CHAPTER 2

NORMED SPACES. BANACH SPACES

Particularly useful and important metric spaces are obtained if we take a vector space and define on it a metric by means of a *norm*. The resulting space is called a *normed space*. If it is a complete metric space, it is called a *Banach space*. The theory of normed spaces, in particular Banach spaces, and the theory of linear operators defined on them are the most highly developed parts of functional analysis. The present chapter is devoted to the basic ideas of those theories.

Important concepts, brief orientation about main content

A *normed space* (cf. 2.2-1) is a *vector space* (cf. 2.1-1) with a metric defined by a *norm* (cf. 2.2-1); the latter generalizes the length of a vector in the plane or in three-dimensional space. A *Banach space* (cf. 2.2-1) is a normed space which is a complete metric space. A normed space has a completion which is a Banach space (cf. 2.3-2). In a normed space we can also define and use infinite series (cf. Sec. 2.3).

A mapping from a normed space X into a normed space Y is called an *operator*. A mapping from X into the scalar field $\mathbf{R}$ or $\mathbf{C}$ is called a *functional*. Of particular importance are so-called *bounded linear operators* (cf. 2.7-1) and *bounded linear functionals* (cf. 2.8-2) since they are continuous and take advantage of the vector space structure. In fact, Theorem 2.7-9 states that a *linear* operator is continuous if and only if it is bounded. This is a fundamental result. And vector spaces are of importance here mainly because of the linear operators and functionals they carry.

It is basic that the set of all bounded linear operators from a given normed space X into a given normed space Y can be made into a normed space (cf. 2.10–1), which is denoted by $B(X, Y)$. Similarly, the set of all bounded linear functionals on X becomes a normed space, which is called the *dual space* X' of X (cf. 2.10-3).

In analysis, infinite dimensional normed spaces are more important than finite dimensional ones. The latter are simpler (cf. Secs. 2.4, 2.5), and operators on them can be represented by matrices (cf. Sec. 2.9).

Remark on notation

We denote spaces by X and Y, operators by capital letters (preferably T), the image of an x under T by Tx (without parentheses), functionals by lowercase letters (preferably f) and the value of f at an x by $f(x)$ (with parentheses). This is a widely used practice.

2.1 Vector Space

Vector spaces play a role in many branches of mathematics and its applications. In fact, in various practical (and theoretical) problems we have a set X whose elements may be vectors in three-dimensional space, or sequences of numbers, or functions, and these elements can be added and multiplied by constants (numbers) in a natural way, the result being again an element of X. Such concrete situations suggest the concept of a vector space as defined below. The definition will involve a general field K, but in functional analysis, K will be $\mathbf{R}$ or $\mathbf{C}$. The elements of K are called *scalars;* hence in our case they will be real or complex numbers.

2.1-1 Definition (Vector space). A *vector space* (or *linear space*) *over a field* K is a nonempty set X of elements $x, y, \cdots$ (called *vectors*) together with two algebraic operations. These operations are called *vector addition* and *multiplication of vectors by scalars,* that is, by elements of K.

Vector addition associates with every ordered pair (x, y) of vectors a vector $x+y$, called the *sum* of x *and* y, in such a way that the following properties hold.[1] Vector addition is commutative and associative, that is, for all vectors we have

$$x+y=y+x$$

$$x+(y+z)=(x+y)+z;$$

furthermore, there exists a vector 0, called the *zero vector,* and for every vector x there exists a vector $-x$, such that for all vectors we

[1] Readers familiar with groups will notice that we can summarize the defining properties of vector addition by saying that X is an additive abelian group.

have

$$x+0=x$$

$$x+(-x)=0.$$

Multiplication by scalars associates with every vector x and scalar α a vector αx (also written $x\alpha$), called the *product* of α and x, in such a way that for all vectors x, y and scalars α, β we have

$$\alpha(\beta x)=(\alpha\beta)x$$

$$1x=x$$

and the distributive laws

$$\alpha(x+y)=\alpha x+\alpha y$$

$$(\alpha+\beta)x=\alpha x+\beta x. \qquad \blacksquare$$

From the definition we see that vector addition is a mapping $X\times X\longrightarrow X$, whereas multiplication by scalars is a mapping $K\times X\longrightarrow X$.

K is called the **scalar field** (or *coefficient field*) of the vector space X, and X is called a **real vector space** if $K=\mathbf{R}$ (the field of real numbers), and a **complex vector space** if $K=\mathbf{C}$ (the field of complex numbers[2]).

The use of 0 for the scalar 0 as well as for the zero vector should cause no confusion, in general. If desirable for clarity, we can denote the zero vector by θ.

The reader may prove that for all vectors and scalars,

$$0x=\theta \tag{1a}$$

$$\alpha\theta=\theta \tag{1b}$$

and

$$(-1)x=-x. \tag{2}$$

[2] Remember that **R** and **C** also denote the real line and the complex plane, respectively (cf. 1.1-2 and 1.1-5), but we need not use other letters here since there is little danger of confusion.

Examples

2.1-2 Space $\mathbf{R}^n$. This is the Euclidean space introduced in 1.1-5, the underlying set being the set of all n-tuples of real numbers, written $x=(\xi_1, \cdots, \xi_n)$, $y=(\eta_1, \cdots, \eta_n)$, etc., and we now see that this is a real vector space with the two algebraic operations defined in the usual fashion

$$x+y=(\xi_1+\eta_1, \cdots, \xi_n+\eta_n)$$

$$\alpha x=(\alpha\xi_1, \cdots, \alpha\xi_n) \qquad (\alpha \in \mathbf{R}).$$

The next examples are of a similar nature because in each of them we shall recognize a previously defined space as a vector space.

2.1-3 Space $\mathbf{C}^n$. This space was defined in 1.1-5. It consists of all ordered n-tuples of complex numbers $x=(\xi_1, \cdots, \xi_n)$, $y=(\eta_1, \cdots, \eta_n)$, etc., and is a complex vector space with the algebraic operations defined as in the previous example, where now $\alpha \in \mathbf{C}$.

2.1-4 Space $C[a, b]$. This space was defined in 1.1-7. Each point of this space is a continuous real-valued function on $[a, b]$. The set of all these functions forms a real vector space with the algebraic operations defined in the usual way:

$$(x+y)(t)=x(t)+y(t)$$

$$(\alpha x)(t)=\alpha x(t) \qquad (\alpha \in \mathbf{R}).$$

In fact, $x+y$ and αx are continuous real-valued functions defined on $[a, b]$ if x and y are such functions and α is real.

Other important vector spaces of functions are (a) the vector space $B(A)$ in 1.2-2, (b) the vector space of all differentiable functions on $\mathbf{R}$, and (c) the vector space of all real-valued functions on $[a, b]$ which are integrable in some sense.

2.1-5 Space l^2. This space was introduced in 1.2-3. It is a vector space with the algebraic operations defined as usual in connection with sequences, that is,

$$(\xi_1, \xi_2, \cdots)+(\eta_1, \eta_2, \cdots)=(\xi_1+\eta_1, \xi_2+\eta_2, \cdots)$$

$$\alpha(\xi_1, \xi_2, \cdots)=(\alpha\xi_1, \alpha\xi_2, \cdots).$$

In fact, $x=(\xi_j)\in l^2$ and $y=(\eta_j)\in l^2$ implies $x+y\in l^2$, as follows readily from the Minkowski inequality (12) in Sec. 1.2; also $\alpha x\in l^2$.

Other vector spaces whose points are sequences are l^∞ in 1.1-6, l^p in 1.2-3, where $1\leqq p<+\infty$, and s in 1.2-1. ∎

A **subspace** of a vector space X is a nonempty subset Y of X such that for all $y_1, y_2\in Y$ and all scalars α, β we have $\alpha y_1+\beta y_2\in Y$. Hence Y is itself a vector space, the two algebraic operations being those induced from X.

A special subspace of X is the *improper subspace* $Y=X$. Every other subspace of X ($\neq\{0\}$) is called *proper.*

Another special subspace of any vector space X is $Y=\{0\}$.

A **linear combination** of vectors $x_1, \cdots, x_m$ of a vector space X is an expression of the form

$$\alpha_1x_1+\alpha_2x_2+\cdots+\alpha_mx_m$$

where the coefficients $\alpha_1, \cdots, \alpha_m$ are any scalars.

For any nonempty subset $M\subset X$ the set of all linear combinations of vectors of M is called the **span** of M, written

$$\text{span } M.$$

Obviously, this is a subspace Y of X, and we say that Y is **spanned** or **generated** by M.

We shall now introduce two important related concepts which will be used over and over again.

2.1-6 Definition (Linear independence, linear dependence). Linear independence and dependence of a given set M of vectors $x_1, \cdots, x_r$ $(r\geqq 1)$ in a vector space X are defined by means of the equation

$$\alpha_1x_1+\alpha_2x_2+\cdots+\alpha_rx_r=0, \tag{3}$$

where $\alpha_1, \cdots, \alpha_r$ are scalars. Clearly, equation (3) holds for $\alpha_1=\alpha_2=\cdots=\alpha_r=0$. If this is the only r-tuple of scalars for which (3) holds, the set M is said to be *linearly independent.* M is said to be *linearly dependent* if M is not linearly independent, that is, if (3) also holds for some r-tuple of scalars, not all zero.

An arbitrary subset M of X is said to be *linearly independent* if every nonempty finite subset of M is linearly independent. M is said to be *linearly dependent* if M is not linearly independent. ∎

A motivation for this terminology results from the fact that if $M=\{x_1, \cdots, x_r\}$ is linearly dependent, at least one vector of M can be written as a linear combination of the others; for instance, if (3) holds with an $\alpha_r \neq 0$, then M is linearly dependent and we may solve (3) for x_r to get

$$x_r = \beta_1 x_1 + \cdots + \beta_{r-1} x_{r-1} \qquad (\beta_j = -\alpha_j/\alpha_r).$$

We can use the concepts of linear dependence and independence to define the dimension of a vector space, starting as follows.

2.1-7 Definition (Finite and infinite dimensional vector spaces). A vector space X is said to be *finite dimensional* if there is a positive integer n such that X contains a linearly independent set of n vectors whereas any set of $n+1$ or more vectors of X is linearly dependent. n is called the *dimension* of X, written $n = \dim X$. By definition, $X=\{0\}$ is finite dimensional and $\dim X = 0$. If X is not finite dimensional, it is said to be *infinite dimensional.* ∎

In analysis, infinite dimensional vector spaces are of greater interest than finite dimensional ones. For instance, $C[a, b]$ and l^2 are infinite dimensional, whereas $\mathbf{R}^n$ and $\mathbf{C}^n$ are n-dimensional.

If $\dim X = n$, a linearly independent n-tuple of vectors of X is called a **basis** *for* X (or a *basis in* X). If $\{e_1, \cdots, e_n\}$ is a basis for X, every $x \in X$ has a unique representation as a linear combination of the basis vectors:

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

For instance, a basis for $\mathbf{R}^n$ is

$$\begin{aligned} e_1 &= (1, 0, 0, \cdots, 0), \\ e_2 &= (0, 1, 0, \cdots, 0), \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ e_n &= (0, 0, 0, \cdots, 1). \end{aligned}$$

This is sometimes called the *canonical basis* for $\mathbf{R}^n$.

More generally, if X is any vector space, not necessarily finite dimensional, and B is a linearly independent subset of X which spans

X, then B is called a **basis** (or **Hamel basis**) for X. Hence if B is a basis for X, then every nonzero $x \in X$ has a unique representation as a linear combination of (finitely many!) elements of B with nonzero scalars as coefficients.

Every vector space $X \neq \{0\}$ has a basis.

In the finite dimensional case this is clear. For arbitrary infinite dimensional vector spaces an existence proof will be given by the use of Zorn's lemma. This lemma involves several concepts whose explanation would take us some time and, since at present a number of other things are more important to us, we do not pause but postpone that existence proof to Sec. 4.1, where we must introduce Zorn's lemma for another purpose.

We mention that all bases for a given (finite or infinite dimensional) vector space X have the same cardinal number. (A proof would require somewhat more advanced tools from set theory; cf. M. M. Day (1973), p. 3.) This number is called the **dimension** of X. Note that this includes and extends Def. 2.1-7.

Later we shall need the following simple

2.1-8 Theorem (Dimension of a subspace). *Let X be an n-dimensional vector space. Then any proper subspace Y of X has dimension less than n.*

Proof. If $n=0$, then $X=\{0\}$ and has no proper subspace. If $\dim Y=0$, then $Y=\{0\}$, and $X \neq Y$ implies $\dim X \geqq 1$. Clearly, $\dim Y \leqq \dim X = n$. If $\dim Y$ were n, then Y would have a basis of n elements, which would also be a basis for X since $\dim X = n$, so that $X = Y$. This shows that any linearly independent set of vectors in Y must have fewer than n elements, and $\dim Y < n$. ∎

Problems

1. Show that the set of all real numbers, with the usual addition and multiplication, constitutes a one-dimensional real vector space, and the set of all complex numbers constitutes a one-dimensional complex vector space.

2. Prove (1) and (2).

3. Describe the span of $M = \{(1, 1, 1), (0, 0, 2)\}$ in $\mathbf{R}^3$.

4. Which of the following subsets of $\mathbf{R}^3$ constitute a subspace of $\mathbf{R}^3$? [Here, $x = (\xi_1, \xi_2, \xi_3)$.]
   (a) All x with $\xi_1 = \xi_2$ and $\xi_3 = 0$.
   (b) All x with $\xi_1 = \xi_2 + 1$.
   (c) All x with positive ξ_1, ξ_2, ξ_3.
   (d) All x with $\xi_1 - \xi_2 + \xi_3 = k = const.$

5. Show that $\{x_1, \cdots, x_n\}$, where $x_j(t) = t^j$, is a linearly independent set in the space $C[a, b]$.

6. Show that in an n-dimensional vector space X, the representation of any x as a linear combination of given basis vectors $e_1, \cdots, e_n$ is unique.

7. Let $\{e_1, \cdots, e_n\}$ be a basis for a complex vector space X. Find a basis for X regarded as a real vector space. What is the dimension of X in either case?

8. If M is a linearly dependent set in a complex vector space X, is M linearly dependent in X, regarded as a real vector space?

9. On a fixed interval $[a, b] \subset \mathbf{R}$, consider the set X consisting of all polynomials with real coefficients and of degree not exceeding a given n, and the polynomial $x = 0$ (for which a degree is not defined in the usual discussion of degree). Show that X, with the usual addition and the usual multiplication by real numbers, is a real vector space of dimension $n+1$. Find a basis for X. Show that we can obtain a complex vector space $\tilde{X}$ in a similar fashion if we let those coefficients be complex. Is X a subspace of $\tilde{X}$?

10. If Y and Z are subspaces of a vector space X, show that $Y \cap Z$ is a subspace of X, but $Y \cup Z$ need not be one. Give examples.

11. If $M \neq \varnothing$ is any subset of a vector space X, show that span M is a subspace of X.

12. Show that the set of all real two-rowed square matrices forms a vector space X. What is the zero vector in X? Determine dim X. Find a basis for X. Give examples of subspaces of X. Do the symmetric matrices $x \in X$ form a subspace? The singular matrices?

13. **(Product)** Show that the Cartesian product $X = X_1 \times X_2$ of two vector

spaces over the same field becomes a vector space if we define the two algebraic operations by

$$(x_1, x_2)+(y_1, y_2)=(x_1+y_1, x_2+y_2),$$

$$\alpha(x_1, x_2)=(\alpha x_1, \alpha x_2).$$

14. (Quotient space, codimension) Let Y be a subspace of a vector space X. The *coset* of an element $x \in X$ with respect to Y is denoted by $x+Y$ and is defined to be the set (see Fig. 12)

$$x+Y=\{v \mid v=x+y, y \in Y\}.$$

Show that the distinct cosets form a partition of X. Show that under algebraic operations defined by (see Figs. 13, 14)

$$(w+Y)+(x+Y)=(w+x)+Y$$

$$\alpha(x+Y)=\alpha x+Y$$

these cosets constitute the elements of a vector space. This space is called the *quotient space* (or sometimes *factor space*) *of X by Y* (or *modulo Y*) and is denoted by X/Y. Its dimension is called the *codimension* of Y and is denoted by codim Y, that is,

$$\operatorname{codim} Y=\dim (X/Y).$$

15. Let $X=\mathbf{R}^3$ and $Y=\{(\xi_1, 0, 0) \mid \xi_1 \in \mathbf{R}\}$. Find X/Y, X/X, $X/\{0\}$.

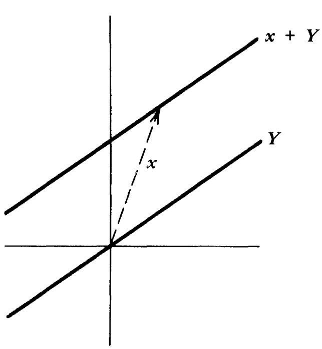

Fig. 12. Illustration of the notation $x+Y$ in Prob. 14

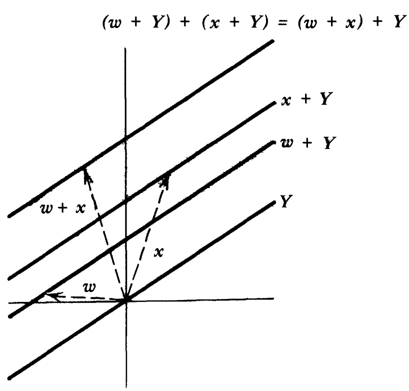

Fig. 13. Illustration of vector addition in a quotient space (cf. Prob. 14)

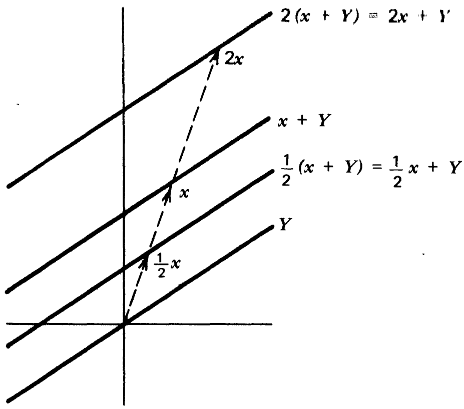

Fig. 14. Illustration of multiplication by scalars in a quotient space (cf. Prob. 14)

2.2 Normed Space. Banach Space

The examples in the last section illustrate that in many cases a vector space X may at the same time be a metric space because a metric d is defined on X. However, if there is no relation between the algebraic structure and the metric, we cannot expect a useful and applicable theory that combines algebraic and metric concepts. To guarantee such a relation between "algebraic" and "geometric" properties of X we define on X a metric d in a special way as follows. We first introduce an auxiliary concept, the *norm* (definition below), which uses the algebraic operations of vector space. Then we employ the norm to obtain a metric d that is of the desired kind. This idea leads to the concept of a *normed space.* It turns out that normed spaces are special enough to provide a basis for a rich and interesting theory, but general enough to include many concrete models of practical importance. In fact, a large number of metric spaces in analysis can be regarded as normed spaces, so that a normed space is probably the most important kind of space in functional analysis, at least from the viewpoint of present-day applications. Here are the definitions:

2.2-1 Definition (Normed space, Banach space). A *normed space*[3] X is a vector space with a norm defined on it. A *Banach space* is a

[3] Also called a *normed vector space* or *normed linear space.* The definition was given (independently) by S. Banach (1922), H. Hahn (1922) and N. Wiener (1922). The theory developed rapidly, as can be seen from the treatise by S. Banach (1932) published only ten years later.

complete normed space (complete in the metric defined by the norm; see (1), below). Here a **norm** on a (real or complex) vector space X is a real-valued function on X whose value at an $x \in X$ is denoted by

$$\|x\| \qquad \text{(read "norm of } x\text{")}$$

and which has the properties

(N1) $$\|x\| \geqq 0$$

(N2) $$\|x\| = 0 \iff x = 0$$

(N3) $$\|\alpha x\| = |\alpha| \, \|x\|$$

(N4) $$\|x+y\| \leqq \|x\| + \|y\| \qquad (\textit{Triangle inequality});$$

here x and y are arbitrary vectors in X and α is any scalar.

A norm on X defines a metric d on X which is given by

(1) $$d(x, y) = \|x - y\| \qquad (x, y \in X)$$

and is called the *metric induced by the norm.* The normed space just defined is denoted by $(X, \|\cdot\|)$ or simply by X. ∎

The defining properties (N1) to (N4) of a norm are suggested and motivated by the length $|x|$ of a vector x in elementary vector algebra, so that in this case we can write $\|x\| = |x|$. In fact, (N1) and (N2) state that all vectors have positive lengths except the zero vector which has length zero. (N3) means that when a vector is multiplied by a scalar, its length is multiplied by the absolute value of the scalar. (N4) is illustrated in Fig. 15. It means that the length of one side of a triangle cannot exceed the sum of the lengths of the two other sides.

It is not difficult to conclude from (N1) to (N4) that (1) does define a metric. Hence normed spaces and Banach spaces are metric spaces.

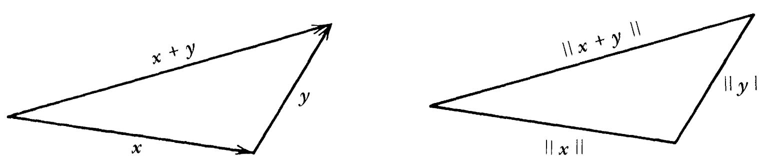

Fig. 15. Illustration of the triangle inequality (N4)

Banach spaces are important because they enjoy certain properties (to be discussed in Chap 4) which are not shared by incomplete normed spaces.

For later use we notice that (N4) implies

$$\big|\|y\|-\|x\|\big| \leqq \|y-x\|, \tag{2}$$

as the reader may readily prove (cf. Prob. 3). Formula (2) implies an important property of the norm:

The norm is continuous, that is, $x \longmapsto \|x\|$ *is a continuous mapping of* $(X, \|\cdot\|)$ *into* $\mathbf{R}$. (Cf. 1.3-3.)

Prototypes of normed spaces are the familiar spaces of all vectors in the plane and in three dimensional space. Further examples result from Secs. 1.1 and 1.2 since some of the metric spaces in those sections can be made into normed spaces in a natural way. However, we shall see later in this section that not every metric on a vector space can be obtained from a norm.

Examples

2.2-2 Euclidean space $\mathbf{R}^n$ and unitary space $\mathbf{C}^n$. These spaces were defined in 1.1-5. They are Banach spaces with norm defined by

$$\|x\| = \left(\sum_{j=1}^{n} |\xi_j|^2\right)^{1/2} = \sqrt{|\xi_1|^2 + \cdots + |\xi_n|^2}. \tag{3}$$

In fact, $\mathbf{R}^n$ and $\mathbf{C}^n$ are complete (cf. 1.5-1), and (3) yields the metric (7) in Sec. 1.1:

$$d(x, y) = \|x-y\| = \sqrt{|\xi_1-\eta_1|^2 + \cdots + |\xi_n-\eta_n|^2}.$$

We note in particular that in $\mathbf{R}^3$ we have

$$\|x\| = |x| = \sqrt{\xi_1^{\,2} + \xi_2^{\,2} + \xi_3^{\,2}}.$$

This confirms our previous remark that the norm generalizes the elementary notion of the length $|x|$ of a vector.

2.2-3 Space l^p. This space was defined in 1.2-3. It is a Banach space with norm given by

$$\|x\| = \left(\sum_{j=1}^{\infty} |\xi_j|^p\right)^{1/p}. \tag{4}$$

In fact, this norm induces the metric in 1.2-3:

$$d(x, y) = \|x - y\| = \left(\sum_{j=1}^{\infty} |\xi_j - \eta_j|^p\right)^{1/p}.$$

Completeness was shown in 1.5-4.

2.2-4 Space l^∞. This space was defined in 1.1-6 and is a Banach space since its metric is obtained from the norm defined by

$$\|x\| = \sup_j |\xi_j|$$

and completeness was shown in 1.5-2.

2.2-5 Space $C[a, b]$. This space was defined in 1.1-7 and is a Banach space with norm given by

$$\|x\| = \max_{t \in J} |x(t)| \tag{5}$$

where $J = [a, b]$. Completeness was shown in 1.5-5.

2.2-6 Incomplete normed spaces. From the incomplete metric spaces in 1.5-7, 1.5-8 and 1.5-9 we may readily obtain incomplete normed spaces. For instance, the metric in 1.5-9 is induced by the norm defined by

$$\|x\| = \int_0^1 |x(t)|\, dt. \tag{6}$$

Can every incomplete normed space be completed? As a metric space certainly by 1.6-2. But what about extending the operations of a vector space and the norm to the completion? We shall see in the next section that the extension is indeed possible.

2.2-7 An incomplete normed space and its completion $L^2[a, b]$. The vector space of all continuous real-valued functions on $[a, b]$ forms a normed space X with norm defined by

$$\|x\| = \left(\int_a^b x(t)^2 \, dt \right)^{1/2}. \tag{7}$$

This space is not complete. For instance, if $[a, b] = [0, 1]$, the sequence in 1.5-9 is also Cauchy in the present space X; this is almost obvious from Fig. 10, Sec. 1.5, and results formally by integration because for $n > m$ we obtain

$$\|x_n - x_m\|^2 = \int_0^1 [x_n(t) - x_m(t)]^2 \, dt = \frac{(n-m)^2}{3mn^2} < \frac{1}{3m} - \frac{1}{3n}.$$

This Cauchy sequence does not converge. The proof is the same as in 1.5-9, with the metric in 1.5-9 replaced by the present metric. For a general interval $[a, b]$ we can construct a similar Cauchy sequence which does not converge in X.

The space X can be completed by Theorem 1.6-2. The completion is denoted by $L^2[a, b]$. This is a Banach space. In fact, the norm on X and the operations of vector space can be extended to the completion of X, as we shall see from Theorem 2.3-2 in the next section.

More generally, for any fixed real number $p \geqq 1$, the Banach space

$$L^p[a, b]$$

is the completion of the normed space which consists of all continuous real-valued functions on $[a, b]$, as before, and the norm defined by

$$\|x\|_p = \left(\int_a^b |x(t)|^p \, dt \right)^{1/p}. \tag{8}$$

The subscript p is supposed to remind us that this norm depends on the choice of p, which is kept fixed. Note that for $p = 2$ this equals (7).

For readers familiar with the Lebesgue integral we want to mention that the space $L^p[a, b]$ can also be obtained in a direct way by the use of the Lebesgue integral and Lebesgue measurable functions x on $[a, b]$ such that the Lebesgue integral of $|x|^p$ over $[a, b]$ exists and is finite. The elements of $L^p[a, b]$ are equivalence classes of those functions, where x is equivalent to y if the Lebesgue integral of $|x - y|^p$

over $[a, b]$ is zero. [Note that this guarantees the validity of axiom (N2).]

Readers without that background should not be disturbed. In fact, this example is not essential to the later development. At any rate, the example illustrates that completion may lead to a new kind of elements and one may have to find out what their nature is.

2.2-8 Space s. Can every metric on a vector space be obtained from a norm? The answer is no. A counterexample is the space s in 1.2-1. In fact, s is a vector space, but its metric d defined by

$$d(x, y) = \sum_{j=1}^{\infty} \frac{1}{2^j} \frac{|\xi_j - \eta_j|}{1 + |\xi_j - \eta_j|}$$

cannot be obtained from a norm. This may immediately be seen from the following lemma which states two basic properties of a metric d obtained from a norm. The first property, as expressed by (9a), is called the *translation invariance* of d.

2.2-9 Lemma (Translation invariance). *A metric d induced by a norm on a normed space X satisfies*

$$\text{(9)}\quad \begin{aligned} &\text{(a)} \quad d(x+a, y+a) = d(x, y) \\ &\text{(b)} \quad d(\alpha x, \alpha y) = |\alpha|\, d(x, y) \end{aligned}$$

for all $x, y, a \in X$ and every scalar α.

Proof. We have

$$d(x+a, y+a) = \|x+a-(y+a)\| = \|x-y\| = d(x, y)$$

and

$$d(\alpha x, \alpha y) = \|\alpha x - \alpha y\| = |\alpha|\, \|x-y\| = |\alpha|\, d(x, y). \qquad \blacksquare$$

Problems

1. Show that the norm $\|x\|$ of x is the distance from x to 0.

2. Verify that the usual length of a vector in the plane or in three dimensional space has the properties (N1) to (N4) of a norm.

3. Prove (2).

4. Show that we may replace (N2) by

$$\|x\|=0 \qquad \Longrightarrow \qquad x=0$$

without altering the concept of a norm. Show that nonnegativity of a norm also follows from (N3) and (N4).

5. Show that (3) defines a norm.

6. Let X be the vector space of all ordered pairs $x=(\xi_1, \xi_2)$, $y=(\eta_1, \eta_2), \cdots$ of real numbers. Show that norms on X are defined by

$$\|x\|_1=|\xi_1|+|\xi_2|$$

$$\|x\|_2=(\xi_1^{\;2}+\xi_2^{\;2})^{1/2}$$

$$\|x\|_\infty=\max\{|\xi_1|, |\xi_2|\}.$$

7. Verify that (4) satisfies (N1) to (N4).

8. There are several norms of practical importance on the vector space of ordered n-tuples of numbers (cf. 2.2-2), notably those defined by

$$\|x\|_1=|\xi_1|+|\xi_2|+\cdots+|\xi_n|$$

$$\|x\|_p=(|\xi_1|^p+|\xi_2|^p+\cdots+|\xi_n|^p)^{1/p} \qquad (1<p<+\infty)$$

$$\|x\|_\infty=\max\{|\xi_1|, \cdots, |\xi_n|\}.$$

In each case, verify that (N1) to (N4) are satisfied.

9. Verify that (5) defines a norm.

10. **(Unit sphere)** The sphere

$$S(0; 1)=\{x \in X \mid \|x\|=1\}$$

in a normed space X is called the *unit sphere.* Show that for the norms in Prob. 6 and for the norm defined by

$$\|x\|_4=(\xi_1^{\;4}+\xi_2^{\;4})^{1/4}$$

the unit spheres look as shown in Fig. 16.

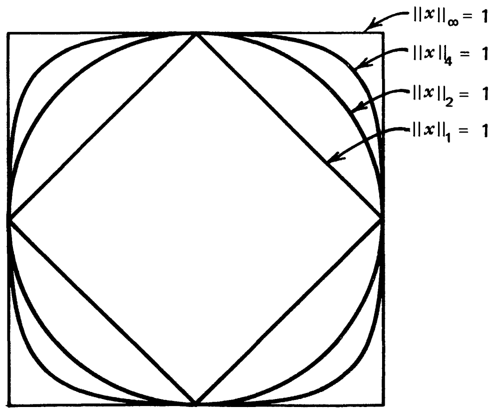

Fig. 16. Unit spheres in Prob. 10

11. (Convex set, segment) A subset A of a vector space X is said to be *convex* if $x, y \in A$ implies

$$M = \{z \in X \mid z = \alpha x + (1-\alpha)y, \quad 0 \leqq \alpha \leqq 1\} \subset A.$$

M is called a *closed segment* with *boundary points* x and y; any other $z \in M$ is called an *interior point* of M. Show that the *closed unit ball*

$$\tilde{B}(0; 1) = \{x \in X \mid \|x\| \leqq 1\}$$

in a normed space X is convex.

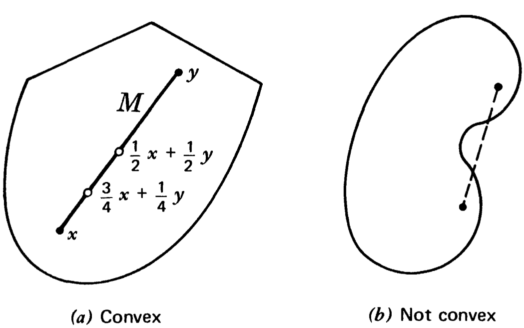

Fig. 17. Illustrative examples of convex and nonconvex sets (cf. Prob. 11)

12. Using Prob. 11, show that

$$\varphi(x) = (\sqrt{|\xi_1|} + \sqrt{|\xi_2|})^2$$

does not define a norm on the vector space of all ordered pairs $x = (\xi_1, \xi_2), \cdots$ of real numbers. Sketch the curve $\varphi(x) = 1$ and compare it with Fig. 18.

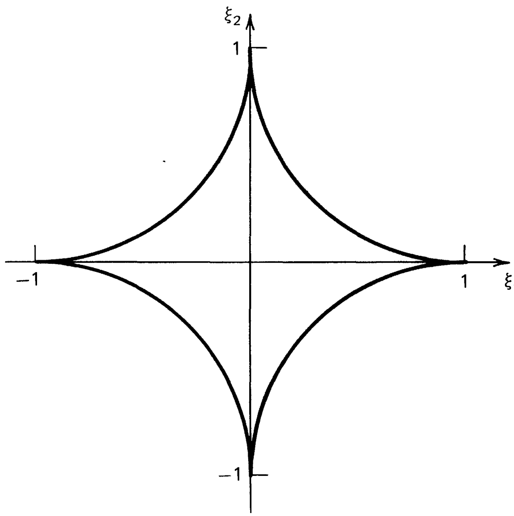

Fig. 18. Curve $\varphi(x) = 1$ in Prob. 12

13. Show that the discrete metric on a vector space $X \neq \{0\}$ cannot be obtained from a norm. (Cf. 1.1-8.)

14. If d is a metric on a vector space $X \neq \{0\}$ which is obtained from a norm, and $\tilde{d}$ is defined by

$$\tilde{d}(x, x) = 0, \qquad \tilde{d}(x, y) = d(x, y) + 1 \qquad (x \neq y),$$

show that $\tilde{d}$ cannot be obtained from a norm.

15. (Bounded set) Show that a subset M in a normed space X is bounded if and only if there is a positive number c such that $\|x\| \leqq c$ for every $x \in M$. (For the definition, see Prob. 6 in Sec. 1.2.)

2.3 Further Properties of Normed Spaces

By definition, a **subspace** Y of a normed space X is a subspace of X considered as a vector space, with the norm obtained by restricting the norm on X to the subset Y. This norm on Y is said to be *induced* by the norm on X. If Y is closed in X, then Y is called a **closed subspace** of X.

By definition, a **subspace** Y of a Banach space X is a subspace of X considered as a normed space. Hence we do *not* require Y to be complete. (Some writers do, so be careful when comparing books.)

In this connection, Theorem 1.4-7 is useful since it yields immediately the following

2.3-1 Theorem (Subspace of a Banach space). *A subspace Y of a Banach space X is complete if and only if the set Y is closed in X.*

Convergence of sequences and related concepts in normed spaces follow readily from the corresponding definitions 1.4-1 and 1.4-3 for metric spaces and the fact that now $d(x, y) = \|x - y\|$:

(*i*) A sequence (x_n) in a normed space X is *convergent* if X contains an x such that

$$\lim_{n\to\infty} \|x_n - x\| = 0.$$

Then we write $x_n \longrightarrow x$ and call x the *limit* of (x_n).

(*ii*) A sequence (x_n) in a normed space X is *Cauchy* if for every $\varepsilon > 0$ there is an N such that

$$\|x_m - x_n\| < \varepsilon \qquad \text{for all } m, n > N. \tag{1}$$

Sequences were available to us even in a general metric space. In a normed space we may go an important step further and use series as follows.

Infinite series can now be defined in a way similar to that in calculus. In fact, if (x_k) is a sequence in a normed space X, we can associate with (x_k) the sequence (s_n) of *partial sums*

$$s_n = x_1 + x_2 + \cdots + x_n$$

where $n = 1, 2, \cdots$. If (s_n) is convergent, say,

$$s_n \longrightarrow s, \qquad \text{that is,} \qquad \|s_n - s\| \longrightarrow 0,$$

then the *infinite series* or, briefly, *series*

$$\sum_{k=1}^{\infty} x_k = x_1 + x_2 + \cdots \tag{2}$$

is said to *converge* or to *be* **convergent**, s is called the *sum* of the series and we write

$$s = \sum_{k=1}^{\infty} x_k = x_1 + x_2 + \cdots .$$

If $\|x_1\| + \|x_2\| + \cdots$ converges, the series (2) is said to be **absolutely convergent**. However, we warn the reader that in a normed space X, absolute convergence implies convergence if and only if X is complete (cf. Probs. 7 to 9).

The concept of convergence of a series can be used to define a "basis" as follows. If a normed space X contains a sequence (e_n) with the property that for every $x \in X$ there is a unique sequence of scalars (α_n) such that

$$\|x - (\alpha_1 e_1 + \cdots + \alpha_n e_n)\| \longrightarrow 0 \qquad (\text{as } n \longrightarrow \infty) \tag{3}$$

then (e_n) is called a **Schauder basis** (or *basis*) for X. The series

$$\sum_{k=1}^{\infty} \alpha_k e_k$$

which has the sum x is then called the *expansion* of x with respect to (e_n), and we write

$$x = \sum_{k=1}^{\infty} \alpha_k e_k.$$

For example, l^p in 2.2-3 has a Schauder basis, namely (e_n), where $e_n = (\delta_{nj})$, that is, e_n is the sequence whose nth term is 1 and all other terms are zero; thus

$$
\begin{aligned}
e_1 &= (1, 0, 0, 0, \cdots) \\
e_2 &= (0, 1, 0, 0, \cdots) \qquad (4) \\
e_3 &= (0, 0, 1, 0, \cdots)
\end{aligned}
$$

etc.

If a normed space X has a Schauder basis, then X is separable (cf. Def. 1.3-5). The proof is simple, so that we can leave it to the reader (Prob. 10). Conversely, does *every* separable Banach space have a Schauder basis? This is a famous question raised by Banach himself about forty years ago. Almost all known separable Banach spaces had been shown to possess a Schauder basis. Nevertheless, the surprising answer to the question is no. It was given only quite recently, by P. Enflo (1973) who was able to construct a separable Banach space which has no Schauder basis.

Let us finally turn to the problem of completing a normed space, which was briefly mentioned in the last section.

2.3-2 Theorem (Completion). *Let $X=(X, \|\cdot\|)$ be a normed space. Then there is a Banach space $\hat{X}$ and an isometry A from X onto a subspace W of $\hat{X}$ which is dense in $\hat{X}$. The space $\hat{X}$ is unique, except for isometries.*

Proof. Theorem 1.6-2 implies the existence of a complete metric space $\hat{X}=(\hat{X}, \hat{d})$ and an isometry $A: X \longrightarrow W=A(X)$, where W is dense in $\hat{X}$ and $\hat{X}$ is unique, except for isometries. (We write A, not T as in 1.6-2, to free the letter T for later applications of the theorem in Sec. 8.2) Consequently, to prove the present theorem, we must make $\hat{X}$ into a vector space and then introduce on $\hat{X}$ a suitable norm.

To define on $\hat{X}$ the two algebraic operations of a vector space, we consider any $\hat{x}, \hat{y} \in \hat{X}$ and any representatives $(x_n) \in \hat{x}$ and $(y_n) \in \hat{y}$. Remember that $\hat{x}$ and $\hat{y}$ are equivalence classes of Cauchy sequences in X. We set $z_n = x_n + y_n$. Then (z_n) is Cauchy in X since

$$\|z_n - z_m\| = \|x_n + y_n - (x_m + y_m)\| \leqq \|x_n - x_m\| + \|y_n - y_m\|.$$

We define the sum $\hat{z} = \hat{x} + \hat{y}$ of $\hat{x}$ and $\hat{y}$ to be the equivalence class for which (z_n) is a representative; thus $(z_n) \in \hat{z}$. This definition is independent of the particular choice of Cauchy sequences belonging to $\hat{x}$ and $\hat{y}$. In fact, (1) in Sec. 1.6 shows that if $(x_n) \sim (x_n')$ and $(y_n) \sim (y_n')$, then

$(x_n+y_n) \sim (x_n'+y_n')$ because

$$\|x_n+y_n-(x_n'+y_n')\| \leqq \|x_n-x_n'\|+\|y_n-y_n'\|.$$

Similarly we define the product $\alpha\hat{x} \in \hat{X}$ of a scalar α and $\hat{x}$ to be the equivalence class for which (αx_n) is a representative. Again, this definition is independent of the particular choice of a representative of $\hat{x}$. The zero element of $\hat{X}$ is the equivalence class containing all Cauchy sequences which converge to zero. It is not difficult to see that those two algebraic operations have all the properties required by the definition, so that $\hat{X}$ is a vector space. From the definition it follows that on W the operations of vector space induced from $\hat{X}$ agree with those induced from X by means of A.

Furthermore, A induces on W a norm $\|\cdot\|_1$, whose value at every $\hat{y}=Ax \in W$ is $\|\hat{y}\|_1=\|x\|$. The corresponding metric on W is the restriction of $\hat{d}$ to W since A is isometric. We can extend the norm $\|\cdot\|_1$ to $\hat{X}$ by setting $\|\hat{x}\|_2=\hat{d}(\hat{0}, \hat{x})$ for every $\hat{x} \in \hat{X}$. In fact, it is obvious that $\|\cdot\|_2$ satisfies (N1) and (N2) in Sec. 2.2, and the other two axioms (N3) and (N4) follow from those for $\|\cdot\|_1$ by a limit process. ▮

Problems

1. Show that $c \subset l^\infty$ is a vector subspace of l^∞ (cf. 1.5-3) and so is c_0, the space of all sequences of scalars converging to zero.

2. Show that c_0 in Prob. 1 is a *closed* subspace of l^∞, so that c_0 is complete by 1.5-2 and 1.4-7.

3. In l^∞, let Y be the subset of all sequences with only finitely many nonzero terms. Show that Y is a subspace of l^∞ but not a closed subspace.

4. (Continuity of vector space operations) Show that in a normed space X, vector addition and multiplication by scalars are continuous operations with respect to the norm; that is, the mappings defined by $(x, y) \longmapsto x+y$ and $(\alpha, x) \longmapsto \alpha x$ are continuous.

5. Show that $x_n \longrightarrow x$ and $y_n \longrightarrow y$ implies $x_n+y_n \longrightarrow x+y$. Show that $\alpha_n \longrightarrow \alpha$ and $x_n \longrightarrow x$ implies $\alpha_n x_n \longrightarrow \alpha x$.

6. Show that the closure $\bar{Y}$ of a subspace Y of a normed space X is again a vector subspace.

7. (Absolute convergence) Show that convergence of $\|y_1\|+\|y_2\|+\|y_3\|+\cdots$ may not imply convergence of $y_1+y_2+y_3+\cdots$. *Hint.* Consider Y in Prob. 3 and (y_n), where $y_n=(\eta_j^{(n)})$, $\eta_n^{(n)}=1/n^2$, $\eta_j^{(n)}=0$ for all $j\neq n$.

8. If in a normed space X, absolute convergence of any series always implies convergence of that series, show that X is complete.

9. Show that in a Banach space, an absolutely convergent series is convergent.

10. (Schauder basis) Show that if a normed space has a Schauder basis, it is separable.

11. Show that (e_n), where $e_n=(\delta_{nj})$, is a Schauder basis for l^p, where $1\leqq p<+\infty$.

12. (Seminorm) A *seminorm* on a vector space X is a mapping $p\colon X\longrightarrow \mathbf{R}$ satisfying (N1), (N3), (N4) in Sec. 2.2. (Some authors call this a *pseudonorm.*) Show that

$$p(0)=0,$$

$$|p(y)-p(x)|\leqq p(y-x).$$

(Hence if $p(x)=0$ implies $x=0$, then p is a norm.)

13. Show that in Prob. 12, the elements $x\in X$ such that $p(x)=0$ form a subspace N of X and a norm on X/N (cf. Prob. 14, Sec. 2.1) is defined by $\|\hat{x}\|_0=p(x)$, where $x\in\hat{x}$ and $\hat{x}\in X/N$.

14. (Quotient space) Let Y be a closed subspace of a normed space $(X, \|\cdot\|)$. Show that a norm $\|\cdot\|_0$ on X/Y (cf. Prob. 14, Sec. 2.1) is defined by

$$\|\hat{x}\|_0=\inf_{x\in\hat{x}}\|x\|$$

where $\hat{x}\in X/Y$, that is, $\hat{x}$ is any coset of Y.

15. (Product of normed spaces) If $(X_1, \|\cdot\|_1)$ and $(X_2, \|\cdot\|_2)$ are normed spaces, show that the product vector space $X=X_1\times X_2$ (cf. Prob. 13, Sec. 2.1) becomes a normed space if we define

$$\|x\|=\max\,(\|x_1\|_1, \|x_2\|_2) \qquad [x=(x_1, x_2)].$$

2.4 Finite Dimensional Normed Spaces and Subspaces

Are finite dimensional normed spaces simpler than infinite dimensional ones? In what respect? These questions are rather natural. They are important since finite dimensional spaces and subspaces play a role in various considerations (for instance, in approximation theory and spectral theory). Quite a number of interesting things can be said in this connection. Hence it is worthwhile to collect some relevant facts, for their own sake and as tools for our further work. This is our program in this section and the next one.

A source for results of the desired type is the following lemma. Very roughly speaking it states that in the case of linear independence of vectors we cannot find a linear combination that involves large scalars but represents a small vector.

2.4-1 Lemma (Linear combinations). *Let $\{x_1, \cdots, x_n\}$ be a linearly independent set of vectors in a normed space X (of any dimension). Then there is a number $c>0$ such that for every choice of scalars $\alpha_1, \cdots, \alpha_n$ we have*

$$\|\alpha_1 x_1 + \cdots + \alpha_n x_n\| \geqq c(|\alpha_1| + \cdots + |\alpha_n|) \qquad (c>0). \tag{1}$$

Proof. We write $s = |\alpha_1| + \cdots + |\alpha_n|$. If $s=0$, all α_j are zero, so that (1) holds for any c. Let $s>0$. Then (1) is equivalent to the inequality which we obtain from (1) by dividing by s and writing $\beta_j = \alpha_j/s$, that is,

$$\|\beta_1 x_1 + \cdots + \beta_n x_n\| \geqq c \qquad \left(\sum_{j=1}^{n} |\beta_j| = 1\right). \tag{2}$$

Hence it suffices to prove the existence of a $c>0$ such that (2) holds for every n-tuple of scalars $\beta_1, \cdots, \beta_n$ with $\sum |\beta_j| = 1$.

Suppose that this is false. Then there exists a sequence (y_m) of vectors

$$y_m = \beta_1^{(m)} x_1 + \cdots + \beta_n^{(m)} x_n \qquad \left(\sum_{j=1}^{n} |\beta_j^{(m)}| = 1\right)$$

such that

$$\|y_m\| \longrightarrow 0 \qquad \text{as } m \longrightarrow \infty.$$

Now we reason as follows. Since $\sum |\beta_j^{(m)}| = 1$, we have $|\beta_j^{(m)}| \leqq 1$. Hence for each fixed j the sequence

$$(\beta_j^{(m)}) = (\beta_j^{(1)}, \beta_j^{(2)}, \cdots)$$

is bounded. Consequently, by the Bolzano-Weierstrass theorem, $(\beta_1^{(m)})$ has a convergent subsequence. Let β_1 denote the limit of that subsequence, and let $(y_{1,m})$ denote the corresponding subsequence of (y_m). By the same argument, $(y_{1,m})$ has a subsequence $(y_{2,m})$ for which the corresponding subsequence of scalars $\beta_2^{(m)}$ converges; let β_2 denote the limit. Continuing in this way, after n steps we obtain a subsequence $(y_{n,m}) = (y_{n,1}, y_{n,2}, \cdots)$ of (y_m) whose terms are of the form

$$y_{n,m} = \sum_{j=1}^{n} \gamma_j^{(m)} x_j \qquad \left(\sum_{j=1}^{n} |\gamma_j^{(m)}| = 1\right)$$

with scalars $\gamma_j^{(m)}$ satisfying $\gamma_j^{(m)} \longrightarrow \beta_j$ as $m \longrightarrow \infty$. Hence, as $m \longrightarrow \infty$,

$$y_{n,m} \longrightarrow y = \sum_{j=1}^{n} \beta_j x_j$$

where $\sum |\beta_j| = 1$, so that not all β_j can be zero. Since $\{x_1, \cdots, x_n\}$ is a linearly independent set, we thus have $y \neq 0$. On the other hand, $y_{n,m} \longrightarrow y$ implies $\|y_{n,m}\| \longrightarrow \|y\|$, by the continuity of the norm. Since $\|y_m\| \longrightarrow 0$ by assumption and $(y_{n,m})$ is a subsequence of (y_m), we must have $\|y_{n,m}\| \longrightarrow 0$. Hence $\|y\| = 0$, so that $y = 0$ by (N2) in Sec. 2.2. This contradicts $y \neq 0$, and the lemma is proved. ∎

As a first application of this lemma, let us prove the basic

2.4-2 Theorem (Completeness). *Every finite dimensional subspace Y of a normed space X is complete. In particular, every finite dimensional normed space is complete.*

Proof. We consider an arbitrary Cauchy sequence (y_m) in Y and show that it is convergent in Y; the limit will be denoted by y. Let $\dim Y = n$ and $\{e_1, \cdots, e_n\}$ any basis for Y. Then each y_m has a unique representation of the form

$$y_m = \alpha_1^{(m)} e_1 + \cdots + \alpha_n^{(m)} e_n.$$

Since (y_m) is a Cauchy sequence, for every $\varepsilon>0$ there is an N such that $\|y_m-y_r\|<\varepsilon$ when $m, r>N$. From this and Lemma 2.4-1 we have for some $c>0$

$$\varepsilon>\|y_m-y_r\|=\left\|\sum_{j=1}^{n}(\alpha_j^{(m)}-\alpha_j^{(r)})e_j\right\|\geqq c\sum_{j=1}^{n}|\alpha_j^{(m)}-\alpha_j^{(r)}|,$$

where $m, r>N$. Division by $c>0$ gives

$$|\alpha_j^{(m)}-\alpha_j^{(r)}|\leqq\sum_{j=1}^{n}|\alpha_j^{(m)}-\alpha_j^{(r)}|<\frac{\varepsilon}{c}\qquad (m, r>N).$$

This shows that each of the n sequences

$$(\alpha_j^{(m)})=(\alpha_j^{(1)}, \alpha_j^{(2)}, \cdots)\qquad j=1,\cdots,n$$

is Cauchy in $\mathbf{R}$ or $\mathbf{C}$. Hence it converges; let α_j denote the limit. Using these n limits $\alpha_1, \cdots, \alpha_n$, we define

$$y=\alpha_1e_1+\cdots+\alpha_ne_n.$$

Clearly, $y\in Y$. Furthermore,

$$\|y_m-y\|=\left\|\sum_{j=1}^{n}(\alpha_j^{(m)}-\alpha_j)e_j\right\|\leqq\sum_{j=1}^{n}|\alpha_j^{(m)}-\alpha_j|\,\|e_j\|.$$

On the right, $\alpha_j^{(m)}\longrightarrow\alpha_j$. Hence $\|y_m-y\|\longrightarrow 0$, that is, $y_m\longrightarrow y$. This shows that (y_m) is convergent in Y. Since (y_m) was an arbitrary Cauchy sequence in Y, this proves that Y is complete. ∎

From this theorem and Theorem 1.4-7 we have

2.4-3 Theorem (Closedness). *Every finite dimensional subspace Y of a normed space X is closed in X.*

We shall need this theorem at several occasions in our further work.

Note that infinite dimensional subspaces need not be closed. *Example.* Let $X=C[0,1]$ and $Y=\operatorname{span}(x_0, x_1, \cdots)$, where $x_j(t)=t^j$, so that Y is the set of all polynomials. Y is not closed in X. (Why?)

Another interesting property of a finite dimensional vector space X is that all norms on X lead to the same topology for X (cf. Sec. 1.3), that is, the open subsets of X are the same, regardless of the particular choice of a norm on X. The details are as follows.

2.4-4 Definition (Equivalent norms). A norm $\|\cdot\|$ on a vector space X is said to be *equivalent* to a norm $\|\cdot\|_0$ on X if there are positive numbers a and b such that for all $x \in X$ we have

$$a\|x\|_0 \leqq \|x\| \leqq b\|x\|_0. \tag{3}$$ ∎

This concept is motivated by the following fact.

Equivalent norms on X define the same topology for X.

Indeed, this follows from (3) and the fact that every nonempty open set is a union of open balls (cf. Prob. 4, Sec. 1.3). We leave the details of a formal proof to the reader (Prob. 4), who may also show that the Cauchy sequences in $(X, \|\cdot\|)$ and $(X, \|\cdot\|_0)$ are the same (Prob. 5).

Using Lemma 2.4-1, we can now prove the following theorem (which does *not* hold for infinite dimensional spaces).

2.4-5 Theorem (Equivalent norms). *On a finite dimensional vector space X, any norm $\|\cdot\|$ is equivalent to any other norm $\|\cdot\|_0$.*

Proof. Let $\dim X = n$ and $\{e_1, \cdots, e_n\}$ any basis for X. Then every $x \in X$ has a unique representation

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

By Lemma 2.4-1 there is a positive constant c such that

$$\|x\| \geqq c(|\alpha_1| + \cdots + |\alpha_n|).$$

On the other hand the triangle inequality gives

$$\|x\|_0 \leqq \sum_{j=1}^{n} |\alpha_j| \|e_j\|_0 \leqq k \sum_{j=1}^{n} |\alpha_j|, \qquad k = \max_j \|e_j\|_0.$$

Together, $a\|x\|_0 \leqq \|x\|$ where $a = c/k > 0$. The other inequality in (3) is now obtained by an interchange of the roles of $\|\cdot\|$ and $\|\cdot\|_0$ in the preceding argument. ∎

This theorem is of considerable practical importance. For instance, it implies that convergence or divergence of a sequence in a finite dimensional vector space does not depend on the particular choice of a norm on that space.

Problems

1. Give examples of subspaces of l^∞ and l^2 which are not closed.

2. What is the largest possible c in (1) if $X = \mathbf{R}^2$ and $x_1 = (1, 0)$, $x_2 = (0, 1)$? If $X = \mathbf{R}^3$ and $x_1 = (1, 0, 0)$, $x_2 = (0, 1, 0)$, $x_3 = (0, 0, 1)$?

3. Show that in Def. 2.4-4 the axioms of an equivalence relation hold (cf. A1.4 in Appendix 1).

4. Show that equivalent norms on a vector space X induce the same topology for X.

5. If $\|\cdot\|$ and $\|\cdot\|_0$ are equivalent norms on X, show that the Cauchy sequences in $(X, \|\cdot\|)$ and $(X, \|\cdot\|_0)$ are the same.

6. Theorem 2.4-5 implies that $\|\cdot\|_2$ and $\|\cdot\|_\infty$ in Prob. 8, Sec. 2.2, are equivalent. Give a direct proof of this fact.

7. Let $\|\cdot\|_2$ be as in Prob. 8, Sec. 2.2, and let $\|\cdot\|$ be any norm on that vector space, call it X. Show directly (without using 2.4-5) that there is a $b > 0$ such that $\|x\| \leqq b\,\|x\|_2$ for all x.

8. Show that the norms $\|\cdot\|_1$ and $\|\cdot\|_2$ in Prob. 8, Sec. 2.2, satisfy

$$\frac{1}{\sqrt{n}}\|x\|_1 \leqq \|x\|_2 \leqq \|x\|_1.$$

9. If two norms $\|\cdot\|$ and $\|\cdot\|_0$ on a vector space X are equivalent, show that (*i*) $\|x_n - x\| \longrightarrow 0$ implies (*ii*) $\|x_n - x\|_0 \longrightarrow 0$ (and vice versa, of course).

10. Show that all complex $m \times n$ matrices $A = (\alpha_{jk})$ with fixed m and n constitute an mn-dimensional vector space Z. Show that all norms on Z are equivalent. What would be the analogues of $\|\cdot\|_1$, $\|\cdot\|_2$ and $\|\cdot\|_\infty$ in Prob. 8, Sec. 2.2, for the present space Z?

2.5 Compactness and Finite Dimension

A few other basic properties of finite dimensional normed spaces and subspaces are related to the concept of compactness. The latter is defined as follows.

2.5-1 Definition (Compactness). A metric space X is said to be *compact*[4] if every sequence in X has a convergent subsequence. A subset M of X is said to be *compact* if M is compact considered as a subspace of X, that is, if every sequence in M has a convergent subsequence whose limit is an element of M. ∎

A general property of compact sets is expressed in

2.5-2 Lemma (Compactness). *A compact subset M of a metric space is closed and bounded.*

Proof. For every $x \in \bar{M}$ there is a sequence (x_n) in M such that $x_n \longrightarrow x$; cf. 1.4-6(*a*). Since M is compact, $x \in M$. Hence M is closed because $x \in \bar{M}$ was arbitrary. We prove that M is bounded. If M were unbounded, it would contain an unbounded sequence (y_n) such that $d(y_n, b) > n$, where b is any fixed element. This sequence could not have a convergent subsequence since a convergent subsequence must be bounded, by Lemma 1.4-2. ∎

The converse of this lemma is in general false.

Proof. To prove this important fact, we consider the sequence (e_n) in l^2, where $e_n = (\delta_{nj})$ has the nth term 1 and all other terms 0; cf. (4), Sec. 2.3. This sequence is bounded since $\|e_n\| = 1$. Its terms constitute a point set which is closed because it has no point of accumulation. For the same reason, that point set is not compact. ∎

However, for a finite dimensional normed space we have

2.5-3 Theorem (Compactness). *In a finite dimensional normed space X, any subset $M \subset X$ is compact if and only if M is closed and bounded.*

[4] More precisely, *sequentially compact;* this is the most important kind of compactness in analysis. We mention that there are two other kinds of compactness, but for metric spaces the three concepts become identical, so that the distinction does not matter in our work. (The interested reader will find some further remarks in A1.5. Appendix 1.)

Proof. Compactness implies closedness and boundedness by Lemma 2.5-2, and we prove the converse. Let M be closed and bounded. Let $\dim X = n$ and $\{e_1, \cdots, e_n\}$ a basis for X. We consider any sequence (x_m) in M. Each x_m has a representation

$$x_m = \xi_1^{(m)} e_1 + \cdots + \xi_n^{(m)} e_n.$$

Since M is bounded, so is (x_m), say, $\|x_m\| \leqq k$ for all m. By Lemma 2.4-1,

$$k \geqq \|x_m\| = \left\| \sum_{j=1}^{n} \xi_j^{(m)} e_j \right\| \geqq c \sum_{j=1}^{n} |\xi_j^{(m)}|$$

where $c > 0$. Hence the sequence of numbers $(\xi_j^{(m)})$ (j fixed) is bounded and, by the Bolzano-Weierstrass theorem, has a point of accumulation ξ_j; here $1 \leqq j \leqq n$. As in the proof of Lemma 2.4-1 we conclude that (x_m) has a subsequence (z_m) which converges to $z = \sum \xi_j e_j$. Since M is closed, $z \in M$. This shows that the arbitrary sequence (x_m) in M has a subsequence which converges in M. Hence M is compact. ∎

Our discussion shows the following. In $\mathbf{R}^n$ (or in any other finite dimensional normed space) the compact subsets are precisely the closed and bounded subsets, so that this property (closedness and boundedness) can be used for *defining* compactness. However, this can no longer be done in the case of an infinite dimensional normed space.

A source of other interesting results is the following lemma by F. Riesz (1918, pp. 75–76).

2.5-4 F. Riesz's Lemma. *Let Y and Z be subspaces of a normed space X (of any dimension), and suppose that Y is closed and is a proper subset of Z. Then for every real number θ in the interval $(0, 1)$ there is a $z \in Z$ such that*

$$\|z\| = 1, \qquad \|z - y\| \geqq \theta \text{ for all } y \in Y.$$

Proof. We consider any $v \in Z - Y$ and denote its distance from Y by a, that is (Fig. 19),

$$a = \inf_{y \in Y} \|v - y\|.$$

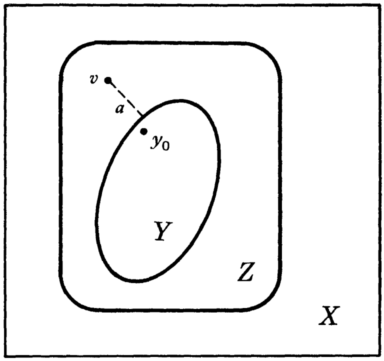

Fig. 19. Notations in the proof of Riesz's lemma

Clearly, $a>0$ since Y is closed. We now take any $\theta\in(0,1)$. By the definition of an infimum there is a $y_0\in Y$ such that

(1) $$a\leqq\|v-y_0\|\leqq\frac{a}{\theta}$$

(note that $a/\theta>a$ since $0<\theta<1$). Let

$$z=c(v-y_0) \qquad \text{where} \qquad c=\frac{1}{\|v-y_0\|}.$$

Then $\|z\|=1$, and we show that $\|z-y\|\geqq\theta$ for every $y\in Y$. We have

$$\begin{aligned}\|z-y\|&=\|c(v-y_0)-y\|\\&=c\,\|v-y_0-c^{-1}y\|\\&=c\,\|v-y_1\|\end{aligned}$$

where

$$y_1=y_0+c^{-1}y.$$

The form of y_1 shows that $y_1\in Y$. Hence $\|v-y_1\|\geqq a$, by the definition of a. Writing c out and using (1), we obtain

$$\|z-y\|=c\,\|v-y_1\|\geqq ca=\frac{a}{\|v-y_0\|}\geqq\frac{a}{a/\theta}=\theta.$$

Since $y\in Y$ was arbitrary, this completes the proof. ∎

In a finite dimensional normed space the closed unit ball is compact by Theorem 2.5-3. Conversely, Riesz's lemma gives the following useful and remarkable

2.5-5 Theorem (Finite dimension). *If a normed space X has the property that the closed unit ball $M=\{x \mid \|x\| \leqq 1\}$ is compact, then X is finite dimensional.*

Proof. We assume that M is compact but $\dim X=\infty$, and show that this leads to a contradiction. We choose any x_1 of norm 1. This x_1 generates a one dimensional subspace X_1 of X, which is closed (cf. 2.4-3) and is a proper subspace of X since $\dim X=\infty$. By Riesz's lemma there is an $x_2 \in X$ of norm 1 such that

$$\|x_2 - x_1\| \geqq \theta = \frac{1}{2}.$$

The elements x_1, x_2 generate a two dimensional proper closed subspace X_2 of X. By Riesz's lemma there is an x_3 of norm 1 such that for all $x \in X_2$ we have

$$\|x_3 - x\| \geqq \frac{1}{2}.$$

In particular,

$$\|x_3 - x_1\| \geqq \frac{1}{2},$$

$$\|x_3 - x_2\| \geqq \frac{1}{2}.$$

Proceeding by induction, we obtain a sequence (x_n) of elements $x_n \in M$ such that

$$\|x_m - x_n\| \geqq \frac{1}{2} \qquad (m \neq n).$$

Obviously, (x_n) cannot have a convergent subsequence. This contradicts the compactness of M. Hence our assumption $\dim X=\infty$ is false, and $\dim X<\infty$. ∎

This theorem has various applications. We shall use it in Chap. 8 as a basic tool in connection with so-called compact operators.

Compact sets are important since they are "well-behaved": they have several basic properties similar to those of finite sets and not shared by noncompact sets. In connection with continuous mappings a fundamental property is that compact sets have compact images, as follows.

2.5-6 Theorem (Continuous mapping). *Let X and Y be metric spaces and $T: X \longrightarrow Y$ a continuous mapping* (cf. 1.3-3). *Then the image of a compact subset M of X under T is compact.*

Proof. By the definition of compactness it suffices to show that every sequence (y_n) in the image $T(M) \subset Y$ contains a subsequence which converges in $T(M)$. Since $y_n \in T(M)$, we have $y_n = Tx_n$ for some $x_n \in M$. Since M is compact, (x_n) contains a subsequence (x_{n_k}) which converges in M. The image of (x_{n_k}) is a subsequence of (y_n) which converges in $T(M)$ by 1.4-8 because T is continuous. Hence $T(M)$ is compact. ∎

From this theorem we conclude that the following property, well-known from calculus for continuous functions, carries over to metric spaces.

2.5-7 Corollary (Maximum and minimum). *A continuous mapping T of a compact subset M of a metric space X into $\mathbf{R}$ assumes a maximum and a minimum at some points of M.*

Proof. $T(M) \subset \mathbf{R}$ is compact by Theorem 2.5-6 and closed and bounded by Lemma 2.5-2 [applied to $T(M)$], so that $\inf T(M) \in T(M)$, $\sup T(M) \in T(M)$, and the inverse images of these two points consist of points of M at which Tx is minimum or maximum, respectively. ∎

Problems

1. Show that $\mathbf{R}^n$ and $\mathbf{C}^n$ are not compact.

2. Show that a discrete metric space X (cf. 1.1-8) consisting of infinitely many points is not compact.

3. Give examples of compact and noncompact curves in the plane $\mathbf{R}^2$.

4. Show that for an infinite subset M in the space s (cf. 2.2-8) to be compact, it is necessary that there are numbers $\gamma_1, \gamma_2, \cdots$ such that for all $x = (\xi_k(x)) \in M$ we have $|\xi_k(x)| \leqq \gamma_k$. (It can be shown that the condition is also sufficient for the compactness of M.)

5. **(Local compactness)** A metric space X is said to be *locally compact* if every point of X has a compact neighborhood. Show that $\mathbf{R}$ and $\mathbf{C}$ and, more generally, $\mathbf{R}^n$ and $\mathbf{C}^n$ are locally compact.

6. Show that a compact metric space X is locally compact.

7. If $\dim Y < \infty$ in Riesz's lemma 2.5-4, show that one can even choose $\theta = 1$.

8. In Prob. 7, Sec. 2.4, show directly (without using 2.4-5) that there is an $a > 0$ such that $a\|x\|_2 \leqq \|x\|$. (Use 2.5-7.)

9. If X is a compact metric space and $M \subset X$ is closed, show that M is compact.

10. Let X and Y be metric spaces, X compact, and $T: X \longrightarrow Y$ bijective and continuous. Show that T is a homeomorphism (cf. Prob. 5, Sec. 1.6).

2.6 Linear Operators

In calculus we consider the real line $\mathbf{R}$ and real-valued functions on $\mathbf{R}$ (or on a subset of $\mathbf{R}$). Obviously, any such function is a mapping[5] of its domain into $\mathbf{R}$. In functional analysis we consider more general spaces, such as metric spaces and normed spaces, and mappings of these spaces.

In the case of vector spaces and, in particular, normed spaces, a mapping is called an **operator**.

Of special interest are operators which "preserve" the two algebraic operations of vector space, in the sense of the following definition.

2.6-1 Definition (Linear operator). A *linear operator* T is an operator such that

(*i*) the domain $\mathcal{D}(T)$ of T is a vector space and the range $\mathcal{R}(T)$ lies in a vector space over the same field,

[5] Some familiarity with the concept of a mapping and simple related concepts is assumed, but a review is included in A1.2; cf. Appendix 1.

(*ii*) for all $x, y \in \mathcal{D}(T)$ and scalars α,

(1)
$$T(x+y) = Tx + Ty$$
$$T(\alpha x) = \alpha Tx.$$ ∎

Observe the **notation;** we write Tx instead of $T(x)$; this simplification is standard in functional analysis. Furthermore, ***for the remainder of the book we shall use the following notations.***

$\mathcal{D}(T)$ denotes the domain of T.

$\mathcal{R}(T)$ denotes the range of T.

$\mathcal{N}(T)$ denotes the null space of T.

By definition, the **null space** of T is the set of all $x \in \mathcal{D}(T)$ such that $Tx = 0$. (Another word for null space is "kernel." We shall not adopt this term since we must reserve the word "kernel" for another purpose in the theory of integral equations.)

We should also say something about the use of arrows in connection with operators. Let $\mathcal{D}(T) \subset X$ and $\mathcal{R}(T) \subset Y$, where X and Y are vector spaces, both real or both complex. Then T is an operator *from* (or mapping *of*) $\mathcal{D}(T)$ **onto** $\mathcal{R}(T)$, written

$$T\colon \mathcal{D}(T) \longrightarrow \mathcal{R}(T),$$

or from $\mathcal{D}(T)$ *into* Y, written

$$T\colon \mathcal{D}(T) \longrightarrow Y.$$

If $\mathcal{D}(T)$ is the whole space X, then—and only then—we write

$$T\colon X \longrightarrow Y.$$

Clearly, (1) is equivalent to

(2)
$$T(\alpha x + \beta y) = \alpha Tx + \beta Ty.$$

By taking $\alpha = 0$ in (1) we obtain the following formula which we shall need many times in our further work:

(3)
$$T0 = 0.$$

Formula (1) expresses the fact that a linear operator T is a **homomorphism** of a vector space (its domain) into another vector space, that is, T preserves the two operations of vector space, in the following sense. In (1) on the left we first apply a vector space operation (addition or multiplication by scalars) and then map the resulting vector into Y, whereas on the right we first map x and y into Y and then perform the vector space operations in Y, the outcome being the same. This property makes linear operators important. In turn, vector spaces are important in functional analysis mainly because of the linear operators defined on them.

We shall now consider some basic examples of linear operators and invite the reader to verify the linearity of the operator in each case.

Examples

2.6-2 Identity operator. The *identity operator* $I_X: X \longrightarrow X$ is defined by $I_X x = x$ for all $x \in X$. We also write simply I for I_X; thus, $Ix = x$.

2.6-3 Zero operator. The *zero operator* $0: X \longrightarrow Y$ is defined by $0x = 0$ for all $x \in X$.

2.6-4 Differentiation. Let X be the vector space of all polynomials on $[a, b]$. We may define a linear operator T on X by setting

$$Tx(t) = x'(t)$$

for every $x \in X$, where the prime denotes differentiation with respect to t. This operator T maps X onto itself.

2.6-5 Integration. A linear operator T from $C[a, b]$ into itself can be defined by

$$Tx(t) = \int_a^t x(\tau)\, d\tau \qquad t \in [a, b].$$

2.6-6 Multiplication by t. Another linear operator from $C[a, b]$ into itself is defined by

$$Tx(t) = tx(t).$$

T plays a role in physics (quantum theory), as we shall see in Chap. 11.

2.6-7 Elementary vector algebra. The *cross product* with one factor kept fixed defines a linear operator T_1: $\mathbf{R}^3 \longrightarrow \mathbf{R}^3$. Similarly, the *dot product* with one fixed factor defines a linear operator T_2: $\mathbf{R}^3 \longrightarrow \mathbf{R}$, say,

$$T_2 x = x \cdot a = \xi_1\alpha_1 + \xi_2\alpha_2 + \xi_3\alpha_3$$

where $a = (\alpha_j) \in \mathbf{R}^3$ is fixed.

2.6-8 Matrices. A *real matrix* $A = (\alpha_{jk})$ with r rows and n columns defines an operator T: $\mathbf{R}^n \longrightarrow \mathbf{R}^r$ by means of

$$y = Ax$$

where $x = (\xi_j)$ has n components and $y = (\eta_j)$ has r components and both vectors are written as column vectors because of the usual convention of matrix multiplication; writing $y = Ax$ out, we have

$$\begin{bmatrix} \eta_1 \\ \eta_2 \\ \cdot \\ \cdot \\ \cdot \\ \eta_r \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \alpha_{r1} & \alpha_{r2} & \cdots & \alpha_{rn} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \cdot \\ \cdot \\ \cdot \\ \xi_n \end{bmatrix}.$$

T is linear because matrix multiplication is a linear operation. If A were complex, it would define a linear operator from $\mathbf{C}^n$ into $\mathbf{C}^r$. A detailed discussion of the role of matrices in connection with linear operators follows in Sec. 2.9. ∎

In these examples we can easily verify that the ranges and null spaces of the linear operators are vector spaces. This fact is typical. Let us prove it, thereby observing how the linearity is used in simple proofs. The theorem itself will have various applications in our further work.

2.6-9 Theorem (Range and null space). *Let T be a linear operator. Then:*

(a) *The range $\mathcal{R}(T)$ is a vector space.*

(b) *If* $\dim \mathcal{D}(T) = n < \infty$, *then* $\dim \mathcal{R}(T) \leqq n$.

(c) *The null space $\mathcal{N}(T)$ is a vector space.*

Proof. **(a)** We take any $y_1, y_2 \in \mathscr{R}(T)$ and show that $\alpha y_1 + \beta y_2 \in \mathscr{R}(T)$ for any scalars α, β. Since $y_1, y_2 \in \mathscr{R}(T)$, we have $y_1 = Tx_1$, $y_2 = Tx_2$ for some $x_1, x_2 \in \mathscr{D}(T)$, and $\alpha x_1 + \beta x_2 \in \mathscr{D}(T)$ because $\mathscr{D}(T)$ is a vector space. The linearity of T yields

$$T(\alpha x_1 + \beta x_2) = \alpha Tx_1 + \beta Tx_2 = \alpha y_1 + \beta y_2.$$

Hence $\alpha y_1 + \beta y_2 \in \mathscr{R}(T)$. Since $y_1, y_2 \in \mathscr{R}(T)$ were arbitrary and so were the scalars, this proves that $\mathscr{R}(T)$ is a vector space.

(b) We choose $n+1$ elements $y_1, \cdots, y_{n+1}$ of $\mathscr{R}(T)$ in an arbitrary fashion. Then we have $y_1 = Tx_1, \cdots, y_{n+1} = Tx_{n+1}$ for some $x_1, \cdots, x_{n+1}$ in $\mathscr{D}(T)$. Since $\dim \mathscr{D}(T) = n$, this set $\{x_1, \cdots, x_{n+1}\}$ must be linearly dependent. Hence

$$\alpha_1 x_1 + \cdots + \alpha_{n+1} x_{n+1} = 0$$

for some scalars $\alpha_1, \cdots, \alpha_{n+1}$, not all zero. Since T is linear and $T0 = 0$, application of T on both sides gives

$$T(\alpha_1 x_1 + \cdots + \alpha_{n+1} x_{n+1}) = \alpha_1 y_1 + \cdots + \alpha_{n+1} y_{n+1} = 0.$$

This shows that $\{y_1, \cdots, y_{n+1}\}$ is a linearly dependent set because the α_j's are not all zero. Remembering that this subset of $\mathscr{R}(T)$ was chosen in an arbitrary fashion, we conclude that $\mathscr{R}(T)$ has no linearly independent subsets of $n+1$ or more elements. By the definition this means that $\dim \mathscr{R}(T) \leqq n$.

(c) We take any $x_1, x_2 \in \mathscr{N}(T)$. Then $Tx_1 = Tx_2 = 0$. Since T is linear, for any scalars α, β we have

$$T(\alpha x_1 + \beta x_2) = \alpha Tx_1 + \beta Tx_2 = 0.$$

This shows that $\alpha x_1 + \beta x_2 \in \mathscr{N}(T)$. Hence $\mathscr{N}(T)$ is a vector space. ∎

An immediate consequence of part (*b*) of the proof is worth noting:

Linear operators preserve linear dependence.

Let us turn to the inverse of a linear operator. We first remember that a mapping $T\colon \mathscr{D}(T) \longrightarrow Y$ is said to be **injective** or **one-to-one** if different points in the domain have different images, that is, if for any $x_1, x_2 \in \mathcal{D}(T)$,

(4) $$x_1 \neq x_2 \quad \Longrightarrow \quad Tx_1 \neq Tx_2;$$

equivalently,

(4*) $$Tx_1 = Tx_2 \quad \Longrightarrow \quad x_1 = x_2.$$

In this case there exists the mapping

(5) $$\begin{aligned} T^{-1}\colon \mathcal{R}(T) &\longrightarrow \mathcal{D}(T) \\ y_0 &\longmapsto x_0 \qquad (y_0 = Tx_0) \end{aligned}$$

which maps every $y_0 \in \mathcal{R}(T)$ onto that $x_0 \in \mathcal{D}(T)$ for which $Tx_0 = y_0$. See Fig. 20. The mapping T^{-1} is called the **inverse**[6] of T.

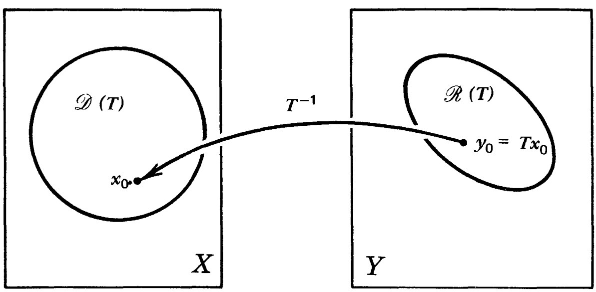

Fig. 20. Notations in connection with the inverse of a mapping; cf. (5)

From (5) we clearly have

$$T^{-1}Tx = x \qquad \text{for all } x \in \mathcal{D}(T)$$

$$TT^{-1}y = y \qquad \text{for all } y \in \mathcal{R}(T).$$

In connection with linear operators on vector spaces the situation is as follows. The inverse of a linear operator exists if and only if the null space of the operator consists of the zero vector only. More

[6] The reader may wish to review the terms "surjective" and "bijective" in A1.2, Appendix 1, which also contains a remark on the use of the term "inverse."

precisely, we have the following useful criterion which we shall apply quite often.

2.6-10 Theorem (Inverse operator). *Let X, Y be vector spaces, both real or both complex. Let $T: \mathfrak{D}(T) \longrightarrow Y$ be a linear operator with domain $\mathfrak{D}(T) \subset X$ and range $\mathfrak{R}(T) \subset Y$. Then:*

(a) *The inverse $T^{-1}: \mathfrak{R}(T) \longrightarrow \mathfrak{D}(T)$ exists if and only if*

$$Tx = 0 \qquad \Longrightarrow \qquad x = 0.$$

(b) *If T^{-1} exists, it is a linear operator.*

(c) *If $\dim \mathfrak{D}(T) = n < \infty$ and T^{-1} exists, then $\dim \mathfrak{R}(T) = \dim \mathfrak{D}(T)$.*

Proof. **(a)** Suppose that $Tx = 0$ implies $x = 0$. Let $Tx_1 = Tx_2$. Since T is linear,

$$T(x_1 - x_2) = Tx_1 - Tx_2 = 0,$$

so that $x_1 - x_2 = 0$ by the hypothesis. Hence $Tx_1 = Tx_2$ implies $x_1 = x_2$, and T^{-1} exists by (4*). Conversely, if T^{-1} exists, then (4*) holds. From (4*) with $x_2 = 0$ and (3) we obtain

$$Tx_1 = T0 = 0 \qquad \Longrightarrow \qquad x_1 = 0.$$

This completes the proof of (a).

(b) We assume that T^{-1} exists and show that T^{-1} is linear. The domain of T^{-1} is $\mathfrak{R}(T)$ and is a vector space by Theorem 2.6-9(a). We consider any $x_1, x_2 \in \mathfrak{D}(T)$ and their images

$$y_1 = Tx_1 \qquad \text{and} \qquad y_2 = Tx_2.$$

Then

$$x_1 = T^{-1}y_1 \qquad \text{and} \qquad x_2 = T^{-1}y_2.$$

T is linear, so that for any scalars α and β we have

$$\alpha y_1 + \beta y_2 = \alpha Tx_1 + \beta Tx_2 = T(\alpha x_1 + \beta x_2).$$

Since $x_j = T^{-1}y_j$, this implies

$$T^{-1}(\alpha y_1 + \beta y_2) = \alpha x_1 + \beta x_2 = \alpha T^{-1}y_1 + \beta T^{-1}y_2$$

and proves that T^{-1} is linear.

(c) We have $\dim \mathscr{R}(T) \leqq \dim \mathscr{D}(T)$ by Theorem 2.6-9(*b*), and $\dim \mathscr{D}(T) \leqq \dim \mathscr{R}(T)$ by the same theorem applied to T^{-1}. ∎

We finally mention a useful formula for the inverse of the composite of linear operators. (The reader may perhaps know this formula for the case of square matrices.)

2.6-11 Lemma (Inverse of product). *Let $T\colon X \longrightarrow Y$ and $S\colon Y \longrightarrow Z$ be bijective linear operators, where X, Y, Z are vector spaces (see* Fig. 21). *Then the inverse $(ST)^{-1}\colon Z \longrightarrow X$ of the product (the composite) ST exists, and*

$$(ST)^{-1} = T^{-1}S^{-1}. \tag{6}$$

Proof. The operator $ST\colon X \longrightarrow Z$ is bijective, so that $(ST)^{-1}$ exists. We thus have

$$ST(ST)^{-1} = I_Z$$

where I_Z is the identity operator on Z. Applying S^{-1} and using $S^{-1}S = I_Y$ (the identity operator on Y), we obtain

$$S^{-1}ST(ST)^{-1} = T(ST)^{-1} = S^{-1}I_Z = S^{-1}.$$

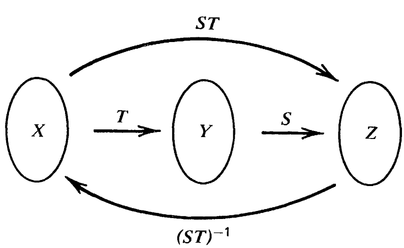

Fig. 21. Notations in Lemma 2.6-11

Applying T^{-1} and using $T^{-1}T = I_X$, we obtain the desired result

$$T^{-1}T(ST)^{-1} = (ST)^{-1} = T^{-1}S^{-1}.$$

This completes the proof. ∎

Problems

1. Show that the operators in 2.6-2, 2.6-3 and 2.6-4 are linear.

2. Show that the operators $T_1, \cdots, T_4$ from $\mathbf{R}^2$ into $\mathbf{R}^2$ defined by

$$(\xi_1, \xi_2) \longmapsto (\xi_1, 0)$$

$$(\xi_1, \xi_2) \longmapsto (0, \xi_2)$$

$$(\xi_1, \xi_2) \longmapsto (\xi_2, \xi_1)$$

$$(\xi_1, \xi_2) \longmapsto (\gamma\xi_1, \gamma\xi_2)$$

respectively, are linear, and interpret these operators geometrically.

3. What are the domain, range and null space of T_1, T_2, T_3 in Prob. 2?

4. What is the null space of T_4 in Prob. 2? Of T_1 and T_2 in 2.6-7? Of T in 2.6-4?

5. Let $T: X \longrightarrow Y$ be a linear operator. Show that the image of a subspace V of X is a vector space, and so is the inverse image of a subspace W of Y.

6. If the product (the composite) of two linear operators exists, show that it is linear.

7. (Commutativity) Let X be any vector space and $S: X \longrightarrow X$ and $T: X \longrightarrow X$ any operators. S and T are said to *commute* if $ST = TS$, that is, $(ST)x = (TS)x$ for all $x \in X$. Do T_1 and T_3 in Prob. 2 commute?

8. Write the operators in Prob. 2 using 2×2 matrices.

9. In 2.6-8, write $y = Ax$ in terms of components, show that T is linear and give examples.

10. Formulate the condition in 2.6-10(a) in terms of the null space of T.

11. Let X be the vector space of all complex 2×2 matrices and define $T\colon X \longrightarrow X$ by $Tx = bx$, where $b \in X$ is fixed and bx denotes the usual product of matrices. Show that T is linear. Under what condition does T^{-1} exist?

12. Does the inverse of T in 2.6-4 exist?

13. Let $T\colon \mathfrak{D}(T) \longrightarrow Y$ be a linear operator whose inverse exists. If $\{x_1, \cdots, x_n\}$ is a linearly independent set in $\mathfrak{D}(T)$, show that the set $\{Tx_1, \cdots, Tx_n\}$ is linearly independent.

14. Let $T\colon X \longrightarrow Y$ be a linear operator and $\dim X = \dim Y = n < \infty$. Show that $\mathfrak{R}(T) = Y$ if and only if T^{-1} exists.

15. Consider the vector space X of all real-valued functions which are defined on $\mathbf{R}$ and have derivatives of all orders everywhere on $\mathbf{R}$. Define $T\colon X \longrightarrow X$ by $y(t) = Tx(t) = x'(t)$. Show that $\mathfrak{R}(T)$ is all of X but T^{-1} does not exist. Compare with Prob. 14 and comment.

2.7 Bounded and Continuous Linear Operators

The reader may have noticed that in the whole last section we did not make any use of norms. We shall now again take norms into account, in the following basic definition.

2.7-1 Definition (Bounded linear operator). Let X and Y be normed spaces and $T\colon \mathfrak{D}(T) \longrightarrow Y$ a linear operator, where $\mathfrak{D}(T) \subset X$. The operator T is said to be *bounded* if there is a real number c such that for all $x \in \mathfrak{D}(T)$,

$$\|Tx\| \leqq c\|x\|. \tag{1}$$ ∎

In (1) the norm on the left is that on Y, and the norm on the right is that on X. For simplicity we have denoted both norms by the same symbol $\|\cdot\|$, without danger of confusion. Distinction by subscripts ($\|x\|_0$, $\|Tx\|_1$, etc.) seems unnecessary here. Formula (1) shows that a bounded linear operator maps bounded sets in $\mathfrak{D}(T)$ onto bounded sets in Y. This motivates the term "bounded operator."

Warning. Note that our present use of the word "bounded" is different from that in calculus, where a bounded function is one whose range is a bounded set. Unfortunately, both terms are standard. But there is little danger of confusion.

What is the smallest possible c such that (1) still holds for all nonzero $x \in \mathcal{D}(T)$? [We can leave out $x=0$ since $Tx=0$ for $x=0$ by (3), Sec. 2.6.] By division,

$$\frac{\|Tx\|}{\|x\|} \leqq c \qquad (x \neq 0)$$

and this shows that c must be at least as big as the supremum of the expression on the left taken over $\mathcal{D}(T)-\{0\}$. Hence the answer to our question is that the smallest possible c in (1) is that supremum. This quantity is denoted by $\|T\|$; thus

$$\|T\| = \sup_{\substack{x \in \mathcal{D}(T) \\ x \neq 0}} \frac{\|Tx\|}{\|x\|}. \tag{2}$$

$\|T\|$ is called the **norm** of the operator T. If $\mathcal{D}(T)=\{0\}$, we define $\|T\|=0$; in this (relatively uninteresting) case, $T=0$ since $T0=0$ by (3), Sec. 2.6.

Note that (1) with $c=\|T\|$ is

$$\|Tx\| \leqq \|T\| \|x\|. \tag{3}$$

This formula will be applied quite frequently.

Of course, we should justify the use of the term "norm" in the present context. This will be done in the following lemma.

2.7-2 Lemma (Norm). *Let T be a bounded linear operator as defined in 2.7-1. Then:*

(a) *An alternative formula for the norm of T is*

$$\|T\| = \sup_{\substack{x \in \mathcal{D}(T) \\ \|x\|=1}} \|Tx\|. \tag{4}$$

(b) *The norm defined by* (2) *satisfies* (N1) *to* (N4) *in Sec.* 2.2.

Proof. **(a)** We write $\|x\|=a$ and set $y=(1/a)x$, where $x\neq 0$. Then $\|y\|=\|x\|/a=1$, and since T is linear, (2) gives

$$\|T\|=\sup_{\substack{x\in\mathcal{D}(T)\\ x\neq 0}}\frac{1}{a}\|Tx\|=\sup_{\substack{x\in\mathcal{D}(T)\\ x\neq 0}}\left\|T\left(\frac{1}{a}x\right)\right\|=\sup_{\substack{y\in\mathcal{D}(T)\\ \|y\|=1}}\|Ty\|.$$

Writing x for y on the right, we have (4).

(b) (N1) is obvious, and so is $\|0\|=0$. From $\|T\|=0$ we have $Tx=0$ for all $x\in\mathcal{D}(T)$, so that $T=0$. Hence (N2) holds. Furthermore, (N3) is obtained from

$$\sup_{\|x\|=1}\|\alpha Tx\|=\sup_{\|x\|=1}|\alpha|\,\|Tx\|=|\alpha|\sup_{\|x\|=1}\|Tx\|$$

where $x\in\mathcal{D}(T)$. Finally, (N4) follows from

$$\sup_{\|x\|=1}\|(T_1+T_2)x\|=\sup_{\|x\|=1}\|T_1x+T_2x\|\leqq\sup_{\|x\|=1}\|T_1x\|+\sup_{\|x\|=1}\|T_2x\|;$$

here, $x\in\mathcal{D}(T)$. ∎

Before we consider general properties of bounded linear operators, let us take a look at some typical examples, so that we get a better feeling for the concept of a bounded linear operator.

Examples

2.7-3 Identity operator. The identity operator $I: X\longrightarrow X$ on a normed space $X\neq\{0\}$ is bounded and has norm $\|I\|=1$. Cf. 2.6-2.

2.7-4 Zero operator. The zero operator $0: X\longrightarrow Y$ on a normed space X is bounded and has norm $\|0\|=0$. Cf. 2.6-3.

2.7-5 Differentiation operator. Let X be the normed space of all polynomials on $J=[0,1]$ with norm given $\|x\|=\max|x(t)|$, $t\in J$. A differentiation operator T is defined on X by

$$Tx(t)=x'(t)$$

where the prime denotes differentiation with respect to t. This operator is linear but not bounded. Indeed, let $x_n(t)=t^n$, where $n\in\mathbf{N}$. Then $\|x_n\|=1$ and

$$Tx_n(t)=x_n'(t)=nt^{n-1}$$

so that $\|Tx_n\|=n$ and $\|Tx_n\|/\|x_n\|=n$. Since $n\in\mathbf{N}$ is arbitrary, this shows that there is no fixed number c such that $\|Tx_n\|/\|x_n\|\leqq c$. From this and (1) we conclude that T is not bounded.

Since differentiation is an important operation, our result seems to indicate that unbounded operators are also of practical importance. This is indeed the case, as we shall see in Chaps. 10 and 11, after a detailed study of the theory and application of bounded operators, which are simpler than unbounded ones.

2.7-6 Integral operator. We can define an integral operator $T\colon C[0,1]\longrightarrow C[0,1]$ by

$$y=Tx \qquad \text{where} \qquad y(t)=\int_0^1 k(t,\tau)x(\tau)\,d\tau.$$

Here k is a given function, which is called the *kernel* of T and is assumed to be continuous on the closed square $G=J\times J$ in the $t\tau$-plane, where $J=[0,1]$. This operator is linear.

T is bounded.

To prove this, we first note that the continuity of k on the closed square implies that k is bounded, say, $|k(t,\tau)|\leqq k_0$ for all $(t,\tau)\in G$, where k_0 is a real number. Furthermore,

$$|x(t)|\leqq \max_{t\in J}|x(t)|=\|x\|.$$

Hence

$$\begin{aligned}\|y\|=\|Tx\|&=\max_{t\in J}\left|\int_0^1 k(t,\tau)x(\tau)\,d\tau\right|\\ &\leqq \max_{t\in J}\int_0^1 |k(t,\tau)|\,|x(\tau)|\,d\tau\\ &\leqq k_0\,\|x\|.\end{aligned}$$

The result is $\|Tx\| \leqq k_0 \|x\|$. This is (1) with $c = k_0$. Hence T is bounded.

2.7-7 Matrix. A real matrix $A = (\alpha_{jk})$ with r rows and n columns defines an operator $T: \mathbf{R}^n \longrightarrow \mathbf{R}^r$ by means of

$$y = Ax \tag{5}$$

where $x = (\xi_j)$ and $y = (\eta_j)$ are column vectors with n and r components, respectively, and we used matrix multiplication, as in 2.6-8. In terms of components, (5) becomes

$$\eta_j = \sum_{k=1}^{n} \alpha_{jk}\xi_k \qquad (j = 1, \cdots, r). \tag{5'}$$

T is linear because matrix multiplication is a linear operation.

T is bounded.

To prove this, we first remember from 2.2-2 that the norm on $\mathbf{R}^n$ is given by

$$\|x\| = \left(\sum_{m=1}^{n} \xi_m^2\right)^{1/2};$$

similarly for $y \in \mathbf{R}^r$. From (5') and the Cauchy-Schwarz inequality (11) in Sec. 1.2 we thus obtain

$$\begin{aligned} \|Tx\|^2 = \sum_{j=1}^{r} \eta_j^2 &= \sum_{j=1}^{r} \left[\sum_{k=1}^{n} \alpha_{jk}\xi_k\right]^2 \\ &\leqq \sum_{j=1}^{r} \left[\left(\sum_{k=1}^{n} \alpha_{jk}^2\right)^{1/2} \left(\sum_{m=1}^{n} \xi_m^2\right)^{1/2}\right]^2 \\ &= \|x\|^2 \sum_{j=1}^{r} \sum_{k=1}^{n} \alpha_{jk}^2. \end{aligned}$$

Noting that the double sum in the last line does not depend on x, we can write our result in the form

$$\|Tx\|^2 \leqq c^2 \|x\|^2 \qquad \text{where} \qquad c^2 = \sum_{j=1}^{r} \sum_{k=1}^{n} \alpha_{jk}^2.$$

This gives (1) and completes the proof that T is bounded. ∎

The role of matrices in connection with linear operators will be studied in a separate section (Sec. 2.9). Boundedness is typical; it is an essential simplification which we always have in the finite dimensional case, as follows.

2.7-8 Theorem (Finite dimension). *If a normed space X is finite dimensional, then every linear operator on X is bounded.*

Proof. Let $\dim X = n$ and $\{e_1, \cdots, e_n\}$ a basis for X. We take any $x = \sum \xi_j e_j$ and consider any linear operator T on X. Since T is linear,

$$\|Tx\| = \left\|\sum \xi_j Te_j\right\| \leqq \sum |\xi_j| \, \|Te_j\| \leqq \max_k \|Te_k\| \sum |\xi_j|$$

(summations from 1 to n). To the last sum we apply Lemma 2.4-1 with $\alpha_j = \xi_j$ and $x_j = e_j$. Then we obtain

$$\sum |\xi_j| \leqq \frac{1}{c} \left\|\sum \xi_j e_j\right\| = \frac{1}{c} \|x\|.$$

Together,

$$\|Tx\| \leqq \gamma \|x\| \qquad \text{where} \qquad \gamma = \frac{1}{c} \max_k \|Te_k\|.$$

From this and (1) we see that T is bounded. ∎

We shall now consider important general properties of bounded linear operators.

Operators are mappings, so that the definition of continuity (cf. 1.3-3) applies to them. It is a fundamental fact that for a *linear* operator, continuity and boundedness become equivalent concepts. The details are as follows.

Let $T\colon \mathfrak{D}(T) \longrightarrow Y$ be any operator, not necessarily linear, where $\mathfrak{D}(T) \subset X$ and X and Y are normed spaces. By Def. 1.3-3, the operator T is *continuous at an* $x_0 \in \mathfrak{D}(T)$ if for every $\varepsilon > 0$ there is a $\delta > 0$ such that

$$\|Tx - Tx_0\| < \varepsilon \qquad \text{for all } x \in \mathfrak{D}(T) \text{ satisfying} \qquad \|x - x_0\| < \delta.$$

T is *continuous* if T is continuous at every $x \in \mathfrak{D}(T)$.

Now, if T is linear, we have the remarkable

2.7-9 Theorem (Continuity and boundedness). *Let* $T\colon \mathfrak{D}(T) \longrightarrow Y$ *be a linear*[7] *operator, where* $\mathfrak{D}(T) \subset X$ *and* X, Y *are normed spaces. Then:*

(a) *T is continuous if and only if T is bounded.*

(b) *If T is continuous at a single point, it is continuous.*

Proof. **(a)** For $T=0$ the statement is trivial. Let $T\neq 0$. Then $\|T\|\neq 0$. We assume T to be bounded and consider any $x_0 \in \mathfrak{D}(T)$. Let any $\varepsilon>0$ be given. Then, since T is linear, for every $x \in \mathfrak{D}(T)$ such that

$$\|x-x_0\|<\delta \qquad \text{where} \qquad \delta=\frac{\varepsilon}{\|T\|}$$

we obtain

$$\|Tx-Tx_0\|=\|T(x-x_0)\|\leqq \|T\|\,\|x-x_0\|<\|T\|\delta=\varepsilon.$$

Since $x_0 \in \mathfrak{D}(T)$ was arbitrary, this shows that T is continuous.

Conversely, assume that T is continuous at an arbitrary $x_0 \in \mathfrak{D}(T)$. Then, given any $\varepsilon>0$, there is a $\delta>0$ such that

$$(6) \qquad \|Tx-Tx_0\|\leqq \varepsilon \qquad \text{for all } x \in \mathfrak{D}(T) \text{ satisfying} \qquad \|x-x_0\|\leqq \delta.$$

We now take any $y\neq 0$ in $\mathfrak{D}(T)$ and set

$$x=x_0+\frac{\delta}{\|y\|}\,y. \qquad \text{Then} \qquad x-x_0=\frac{\delta}{\|y\|}\,y.$$

Hence $\|x-x_0\|=\delta$, so that we may use (6). Since T is linear, we have

$$\|Tx-Tx_0\|=\|T(x-x_0)\|=\left\|T\left(\frac{\delta}{\|y\|}\,y\right)\right\|=\frac{\delta}{\|y\|}\,\|Ty\|$$

[7] **Warning.** Unfortunately, continuous linear operators are called "linear operators" by some authors. We shall not adopt this terminology; in fact, there are linear operators of practical importance which are not continuous. A first example is given in 2.7-5 and further operators of that type will be considered in Chaps. 10 and 11.

and (6) implies

$$\frac{\delta}{\|y\|}\|Ty\| \leqq \varepsilon. \qquad \text{Thus} \qquad \|Ty\| \leqq \frac{\varepsilon}{\delta}\|y\|.$$

This can be written $\|Ty\| \leqq c\|y\|$, where $c = \varepsilon/\delta$, and shows that T is bounded.

(b) Continuity of T at a point implies boundedness of T by the second part of the proof of (a), which in turn implies continuity of T by (a). ▮

2.7-10 Corollary (Continuity, null space). *Let T be a bounded linear operator. Then:*

(a) $x_n \longrightarrow x$ *[where $x_n, x \in \mathcal{D}(T)$] implies $Tx_n \longrightarrow Tx$.*

(b) *The null space $\mathcal{N}(T)$ is closed.*

Proof. **(a)** follows from Theorems 2.7-9(a) and 1.4-8 or directly from (3) because, as $n \longrightarrow \infty$,

$$\|Tx_n - Tx\| = \|T(x_n - x)\| \leqq \|T\| \|x_n - x\| \longrightarrow 0.$$

(b) For every $x \in \overline{\mathcal{N}(T)}$ there is a sequence (x_n) in $\mathcal{N}(T)$ such that $x_n \longrightarrow x$; cf. 1.4-6(a). Hence $Tx_n \longrightarrow Tx$ by part (a) of this Corollary. Also $Tx = 0$ since $Tx_n = 0$, so that $x \in \mathcal{N}(T)$. Since $x \in \overline{\mathcal{N}(T)}$ was arbitrary, $\mathcal{N}(T)$ is closed. ▮

It is worth noting that the range of a bounded linear operator may not be closed. Cf. Prob. 6.

The reader may give the simple proof of another useful formula, namely,

$$\text{(7)} \qquad \|T_1 T_2\| \leqq \|T_1\| \|T_2\|, \qquad \|T^n\| \leqq \|T\|^n \qquad (n \in \mathbf{N})$$

valid for bounded linear operators $T_2\colon X \longrightarrow Y$, $T_1\colon Y \longrightarrow Z$ and $T\colon X \longrightarrow X$, where X, Y, Z are normed spaces.

Operators are mappings, and some concepts related to mappings[8] have been discussed, notably the domain, range and null space of an

[8] A review of some of these concepts is given in A1.2; cf. Appendix 1.

operator. Two further concepts (restriction and extension) will now be added. We could have done this earlier, but we prefer to do it here, where we can immediately give an interesting application (Theorem 2.7-11, below). Let us begin by defining equality of operators as follows.

Two operators T_1 and T_2 are defined to be **equal**, written

$$T_1 = T_2,$$

if they have the same domain $\mathscr{D}(T_1) = \mathscr{D}(T_2)$ and if $T_1x = T_2x$ for all $x \in \mathscr{D}(T_1) = \mathscr{D}(T_2)$.

The **restriction** of an operator $T\colon \mathscr{D}(T) \longrightarrow Y$ to a subset $B \subset \mathscr{D}(T)$ is denoted by

$$T|_B$$

and is the operator defined by

$$T|_B\colon B \longrightarrow Y, \qquad T|_B x = Tx \text{ for all } x \in B.$$

An **extension** of T to a set $M \supset \mathscr{D}(T)$ is an operator

$$\tilde{T}\colon M \longrightarrow Y \qquad \text{such that} \qquad \tilde{T}|_{\mathscr{D}(T)} = T,$$

that is, $\tilde{T}x = Tx$ for all $x \in \mathscr{D}(T)$. [Hence T is the restriction of $\tilde{T}$ to $\mathscr{D}(T)$.]

If $\mathscr{D}(T)$ is a proper subset of M, then a given T has many extensions. Of practical interest are usually those extensions which preserve some basic property, for instance linearity (if T happens to be linear) or boundedness (if $\mathscr{D}(T)$ lies in a normed space and T is bounded). The following important theorem is typical in that respect. It concerns an extension of a bounded linear operator T to the closure $\overline{\mathscr{D}(T)}$ of the domain such that the extended operator is again bounded and linear, and even has the same norm. This includes the case of an extension from a dense set in a normed space X to all of X. It also includes the case of an extension from a normed space X to its completion (cf. 2.3-2).

2.7-11 Theorem (Bounded linear extension). *Let*

$$T: \mathscr{D}(T) \longrightarrow Y$$

be a bounded linear operator, where $\mathscr{D}(T)$ lies in a normed space X and Y is a Banach space. Then T has an extension

$$\tilde{T}: \overline{\mathscr{D}(T)} \longrightarrow Y$$

where $\tilde{T}$ is a bounded linear operator of norm

$$\|\tilde{T}\| = \|T\|.$$

Proof. We consider any $x \in \overline{\mathscr{D}(T)}$. By Theorem 1.4-6(*a*) there is a sequence (x_n) in $\mathscr{D}(T)$ such that $x_n \longrightarrow x$. Since T is linear and bounded, we have

$$\|Tx_n - Tx_m\| = \|T(x_n - x_m)\| \leqq \|T\| \|x_n - x_m\|.$$

This shows that (Tx_n) is Cauchy because (x_n) converges. By assumption, Y is complete, so that (Tx_n) converges, say,

$$Tx_n \longrightarrow y \in Y.$$

We define $\tilde{T}$ by

$$\tilde{T}x = y.$$

We show that this definition is independent of the particular choice of a sequence in $\mathscr{D}(T)$ converging to x. Suppose that $x_n \longrightarrow x$ and $z_n \longrightarrow x$. Then $v_m \longrightarrow x$, where (v_m) is the sequence

$$(x_1, z_1, x_2, z_2, \cdots).$$

Hence (Tv_m) converges by 2.7-10(*a*), and the two subsequences (Tx_n) and (Tz_n) of (Tv_m) must have the same limit. This proves that $\tilde{T}$ is uniquely defined at every $x \in \overline{\mathscr{D}(T)}$.

Clearly, $\tilde{T}$ is linear and $\tilde{T}x = Tx$ for every $x \in \mathscr{D}(T)$, so that $\tilde{T}$ is an extension of T. We now use

$$\|Tx_n\| \leqq \|T\| \|x_n\|$$

and let $n \longrightarrow \infty$. Then $Tx_n \longrightarrow y = \tilde{T}x$. Since $x \longmapsto \|x\|$ defines a continuous mapping (cf. Sec. 2.2), we thus obtain

$$\|\tilde{T}x\| \leqq \|T\| \|x\|.$$

Hence $\tilde{T}$ is bounded and $\|\tilde{T}\| \leqq \|T\|$. Of course, $\|\tilde{T}\| \geqq \|T\|$ because the norm, being defined by a supremum, cannot decrease in an extension. Together we have $\|\tilde{T}\| = \|T\|$. ∎

Problems

1. Prove (7).

2. Let X and Y be normed spaces. Show that a linear operator $T: X \longrightarrow Y$ is bounded if and only if T maps bounded sets in X into bounded sets in Y.

3. If $T \neq 0$ is a bounded linear operator, show that for any $x \in \mathscr{D}(T)$ such that $\|x\| < 1$ we have the strict inequality $\|Tx\| < \|T\|$.

4. Give a direct proof of 2.7-9(b), without using 2.7-9(a).

5. Show that the operator $T: l^\infty \longrightarrow l^\infty$ defined by $y = (\eta_j) = Tx$, $\eta_j = \xi_j/j$, $x = (\xi_j)$, is linear and bounded.

6. **(Range)** Show that the range $\mathscr{R}(T)$ of a bounded linear operator $T: X \longrightarrow Y$ need not be closed in Y. *Hint.* Use T in Prob. 5.

7. **(Inverse operator)** Let T be a bounded linear operator from a normed space X onto a normed space Y. If there is a positive b such that

$$\|Tx\| \geqq b\|x\| \qquad \text{for all } x \in X,$$

show that then $T^{-1}: Y \longrightarrow X$ exists and is bounded.

8. Show that the inverse $T^{-1}: \mathscr{R}(T) \longrightarrow X$ of a bounded linear operator $T: X \longrightarrow Y$ need not be bounded. *Hint.* Use T in Prob. 5.

9. Let $T: C[0, 1] \longrightarrow C[0, 1]$ be defined by

$$y(t) = \int_0^t x(\tau)\, d\tau.$$

Find $\mathscr{R}(T)$ and $T^{-1}: \mathscr{R}(T) \longrightarrow C[0, 1]$. Is T^{-1} linear and bounded?

10. On $C[0, 1]$ define S and T by

$$y(s) = s\int_0^1 x(t)\,dt, \qquad y(s) = sx(s),$$

respectively. Do S and T commute? Find $\|S\|$, $\|T\|$, $\|ST\|$ and $\|TS\|$.

11. Let X be the normed space of all bounded real-valued functions on $\mathbf{R}$ with norm defined by

$$\|x\| = \sup_{t \in \mathbf{R}} |x(t)|,$$

and let $T\colon X \longrightarrow X$ be defined by

$$y(t) = Tx(t) = x(t - \Delta)$$

where $\Delta > 0$ is a constant. (This is a model of a *delay line*, which is an electric device whose output y is a delayed version of the input x, the time delay being Δ; see Fig. 22.) Is T linear? Bounded?

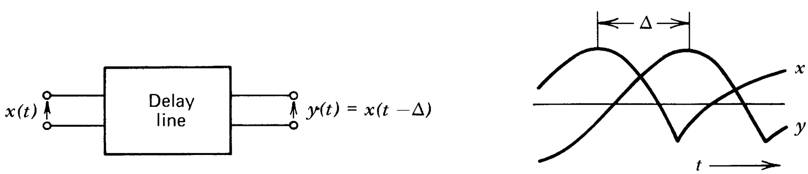

Fig. 22. Electric delay line

12. **(Matrices)** From 2.7-7 we know that an $r \times n$ matrix $A = (\alpha_{jk})$ defines a linear operator from the vector space X of all ordered n-tuples of numbers into the vector space Y of all ordered r-tuples of numbers. Suppose that any norm $\|\cdot\|_1$ is given on X and any norm $\|\cdot\|_2$ is given on Y. Remember from Prob. 10, Sec. 2.4, that there are various norms on the space Z of all those matrices (r and n fixed). A norm $\|\cdot\|$ on Z is said to be *compatible* with $\|\cdot\|_1$ and $\|\cdot\|_2$ if

$$\|Ax\|_2 \leqq \|A\|\,\|x\|_1.$$

Show that the norm defined by

$$\|A\| = \sup_{\substack{x \in X \\ x \neq 0}} \frac{\|Ax\|_2}{\|x\|_1}$$

is compatible with $\|\cdot\|_1$ and $\|\cdot\|_2$. This norm is often called the *natural norm* defined by $\|\cdot\|_1$ and $\|\cdot\|_2$. If we choose $\|x\|_1 = \max_j |\xi_j|$ and $\|y\|_2 = \max_j |\eta_j|$, show that the natural norm is

$$\|A\| = \max_j \sum_{k=1}^{n} |\alpha_{jk}|.$$

13. Show that in 2.7-7 with $r = n$, a compatible norm is defined by

$$\|A\| = \left(\sum_{j=1}^{n} \sum_{k=1}^{n} \alpha_{jk}{}^2 \right)^{1/2},$$

but for $n > 1$ this is *not* the natural norm defined by the Euclidean norm on $\mathbf{R}^n$.

14. If in Prob. 12 we choose

$$\|x\|_1 = \sum_{k=1}^{n} |\xi_k|, \qquad \|y\|_2 = \sum_{j=1}^{r} |\eta_j|,$$

show that a compatible norm is defined by

$$\|A\| = \max_k \sum_{j=1}^{r} |\alpha_{jk}|.$$

15. Show that for $r = n$, the norm in Prob. 14 is the natural norm corresponding to $\|\cdot\|_1$ and $\|\cdot\|_2$ as defined in that problem.

2.8 Linear Functionals

A **functional** is an operator whose range lies on the real line **R** or in the complex plane **C**. And *functional analysis* was initially the analysis of functionals. The latter appear so frequently that special notations are used. We denote functionals by lowercase letters f, g, h, $\cdots$, the domain of f by $\mathfrak{D}(f)$, the range by $\mathfrak{R}(f)$ and the value of f at an $x \in \mathfrak{D}(f)$ by $f(x)$, with parentheses.

Functionals are operators, so that previous definitions apply. We shall need in particular the following two definitions because most of the functionals to be considered will be linear and bounded.

2.8-1 Definition (Linear functional). A *linear functional* f is a linear operator with domain in a vector space X and range in the scalar field K of X; thus,

$$f: \mathfrak{D}(f) \longrightarrow K,$$

where $K = \mathbf{R}$ if X is real and $K = \mathbf{C}$ if X is complex. ∎

2.8-2 Definition (Bounded linear functional). A *bounded linear functional* f is a bounded linear operator (cf. Def. 2.7-1) with range in the scalar field of the normed space X in which the domain $\mathfrak{D}(f)$ lies. Thus there exists a real number c such that for all $x \in \mathfrak{D}(f)$,

(1) $$|f(x)| \leqq c\,\|x\|.$$

Furthermore, the *norm* of f is [cf. (2) in Sec. 2.7]

(2a) $$\|f\| = \sup_{\substack{x \in \mathfrak{D}(f) \\ x \neq 0}} \frac{|f(x)|}{\|x\|}$$

or

(2b) $$\|f\| = \sup_{\substack{x \in \mathfrak{D}(f) \\ \|x\|=1}} |f(x)|.$$ ∎

Formula (3) in Sec. 2.7 now implies

(3) $$|f(x)| \leqq \|f\|\,\|x\|,$$

and a special case of Theorem 2.7-9 is

2.8-3 Theorem (Continuity and boundedness). *A linear functional f with domain $\mathfrak{D}(f)$ in a normed space is continuous if and only if f is bounded.*

Examples

2.8-4 Norm. The *norm* $\|\cdot\|: X \longrightarrow \mathbf{R}$ on a normed space $(X, \|\cdot\|)$ is a functional on X which is not linear.

2.8-5 Dot product. The familiar *dot product* with one factor kept fixed defines a functional $f: \mathbf{R}^3 \longrightarrow \mathbf{R}$ by means of

$$f(x) = x \cdot a = \xi_1\alpha_1 + \xi_2\alpha_2 + \xi_3\alpha_3,$$

where $a = (\alpha_j) \in \mathbf{R}^3$ is fixed.

f is linear. f is bounded. In fact,

$$|f(x)| = |x \cdot a| \leqq \|x\|\|a\|,$$

so that $\|f\| \leqq \|a\|$ follows from (2b) if we take the supremum over all x of norm one. On the other hand, by taking $x = a$ and using (3) we obtain

$$\|f\| \geqq \frac{|f(a)|}{\|a\|} = \frac{\|a\|^2}{\|a\|} = \|a\|.$$

Hence the norm of f is $\|f\| = \|a\|$.

2.8-6 Definite integral. The *definite integral* is a number if we consider it for a single function, as we do in calculus most of the time. However, the situation changes completely if we consider that integral for all functions in a certain function space. Then the integral becomes a functional on that space, call it f. As a space let us choose $C[a, b]$; cf. 2.2-5. Then f is defined by

$$f(x) = \int_a^b x(t)\,dt \qquad x \in C[a, b].$$

f is linear. We prove that f is bounded and has norm $\|f\| = b - a$.

In fact, writing $J = [a, b]$ and remembering the norm on $C[a, b]$, we obtain

$$|f(x)| = \left|\int_a^b x(t)\,dt\right| \leqq (b-a) \max_{t \in J} |x(t)| = (b-a)\,\|x\|.$$

Taking the supremum over all x of norm 1, we obtain $\|f\| \leqq b-a$. To get $\|f\| \geqq b-a$, we choose the particular $x=x_0=1$, note that $\|x_0\|=1$ and use (3):

$$\|f\| \geqq \frac{|f(x_0)|}{\|x_0\|} = |f(x_0)| = \int_a^b dt = b-a.$$

2.8-7 Space $C[a, b]$. Another practically important functional on $C[a, b]$ is obtained if we choose a fixed $t_0 \in J = [a, b]$ and set

$$f_1(x) = x(t_0) \qquad\qquad x \in C[a, b].$$

f_1 is linear. f_1 is bounded and has norm $\|f_1\| = 1$. In fact, we have

$$|f_1(x)| = |x(t_0)| \leqq \|x\|,$$

and this implies $\|f_1\| \leqq 1$ by (2). On the other hand, for $x_0 = 1$ we have $\|x_0\| = 1$ and obtain from (3)

$$\|f_1\| \geqq |f_1(x_0)| = 1.$$

2.8-8 Space l^2. We can obtain a linear functional f on the Hilbert space l^2 (cf. 1.2-3) by choosing a fixed $a = (\alpha_j) \in l^2$ and setting

$$f(x) = \sum_{j=1}^{\infty} \xi_j \alpha_j$$

where $x = (\xi_j) \in l^2$. This series converges absolutely and f is bounded, since the Cauchy-Schwarz inequality (11) in Sec. 1.2 gives (summation over j from 1 to ∞)

$$|f(x)| = \left|\sum \xi_j \alpha_j\right| \leqq \sum |\xi_j \alpha_j| \leqq \sqrt{\sum |\xi_j|^2} \sqrt{\sum |\alpha_j|^2} = \|x\| \, \|a\|. \qquad ∎$$

It is of basic importance that the set of all linear functionals defined on a vector space X can itself be made into a vector space. This space is denoted by X^* and is called the **algebraic**[9] **dual space** of X. Its algebraic operations of vector space are defined in a natural way

[9] Note that this definition does not involve a norm. The so-called *dual space* X' consisting of all *bounded* linear functionals on X will be considered in Sec. 2.10.

as follows. The *sum* f_1+f_2 of two functionals f_1 and f_2 is the functional s whose value at every $x \in X$ is

$$s(x)=(f_1+f_2)(x)=f_1(x)+f_2(x);$$

the *product* αf of a scalar α and a functional f is the functional p whose value at $x \in X$ is

$$p(x)=(\alpha f)(x)=\alpha f(x).$$

Note that this agrees with the usual way of adding functions and multiplying them by constants.

We may go a step further and consider the algebraic dual $(X^*)^*$ of X^*, whose elements are the linear functionals defined on X^*. We denote $(X^*)^*$ by X^{**} and call it the **second algebraic dual space** of X.

Why do we consider X^{**}? The point is that we can obtain an interesting and important relation between X and X^{**}, as follows. We choose the notations:

Space	General element	Value at a point
X	x	—
X^*	f	$f(x)$
X^{**}	g	$g(f)$

We can obtain a $g \in X^{**}$, which is a linear functional defined on X^*, by choosing a *fixed* $x \in X$ and setting

$$g(f)=g_x(f)=f(x) \qquad (x \in X \text{ fixed, } f \in X^* \text{ variable}). \tag{4}$$

The subscript x is a little reminder that we got g by the use of a certain $x \in X$. The reader should observe carefully that here f is the variable whereas x is fixed. Keeping this in mind, he should not have difficulties in understanding our present consideration.

g_x as defined by (4) is linear. This can be seen from

$$g_x(\alpha f_1+\beta f_2)=(\alpha f_1+\beta f_2)(x)=\alpha f_1(x)+\beta f_2(x)=\alpha g_x(f_1)+\beta g_x(f_2).$$

Hence g_x is an element of X^{**}, by the definition of X^{**}.

To each $x \in X$ there corresponds a $g_x \in X^{**}$. This defines a mapping

$$C: X \longrightarrow X^{**}$$

$$x \longmapsto g_x.$$

C is called the **canonical mapping** of X into X^{**}.

C is linear since its domain is a vector space and we have

$$\begin{aligned}(C(\alpha x+\beta y))(f) &= g_{\alpha x+\beta y}(f) \\ &= f(\alpha x+\beta y) \\ &= \alpha f(x)+\beta f(y) \\ &= \alpha g_x(f)+\beta g_y(f) \\ &= \alpha(Cx)(f)+\beta(Cy)(f).\end{aligned}$$

C is also called the *canonical embedding* of X into X^{**}. To understand and motivate this term, we first explain the concept of "isomorphism," which is of general interest.

In our work we are concerned with various spaces. Common to all of them is that they consist of a set, call it X, and a "structure" defined on X. For a metric space, this is the metric. For a vector space, the two algebraic operations form the structure. And for a normed space the structure consists of those two algebraic operations and the norm.

Given two spaces X and $\tilde{X}$ of the same kind (for instance, two vector spaces), it is of interest to know whether X and $\tilde{X}$ are "essentially identical," that is, whether they differ at most by the nature of their points. Then we can regard X and $\tilde{X}$ as identical—as two copies of the same "abstract" space—whenever the structure is the primary object of study, whereas the nature of the points does not matter. This situation occurs quite often. It suggests the concept of an **isomorphism.** By definition, this is a bijective mapping of X onto $\tilde{X}$ which preserves the structure.

Accordingly, an *isomorphism T of a metric space $X=(X, d)$ onto a metric space* $\tilde{X}=(\tilde{X}, \tilde{d})$ is a bijective mapping which preserves distance, that is, for all $x, y \in X$,

$$\tilde{d}(Tx, Ty)=d(x, y).$$

$\tilde{X}$ is then called *isomorphic* with X. This is nothing new to us but merely another name for a bijective isometry as introduced in Def. 1.6-1. New is the following.

An *isomorphism* T *of a vector space* X onto a vector space $\tilde{X}$ over the same field is a bijective mapping which preserves the two algebraic operations of vector space; thus, for all $x, y \in X$ and scalars α,

$$T(x+y) = Tx + Ty, \qquad T(\alpha x) = \alpha Tx,$$

that is, $T\colon X \longrightarrow \tilde{X}$ is a bijective linear operator. $\tilde{X}$ is then called *isomorphic* with X, and X and $\tilde{X}$ are called *isomorphic vector spaces.*

Isomorphisms for normed spaces are vector space isomorphisms which also preserve norms. Details follow in Sec. 2.10 where we need such isomorphisms. At present we can apply vector space isomorphisms as follows.

It can be shown that the canonical mapping C is injective. Since C is linear (see before), it is an isomorphism of X onto the range $\mathcal{R}(C) \subset X^{**}$.

If X is isomorphic with a subspace of a vector space Y, we say that X is **embeddable** in Y. Hence X is embeddable in X^{**}, and C is also called the *canonical embedding* of X into X^{**}.

If C is surjective (hence bijective), so that $\mathcal{R}(C) = X^{**}$, then X is said to be **algebraically reflexive.** We shall prove in the next section that if X is finite dimensional, then X is algebraically reflexive.

A similar discussion involving norms and leading to the concept of *reflexivity* of a *normed* space will be presented later (in Sec. 4.6), after the development of suitable tools (in particular, the famous Hahn-Banach theorem).

Problems

1. Show that the functionals in 2.8-7 and 2.8-8 are linear.

2. Show that the functionals defined on $C[a, b]$ by

$$f_1(x) = \int_a^b x(t)y_0(t)\,dt \qquad (y_0 \in C[a, b])$$

$$f_2(x) = \alpha x(a) + \beta x(b) \qquad (\alpha, \beta \text{ fixed})$$

are linear and bounded.

3. Find the norm of the linear functional f defined on $C[-1, 1]$ by

$$f(x)=\int_{-1}^{0} x(t)\,dt-\int_{0}^{1} x(t)\,dt.$$

4. Show that

$$f_1(x)=\max_{t\in J} x(t)$$

$$J=[a, b]$$

$$f_2(x)=\min_{t\in J} x(t)$$

define functionals on $C[a, b]$. Are they linear? Bounded?

5. Show that on any sequence space X we can define a linear functional f by setting $f(x)=\xi_n$ (n fixed), where $x=(\xi_j)$. Is f bounded if $X=l^\infty$?

6. (Space $C'[a, b]$) The space $C^1[a, b]$ or $C'[a, b]$ is the normed space of all continuously differentiable functions on $J=[a, b]$ with norm defined by

$$\|x\|=\max_{t\in J}|x(t)|+\max_{t\in J}|x'(t)|.$$

Show that the axioms of a norm are satisfied. Show that $f(x)=x'(c)$, $c=(a+b)/2$, defines a bounded linear functional on $C'[a, b]$. Show that f is not bounded, considered as a functional on the subspace of $C[a, b]$ which consists of all continuously differentiable functions.

7. If f is a bounded linear functional on a complex normed space, is $\bar{f}$ bounded? Linear? (The bar denotes the complex conjugate.)

8. (Null space) The *null space* $N(M^*)$ of a set $M^*\subset X^*$ is defined to be the set of all $x\in X$ such that $f(x)=0$ for all $f\in M^*$. Show that $N(M^*)$ is a vector space.

9. Let $f\neq 0$ be any linear functional on a vector space X and x_0 any fixed element of $X-\mathcal{N}(f)$, where $\mathcal{N}(f)$ is the null space of f. Show that any $x\in X$ has a unique representation $x=\alpha x_0+y$, where $y\in\mathcal{N}(f)$.

10. Show that in Prob. 9, two elements $x_1, x_2\in X$ belong to the same element of the quotient space $X/\mathcal{N}(f)$ if and only if $f(x_1)=f(x_2)$; show that $\operatorname{codim}\mathcal{N}(f)=1$. (Cf. Sec. 2.1, Prob. 14.)

11. Show that two linear functionals $f_1\neq 0$ and $f_2\neq 0$ which are defined on the same vector space and have the same null space are proportional.

12. **(Hyperplane)** If Y is a subspace of a vector space X and codim $Y=1$ (cf. Sec. 2.1, Prob. 14), then every element of X/Y is called a *hyperplane parallel to* Y. Show that for any linear functional $f\neq 0$ on X, the set $H_1=\{x\in X \mid f(x)=1\}$ is a hyperplane parallel to the null space $\mathcal{N}(f)$ of f.

13. If Y is a subspace of a vector space X and f is a linear functional on X such that $f(Y)$ is not the whole scalar field of X, show that $f(y)=0$ for all $y\in Y$.

14. Show that the norm $\|f\|$ of a bounded linear functional $f\neq 0$ on a normed space X can be interpreted geometrically as the reciprocal of the distance $\tilde{d}=\inf\{\|x\| \mid f(x)=1\}$ of the hyperplane $H_1=\{x\in X \mid f(x)=1\}$ from the orgin.

15. **(Half space)** Let $f\neq 0$ be a bounded linear functional on a real normed space X. Then for any scalar c we have a hyperplane $H_c=\{x\in X \mid f(x)=c\}$, and H_c determines the two *half spaces*

$$X_{c1}=\{x \mid f(x)\leqq c\} \qquad \text{and} \qquad X_{c2}=\{x \mid f(x)\geqq c\}.$$

Show that the closed unit ball lies in X_{c1} where $c=\|f\|$, but for no $\varepsilon>0$, the half space X_{c1} with $c=\|f\|-\varepsilon$ contains that ball.

2.9 Linear Operators and Functionals on Finite Dimensional Spaces

Finite dimensional vector spaces are simpler than infinite dimensional ones, and it is natural to ask what simplification this entails with respect to linear operators and functionals defined on such a space. This is the question to be considered, and the answer will clarify the role of (finite) matrices in connection with linear operators as well as the structure of the algebraic dual X^* (Sec. 2.8) of a finite dimensional vector space X.

Linear operators on finite dimensional vector spaces can be represented in terms of matrices, as explained below. In this way, matrices become the most important tools for studying linear operators in the finite dimensional case. In this connection we should also remember Theorem 2.7-8 to understand the full significance of our present consideration. The details are as follows.

Let X and Y be finite dimensional vector spaces over the same field and $T\colon X \longrightarrow Y$ a linear operator. We choose a basis $E=\{e_1, \cdots, e_n\}$ for X and a basis $B=\{b_1, \cdots, b_r\}$ for Y, with the vectors arranged in a definite order which we keep fixed. Then every $x \in X$ has a unique representation

$$x = \xi_1 e_1 + \cdots + \xi_n e_n. \tag{1}$$

Since T is linear, x has the image

$$y = Tx = T\left(\sum_{k=1}^{n} \xi_k e_k\right) = \sum_{k=1}^{n} \xi_k T e_k. \tag{2}$$

Since the representation (1) is unique, we have our first result:

T is uniquely determined if the images $y_k = Te_k$ of the n basis vectors $e_1, \cdots, e_n$ are prescribed.

Since y and $y_k = Te_k$ are in Y, they have unique representations of the form

(3)

(a) $$y = \sum_{j=1}^{r} \eta_j b_j$$

(b) $$Te_k = \sum_{j=1}^{r} \tau_{jk} b_j.$$

Substitution into (2) gives

$$y = \sum_{j=1}^{r} \eta_j b_j = \sum_{k=1}^{n} \xi_k T e_k = \sum_{k=1}^{n} \xi_k \sum_{j=1}^{r} \tau_{jk} b_j = \sum_{j=1}^{r} \left(\sum_{k=1}^{n} \tau_{jk} \xi_k\right) b_j.$$

Since the b_j's form a linearly independent set, the coefficients of each b_j on the left and on the right must be the same, that is,

$$\eta_j = \sum_{k=1}^{n} \tau_{jk} \xi_k \qquad j = 1, \cdots, r. \tag{4}$$

This yields our next result:

The image $y = Tx = \sum \eta_j b_j$ of $x = \sum \xi_k e_k$ can be obtained from (4).

Note the unusual position of the summation index j of τ_{jk} in (3b), which is necessary in order to arrive at the usual position of the summation index in (4).

The coefficients in (4) form a matrix

$$T_{EB} = (\tau_{jk})$$

with r rows and n columns. If a basis E for X and a basis B for Y are given, with the elements of E and B arranged in some definite order (which is arbitrary but fixed), then the matrix T_{EB} is uniquely determined by the linear operator T. We say that the matrix T_{EB} **represents** the operator T with respect to those bases.

By introducing the column vectors $\tilde{x} = (\xi_k)$ and $\tilde{y} = (\eta_j)$ we can write (4) in matrix notation:

(4′) $$\tilde{y} = T_{EB}\tilde{x}.$$

Similarly, (3b) can also be written in matrix notation

(3b′) $$Te = T_{EB}{}^{\mathsf{T}} b$$

where Te is the column vector with components $Te_1, \cdots, Te_n$ (which are themselves vectors) and b is the column vector with components $b_1, \cdots, b_r$, and we have to use the transpose $T_{EB}{}^{\mathsf{T}}$ of T_{EB} because in (3b) we sum over j, which is the first subscript, whereas in (4) we sum over k, which is the second subscript.

Our consideration shows that a linear operator T determines a uniqe matrix representing T with respect to a given basis for X and a given basis for Y, where the vectors of each of the bases are assumed to be arranged in a fixed order. Conversely, any matrix with r rows and n columns determines a linear operator which it represents with respect to given bases for X and Y. (Cf. also 2.6-8 and 2.7-7.)

Let us now turn to **linear functionals** on X, where $\dim X = n$ and $\{e_1, \cdots, e_n\}$ is a basis for X, as before. These functionals constitute the algebraic dual space X^* of X, as we know from the previous section For every such functional f and every $x = \sum \xi_j e_j \in X$ we have

(5a) $$f(x) = f\left(\sum_{j=1}^{n} \xi_j e_j\right) = \sum_{j=1}^{n} \xi_j f(e_j) = \sum_{j=1}^{n} \xi_j \alpha_j$$

where

(5b) $$\alpha_j = f(e_j) \qquad j=1, \cdots, n,$$

and f is uniquely determined by its values α_j at the n basis vectors of X.

Conversely, every n-tuple of scalars $\alpha_1, \cdots, \alpha_n$ determines a linear functional on X by (5). In particular, let us take the n-tuples

$$\begin{array}{cccccc} (1, & 0, & 0, & \cdots & 0, & 0) \\ (0, & 1, & 0, & \cdots & 0, & 0) \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ (0, & 0, & 0, & \cdots & 0, & 1). \end{array}$$

By (5) this gives n functionals, which we denote by $f_1, \cdots, f_n$, with values

(6) $$f_k(e_j) = \delta_{jk} = \begin{cases} 0 & \text{if } j \neq k, \\ 1 & \text{if } j = k; \end{cases}$$

that is, f_k has the value 1 at the kth basis vector and 0 at the $n-1$ other basis vectors. δ_{jk} is called the *Kronecker delta.* $\{f_1, \cdots, f_n\}$ is called the **dual basis** of the basis $\{e_1, \cdots, e_n\}$ for X. This is justified by the following theorem.

2.9-1 Theorem (Dimension of X^*). *Let X be an n-dimensional vector space and $E = \{e_1, \cdots, e_n\}$ a basis for X. Then $F = \{f_1, \cdots, f_n\}$ given by (6) is a basis for the algebraic dual X^* of X, and* dim $X^* =$ dim $X = n$.

Proof. F is a linearly independent set since

(7) $$\sum_{k=1}^{n} \beta_k f_k(x) = 0 \qquad (x \in X)$$

with $x = e_j$ gives

$$\sum_{k=1}^{n} \beta_k f_k(e_j) = \sum_{k=1}^{n} \beta_k \delta_{jk} = \beta_j = 0,$$

so that all the β_k's in (7) are zero. We show that every $f \in X^*$ can be represented as a linear combination of the elements of F in a unique way. We write $f(e_j) = \alpha_j$ as in (5b). By (5a),

$$f(x) = \sum_{j=1}^{n} \xi_j \alpha_j$$

for every $x \in X$. On the other hand, by (6) we obtain

$$f_j(x) = f_j(\xi_1 e_1 + \cdots + \xi_n e_n) = \xi_j.$$

Together,

$$f(x) = \sum_{j=1}^{n} \alpha_j f_j(x).$$

Hence the unique representation of the arbitrary linear functional f on X in terms of the functionals $f_1, \cdots, f_n$ is

$$f = \alpha_1 f_1 + \cdots + \alpha_n f_n. \qquad \blacksquare$$

To prepare for an interesting application of this theorem, we first prove the following lemma. (A similar lemma for arbitrary normed spaces will be given later, in 4.3-4.)

2.9-2 Lemma (Zero vector). *Let X be a finite dimensional vector space. If $x_0 \in X$ has the property that $f(x_0) = 0$ for all $f \in X^*$, then $x_0 = 0$.*

Proof. Let $\{e_1, \cdots, e_n\}$ be a basis for X and $x_0 = \sum \xi_{0j} e_j$. Then (5) becomes

$$f(x_0) = \sum_{j=1}^{n} \xi_{0j} \alpha_j.$$

By assumption this is zero for every $f \in X^*$, that is, for every choice of $\alpha_1, \cdots, \alpha_n$. Hence all ξ_{0j} must be zero. $\blacksquare$

Using this lemma, we can now obtain

2.9-3 Theorem (Algebraic reflexivity). *A finite dimensional vector space is algebraically reflexive.*

Proof. The canonical mapping $C: X \longrightarrow X^{**}$ considered in the previous section is linear. $Cx_0 = 0$ means that for all $f \in X^*$ we have

$$(Cx_0)(f) = g_{x_0}(f) = f(x_0) = 0,$$

by the definition of C. This implies $x_0 = 0$ by Lemma 2.9-2. Hence from Theorem 2.6-10 it follows that the mapping C has an inverse $C^{-1}: \mathscr{R}(C) \longrightarrow X$, where $\mathscr{R}(C)$ is the range of C. We also have $\dim \mathscr{R}(C) = \dim X$ by the same theorem. Now by Theorem 2.9-1,

$$\dim X^{**} = \dim X^* = \dim X.$$

Together, $\dim \mathscr{R}(C) = \dim X^{**}$. Hence $\mathscr{R}(C) = X^{**}$ because $\mathscr{R}(C)$ is a vector space (cf. 2.6-9) and a proper subspace of X^{**} has dimension less than $\dim X^{**}$, by Theorem 2.1-8. By the definition, this proves algebraic reflexivity. ∎

Problems

1. Determine the null space of the operator $T: \mathbf{R}^3 \longrightarrow \mathbf{R}^2$ represented by

$$\begin{bmatrix} 1 & 3 & 2 \\ -2 & 1 & 0 \end{bmatrix}.$$

2. Let $T: \mathbf{R}^3 \longrightarrow \mathbf{R}^3$ be defined by $(\xi_1, \xi_2, \xi_3) \longmapsto (\xi_1, \xi_2, -\xi_1 - \xi_2)$. Find $\mathscr{R}(T)$, $\mathscr{N}(T)$ and a matrix which represents T.

3. Find the dual basis of the basis $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ for $\mathbf{R}^3$.

4. Let $\{f_1, f_2, f_3\}$ be the dual basis of $\{e_1, e_2, e_3\}$ for $\mathbf{R}^3$, where $e_1 = (1, 1, 1)$, $e_2 = (1, 1, -1)$, $e_3 = (1, -1, -1)$. Find $f_1(x)$, $f_2(x)$, $f_3(x)$, where $x = (1, 0, 0)$.

5. If f is a linear functional on an n-dimensional vector space X, what dimension can the null space $\mathscr{N}(f)$ have?

6. Find a basis for the null space of the functional f defined on $\mathbf{R}^3$ by $f(x) = \xi_1 + \xi_2 - \xi_3$, where $x = (\xi_1, \xi_2, \xi_3)$.

7. Same task as in Prob. 6, if $f(x) = \alpha_1\xi_1 + \alpha_2\xi_2 + \alpha_3\xi_3$, where $\alpha_1 \neq 0$.

8. If Z is an $(n-1)$-dimensional subspace of an n-dimensional vector space X, show that Z is the null space of a suitable linear functional f on X, which is uniquely determined to within a scalar multiple.

9. Let X be the vector space of all real polynomials of a real variable and of degree less than a given n, together with the polynomial $x=0$ (whose degree is left undefined in the usual discussion of degree). Let $f(x)=x^{(k)}(a)$, the value of the kth derivative (k fixed) of $x \in X$ at a fixed $a \in \mathbf{R}$. Show that f is a linear functional on X.

10. Let Z be a proper subspace of an n-dimensional vector space X, and let $x_0 \in X - Z$. Show that there is a linear functional f on X such that $f(x_0)=1$ and $f(x)=0$ for all $x \in Z$.

11. If x and y are different vectors in a finite dimensional vector space X, show that there is a linear functional f on X such that $f(x) \neq f(y)$.

12. If $f_1, \cdots, f_p$ are linear functionals on an n-dimensional vector space X, where $p<n$, show that there is a vector $x \neq 0$ in X such that $f_1(x)=0, \cdots, f_p(x)=0$. What consequences does this result have with respect to linear equations?

13. (Linear extension) Let Z be a proper subspace of an n-dimensional vector space X, and let f be a linear functional on Z. Show that f can be *extended linearly* to X, that is, there is a linear functional $\tilde{f}$ on X such that $\tilde{f}|_Z = f$.

14. Let the functional f on $\mathbf{R}^2$ be defined by $f(x)=4\xi_1-3\xi_2$, where $x=(\xi_1, \xi_2)$. Regard $\mathbf{R}^2$ as the subspace of $\mathbf{R}^3$ given by $\xi_3=0$. Determine all linear extensions $\tilde{f}$ of f from $\mathbf{R}^2$ to $\mathbf{R}^3$.

15. Let $Z \subset \mathbf{R}^3$ be the subspace represented by $\xi_2=0$ and let f on Z be defined by $f(x)=(\xi_1-\xi_3)/2$. Find a linear extension $\tilde{f}$ of f to $\mathbf{R}^3$ such that $\tilde{f}(x_0)=k$ (a given constant), where $x_0=(1, 1, 1)$. Is $\tilde{f}$ unique?

2.10 Normed Spaces of Operators. Dual Space

In Sec. 2.7 we defined the concept of a bounded linear operator and illustrated it by basic examples which gave the reader a first impression of the importance of these operators. In the present section our goal is as follows. We take any two normed spaces X and Y (both real or both complex) and consider the set

$$B(X, Y)$$

consisting of all bounded linear operators from X into Y, that is, each such operator is defined on all of X and its range lies in Y. We want to show that $B(X, Y)$ can itself be made into a normed space.[10]

The whole matter is quite simple. First of all, $B(X, Y)$ becomes a vector space if we define the sum T_1+T_2 of two operators T_1, $T_2 \in B(X, Y)$ in a natural way by

$$(T_1+T_2)x = T_1x + T_2x$$

and the product αT of $T \in B(X, Y)$ and a scalar α by

$$(\alpha T)x = \alpha Tx.$$

Now we remember Lemma 2.7-2(b) and have at once the desired result:

2.10-1 Theorem (Space $B(X, Y)$). *The vector space $B(X, Y)$ of all bounded linear operators from a normed space X into a normed space Y is itself a normed space with norm defined by*

$$\|T\| = \sup_{\substack{x \in X \\ x \neq 0}} \frac{\|Tx\|}{\|x\|} = \sup_{\substack{x \in X \\ \|x\|=1}} \|Tx\|. \tag{1}$$

In what case will $B(X, Y)$ be a Banach space? This is a central question, which is answered in the following theorem. It is remarkable that the condition in the theorem does not involve X; that is, X may or may not be complete:

2.10-2 Theorem (Completeness). *If Y is a Banach space, then $B(X, Y)$ is a Banach space.*

Proof. We consider an arbitrary Cauchy sequence (T_n) in $B(X, Y)$ and show that (T_n) converges to an operator $T \in B(X, Y)$.

[10] B in $B(X, Y)$ suggests "bounded." Another notation for $B(X, Y)$ is $L(X, Y)$, where L suggests "linear." Both notations are common. We use $B(X, Y)$ throughout.

Since (T_n) is Cauchy, for every $\varepsilon > 0$ there is an N such that

$$\|T_n - T_m\| < \varepsilon \qquad (m, n > N).$$

For all $x \in X$ and $m, n > N$ we thus obtain [cf. (3) in Sec. 2.7]

$$(2) \qquad \|T_n x - T_m x\| = \|(T_n - T_m)x\| \leqq \|T_n - T_m\| \, \|x\| < \varepsilon \|x\|.$$

Now for any fixed x and given $\tilde{\varepsilon}$ we may choose $\varepsilon = \varepsilon_x$ so that $\varepsilon_x \|x\| < \tilde{\varepsilon}$. Then from (2) we have $\|T_n x - T_m x\| < \tilde{\varepsilon}$ and see that $(T_n x)$ is Cauchy in Y. Since Y is complete, $(T_n x)$ converges, say, $T_n x \longrightarrow y$. Clearly, the limit $y \in Y$ depends on the choice of $x \in X$. This defines an operator $T: X \longrightarrow Y$, where $y = Tx$. The operator T is linear since

$$\lim T_n(\alpha x + \beta z) = \lim (\alpha T_n x + \beta T_n z) = \alpha \lim T_n x + \beta \lim T_n z.$$

We prove that T is bounded and $T_n \longrightarrow T$, that is, $\|T_n - T\| \longrightarrow 0$.

Since (2) holds for every $m > N$ and $T_m x \longrightarrow Tx$, we may let $m \longrightarrow \infty$. Using the continuity of the norm, we then obtain from (2) for every $n > N$ and all $x \in X$

$$(3) \qquad \|T_n x - Tx\| = \|T_n x - \lim_{m \to \infty} T_m x\| = \lim_{m \to \infty} \|T_n x - T_m x\| \leqq \varepsilon \|x\|.$$

This shows that $(T_n - T)$ with $n > N$ is a bounded linear operator. Since T_n is bounded, $T = T_n - (T_n - T)$ is bounded, that is, $T \in B(X, Y)$. Furthermore, if in (3) we take the supremum over all x of norm 1, we obtain

$$\|T_n - T\| \leqq \varepsilon \qquad (n > N).$$

Hence $\|T_n - T\| \longrightarrow 0$. ∎

This theorem has an important consequence with respect to the dual space X' of X, which is defined as follows.

2.10-3 Definition (Dual space X'). Let X be a normed space. Then the set of all bounded linear functionals on X constitutes a normed space with norm defined by

$$\|f\| = \sup_{\substack{x \in X \\ x \neq 0}} \frac{|f(x)|}{\|x\|} = \sup_{\substack{x \in X \\ \|x\|=1}} |f(x)| \tag{4}$$

[cf. (2) in Sec. 2.8] which is called the *dual space*[11] of X and is denoted by X'. ∎

Since a linear functional on X maps X into $\mathbf{R}$ or $\mathbf{C}$ (the scalar field of X), and since $\mathbf{R}$ or $\mathbf{C}$, taken with the usual metric, is complete, we see that X' is $B(X, Y)$ with the complete space $Y = \mathbf{R}$ or $\mathbf{C}$. Hence Theorem 2.10-2 is applicable and implies the basic.

2.10-4 Theorem (Dual space). *The dual space X' of a normed space X is a Banach space (whether or not X is).*

It is a fundamental principle of functional analysis that investigations of spaces are often combined with those of the dual spaces. For this reason it is worthwhile to consider some of the more frequently occurring spaces and find out what their duals look like. In this connection the concept of an isomorphism will be helpful in understanding the present discussion. Remembering our consideration in Sec. 2.8, we give the following definition.

An **isomorphism** of a normed space X onto a normed space $\tilde{X}$ is a bijective linear operator $T: X \longrightarrow \tilde{X}$ which preserves the norm, that is, for all $x \in X$,

$$\|Tx\| = \|x\|.$$

(Hence T is isometric.) X is then called *isomorphic* with $\tilde{X}$, and X and $\tilde{X}$ are called *isomorphic normed spaces.*—From an abstract point of view, X and $\tilde{X}$ are then identical, the isomorphism merely amounting to renaming of the elements (attaching a "tag" T to each point).

Our first example shows that the dual space of $\mathbf{R}^n$ is isomorphic with $\mathbf{R}^n$; we express this more briefly by saying that the dual space of $\mathbf{R}^n$ *is* $\mathbf{R}^n$; similarly for the other examples.

[11] Other terms are *dual, adjoint space* and *conjugate space.* Remember from Sec. 2.8 that the *algebraic* dual space X^* of X is the vector space of *all* linear functionals on X.

Examples

2.10-5 Space $\mathbf{R}^n$. *The dual space of $\mathbf{R}^n$ is $\mathbf{R}^n$.*

Proof. We have $\mathbf{R}^{n\prime} = \mathbf{R}^{n*}$ by Theorem 2.7-8, and every $f \in \mathbf{R}^{n*}$ has a representation (5), Sec. 2.9:

$$f(x) = \sum \xi_k \gamma_k \qquad \gamma_k = f(e_k)$$

(sum from 1 to n). By the Cauchy-Schwarz inequality (Sec. 1.2),

$$|f(x)| \leqq \sum |\xi_k \gamma_k| \leqq \left(\sum \xi_j^2\right)^{1/2} \left(\sum \gamma_k^2\right)^{1/2} = \|x\| \left(\sum \gamma_k^2\right)^{1/2}.$$

Taking the supremum over all x of norm 1 we obtain

$$\|f\| \leqq \left(\sum \gamma_k^2\right)^{1/2}.$$

However, since for $x = (\gamma_1, \cdots, \gamma_n)$ equality is achieved in the Cauchy-Schwarz inequality, we must in fact have

$$\|f\| = \left(\sum_{k=1}^{n} \gamma_k^2\right)^{1/2}.$$

This proves that the norm of f is the Euclidean norm, and $\|f\| = \|c\|$, where $c = (\gamma_k) \in \mathbf{R}^n$. Hence the mapping of $\mathbf{R}^{n\prime}$ onto $\mathbf{R}^n$ defined by $f \longmapsto c = (\gamma_k)$, $\gamma_k = f(e_k)$, is norm preserving and, since it is linear and bijective, it is an isomorphism. ∎

2.10-6 Space l^1. *The dual space of l^1 is l^∞.*

Proof. A Schauder basis (Sec. 2.3) for l^1 is (e_k), where $e_k = (\delta_{kj})$ has 1 in the kth place and zeros otherwise. Then every $x \in l^1$ has a unique representation

$$x = \sum_{k=1}^{\infty} \xi_k e_k. \tag{5}$$

We consider any $f \in l^{1\prime}$, where $l^{1\prime}$ is the dual space of l^1. Since f is linear and bounded,

$$f(x) = \sum_{k=1}^{\infty} \xi_k \gamma_k \qquad \gamma_k = f(e_k) \tag{6}$$

where the numbers $\gamma_k = f(e_k)$ are uniquely determined by f. Also $\|e_k\| = 1$ and

(7) $$|\gamma_k| = |f(e_k)| \leqq \|f\| \, \|e_k\| = \|f\|, \qquad \sup_k |\gamma_k| \leqq \|f\|.$$

Hence $(\gamma_k) \in l^\infty$.

On the other hand, for every $b = (\beta_k) \in l^\infty$ we can obtain a corresponding bounded linear functional g on l^1. In fact, we may define g on l^1 by

$$g(x) = \sum_{k=1}^{\infty} \xi_k \beta_k$$

where $x = (\xi_k) \in l^1$. Then g is linear, and boundedness follows from

$$|g(x)| \leqq \sum |\xi_k \beta_k| \leqq \sup_j |\beta_j| \sum |\xi_k| = \|x\| \sup_j |\beta_j|$$

(sum from 1 to ∞). Hence $g \in l^{1\prime}$.

We finally show that the norm of f is the norm on the space l^∞. From (6) we have

$$|f(x)| = \left| \sum \xi_k \gamma_k \right| \leqq \sup_j |\gamma_j| \sum |\xi_k| = \|x\| \sup_j |\gamma_j|.$$

Taking the supremum over all x of norm 1, we see that

$$\|f\| \leqq \sup_j |\gamma_j|.$$

From this and (7),

(8) $$\|f\| = \sup_j |\gamma_j|,$$

which is the norm on l^∞. Hence this formula can be written $\|f\| = \|c\|_\infty$, where $c = (\gamma_j) \in l^\infty$. It shows that the bijective linear mapping of $l^{1\prime}$ onto l^∞ defined by $f \longmapsto c = (\gamma_j)$ is an isomorphism. ∎

2.10-7 Space l^p. *The dual space of l^p is l^q; here, $1 < p < +\infty$ and q is the conjugate of p, that is, $1/p + 1/q = 1$.*

Proof. A Schauder basis for l^p is (e_k), where $e_k=(\delta_{kj})$ as in the preceding example. Then every $x\in l^p$ has a unique representation

$$x=\sum_{k=1}^{\infty}\xi_k e_k. \tag{9}$$

We consider any $f\in l^{p\prime}$, where $l^{p\prime}$ is the dual space of l^p. Since f is linear and bounded,

$$f(x)=\sum_{k=1}^{\infty}\xi_k\gamma_k \qquad\qquad \gamma_k=f(e_k). \tag{10}$$

Let q be the conjugate of p (cf. 1.2-3) and consider $x_n=(\xi_k^{(n)})$ with

$$\xi_k^{(n)}=\begin{cases}|\gamma_k|^q/\gamma_k & \text{if } k\leqq n \text{ and } \gamma_k\neq 0,\\ 0 & \text{if } k>n \text{ or } \gamma_k=0.\end{cases} \tag{11}$$

By substituting this into (10) we obtain

$$f(x_n)=\sum_{k=1}^{\infty}\xi_k^{(n)}\gamma_k=\sum_{k=1}^{n}|\gamma_k|^q.$$

We also have, using (11) and $(q-1)p=q$,

$$\begin{aligned} f(x_n)\leqq \|f\|\,\|x_n\| &= \|f\|\Big(\sum|\xi_k^{(n)}|^p\Big)^{1/p}\\ &= \|f\|\Big(\sum|\gamma_k|^{(q-1)p}\Big)^{1/p}\\ &= \|f\|\Big(\sum|\gamma_k|^q\Big)^{1/p}\end{aligned}$$

(sum from 1 to n). Together,

$$f(x_n)=\sum|\gamma_k|^q\leqq\|f\|\Big(\sum|\gamma_k|^q\Big)^{1/p}.$$

Dividing by the last factor and using $1-1/p=1/q$, we get

$$\Big(\sum_{k=1}^{n}|\gamma_k|^q\Big)^{1-1/p}=\Big(\sum_{k=1}^{n}|\gamma_k|^q\Big)^{1/q}\leqq\|f\|.$$

Since n is arbitrary, letting $n \longrightarrow \infty$, we obtain

$$\left(\sum_{k=1}^{\infty} |\gamma_k|^q\right)^{1/q} \leqq \|f\|. \tag{12}$$

Hence $(\gamma_k) \in l^q$.

Conversely, for any $b = (\beta_k) \in l^q$ we can get a corresponding bounded linear functional g on l^p. In fact, we may define g on l^p by setting

$$g(x) = \sum_{k=1}^{\infty} \xi_k \beta_k$$

where $x = (\xi_k) \in l^p$. Then g is linear, and boundedness follows from the Hölder inequality (10), Sec. 1.2. Hence $g \in l^{p\prime}$.

We finally prove that the norm of f is the norm on the space l^q. From (10) and the Hölder inequality we have

$$\begin{aligned} |f(x)| = \left|\sum \xi_k \gamma_k\right| &\leqq \left(\sum |\xi_k|^p\right)^{1/p} \left(\sum |\gamma_k|^q\right)^{1/q} \\ &= \|x\| \left(\sum |\gamma_k|^q\right)^{1/q} \end{aligned}$$

(sum from 1 to ∞); hence by taking the supremum over all x of norm 1 we obtain

$$\|f\| \leqq \left(\sum |\gamma_k|^q\right)^{1/q}.$$

From (12) we see that the equality sign must hold, that is,

$$\|f\| = \left(\sum_{k=1}^{\infty} |\gamma_k|^q\right)^{1/q}. \tag{13}$$

This can be written $\|f\| = \|c\|_q$, where $c = (\gamma_k) \in l^q$ and $\gamma_k = f(e_k)$. The mapping of $l^{p\prime}$ onto l^q defined by $f \longmapsto c$ is linear and bijective, and from (13) we see that it is norm preserving, so that it is an isomorphism. ∎

What is the significance of these and similar examples? In applications it is frequently quite useful to know the general form of bounded linear functionals on spaces of practical importance, and many spaces have been investigated in that respect. Our examples give general representations of bounded linear functionals on $\mathbf{R}^n$, l^1 and l^p with $p>1$. The space $C[a, b]$ will be considered later, in Sec. 4.4, since this will require additional tools (in particular the so-called Hahn-Banach theorem).

Furthermore, remembering the discussion of the second algebraic dual space X^{**} in Sec. 2.8, we may ask whether it is worthwhile to consider $X''=(X')'$, the second dual space of X. The answer is in the affirmative, but we have to postpone this discussion until Sec. 4.6 where we develop suitable tools for obtaining substantial results in that direction. At present let us turn to matters which are somewhat simpler, namely, to inner product and Hilbert spaces. We shall see that these are special normed spaces which are of great importance in applications.

Problems

1. What is the zero element of the vector space $B(X, Y)$? The inverse of a $T \in B(X, Y)$ in the sense of Def. 2.1-1?

2. The operators and functionals considered in the text are defined on the entire space X. Show that without that assumption, in the case of functionals we still have the following theorem. If f and g are bounded linear functionals with domains in a normed space X, then for any nonzero scalars α and β the linear combination $h=\alpha f+\beta g$ is a bounded linear functional with domain $\mathfrak{D}(h)=\mathfrak{D}(f)\cap\mathfrak{D}(g)$.

3. Extend the theorem in Prob. 2 to bounded linear operators T_1 and T_2.

4. Let X and Y be normed spaces and $T_n: X \longrightarrow Y$ $(n=1, 2, \cdots)$ bounded linear operators. Show that convergence $T_n \longrightarrow T$ implies that for every $\varepsilon>0$ there is an N such that for all $n>N$ and all x in any given closed ball we have $\|T_nx-Tx\|<\varepsilon$.

5. Show that 2.8-5 is in agreement with 2.10-5.

6. If X is the space of ordered n-tuples of real numbers and $\|x\|=\max_j |\xi_j|$, where $x=(\xi_1, \cdots, \xi_n)$, what is the corresponding norm on the dual space X'?

7. What conclusion can we draw from 2.10-6 with respect to the space X of all ordered n-tuples of real numbers?

8. Show that the dual space of the space c_0 is l^1. (Cf. Prob. 1 in Sec. 2.3.)

9. Show that a linear functional f on a vector space X is uniquely determined by its values on a Hamel basis for X. (Cf. Sec. 2.1.)

10. Let X and $Y \neq \{0\}$ be normed spaces, where $\dim X = \infty$. Show that there is at least one unbounded linear operator $T: X \longrightarrow Y$. (Use a Hamel basis.)

11. If X is a normed space and $\dim X = \infty$, show that the dual space X' is not identical with the algebraic dual space X^*.

12. **(Completeness)** The examples in the text can be used to prove completeness of certain spaces. How? For what spaces?

13. **(Annihilator)** Let $M \neq \varnothing$ be any subset of a normed space X. The *annihilator* M^a of M is defined to be the set of all bounded linear functionals on X which are zero everywhere on M. Thus M^a is a subset of the dual space X' of X. Show that M^a is a vector subspace of X' and is closed. What are X^a and $\{0\}^a$?

14. If M is an m-dimensional subspace of an n-dimensional normed space X, show that M^a is an $(n-m)$-dimensional subspace of X'. Formulate this as a theorem about solutions of a system of linear equations.

15. Let $M = \{(1, 0, -1), (1, -1, 0), (0, 1, -1)\} \subset \mathbf{R}^3$. Find a basis for M^a.