}$$

The inner products $\langle x, e_k\rangle$ in (12) are called the **Fourier coefficients** of x with respect to the orthonormal sequence (e_k).

Note that if X is finite dimensional, then every orthonormal set in X must be finite because it is linearly independent by 3.4-2. Hence in (12) we then have a finite sum.

We have seen that orthonormal sequences are very convenient to work with. The remaining practical problem is how to obtain an orthonormal sequence if an arbitrary linearly independent sequence is given. This is accomplished by a constructive procedure, the **Gram-Schmidt process** for orthonormalizing a linearly independent sequence (x_j) in an inner product space. The resulting orthonormal sequence (e_j) has the property that for every n,

$$\operatorname{span}\{e_1, \cdots, e_n\} = \operatorname{span}\{x_1, \cdots, x_n\}.$$

The process is as follows.

1st step. The first element of (e_k) *is*

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

2nd step. x_2 can be written

$$x_2 = \langle x_2, e_1\rangle e_1 + v_2.$$

Then (Fig. 32)

$$v_2 = x_2 - \langle x_2, e_1\rangle e_1$$

is not the zero vector since (x_j) is linearly independent; also $v_2 \perp e_1$ since $\langle v_2, e_1\rangle = 0$, so that we can take

$$e_2 = \frac{1}{\|v_2\|} v_2.$$

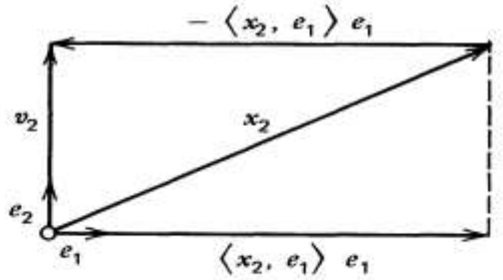

Fig. 32. Gram-Schmidt process, 2nd step

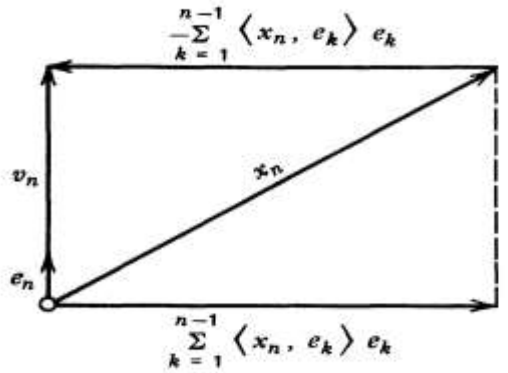

Fig. 33. Gram-Schmidt process, nth step

3rd step. The vector

$$v_3 = x_3 - \langle x_3, e_1\rangle e_1 - \langle x_3, e_2\rangle e_2$$

is not the zero vector, and $v_3 \perp e_1$ as well as $v_3 \perp e_2$. We take

$$e_3 = \frac{1}{\|v_3\|} v_3.$$

***n*th step.** The vector (see Fig. 33)

$$v_n = x_n - \sum_{k=1}^{n-1} \langle x_n, e_k\rangle e_k \tag{13}$$

is not the zero vector and is orthogonal to $e_1, \cdots, e_{n-1}$. From it we obtain

$$e_n = \frac{1}{\|v_n\|} v_n. \tag{14}$$

These are the general formulas for the Gram-Schmidt process, which was designed by E. Schmidt (1907). Cf. also J. P. Gram (1883). Note that the sum which is subtracted on the right-hand side of (13) is the projection of x_n on span $\{e_1, \cdots, e_{n-1}\}$. In other words, in each step we subtract from x_n its "components" in the directions of the previously orthogonalized vectors. This gives v_n, which is then multiplied by $1/\|v_n\|$, so that we get a vector of norm one. v_n cannot be the zero vector for any n. In fact, if n were the smallest subscript for which $v_n = 0$, then (13) shows that x_n would be a linear combination of $e_1, \cdots, e_{n-1}$, hence a linear combination of $x_1, \cdots, x_{n-1}$, contradicting the assumption that $\{x_1, \cdots, x_n\}$ is linearly independent.

Problems

1. Show that an inner product space of finite dimension n has a basis $\{b_1, \cdots, b_n\}$ of orthonormal vectors. (The infinite dimensional case will be considered in Sec. 3.6.)

2. How can we interpret (12*) geometrically in $\mathbf{R}^r$, where $r \geqq n$?

3. Obtain the Schwarz inequality (Sec. 3.2) from (12*).

4. Give an example of an $x \in l^2$ such that we have strict inequality in (12).

5. If (e_k) is an orthonormal sequence in an inner product space X, and $x \in X$, show that $x - y$ with y given by

$$y = \sum_{k=1}^{n} \alpha_k e_k \qquad \alpha_k = \langle x, e_k \rangle$$

is orthogonal to the subspace $Y_n = \operatorname{span}\{e_1, \cdots e_n\}$.

6. (Minimum property of Fourier coefficients) Let $\{e_1, \cdots, e_n\}$ be an orthonormal set in an inner product space X, where n is fixed. Let $x \in X$ be any fixed element and $y = \beta_1 e_1 + \cdots + \beta_n e_n$. Then $\|x - y\|$ depends on $\beta_1, \cdots, \beta_n$. Show by direct calculation that $\|x - y\|$ is minimum if and only if $\beta_j = \langle x, e_j \rangle$, where $j = 1, \cdots, n$.

7. Let (e_k) be any orthonormal sequence in an inner product space X. Show that for any $x, y \in X$,

$$\sum_{k=1}^{\infty} |\langle x, e_k \rangle \langle y, e_k \rangle| \leqq \|x\| \, \|y\|.$$

8. Show that an element x of an inner product space X cannot have "too many" Fourier coefficients $\langle x, e_k \rangle$ which are "big"; here, (e_k) is a given orthonormal sequence; more precisely, show that the number n_m of $\langle x, e_k \rangle$ such that $|\langle x, e_k \rangle| > 1/m$ must satisfy $n_m < m^2 \|x\|^2$.

9. Orthonormalize the first three terms of the sequence $(x_0, x_1, x_2, \cdots)$, where $x_j(t) = t^j$, on the interval $[-1, 1]$, where

$$\langle x, y\rangle = \int_{-1}^{1} x(t)y(t)\,dt.$$

10. Let $x_1(t) = t^2$, $x_2(t) = t$ and $x_3(t) = 1$. Orthonormalize x_1, x_2, x_3, in this order, on the interval $[-1, 1]$ with respect to the inner product given in Prob. 9. Compare with Prob. 9 and comment.

3.5 Series Related to Orthonormal Sequences and Sets

There are some facts and questions that arise in connection with the Bessel inequality. In this section we first motivate the term "Fourier coefficients," then consider infinite series related to orthonormal sequences, and finally take a first look at orthonormal sets which are uncountable.

3.5-1 Example (Fourier series). A *trigonometric series* is a series of the form

$$a_0 + \sum_{k=1}^{\infty} (a_k \cos kt + b_k \sin kt). \tag{1*}$$

A real-valued function x on $\mathbf{R}$ is said to be *periodic* if there is a positive number p (called a *period* of x) such that $x(t+p) = x(t)$ for all $t \in \mathbf{R}$.

Let x be of period 2π and continuous. By definition, the *Fourier series* of x is the trigonometric series (1*) with coefficients a_k and b_k given by the *Euler formulas*

$$\begin{aligned} a_0 &= \frac{1}{2\pi}\int_0^{2\pi} x(t)\,dt \\ a_k &= \frac{1}{\pi}\int_0^{2\pi} x(t)\cos kt\,dt \qquad k = 1, 2, \cdots, \\ b_k &= \frac{1}{\pi}\int_0^{2\pi} x(t)\sin kt\,dt \qquad k = 1, 2, \cdots. \end{aligned} \tag{2}$$

These coefficients are called the *Fourier coefficients of x*.