§ 13. Dual spaces

DEFINITION. A *linear functional* on a vector space $\mathcal{V}$ is a scalar-valued function y defined for every vector x, with the property that (identically in the vectors x_1 and x_2 and the scalars α_1 and α_2)

$$y(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 y(x_1) + \alpha_2 y(x_2).$$

Let us look at some examples of linear functionals.

(1) For $x = (\xi_1, \cdots, \xi_n)$ in $\mathcal{C}^n$, write $y(x) = \xi_1$. More generally, let $\alpha_1, \cdots, \alpha_n$ be any n scalars and write

$$y(x) = \alpha_1 \xi_1 + \cdots + \alpha_n \xi_n.$$

We observe that for any linear functional y on any vector space

$$y(0) = y(0 \cdot 0) = 0 \cdot y(0) = 0;$$

for this reason a linear functional, as we defined it, is sometimes called *homogeneous*. In particular in $\mathcal{C}^n$, if y is defined by

$$y(x) = \alpha_1 \xi_1 + \cdots + \alpha_n \xi_n + \beta,$$

then y is not a linear functional unless $\beta = 0$.

(2) For any polynomial x in $\mathcal{P}$, write $y(x) = x(0)$. More generally, let $\alpha_1, \cdots, \alpha_n$ be any n scalars, let $t_1, \cdots, t_n$ be any n real numbers, and write

$$y(x) = \alpha_1 x(t_1) + \cdots + \alpha_n x(t_n).$$

Another example, in a sense a limiting case of the one just given, is obtained as follows. Let (a, b) be any finite interval on the real t-axis, and let α be any complex-valued integrable function defined on (a, b); define y by

$$y(x) = \int_a^b \alpha(t) x(t)\, dt.$$

(3) On an arbitrary vector space $\mathcal{V}$, define y by writing

$$y(x) = 0$$

for every x in $\mathcal{V}$.

The last example is the first hint of a general situation. Let $\mathcal{V}$ be any vector space and let $\mathcal{V}'$ be the collection of all linear functionals on $\mathcal{V}$. Let us denote by 0 the linear functional defined in (3) (compare the comment at the end of § 4). If y_1 and y_2 are linear functionals on $\mathcal{V}$ and if α_1 and α_2 are scalars, let us write y for the function defined by

$$y(x) = \alpha_1 y_1(x) + \alpha_2 y_2(x).$$

It is easy to see that y is a linear functional; we denote it by $\alpha_1 y_1 + \alpha_2 y_2$. With these definitions of the linear concepts (zero, addition, scalar multiplication), the set $\mathcal{V}'$ forms a vector space, the *dual space* of $\mathcal{V}$.

§ 14. Brackets

Before studying linear functionals and dual spaces in more detail, we wish to introduce a notation that may appear weird at first sight but that will clarify many situations later on. Usually we denote a linear functional by a single letter such as y. Sometimes, however, it is necessary to use the function notation fully and to indicate somehow that if y is a linear functional on $\mathcal{V}$ and if x is a vector in $\mathcal{V}$, then $y(x)$ is a particular scalar. According to the notation we propose to adopt here, we shall not write y followed by x in parentheses, but, instead, we shall write x and y enclosed between square brackets and separated by a comma. Because of the unusual nature of this notation, we shall expend on it some further verbiage.

As we have just pointed out $[x, y]$ is a substitute for the ordinary function symbol $y(x)$; both these symbols denote the scalar we obtain if we take the value of the linear function y at the vector x. Let us take an analogous situation (concerned with functions that are, however, not linear). Let y be the real function of a real variable defined for each real number x by $y(x) = x^2$. The notation $[x, y]$ is a symbolic way of writing down the recipe for actual operations performed; it corresponds to the sentence [take a number, and square it].

Using this notation, we may sum up: to every vector space $\mathcal{V}$ we make correspond the dual space $\mathcal{V}'$ consisting of all linear functionals on $\mathcal{V}$; to every pair, x and y, where x is a vector in $\mathcal{V}$ and y is a linear functional in $\mathcal{V}'$, we make correspond the scalar $[x, y]$ defined to be the value of y at x. In terms of the symbol $[x, y]$ the defining property of a linear functional is

(1) $$[\alpha_1 x_1 + \alpha_2 x_2, y] = \alpha_1[x_1, y] + \alpha_2[x_2, y],$$

and the definition of the linear operations for linear functionals is

(2) $$[x, \alpha_1 y_1 + \alpha_2 y_2] = \alpha_1[x, y_1] + \alpha_2[x, y_2].$$

The two relations together are expressed by saying that $[x, y]$ is a *bilinear functional* of the vectors x in $\mathcal{V}$ and y in $\mathcal{V}'$.

EXERCISES

1. Consider the set $\mathcal{C}$ of complex numbers as a real vector space (as in § 3, (9)). Suppose that for each $x = \xi_1 + i\xi_2$ in $\mathcal{C}$ (where ξ_1 and ξ_2 are real numbers and $i = \sqrt{-1}$) the function y is defined by
(a) $y(x) = \xi_1$,
(b) $y(x) = \xi_2$,
(c) $y(x) = \xi_1^2$,
(d) $y(x) = \xi_1 - i\xi_2$,
(e) $y(x) = \sqrt{\xi_1^2 + \xi_2^2}$. (The square root sign attached to a positive number always denotes the positive square root of that number.)
In which of these cases is y a linear functional?

2. Suppose that for each $x = (\xi_1, \xi_2, \xi_3)$ in $\mathcal{C}^3$ the function y is defined by
(a) $y(x) = \xi_1 + \xi_2$,
(b) $y(x) = \xi_1 - \xi_3^2$,
(c) $y(x) = \xi_1 + 1$,
(d) $y(x) = \xi_1 - 2\xi_2 + 3\xi_3$.
In which of these cases is y a linear functional?

3. Suppose that for each x in $\mathcal{P}$ the function y is defined by

(a) $y(x) = \int_{-1}^{+2} x(t)\,dt$,

(b) $y(x) = \int_0^2 (x(t))^2\,dt$,

(c) $y(x) = \int_0^1 t^2 x(t)\,dt$,

(d) $y(x) = \int_0^1 x(t^2)\,dt$,

(e) $y(x) = \dfrac{dx}{dt}$,

(f) $y(x) = \left.\dfrac{d^2x}{dt^2}\right|_{t=1}$.

In which of these cases is y a linear functional?

4. If $(\alpha_0, \alpha_1, \alpha_2, \cdots)$ is an arbitrary sequence of complex numbers, and if x is an element of $\mathcal{P}$, $x(t) = \sum_{i=0}^n \xi_i t^i$, write $y(x) = \sum_{i=0}^n \xi_i\alpha_i$. Prove that y is an element of $\mathcal{P}'$ and that every element of $\mathcal{P}'$ can be obtained in this manner by a suitable choice of the α's.

5. If y is a non-zero linear functional on a vector space $\mathcal{V}$, and if α is an arbitrary scalar, does there necessarily exist a vector x in $\mathcal{V}$ such that $[x, y] = \alpha$?

6. Prove that if y and z are linear functionals (on the same vector space) such that $[x, y] = 0$ whenever $[x, z] = 0$, then there exists a scalar α such that $y = \alpha z$. (Hint: if $[x_0, z] \neq 0$, write $\alpha = [x_0, y]/[x_0, z]$.)

§ 15. Dual bases

One more word before embarking on the proofs of the important theorems. The concept of dual space was defined without any reference to coordinate systems; a glance at the following proofs will show a superabundance of coordinate systems. We wish to point out that this phenomenon is inevitable; we shall be establishing results concerning dimension, and dimension is the one concept (so far) whose very definition is given in terms of a basis.

THEOREM 1. *If* $\mathcal{V}$ *is an n-dimensional vector space, if* $\{x_1, \cdots, x_n\}$ *is a basis in* $\mathcal{V}$, *and if* $\{\alpha_1, \cdots, \alpha_n\}$ *is any set of n scalars, then there is one and only one linear functional y on* $\mathcal{V}$ *such that* $[x_i, y] = \alpha_i$ *for* $i = 1, \cdots, n$.

PROOF. Every x in $\mathcal{V}$ may be written in the form $x = \xi_1 x_1 + \cdots + \xi_n x_n$ in one and only one way; if y is any linear functional, then

$$[x, y] = \xi_1[x_1, y] + \cdots + \xi_n[x_n, y].$$

From this relation the uniqueness of y is clear; if $[x_i, y] = \alpha_i$, then the value of $[x, y]$ is determined, for every x, by $[x, y] = \sum_i \xi_i\alpha_i$. The argument can also be turned around; if we define y by

$$[x, y] = \xi_1\alpha_1 + \cdots + \xi_n\alpha_n,$$

then y is indeed a linear functional, and $[x_i, y] = \alpha_i$.

THEOREM 2. *If* $\mathcal{V}$ *is an n-dimensional vector space and if* $\mathcal{X} = \{x_1, \cdots, x_n\}$ *is a basis in* $\mathcal{V}$, *then there is a uniquely determined basis* $\mathcal{X}'$ *in* $\mathcal{V}'$, $\mathcal{X}' = \{y_1, \cdots, y_n\}$, *with the property that* $[x_i, y_j] = \delta_{ij}$. *Consequently the dual space of an n-dimensional space is n-dimensional.*

The basis $\mathcal{X}'$ is called the *dual basis* of $\mathcal{X}$.

PROOF. It follows from Theorem 1 that, for each $j = 1, \cdots, n$, a unique y_j in $\mathcal{V}'$ can be found so that $[x_i, y_j] = \delta_{ij}$; we have only to prove that the set $\mathcal{X}' = \{y_1, \cdots, y_n\}$ is a basis in $\mathcal{V}'$.

In the first place, $\mathcal{X}'$ is a linearly independent set, for if we had $\alpha_1 y_1 + \cdots + \alpha_n y_n = 0$, in other words, if

$$[x, \alpha_1 y_1 + \cdots + \alpha_n y_n] = \alpha_1[x, y_1] + \cdots + \alpha_n[x, y_n] = 0$$

for all x, then we should have, for $x = x_i$,

$$0 = \sum_j \alpha_j[x_i, y_j] = \sum_j \alpha_j\delta_{ij} = \alpha_i.$$

In the second place, every y in $\mathcal{V}'$ is a linear combination of $y_1, \cdots, y_n$. To prove this, write $[x_i, y] = \alpha_i$; then, for $x = \sum_i \xi_i x_i$, we have

$$[x, y] = \xi_1\alpha_1 + \cdots + \xi_n\alpha_n.$$

On the other hand

$$[x, y_j] = \sum_i \xi_i[x_i, y_j] = \xi_j,$$

so that, substituting in the preceding equation, we get

$$[x, y] = \alpha_1[x, y_1] + \cdots + \alpha_n[x, y_n]$$

$$= [x, \alpha_1 y_1 + \cdots + \alpha_n y_n].$$

Consequently $y = \alpha_1 y_1 + \cdots + \alpha_n y_n$, and the proof of the theorem is complete.

We shall need also the following easy consequence of Theorem 2.

THEOREM 3. *If u and v are any two different vectors of the n-dimensional vector space $\mathcal{V}$, then there exists a linear functional y on $\mathcal{V}$ such that $[u, y] \neq [v, y]$; or, equivalently, to any non-zero vector x in $\mathcal{V}$ there corresponds a y in $\mathcal{V}'$ such that $[x, y] \neq 0$.*

PROOF. That the two statements in the theorem are indeed equivalent is seen by considering $x = u - v$. We shall, accordingly, prove the latter statement only.

Let $\mathfrak{X} = \{x_1, \cdots, x_n\}$ be any basis in $\mathcal{V}$, and let $\mathfrak{X}' = \{y_1, \cdots, y_n\}$ be the dual basis in $\mathcal{V}'$. If $x = \sum_i \xi_i x_i$, then (as above) $[x, y_j] = \xi_j$. Hence if $[x, y] = 0$ for all y, and, in particular, if $[x, y_j] = 0$ for $j = 1, \cdots, n$, then $x = 0$.

§ 16. Reflexivity

It is natural to think that if the dual space $\mathcal{V}'$ of a vector space $\mathcal{V}$, and the relations between a space and its dual, are of any interest at all for $\mathcal{V}$, then they are of just as much interest for $\mathcal{V}'$. In other words, we propose now to form the dual space $(\mathcal{V}')'$ of $\mathcal{V}'$; for simplicity of notation we shall denote it by $\mathcal{V}''$. The verbal description of an element of $\mathcal{V}''$ is clumsy: such an element is a linear functional of linear functionals. It is, however, at this point that the greatest advantage of the notation $[x, y]$ appears; by means of it, it is easy to discuss $\mathcal{V}$ and its relation to $\mathcal{V}''$.

If we consider the symbol $[x, y]$ for some fixed $y = y_0$, we obtain nothing new: $[x, y_0]$ is merely another way of writing the value $y_0(x)$ of the function y_0 at the vector x. If, however, we consider the symbol $[x, y]$ for some fixed $x = x_0$, then we observe that the function of the vectors in $\mathcal{V}'$, whose value at y is $[x_0, y]$, is a scalar-valued function that happens to be linear (see § 14, (2)); in other words, $[x_0, y]$ defines a linear functional on $\mathcal{V}'$, and, consequently, an element of $\mathcal{V}''$.

By this method we have exhibited *some* linear functionals on $\mathcal{V}'$; have we exhibited them all? For the finite-dimensional case the following theorem furnishes the affirmative answer.

THEOREM. *If $\mathcal{V}$ is a finite-dimensional vector space, then corresponding to every linear functional z_0 on $\mathcal{V}'$ there is a vector x_0 in $\mathcal{V}$ such that $z_0(y) = [x_0, y] = y(x_0)$ for every y in $\mathcal{V}'$; the correspondence $z_0 \rightleftarrows x_0$ between $\mathcal{V}''$ and $\mathcal{V}$ is an isomorphism.*

The correspondence described in this statement is called the *natural correspondence* between $\mathcal{V}''$ and $\mathcal{V}$.

PROOF. Let us view the correspondence from the standpoint of going from $\mathcal{V}$ to $\mathcal{V}''$; in other words, to every x_0 in $\mathcal{V}$ we make correspond a vector z_0 in $\mathcal{V}''$ defined by $z_0(y) = y(x_0)$ for every y in $\mathcal{V}'$. Since $[x, y]$ depends linearly on x, the transformation $x_0 \to z_0$ is linear.

We shall show that this transformation is one-to-one, as far as it goes. We assert, in other words, that if x_1 and x_2 are in $\mathcal{V}$, and if z_1 and z_2 are the corresponding vectors in $\mathcal{V}''$ (so that $z_1(y) = [x_1, y]$ and $z_2(y) = [x_2, y]$ for all y in $\mathcal{V}'$), and if $z_1 = z_2$, then $x_1 = x_2$. To say that $z_1 = z_2$ means that $[x_1, y] = [x_2, y]$ for every y in $\mathcal{V}'$; the desired conclusion follows from § 15, Theorem 3.

The last two paragraphs together show that the set of those linear functionals z on $\mathcal{V}'$ (that is, elements of $\mathcal{V}''$) that do have the desired form (that is, $z(y)$ is identically equal to $[x, y]$ for a suitable x in $\mathcal{V}$) is a subspace of $\mathcal{V}''$ which is isomorphic to $\mathcal{V}$ and which is, therefore, n-dimensional. But the n-dimensionality of $\mathcal{V}$ implies that of $\mathcal{V}'$, which in turn implies that $\mathcal{V}''$ is n-dimensional. It follows that $\mathcal{V}''$ must coincide with the n-dimensional subspace just described, and the proof of the theorem is complete.

It is important to observe that the theorem shows not only that $\mathcal{V}$ and $\mathcal{V}''$ are isomorphic—this much is trivial from the fact that they have the same dimension—but that the natural correspondence is an isomorphism. This property of vector spaces is called *reflexivity;* every finite-dimensional vector space is reflexive.

It is frequently convenient to be mildly sloppy about $\mathcal{V}''$: for finite-dimensional vector spaces we shall identify $\mathcal{V}''$ with $\mathcal{V}$ (by the natural isomorphism), and we shall say that the element z_0 of $\mathcal{V}''$ is the *same* as the element x_0 of $\mathcal{V}$ whenever $z_0(y) = [x_0, y]$ for all y in $\mathcal{V}'$. In this language it is very easy to express the relation between a basis $\mathfrak{X}$, in $\mathcal{V}$, and the dual basis of its dual basis, in $\mathcal{V}''$; the symmetry of the relation $[x_i, y_j] = \delta_{ij}$ shows that $\mathfrak{X}'' = \mathfrak{X}$.