2

Linear Transformations and Matrices

In Chapter 1, we developed the theory of abstract vector spaces in considerable detail. It is now natural to consider those functions defined on vector spaces that in some sense "preserve" the structure. These special functions are called *linear transformations*, and they abound in both pure and applied mathematics. In calculus, the operations of differentiation and integration provide us with two of the most important examples of linear transformations (see Examples 6 and 7 of Section 2.1). These two examples allow us to reformulate many of the problems in differential and integral equations in terms of linear transformations on particular vector spaces (see Sections 2.7 and 5.2).

In geometry, rotations, reflections, and projections (see Examples 2, 3, and 4 of Section 2.1) provide us with another class of linear transformations. Later we use these transformations to study rigid motions in R^n (Section 6.10).

In the remaining chapters, we see further examples of linear transformations occurring in both the physical and the social sciences. Throughout this chapter, we assume that all vector spaces are over a common field F.

2.1 LINEAR TRANSFORMATIONS, NULL SPACES, AND RANGES

In this section, we consider a number of examples of linear transformations. Many of these transformations are studied in more detail in later sections. Recall that a function T with domain V and codomain W is denoted by

$\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$. (See Appendix B.)

Definition. *Let* V *and* W *be vector spaces (over* F*). We call a function* $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ *a* ***linear transformation from*** V ***to*** W *if, for all* $x, y \in \mathsf{V}$ *and* $c \in F$, *we have*

(a) $\mathsf{T}(x + y) = \mathsf{T}(x) + \mathsf{T}(y)$ *and*
(b) $\mathsf{T}(cx) = c\mathsf{T}(x)$.

If the underlying field F is the field of rational numbers, then (a) implies (b) (see Exercise 37), but, in general (a) and (b) are logically independent. See Exercises 38 and 39.

We often simply call T **linear**. The reader should verify the following properties of a function $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$. (See Exercise 7.)

1. If T is linear, then $\mathsf{T}(0) = 0$.
2. T is linear if and only if $\mathsf{T}(cx + y) = c\mathsf{T}(x) + \mathsf{T}(y)$ for all $x, y \in \mathsf{V}$ and $c \in F$.
3. If T is linear, then $\mathsf{T}(x - y) = \mathsf{T}(x) - \mathsf{T}(y)$ for all $x, y \in \mathsf{V}$.
4. T is linear if and only if, for $x_1, x_2, \ldots, x_n \in \mathsf{V}$ and $a_1, a_2, \ldots, a_n \in F$, we have

$$\mathsf{T}\left(\sum_{i=1}^{n} a_i x_i\right) = \sum_{i=1}^{n} a_i \mathsf{T}(x_i).$$

We generally use property 2 to prove that a given transformation is linear.

Example 1

Define

$$\mathsf{T}\colon \mathsf{R}^2 \to \mathsf{R}^2 \text{ by } \mathsf{T}(a_1, a_2) = (2a_1 + a_2, a_1).$$

To show that T is linear, let $c \in R$ and $x, y \in \mathsf{R}^2$, where $x = (b_1, b_2)$ and $y = (d_1, d_2)$. Since

$$cx + y = (cb_1 + d_1, cb_2 + d_2),$$

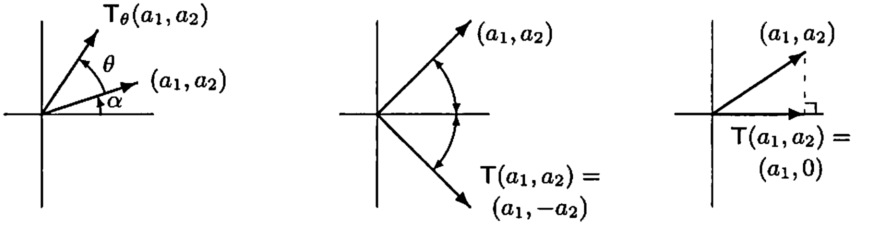
we have

$$\mathsf{T}(cx + y) = (2(cb_1 + d_1) + cb_2 + d_2, cb_1 + d_1).$$

Also

$$\begin{aligned} c\mathsf{T}(x) + \mathsf{T}(y) &= c(2b_1 + b_2, b_1) + (2d_1 + d_2, d_1) \\ &= (2cb_1 + cb_2 + 2d_1 + d_2, cb_1 + d_1) \\ &= (2(cb_1 + d_1) + cb_2 + d_2, cb_1 + d_1). \end{aligned}$$

So T is linear. ♦



(a) Rotation (b) Reflection (c) Projection

Figure 2.1

As we will see in Chapter 6, the applications of linear algebra to geometry are wide and varied. The main reason for this is that most of the important geometrical transformations are linear. Three particular transformations that we now consider are rotation, reflection, and projection. We leave the proofs of linearity to the reader.

Example 2

For any angle θ, define $\mathsf{T}_\theta\colon \mathsf{R}^2 \to \mathsf{R}^2$ by the rule: $\mathsf{T}_\theta(a_1, a_2)$ is the vector obtained by rotating (a_1, a_2) counterclockwise by θ if $(a_1, a_2) \neq (0, 0)$, and $\mathsf{T}_\theta(0, 0) = (0, 0)$. Then $\mathsf{T}_\theta\colon \mathsf{R}^2 \to \mathsf{R}^2$ is a linear transformation that is called the **rotation by** θ.

We determine an explicit formula for T_θ. Fix a nonzero vector $(a_1, a_2) \in \mathsf{R}^2$. Let α be the angle that (a_1, a_2) makes with the positive x-axis (see Figure 2.1(a)), and let $r = \sqrt{a_1^2 + a_2^2}$. Then $a_1 = r\cos\alpha$ and $a_2 = r\sin\alpha$. Also, $\mathsf{T}_\theta(a_1, a_2)$ has length r and makes an angle $\alpha + \theta$ with the positive x-axis. It follows that

$$\begin{aligned}
\mathsf{T}_\theta(a_1, a_2) &= (r\cos(\alpha + \theta), r\sin(\alpha + \theta)) \\
&= (r\cos\alpha\cos\theta - r\sin\alpha\sin\theta, r\cos\alpha\sin\theta + r\sin\alpha\cos\theta) \\
&= (a_1\cos\theta - a_2\sin\theta, a_1\sin\theta + a_2\cos\theta).
\end{aligned}$$

Finally, observe that this same formula is valid for $(a_1, a_2) = (0, 0)$.

It is now easy to show, as in Example 1, that T_θ is linear. ♦

Example 3

Define $\mathsf{T}\colon \mathsf{R}^2 \to \mathsf{R}^2$ by $\mathsf{T}(a_1, a_2) = (a_1, -a_2)$. T is called the **reflection about the** $\boldsymbol{x}$ **-axis.** (See Figure 2.1(b).) ♦

Example 4

Define $\mathsf{T}\colon \mathsf{R}^2 \to \mathsf{R}^2$ by $\mathsf{T}(a_1, a_2) = (a_1, 0)$. T is called the **projection on the** $\boldsymbol{x}$**-axis.** (See Figure 2.1(c).) ♦

We now look at some additional examples of linear transformations.

Example 5

Define $\mathsf{T}\colon \mathsf{M}_{m\times n}(F) \to \mathsf{M}_{n\times m}(F)$ by $\mathsf{T}(A) = A^t$, where A^t is the transpose of A, defined in Section 1.3. Then T is a linear transformation by Exercise 3 of Section 1.3. ♦

Example 6

Define $\mathsf{T}\colon \mathsf{P}_n(R) \to \mathsf{P}_{n-1}(R)$ by $\mathsf{T}(f(x)) = f'(x)$, where $f'(x)$ denotes the derivative of $f(x)$. To show that T is linear, let $g(x), h(x) \in \mathsf{P}_n(R)$ and $a \in R$. Now

$$\mathsf{T}(ag(x) + h(x)) = (ag(x) + h(x))' = ag'(x) + h'(x) = a\mathsf{T}(g(x)) + \mathsf{T}(h(x)).$$

So by property 2 above, T is linear. ♦

Example 7

Let $\mathsf{V} = \mathsf{C}(R)$, the vector space of continuous real-valued functions on R. Let $a, b \in R$, $a < b$. Define $\mathsf{T}\colon \mathsf{V} \to R$ by

$$\mathsf{T}(f) = \int_a^b f(t)\,dt$$

for all $f \in \mathsf{V}$. Then T is a linear transformation because the definite integral of a linear combination of functions is the same as the linear combination of the definite integrals of the functions. ♦

Two very important examples of linear transformations that appear frequently in the remainder of the book, and therefore deserve their own notation, are the identity and zero transformations.

For vector spaces V and W (over F), we define the **identity transformation** $\mathsf{I_V}\colon \mathsf{V} \to \mathsf{V}$ by $\mathsf{I_V}(x) = x$ for all $x \in \mathsf{V}$ and the **zero transformation** $\mathsf{T_0}\colon \mathsf{V} \to \mathsf{W}$ by $\mathsf{T_0}(x) = 0$ for all $x \in \mathsf{V}$. It is clear that both of these transformations are linear. We often write I instead of $\mathsf{I_V}$.

We now turn our attention to two very important sets associated with linear transformations: the *range* and *null space*. The determination of these sets allows us to examine more closely the intrinsic properties of a linear transformation.

Definitions. *Let* V *and* W *be vector spaces, and let* $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ *be linear. We define the* ***null space*** *(or* ***kernel****)* $\mathsf{N}(\mathsf{T})$ *of* T *to be the set of all vectors* x *in* V *such that* $\mathsf{T}(x) = 0$*; that is,* $\mathsf{N}(\mathsf{T}) = \{x \in \mathsf{V}\colon \mathsf{T}(x) = 0\}$.

We define the ***range*** *(or* ***image****)* $\mathsf{R}(\mathsf{T})$ *of* T *to be the subset of* W *consisting of all images (under* T*) of vectors in* V*; that is,* $\mathsf{R}(\mathsf{T}) = \{\mathsf{T}(x)\colon x \in \mathsf{V}\}$.

Example 8

Let V and W be vector spaces, and let $\mathsf{I}\colon \mathsf{V} \to \mathsf{V}$ and $\mathsf{T}_0\colon \mathsf{V} \to \mathsf{W}$ be the identity and zero transformations, respectively. Then $\mathsf{N}(\mathsf{I}) = \{0\}$, $\mathsf{R}(\mathsf{I}) = \mathsf{V}$, $\mathsf{N}(\mathsf{T}_0) = \mathsf{V}$, and $\mathsf{R}(\mathsf{T}_0) = \{0\}$. ♦

Example 9

Let $\mathsf{T}\colon \mathsf{R}^3 \to \mathsf{R}^2$ be the linear transformation defined by

$$\mathsf{T}(a_1, a_2, a_3) = (a_1 - a_2, 2a_3).$$

It is left as an exercise to verify that

$$\mathsf{N}(\mathsf{T}) = \{(a, a, 0)\colon a \in R\} \qquad \text{and} \qquad \mathsf{R}(\mathsf{T}) = \mathsf{R}^2. \quad ♦$$

In Examples 8 and 9, we see that the range and null space of each of the linear transformations is a subspace. The next result shows that this is true in general.

Theorem 2.1. *Let V and W be vector spaces and $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ be linear. Then $\mathsf{N}(\mathsf{T})$ and $\mathsf{R}(\mathsf{T})$ are subspaces of V and W, respectively.*

Proof. To clarify the notation, we use the symbols 0_V and 0_W to denote the zero vectors of V and W, respectively.

Since $\mathsf{T}(0_\mathsf{V}) = 0_\mathsf{W}$, we have that $0_\mathsf{V} \in \mathsf{N}(\mathsf{T})$. Let $x, y \in \mathsf{N}(\mathsf{T})$ and $c \in F$. Then $\mathsf{T}(x+y) = \mathsf{T}(x) + \mathsf{T}(y) = 0_\mathsf{W} + 0_\mathsf{W} = 0_\mathsf{W}$, and $\mathsf{T}(cx) = c\mathsf{T}(x) = c0_\mathsf{W} = 0_\mathsf{W}$. Hence $x + y \in \mathsf{N}(\mathsf{T})$ and $cx \in \mathsf{N}(\mathsf{T})$, so that $\mathsf{N}(\mathsf{T})$ is a subspace of V.

Because $\mathsf{T}(0_\mathsf{V}) = 0_\mathsf{W}$, we have that $0_\mathsf{W} \in \mathsf{R}(\mathsf{T})$. Now let $x, y \in \mathsf{R}(\mathsf{T})$ and $c \in F$. Then there exist v and w in V such that $\mathsf{T}(v) = x$ and $\mathsf{T}(w) = y$. So $\mathsf{T}(v+w) = \mathsf{T}(v) + \mathsf{T}(w) = x + y$, and $\mathsf{T}(cv) = c\mathsf{T}(v) = cx$. Thus $x + y \in \mathsf{R}(\mathsf{T})$ and $cx \in \mathsf{R}(\mathsf{T})$, so $\mathsf{R}(\mathsf{T})$ is a subspace of W. ∎

The next theorem provides a method for finding a spanning set for the range of a linear transformation. With this accomplished, a basis for the range is easy to discover using the technique of Example 6 of Section 1.6.

Theorem 2.2. *Let V and W be vector spaces, and let $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ be linear. If $\beta = \{v_1, v_2, \ldots, v_n\}$ is a basis for V, then*

$$\mathsf{R}(\mathsf{T}) = \operatorname{span}(\mathsf{T}(\beta)) = \operatorname{span}(\{\mathsf{T}(v_1), \mathsf{T}(v_2), \ldots, \mathsf{T}(v_n)\}).$$

Proof. Clearly $\mathsf{T}(v_i) \in \mathsf{R}(\mathsf{T})$ for each i. Because $\mathsf{R}(\mathsf{T})$ is a subspace, $\mathsf{R}(\mathsf{T})$ contains $\operatorname{span}(\{\mathsf{T}(v_1), \mathsf{T}(v_2), \ldots, \mathsf{T}(v_n)\}) = \operatorname{span}(\mathsf{T}(\beta))$ by Theorem 1.5 (p. 30).

Now suppose that $w \in \mathsf{R}(\mathsf{T})$. Then $w = \mathsf{T}(v)$ for some $v \in \mathsf{V}$. Because β is a basis for V, we have

$$v = \sum_{i=1}^{n} a_i v_i \quad \text{for some } a_1, a_2, \ldots, a_n \in F.$$

Since T is linear, it follows that

$$w = \mathsf{T}(v) = \sum_{i=1}^{n} a_i \mathsf{T}(v_i) \in \operatorname{span}(\mathsf{T}(\beta)).$$

So $\mathsf{R}(\mathsf{T})$ is contained in $\operatorname{span}(\mathsf{T}(\beta))$. ∎

It should be noted that Theorem 2.2 is true if β is infinite, that is, $\mathsf{R}(\mathsf{T}) = \operatorname{span}(\{\mathsf{T}(v) : v \in \beta\})$. (See Exercise 33.)

The next example illustrates the usefulness of Theorem 2.2.

Example 10

Define the linear transformation $\mathsf{T}: \mathsf{P}_2(R) \to \mathsf{M}_{2\times 2}(R)$ by

$$\mathsf{T}(f(x)) = \begin{pmatrix} f(1) - f(2) & 0 \\ 0 & f(0) \end{pmatrix}.$$

Since $\beta = \{1, x, x^2\}$ is a basis for $\mathsf{P}_2(R)$, we have

$$\begin{aligned} \mathsf{R}(\mathsf{T}) &= \operatorname{span}(\mathsf{T}(\beta)) = \operatorname{span}(\{\mathsf{T}(1), \mathsf{T}(x), \mathsf{T}(x^2)\}) \\ &= \operatorname{span}\left(\left\{\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} -3 & 0 \\ 0 & 0 \end{pmatrix}\right\}\right) \\ &= \operatorname{span}\left(\left\{\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}\right\}\right). \end{aligned}$$

Thus we have found a basis for $\mathsf{R}(\mathsf{T})$, and so $\dim(\mathsf{R}(\mathsf{T})) = 2$. ♦

As in Chapter 1, we measure the "size" of a subspace by its dimension. The null space and range are so important that we attach special names to their respective dimensions.

Definitions. *Let* V *and* W *be vector spaces, and let* $\mathsf{T}: \mathsf{V} \to \mathsf{W}$ *be linear. If* $\mathsf{N}(\mathsf{T})$ *and* $\mathsf{R}(\mathsf{T})$ *are finite-dimensional, then we define the* **nullity** *of* T*, denoted* $\operatorname{nullity}(\mathsf{T})$*, and the* **rank** *of* T*, denoted* $\operatorname{rank}(\mathsf{T})$*, to be the dimensions of* $\mathsf{N}(\mathsf{T})$ *and* $\mathsf{R}(\mathsf{T})$*, respectively.*

Reflecting on the action of a linear transformation, we see intuitively that the larger the nullity, the smaller the rank. In other words, the more vectors that are carried into $\mathit{0}$, the smaller the range. The same heuristic reasoning tells us that the larger the rank, the smaller the nullity. This balance between rank and nullity is made precise in the next theorem, appropriately called the *dimension theorem*.

Theorem 2.3 (Dimension Theorem). *Let* V *and* W *be vector spaces, and let* $\mathsf{T}: \mathsf{V} \to \mathsf{W}$ *be linear. If* V *is finite-dimensional, then*

$$\text{nullity}(\mathsf{T}) + \text{rank}(\mathsf{T}) = \dim(\mathsf{V}).$$

Proof. Suppose that $\dim(\mathsf{V}) = n$, $\dim(\mathsf{N}(\mathsf{T})) = k$, and $\{v_1, v_2, \ldots, v_k\}$ is a basis for $\mathsf{N}(\mathsf{T})$. By the corollary to Theorem 1.11 (p. 51), we may extend $\{v_1, v_2, \ldots, v_k\}$ to a basis $\beta = \{v_1, v_2, \ldots, v_n\}$ for V. We claim that $S = \{\mathsf{T}(v_{k+1}), \mathsf{T}(v_{k+2}), \ldots, \mathsf{T}(v_n)\}$ is a basis for $\mathsf{R}(\mathsf{T})$.

First we prove that S generates $\mathsf{R}(\mathsf{T})$. Using Theorem 2.2 and the fact that $\mathsf{T}(v_i) = \mathit{0}$ for $1 \le i \le k$, we have

$$\begin{aligned}\mathsf{R}(\mathsf{T}) &= \text{span}(\{\mathsf{T}(v_1), \mathsf{T}(v_2), \ldots, \mathsf{T}(v_n)\}\\ &= \text{span}(\{\mathsf{T}(v_{k+1}), \mathsf{T}(v_{k+2}), \ldots, \mathsf{T}(v_n)\} = \text{span}(S).\end{aligned}$$

Now we prove that S is linearly independent. Suppose that

$$\sum_{i=k+1}^{n} b_i \mathsf{T}(v_i) = \mathit{0} \quad \text{for } b_{k+1}, b_{k+2}, \ldots, b_n \in F.$$

Using the fact that T is linear, we have

$$\mathsf{T}\left(\sum_{i=k+1}^{n} b_i v_i\right) = \mathit{0}.$$

So

$$\sum_{i=k+1}^{n} b_i v_i \in \mathsf{N}(\mathsf{T}).$$

Hence there exist $c_1, c_2, \ldots, c_k \in F$ such that

$$\sum_{i=k+1}^{n} b_i v_i = \sum_{i=1}^{k} c_i v_i \quad \text{or} \quad \sum_{i=1}^{k}(-c_i)v_i + \sum_{i=k+1}^{n} b_i v_i = \mathit{0}.$$

Since β is a basis for V, we have $b_i = 0$ for all i. Hence S is linearly independent. Notice that this argument also shows that $\mathsf{T}(v_{k+1}), \mathsf{T}(v_{k+2}), \ldots, \mathsf{T}(v_n)$ are distinct; therefore $\text{rank}(\mathsf{T}) = n - k$. ∎

If we apply the dimension theorem to the linear transformation T in Example 9, we have that $\text{nullity}(\mathsf{T}) + 2 = 3$, so $\text{nullity}(\mathsf{T}) = 1$.

The reader should review the concepts of "one-to-one" and "onto" presented in Appendix B. Interestingly, for a linear transformation, both of these concepts are intimately connected to the rank and nullity of the transformation. This is demonstrated in the next two theorems.

Theorem 2.4. *Let* V *and* W *be vector spaces, and let* $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ *be linear. Then* T *is one-to-one if and only if* $\mathsf{N}(\mathsf{T}) = \{\mathit{0}\}$.

Proof. Suppose that T is one-to-one and $x \in \mathsf{N}(\mathsf{T})$. Then $\mathsf{T}(x) = \mathit{0} = \mathsf{T}(\mathit{0})$. Since T is one-to-one, we have $x = \mathit{0}$. Hence $\mathsf{N}(\mathsf{T}) = \{\mathit{0}\}$.

Now assume that $\mathsf{N}(\mathsf{T}) = \{\mathit{0}\}$, and suppose that $\mathsf{T}(x) = \mathsf{T}(y)$. Then $\mathit{0} = \mathsf{T}(x) - \mathsf{T}(y) = \mathsf{T}(x-y)$ by property 3 on page 65. Therefore $x - y \in \mathsf{N}(\mathsf{T}) = \{\mathit{0}\}$. So $x - y = \mathit{0}$, or $x = y$. This means that T is one-to-one. ∎

The reader should observe that Theorem 2.4 allows us to conclude that the transformation defined in Example 9 is not one-to-one.

Surprisingly, the conditions of one-to-one and onto are equivalent in an important special case.

Theorem 2.5. *Let* V *and* W *be vector spaces of equal (finite) dimension, and let* $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ *be linear. Then the following are equivalent.*

(a) T *is one-to-one.*
(b) T *is onto.*
(c) $\operatorname{rank}(\mathsf{T}) = \dim(\mathsf{V})$.

Proof. From the dimension theorem, we have

$$\operatorname{nullity}(\mathsf{T}) + \operatorname{rank}(\mathsf{T}) = \dim(\mathsf{V}).$$

Now, with the use of Theorem 2.4, we have that T is one-to-one if and only if $\mathsf{N}(\mathsf{T}) = \{\mathit{0}\}$, if and only if $\operatorname{nullity}(\mathsf{T}) = 0$, if and only if $\operatorname{rank}(\mathsf{T}) = \dim(\mathsf{V})$, if and only if $\operatorname{rank}(\mathsf{T}) = \dim(\mathsf{W})$, and if and only if $\dim(\mathsf{R}(\mathsf{T})) = \dim(\mathsf{W})$. By Theorem 1.11 (p. 50), this equality is equivalent to $\mathsf{R}(\mathsf{T}) = \mathsf{W}$, the definition of T being onto. ∎

We note that if V is not finite-dimensional and $\mathsf{T}\colon \mathsf{V} \to \mathsf{V}$ is linear, then it does *not* follow that one-to-one and onto are equivalent. (See Exercises 15, 16, and 21.)

The linearity of T in Theorems 2.4 and 2.5 is essential, for it is easy to construct examples of functions from R into R that are not one-to-one, but are onto, and vice versa.

The next two examples make use of the preceding theorems in determining whether a given linear transformation is one-to-one or onto.

Example 11

Let $\mathsf{T}\colon \mathsf{P}_2(R) \to \mathsf{P}_3(R)$ be the linear transformation defined by

$$\mathsf{T}(f(x)) = 2f'(x) + \int_0^x 3f(t)\,dt.$$

Now

$$\mathsf{R}(\mathsf{T}) = \operatorname{span}(\{\mathsf{T}(1), \mathsf{T}(x), \mathsf{T}(x^2)\}) = \operatorname{span}(\{3x, 2 + \frac{3}{2}x^2, 4x + x^3\}).$$

Since $\{3x, 2 + \frac{3}{2}x^2, 4x + x^3\}$ is linearly independent, $\operatorname{rank}(\mathsf{T}) = 3$. Since $\dim(\mathsf{P}_3(R)) = 4$, T is not onto. From the dimension theorem, $\operatorname{nullity}(\mathsf{T}) + 3 = 3$. So $\operatorname{nullity}(\mathsf{T}) = 0$, and therefore, $\mathsf{N}(\mathsf{T}) = \{\mathit{0}\}$. We conclude from Theorem 2.4 that T is one-to-one. ♦

Example 12

Let $\mathsf{T}\colon \mathsf{F}^2 \to \mathsf{F}^2$ be the linear transformation defined by

$$\mathsf{T}(a_1, a_2) = (a_1 + a_2, a_1).$$

It is easy to see that $\mathsf{N}(\mathsf{T}) = \{\mathit{0}\}$; so T is one-to-one. Hence Theorem 2.5 tells us that T must be onto. ♦

In Exercise 14, it is stated that if T is linear and one-to-one, then a subset S is linearly independent if and only if $\mathsf{T}(S)$ is linearly independent. Example 13 illustrates the use of this result.

Example 13

Let $\mathsf{T}\colon \mathsf{P}_2(R) \to \mathsf{R}^3$ be the linear transformation defined by

$$\mathsf{T}(a_0 + a_1x + a_2x^2) = (a_0, a_1, a_2).$$

Clearly T is linear and one-to-one. Let $S = \{2 - x + 3x^2, x + x^2, 1 - 2x^2\}$. Then S is linearly independent in $\mathsf{P}_2(R)$ because

$$\mathsf{T}(S) = \{(2, -1, 3), (0, 1, 1), (1, 0, -2)\}$$

is linearly independent in R^3. ♦

In Example 13, we transferred a property from the vector space of polynomials to a property in the vector space of 3-tuples. This technique is exploited more fully later.

One of the most important properties of a linear transformation is that it is completely determined by its action on a basis. This result, which follows from the next theorem and corollary, is used frequently throughout the book.

Theorem 2.6. *Let* V *and* W *be vector spaces over* F, *and suppose that* $\{v_1, v_2, \ldots, v_n\}$ *is a basis for* V. *For* $w_1, w_2, \ldots, w_n$ *in* W, *there exists exactly one linear transformation* $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ *such that* $\mathsf{T}(v_i) = w_i$ *for* $i = 1, 2, \ldots, n$.

Proof. Let $x \in \mathsf{V}$. Then

$$x = \sum_{i=1}^{n} a_i v_i,$$

where $a_1 a_2, \ldots, a_n$ are unique scalars. Define

$$\mathsf{T}\colon \mathsf{V} \to \mathsf{W} \quad \text{by} \quad \mathsf{T}(x) = \sum_{i=1}^{n} a_i w_i.$$

(a) T is linear: Suppose that $u, v \in \mathsf{V}$ and $d \in F$. Then we may write

$$u = \sum_{i=1}^{n} b_i v_i \quad \text{and} \quad v = \sum_{i=1}^{n} c_i v_i$$

for some scalars $b_1, b_2, \ldots, b_n, c_1, c_2, \ldots, c_n$. Thus

$$du + v = \sum_{i=1}^{n} (db_i + c_i) v_i.$$

So

$$\mathsf{T}(du + v) = \sum_{i=1}^{n} (db_i + c_i) w_i = d \sum_{i=1}^{n} b_i w_i + \sum_{i=1}^{n} c_i w_i = d\mathsf{T}(u) + \mathsf{T}(v).$$

(b) Clearly

$$\mathsf{T}(v_i) = w_i \quad \text{for } i = 1, 2, \ldots, n.$$

(c) T is unique: Suppose that $\mathsf{U}\colon \mathsf{V} \to \mathsf{W}$ is linear and $\mathsf{U}(v_i) = w_i$ for $i = 1, 2, \ldots, n$. Then for $x \in \mathsf{V}$ with

$$x = \sum_{i=1}^{n} a_i v_i,$$

we have

$$\mathsf{U}(x) = \sum_{i=1}^{n} a_i \mathsf{U}(v_i) = \sum_{i=1}^{n} a_i w_i = \mathsf{T}(x).$$

Hence $\mathsf{U} = \mathsf{T}$. ∎

Corollary. *Let* V *and* W *be vector spaces, and suppose that* V *has a finite basis* $\{v_1, v_2, \ldots, v_n\}$. *If* $\mathsf{U}, \mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ *are linear and* $\mathsf{U}(v_i) = \mathsf{T}(v_i)$ *for* $i = 1, 2, \ldots, n$, *then* $\mathsf{U} = \mathsf{T}$.

Example 14

Let $\mathsf{T}: \mathsf{R}^2 \to \mathsf{R}^2$ be the linear transformation defined by

$$\mathsf{T}(a_1, a_2) = (2a_2 - a_1, 3a_1),$$

and suppose that $\mathsf{U}: \mathsf{R}^2 \to \mathsf{R}^2$ is linear. If we know that $\mathsf{U}(1,2) = (3,3)$ and $\mathsf{U}(1,1) = (1,3)$, then $\mathsf{U} = \mathsf{T}$. This follows from the corollary and from the fact that $\{(1,2),(1,1)\}$ is a basis for R^2. ♦

EXERCISES

1. Label the following statements as true or false. In each part, V and W are finite-dimensional vector spaces (over F), and T is a function from V to W.

   (a) If T is linear, then T preserves sums and scalar products.
   (b) If $\mathsf{T}(x+y) = \mathsf{T}(x) + \mathsf{T}(y)$, then T is linear.
   (c) T is one-to-one if and only if the only vector x such that $\mathsf{T}(x) = 0$ is $x = 0$.
   (d) If T is linear, then $\mathsf{T}(0_\mathsf{V}) = 0_\mathsf{W}$.
   (e) If T is linear, then nullity(T) + rank(T) = dim(W).
   (f) If T is linear, then T carries linearly independent subsets of V onto linearly independent subsets of W.
   (g) If $\mathsf{T}, \mathsf{U}: \mathsf{V} \to \mathsf{W}$ are both linear and agree on a basis for V, then $\mathsf{T} = \mathsf{U}$.
   (h) Given $x_1, x_2 \in \mathsf{V}$ and $y_1, y_2 \in \mathsf{W}$, there exists a linear transformation $\mathsf{T}: \mathsf{V} \to \mathsf{W}$ such that $\mathsf{T}(x_1) = y_1$ and $\mathsf{T}(x_2) = y_2$.

For Exercises 2 through 6, prove that T is a linear transformation, and find bases for both N(T) and R(T). Then compute the nullity and rank of T, and verify the dimension theorem. Finally, use the appropriate theorems in this section to determine whether T is one-to-one or onto.

2. $\mathsf{T}: \mathsf{R}^3 \to \mathsf{R}^2$ defined by $\mathsf{T}(a_1, a_2, a_3) = (a_1 - a_2, 2a_3)$.

3. $\mathsf{T}: \mathsf{R}^2 \to \mathsf{R}^3$ defined by $\mathsf{T}(a_1, a_2) = (a_1 + a_2, 0, 2a_1 - a_2)$.

4. $\mathsf{T}: \mathsf{M}_{2\times 3}(F) \to \mathsf{M}_{2\times 2}(F)$ defined by

$$\mathsf{T}\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} = \begin{pmatrix} 2a_{11} - a_{12} & a_{13} + 2a_{12} \\ 0 & 0 \end{pmatrix}.$$

5. $\mathsf{T}: \mathsf{P}_2(R) \to \mathsf{P}_3(R)$ defined by $\mathsf{T}(f(x)) = xf(x) + f'(x)$.

6. $\mathsf{T}\colon \mathsf{M}_{n\times n}(F) \to F$ defined by $\mathsf{T}(A) = \operatorname{tr}(A)$. Recall (Example 4, Section 1.3) that

$$\operatorname{tr}(A) = \sum_{i=1}^{n} A_{ii}.$$

7. Prove properties 1, 2, 3, and 4 on page 65.

8. Prove that the transformations in Examples 2 and 3 are linear.

9. In this exercise, $\mathsf{T}\colon \mathsf{R}^2 \to \mathsf{R}^2$ is a function. For each of the following parts, state why T is *not* linear.
   (a) $\mathsf{T}(a_1, a_2) = (1, a_2)$
   (b) $\mathsf{T}(a_1, a_2) = (a_1, a_1^2)$
   (c) $\mathsf{T}(a_1, a_2) = (\sin a_1, 0)$
   (d) $\mathsf{T}(a_1, a_2) = (|a_1|, a_2)$
   (e) $\mathsf{T}(a_1, a_2) = (a_1 + 1, a_2)$

10. Suppose that $\mathsf{T}\colon \mathsf{R}^2 \to \mathsf{R}^2$ is linear, $\mathsf{T}(1,0) = (1,4)$, and $\mathsf{T}(1,1) = (2,5)$. What is $\mathsf{T}(2,3)$? Is T one-to-one?

11. Prove that there exists a linear transformation $\mathsf{T}\colon \mathsf{R}^2 \to \mathsf{R}^3$ such that $\mathsf{T}(1,1) = (1,0,2)$ and $\mathsf{T}(2,3) = (1,-1,4)$. What is $\mathsf{T}(8,11)$?

12. Is there a linear transformation $\mathsf{T}\colon \mathsf{R}^3 \to \mathsf{R}^2$ such that $\mathsf{T}(1,0,3) = (1,1)$ and $\mathsf{T}(-2,0,-6) = (2,1)$?

13. Let V and W be vector spaces, let $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ be linear, and let $\{w_1, w_2, \ldots, w_k\}$ be a linearly independent subset of $\mathsf{R}(\mathsf{T})$. Prove that if $S = \{v_1, v_2, \ldots, v_k\}$ is chosen so that $\mathsf{T}(v_i) = w_i$ for $i = 1, 2, \ldots, k$, then S is linearly independent.

14. Let V and W be vector spaces and $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ be linear.
   (a) Prove that T is one-to-one if and only if T carries linearly independent subsets of V onto linearly independent subsets of W.
   (b) Suppose that T is one-to-one and that S is a subset of V. Prove that S is linearly independent if and only if $\mathsf{T}(S)$ is linearly independent.
   (c) Suppose $\beta = \{v_1, v_2, \ldots, v_n\}$ is a basis for V and T is one-to-one and onto. Prove that $\mathsf{T}(\beta) = \{\mathsf{T}(v_1), \mathsf{T}(v_2), \ldots, \mathsf{T}(v_n)\}$ is a basis for W.

15. Recall the definition of $\mathsf{P}(R)$ on page 10. Define

$$\mathsf{T}\colon \mathsf{P}(R) \to \mathsf{P}(R) \quad \text{by} \quad \mathsf{T}(f(x)) = \int_0^x f(t)\,dt.$$

Prove that T linear and one-to-one, but not onto.

16. Let $\mathsf{T}\colon \mathsf{P}(R) \to \mathsf{P}(R)$ be defined by $\mathsf{T}(f(x)) = f'(x)$. Recall that T is linear. Prove that T is onto, but not one-to-one.

17. Let V and W be finite-dimensional vector spaces and $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ be linear.

(a) Prove that if $\dim(\mathsf{V}) < \dim(\mathsf{W})$, then T cannot be onto.
(b) Prove that if $\dim(\mathsf{V}) > \dim(\mathsf{W})$, then T cannot be one-to-one.

18. Give an example of a linear transformation $\mathsf{T}\colon \mathsf{R}^2 \to \mathsf{R}^2$ such that $\mathsf{N}(\mathsf{T}) = \mathsf{R}(\mathsf{T})$.

19. Give an example of distinct linear transformations T and U such that $\mathsf{N}(\mathsf{T}) = \mathsf{N}(\mathsf{U})$ and $\mathsf{R}(\mathsf{T}) = \mathsf{R}(\mathsf{U})$.

20. Let V and W be vector spaces with subspaces V_1 and W_1, respectively. If $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ is linear, prove that $\mathsf{T}(\mathsf{V}_1)$ is a subspace of W and that $\{x \in \mathsf{V}\colon \mathsf{T}(x) \in \mathsf{W}_1\}$ is a subspace of V.

21. Let V be the vector space of sequences described in Example 5 of Section 1.2. Define the functions $\mathsf{T}, \mathsf{U}\colon \mathsf{V} \to \mathsf{V}$ by

$$\mathsf{T}(a_1, a_2, \ldots) = (a_2, a_3, \ldots) \quad \text{and} \quad \mathsf{U}(a_1, a_2, \ldots) = (0, a_1, a_2, \ldots).$$

T and U are called the **left shift** and **right shift** operators on V, respectively.

(a) Prove that T and U are linear.
(b) Prove that T is onto, but not one-to-one.
(c) Prove that U is one-to-one, but not onto.

22. Let $\mathsf{T}\colon \mathsf{R}^3 \to R$ be linear. Show that there exist scalars a, b, and c such that $\mathsf{T}(x, y, z) = ax + by + cz$ for all $(x, y, z) \in \mathsf{R}^3$. Can you generalize this result for $\mathsf{T}\colon \mathsf{F}^n \to F$? State and prove an analogous result for $\mathsf{T}\colon \mathsf{F}^n \to \mathsf{F}^m$.

23. Let $\mathsf{T}\colon \mathsf{R}^3 \to R$ be linear. Describe geometrically the possibilities for the null space of T. *Hint:* Use Exercise 22.

The following definition is used in Exercises 24–27 and in Exercise 30.

Definition. *Let V be a vector space and W_1 and W_2 be subspaces of V such that $\mathsf{V} = \mathsf{W}_1 \oplus \mathsf{W}_2$. (Recall the definition of direct sum given in the exercises of Section 1.3.) A function $\mathsf{T}\colon \mathsf{V} \to \mathsf{V}$ is called the* ***projection on*** *W_1* ***along*** *W_2 if, for $x = x_1 + x_2$ with $x_1 \in \mathsf{W}_1$ and $x_2 \in \mathsf{W}_2$, we have $\mathsf{T}(x) = x_1$.*

24. Let $\mathsf{T}\colon \mathsf{R}^2 \to \mathsf{R}^2$. Include figures for each of the following parts.

(a) Find a formula for $\mathsf{T}(a, b)$, where T represents the projection on the y-axis along the x-axis.
(b) Find a formula for $\mathsf{T}(a, b)$, where T represents the projection on the y-axis along the line $L = \{(s, s) \colon s \in R\}$.

25. Let $\mathsf{T} \colon \mathsf{R}^3 \to \mathsf{R}^3$.

(a) If $\mathsf{T}(a, b, c) = (a, b, 0)$, show that T is the projection on the xy-plane along the z-axis.
(b) Find a formula for $\mathsf{T}(a, b, c)$, where T represents the projection on the z-axis along the xy-plane.
(c) If $\mathsf{T}(a, b, c) = (a - c, b, 0)$, show that T is the projection on the xy-plane along the line $L = \{(a, 0, a) \colon a \in R\}$.

26. Using the notation in the definition above, assume that $\mathsf{T} \colon \mathsf{V} \to \mathsf{V}$ is the projection on W_1 along W_2.

(a) Prove that T is linear and $\mathsf{W}_1 = \{x \in \mathsf{V} \colon \mathsf{T}(x) = x\}$.
(b) Prove that $\mathsf{W}_1 = \mathsf{R}(\mathsf{T})$ and $\mathsf{W}_2 = \mathsf{N}(\mathsf{T})$.
(c) Describe T if $\mathsf{W}_1 = \mathsf{V}$.
(d) Describe T if W_1 is the zero subspace.

27. Suppose that W is a subspace of a finite-dimensional vector space V.

(a) Prove that there exists a subspace W' and a function $\mathsf{T} \colon \mathsf{V} \to \mathsf{V}$ such that T is a projection on W along W'.
(b) Give an example of a subspace W of a vector space V such that there are *two* projections on W along two (distinct) subspaces.

The following definitions are used in Exercises 28-32.

Definitions. *Let* V *be a vector space, and let* $\mathsf{T} \colon \mathsf{V} \to \mathsf{V}$ *be linear. A subspace* W *of* V *is said to be* **T-invariant** *if* $\mathsf{T}(x) \in \mathsf{W}$ *for every* $x \in \mathsf{W}$, *that is,* $\mathsf{T}(\mathsf{W}) \subseteq \mathsf{W}$. *If* W *is* T*-invariant, we define the* **restriction of** T **on** W *to be the function* $\mathsf{T}_\mathsf{W} \colon \mathsf{W} \to \mathsf{W}$ *defined by* $\mathsf{T}_\mathsf{W}(x) = \mathsf{T}(x)$ *for all* $x \in \mathsf{W}$.

Exercises 28-32 assume that W is a subspace of a vector space V and that $\mathsf{T} \colon \mathsf{V} \to \mathsf{V}$ is linear. *Warning:* Do not assume that W is T-invariant or that T is a projection unless explicitly stated.

28. Prove that the subspaces $\{\mathit{0}\}$, V, $\mathsf{R}(\mathsf{T})$, and $\mathsf{N}(\mathsf{T})$ are all T-invariant.

29. If W is T-invariant, prove that T_W is linear.

30. Suppose that T is the projection on W along some subspace W'. Prove that W is T-invariant and that $\mathsf{T}_\mathsf{W} = \mathsf{I}_\mathsf{W}$.

31. Suppose that $\mathsf{V} = \mathsf{R}(\mathsf{T}) \oplus \mathsf{W}$ and W is T-invariant. (Recall the definition of *direct sum* given in the exercises of Section 1.3.)

(a) Prove that $\mathsf{W} \subseteq \mathsf{N}(\mathsf{T})$.
(b) Show that if V is finite-dimensional, then $\mathsf{W} = \mathsf{N}(\mathsf{T})$.
(c) Show by example that the conclusion of (b) is not necessarily true if V is not finite-dimensional.

32. Suppose that W is T-invariant. Prove that $\mathsf{N}(\mathsf{T}_\mathsf{W}) = \mathsf{N}(\mathsf{T}) \cap \mathsf{W}$ and $\mathsf{R}(\mathsf{T}_\mathsf{W}) = \mathsf{T}(\mathsf{W})$.

33. Prove Theorem 2.2 for the case that β is infinite, that is, $\mathsf{R}(\mathsf{T}) = \mathrm{span}(\{\mathsf{T}(v) : v \in \beta\})$.

34. Prove the following generalization of Theorem 2.6: Let V and W be vector spaces over a common field, and let β be a basis for V. Then for any function $f : \beta \to \mathsf{W}$ there exists exactly one linear transformation $\mathsf{T} : \mathsf{V} \to \mathsf{W}$ such that $\mathsf{T}(x) = f(x)$ for all $x \in \beta$.

Exercises 35 and 36 assume the definition of *direct sum* given in the exercises of Section 1.3.

35. Let V be a finite-dimensional vector space and $\mathsf{T} : \mathsf{V} \to \mathsf{V}$ be linear.

(a) Suppose that $\mathsf{V} = \mathsf{R}(\mathsf{T}) + \mathsf{N}(\mathsf{T})$. Prove that $\mathsf{V} = \mathsf{R}(\mathsf{T}) \oplus \mathsf{N}(\mathsf{T})$.
(b) Suppose that $\mathsf{R}(\mathsf{T}) \cap \mathsf{N}(\mathsf{T}) = \{0\}$. Prove that $\mathsf{V} = \mathsf{R}(\mathsf{T}) \oplus \mathsf{N}(\mathsf{T})$.

Be careful to say in each part where finite-dimensionality is used.

36. Let V and T be as defined in Exercise 21.

(a) Prove that $\mathsf{V} = \mathsf{R}(\mathsf{T}) + \mathsf{N}(\mathsf{T})$, but V is not a direct sum of these two spaces. Thus the result of Exercise 35(a) above cannot be proved without assuming that V is finite-dimensional.
(b) Find a linear operator T_1 on V such that $\mathsf{R}(\mathsf{T}_1) \cap \mathsf{N}(\mathsf{T}_1) = \{0\}$ but V is not a direct sum of $\mathsf{R}(\mathsf{T}_1)$ and $\mathsf{N}(\mathsf{T}_1)$. Conclude that V being finite-dimensional is also essential in Exercise 35(b).

37. A function $\mathsf{T} : \mathsf{V} \to \mathsf{W}$ between vector spaces V and W is called **additive** if $\mathsf{T}(x + y) = \mathsf{T}(x) + \mathsf{T}(y)$ for all $x, y \in \mathsf{V}$. Prove that if V and W are vector spaces over the field of rational numbers, then any additive function from V into W is a linear transformation.

38. Let $\mathsf{T} : C \to C$ be the function defined by $\mathsf{T}(z) = \overline{z}$. Prove that T is additive (as defined in Exercise 37) but not linear.

39. Prove that there is an additive function $\mathsf{T} : R \to R$ (as defined in Exercise 37) that is not linear. *Hint:* Let V be the set of real numbers regarded as a vector space over the field of rational numbers. By the corollary to Theorem 1.13 (p. 60), V has a basis β. Let x and y be two distinct vectors in β, and define $f : \beta \to \mathsf{V}$ by $f(x) = y$, $f(y) = x$, and $f(z) = z$ otherwise. By Exercise 34, there exists a linear transformation

$\mathsf{T}: \mathsf{V} \to \mathsf{V}$ such that $\mathsf{T}(u) = f(u)$ for all $u \in \beta$. Then T is additive, but for $c = y/x$, $\mathsf{T}(cx) \neq c\mathsf{T}(x)$.

The following exercise requires familiarity with the definition of *quotient space* given in Exercise 31 of Section 1.3.

40. Let V be a vector space and W be a subspace of V. Define the mapping $\eta: \mathsf{V} \to \mathsf{V}/\mathsf{W}$ by $\eta(v) = v + \mathsf{W}$ for $v \in \mathsf{V}$.

- **(a)** Prove that η is a linear transformation from V onto V/W and that $\mathsf{N}(\eta) = \mathsf{W}$.
- **(b)** Suppose that V is finite-dimensional. Use (a) and the dimension theorem to derive a formula relating $\dim(\mathsf{V})$, $\dim(\mathsf{W})$, and $\dim(\mathsf{V}/\mathsf{W})$.
- **(c)** Read the proof of the dimension theorem. Compare the method of solving (b) with the method of deriving the same result as outlined in Exercise 35 of Section 1.6.

2.2 THE MATRIX REPRESENTATION OF A LINEAR TRANSFORMATION

Until now, we have studied linear transformations by examining their ranges and null spaces. In this section, we embark on one of the most useful approaches to the analysis of a linear transformation on a finite-dimensional vector space: the representation of a linear transformation by a matrix. In fact, we develop a one-to-one correspondence between matrices and linear transformations that allows us to utilize properties of one to study properties of the other.

We first need the concept of an *ordered basis* for a vector space.

Definition. *Let* V *be a finite-dimensional vector space. An* **ordered basis** *for* V *is a basis for* V *endowed with a specific order; that is, an ordered basis for* V *is a finite sequence of linearly independent vectors in* V *that generates* V.

Example 1

In F^3, $\beta = \{e_1, e_2, e_3\}$ can be considered an ordered basis. Also $\gamma = \{e_2, e_1, e_3\}$ is an ordered basis, but $\beta \neq \gamma$ as ordered bases. ♦

For the vector space F^n, we call $\{e_1, e_2, \ldots, e_n\}$ the **standard ordered basis** for F^n. Similarly, for the vector space $\mathsf{P}_n(F)$, we call $\{1, x, \ldots, x^n\}$ the **standard ordered basis** for $\mathsf{P}_n(F)$.

Now that we have the concept of ordered basis, we can identify abstract vectors in an n-dimensional vector space with n-tuples. This identification is provided through the use of *coordinate vectors*, as introduced next.

Definition. *Let $\beta = \{u_1, u_2, \ldots, u_n\}$ be an ordered basis for a finite-dimensional vector space* V. *For $x \in \mathsf{V}$, let $a_1, a_2, \ldots, a_n$ be the unique scalars such that*

$$x = \sum_{i=1}^{n} a_i u_i.$$

We define the ***coordinate vector of x relative to β****, denoted $[x]_\beta$, by*

$$[x]_\beta = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}.$$

Notice that $[u_i]_\beta = e_i$ in the preceding definition. It is left as an exercise to show that the correspondence $x \to [x]_\beta$ provides us with a linear transformation from V to F^n. We study this transformation in Section 2.4 in more detail.

Example 2

Let $\mathsf{V} = \mathsf{P}_2(R)$, and let $\beta = \{1, x, x^2\}$ be the standard ordered basis for V. If $f(x) = 4 + 6x - 7x^2$, then

$$[f]_\beta = \begin{pmatrix} 4 \\ 6 \\ -7 \end{pmatrix}. \qquad \blacklozenge$$

Let us now proceed with the promised matrix representation of a linear transformation. Suppose that V and W are finite-dimensional vector spaces with ordered bases $\beta = \{v_1, v_2, \ldots, v_n\}$ and $\gamma = \{w_1, w_2, \ldots, w_m\}$, respectively. Let $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ be linear. Then for each j, $1 \le j \le n$, there exist unique scalars $a_{ij} \in F$, $1 \le i \le m$, such that

$$\mathsf{T}(v_j) = \sum_{i=1}^{m} a_{ij} w_i \quad \text{for } 1 \le j \le n.$$

Definition. *Using the notation above, we call the $m \times n$ matrix A defined by $A_{ij} = a_{ij}$ the* ***matrix representation of* T *in the ordered bases* β *and*** *γ and write $A = [\mathsf{T}]_\beta^\gamma$. If $\mathsf{V} = \mathsf{W}$ and $\beta = \gamma$, then we write $A = [\mathsf{T}]_\beta$.*

Notice that the jth column of A is simply $[\mathsf{T}(v_j)]_\gamma$. Also observe that if $\mathsf{U}\colon \mathsf{V} \to \mathsf{W}$ is a linear transformation such that $[\mathsf{U}]_\beta^\gamma = [\mathsf{T}]_\beta^\gamma$, then $\mathsf{U} = \mathsf{T}$ by the corollary to Theorem 2.6 (p. 73).

We illustrate the computation of $[\mathsf{T}]_\beta^\gamma$ in the next several examples.

Example 3

Let $\mathsf{T}\colon \mathsf{R}^2 \to \mathsf{R}^3$ be the linear transformation defined by

$$\mathsf{T}(a_1, a_2) = (a_1 + 3a_2, 0, 2a_1 - 4a_2).$$

Let β and γ be the standard ordered bases for R^2 and R^3, respectively. Now

$$\mathsf{T}(1,0) = (1,0,2) = 1e_1 + 0e_2 + 2e_3$$

and

$$\mathsf{T}(0,1) = (3,0,-4) = 3e_1 + 0e_2 - 4e_3.$$

Hence

$$[\mathsf{T}]_\beta^\gamma = \begin{pmatrix} 1 & 3 \\ 0 & 0 \\ 2 & -4 \end{pmatrix}.$$

If we let $\gamma' = \{e_3, e_2, e_1\}$, then

$$[\mathsf{T}]_\beta^{\gamma'} = \begin{pmatrix} 2 & -4 \\ 0 & 0 \\ 1 & 3 \end{pmatrix}. \quad \blacklozenge$$

Example 4

Let $\mathsf{T}\colon \mathsf{P}_3(R) \to \mathsf{P}_2(R)$ be the linear transformation defined by $\mathsf{T}(f(x)) = f'(x)$. Let β and γ be the standard ordered bases for $\mathsf{P}_3(R)$ and $\mathsf{P}_2(R)$, respectively. Then

$$\begin{aligned} \mathsf{T}(1) &= 0\cdot 1 + 0\cdot x + 0\cdot x^2 \\ \mathsf{T}(x) &= 1\cdot 1 + 0\cdot x + 0\cdot x^2 \\ \mathsf{T}(x^2) &= 0\cdot 1 + 2\cdot x + 0\cdot x^2 \\ \mathsf{T}(x^3) &= 0\cdot 1 + 0\cdot x + 3\cdot x^2. \end{aligned}$$

So

$$[\mathsf{T}]_\beta^\gamma = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Note that when $\mathsf{T}(x^j)$ is written as a linear combination of the vectors of γ, its coefficients give the entries of column $j+1$ of $[\mathsf{T}]_\beta^\gamma$. $\blacklozenge$

Now that we have defined a procedure for associating matrices with linear transformations, we show in Theorem 2.8 that this association "preserves" addition and scalar multiplication. To make this more explicit, we need some preliminary discussion about the addition and scalar multiplication of linear transformations.

Definition. *Let* $\mathsf{T}, \mathsf{U}\colon \mathsf{V} \to \mathsf{W}$ *be arbitrary functions, where* V *and* W *are vector spaces over* F, *and let* $a \in F$. *We define* $\mathsf{T}+\mathsf{U}\colon \mathsf{V} \to \mathsf{W}$ *by* $(\mathsf{T}+\mathsf{U})(x) = \mathsf{T}(x)+\mathsf{U}(x)$ *for all* $x \in \mathsf{V}$, *and* $a\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ *by* $(a\mathsf{T})(x) = a\mathsf{T}(x)$ *for all* $x \in \mathsf{V}$.

Of course, these are just the usual definitions of addition and scalar multiplication of functions. We are fortunate, however, to have the result that both sums and scalar multiples of linear transformations are also linear.

Theorem 2.7. *Let* V *and* W *be vector spaces over a field* F, *and let* $\mathsf{T}, \mathsf{U}\colon \mathsf{V} \to \mathsf{W}$ *be linear.*

(a) *For all* $a \in F$, $a\mathsf{T}+\mathsf{U}$ *is linear.*

(b) *Using the operations of addition and scalar multiplication in the preceding definition, the collection of all linear transformations from* V *to* W *is a vector space over* F.

Proof. (a) Let $x, y \in \mathsf{V}$ and $c \in F$. Then

$$\begin{aligned}(a\mathsf{T}+\mathsf{U})(cx+y) &= a\mathsf{T}(cx+y)+\mathsf{U}(cx+y)\\ &= a[\mathsf{T}(cx+y)]+c\mathsf{U}(x)+\mathsf{U}(y)\\ &= a[c\mathsf{T}(x)+\mathsf{T}(y)]+c\mathsf{U}(x)+\mathsf{U}(y)\\ &= ac\mathsf{T}(x)+c\mathsf{U}(x)+a\mathsf{T}(y)+\mathsf{U}(y)\\ &= c(a\mathsf{T}+\mathsf{U})(x)+(a\mathsf{T}+\mathsf{U})(y).\end{aligned}$$

So $a\mathsf{T}+\mathsf{U}$ is linear.

(b) Noting that T_0, the zero transformation, plays the role of the zero vector, it is easy to verify that the axioms of a vector space are satisfied, and hence that the collection of all linear transformations from V into W is a vector space over F. ∎

Definitions. *Let* V *and* W *be vector spaces over* F. *We denote the vector space of all linear transformations from* V *into* W *by* $\mathcal{L}(\mathsf{V},\mathsf{W})$. *In the case that* $\mathsf{V}=\mathsf{W}$, *we write* $\mathcal{L}(\mathsf{V})$ *instead of* $\mathcal{L}(\mathsf{V},\mathsf{W})$.

In Section 2.4, we see a complete identification of $\mathcal{L}(\mathsf{V},\mathsf{W})$ with the vector space $\mathsf{M}_{m\times n}(F)$, where n and m are the dimensions of V and W, respectively. This identification is easily established by the use of the next theorem.

Theorem 2.8. *Let* V *and* W *be finite-dimensional vector spaces with ordered bases* β *and* γ, *respectively, and let* $\mathsf{T}, \mathsf{U}\colon \mathsf{V} \to \mathsf{W}$ *be linear transformations. Then*

(a) $[\mathsf{T}+\mathsf{U}]_\beta^\gamma = [\mathsf{T}]_\beta^\gamma + [\mathsf{U}]_\beta^\gamma$ *and*

(b) $[a\mathsf{T}]_\beta^\gamma = a[\mathsf{T}]_\beta^\gamma$ *for all scalars a.*

Proof. Let $\beta = \{v_1, v_2, \ldots, v_n\}$ and $\gamma = \{w_1, w_2, \ldots, w_m\}$. There exist unique scalars a_{ij} and b_{ij} $(1 \le i \le m, 1 \le j \le n)$ such that

$$\mathsf{T}(v_j) = \sum_{i=1}^{m} a_{ij} w_i \quad \text{and} \quad \mathsf{U}(v_j) = \sum_{i=1}^{m} b_{ij} w_i \quad \text{for } 1 \le j \le n.$$

Hence

$$(\mathsf{T}+\mathsf{U})(v_j) = \sum_{i=1}^{m} (a_{ij} + b_{ij}) w_i.$$

Thus

$$([\mathsf{T}+\mathsf{U}]_\beta^\gamma)_{ij} = a_{ij} + b_{ij} = ([\mathsf{T}]_\beta^\gamma + [\mathsf{U}]_\beta^\gamma)_{ij}.$$

So (a) is proved, and the proof of (b) is similar. ∎

Example 5

Let $\mathsf{T}\colon \mathsf{R}^2 \to \mathsf{R}^3$ and $\mathsf{U}\colon \mathsf{R}^2 \to \mathsf{R}^3$ be the linear transformations respectively defined by

$\mathsf{T}(a_1, a_2) = (a_1 + 3a_2, 0, 2a_1 - 4a_2)$ and $\mathsf{U}(a_1, a_2) = (a_1 - a_2, 2a_1, 3a_1 + 2a_2)$.

Let β and γ be the standard ordered bases of R^2 and R^3, respectively. Then

$$[\mathsf{T}]_\beta^\gamma = \begin{pmatrix} 1 & 3 \\ 0 & 0 \\ 2 & -4 \end{pmatrix},$$

(as computed in Example 3), and

$$[\mathsf{U}]_\beta^\gamma = \begin{pmatrix} 1 & -1 \\ 2 & 0 \\ 3 & 2 \end{pmatrix}.$$

If we compute $\mathsf{T}+\mathsf{U}$ using the preceding definitions, we obtain

$$(\mathsf{T}+\mathsf{U})(a_1, a_2) = (2a_1 + 2a_2, 2a_1, 5a_1 - 2a_2).$$

So

$$[\mathsf{T}+\mathsf{U}]_\beta^\gamma = \begin{pmatrix} 2 & 2 \\ 2 & 0 \\ 5 & -2 \end{pmatrix},$$

which is simply $[\mathsf{T}]_\beta^\gamma + [\mathsf{U}]_\beta^\gamma$, illustrating Theorem 2.8. ♦

EXERCISES

1. Label the following statements as true or false. Assume that V and W are finite-dimensional vector spaces with ordered bases β and γ, respectively, and $\mathsf{T}, \mathsf{U}\colon \mathsf{V} \to \mathsf{W}$ are linear transformations.
   (a) For any scalar a, $a\mathsf{T} + \mathsf{U}$ is a linear transformation from V to W.
   (b) $[\mathsf{T}]_\beta^\gamma = [\mathsf{U}]_\beta^\gamma$ implies that $\mathsf{T} = \mathsf{U}$.
   (c) If $m = \dim(\mathsf{V})$ and $n = \dim(\mathsf{W})$, then $[\mathsf{T}]_\beta^\gamma$ is an $m \times n$ matrix.
   (d) $[\mathsf{T} + \mathsf{U}]_\beta^\gamma = [\mathsf{T}]_\beta^\gamma + [\mathsf{U}]_\beta^\gamma$.
   (e) $\mathcal{L}(\mathsf{V}, \mathsf{W})$ is a vector space.
   (f) $\mathcal{L}(\mathsf{V}, \mathsf{W}) = \mathcal{L}(\mathsf{W}, \mathsf{V})$.

2. Let β and γ be the standard ordered bases for R^n and R^m, respectively. For each linear transformation $\mathsf{T}\colon \mathsf{R}^n \to \mathsf{R}^m$, compute $[\mathsf{T}]_\beta^\gamma$.
   (a) $\mathsf{T}\colon \mathsf{R}^2 \to \mathsf{R}^3$ defined by $\mathsf{T}(a_1, a_2) = (2a_1 - a_2, 3a_1 + 4a_2, a_1)$.
   (b) $\mathsf{T}\colon \mathsf{R}^3 \to \mathsf{R}^2$ defined by $\mathsf{T}(a_1, a_2, a_3) = (2a_1 + 3a_2 - a_3, a_1 + a_3)$.
   (c) $\mathsf{T}\colon \mathsf{R}^3 \to R$ defined by $\mathsf{T}(a_1, a_2, a_3) = 2a_1 + a_2 - 3a_3$.
   (d) $\mathsf{T}\colon \mathsf{R}^3 \to \mathsf{R}^3$ defined by
   $$\mathsf{T}(a_1, a_2, a_3) = (2a_2 + a_3, -a_1 + 4a_2 + 5a_3, a_1 + a_3).$$
   (e) $\mathsf{T}\colon \mathsf{R}^n \to \mathsf{R}^n$ defined by $\mathsf{T}(a_1, a_2, \ldots, a_n) = (a_1, a_1, \ldots, a_1)$.
   (f) $\mathsf{T}\colon \mathsf{R}^n \to \mathsf{R}^n$ defined by $\mathsf{T}(a_1, a_2, \ldots, a_n) = (a_n, a_{n-1}, \ldots, a_1)$.
   (g) $\mathsf{T}\colon \mathsf{R}^n \to R$ defined by $\mathsf{T}(a_1, a_2, \ldots, a_n) = a_1 + a_n$.

3. Let $\mathsf{T}\colon \mathsf{R}^2 \to \mathsf{R}^3$ be defined by $\mathsf{T}(a_1, a_2) = (a_1 - a_2, a_1, 2a_1 + a_2)$. Let β be the standard ordered basis for R^2 and $\gamma = \{(1,1,0), (0,1,1), (2,2,3)\}$. Compute $[\mathsf{T}]_\beta^\gamma$. If $\alpha = \{(1,2), (2,3)\}$, compute $[\mathsf{T}]_\alpha^\gamma$.

4. Define
   $$\mathsf{T}\colon \mathsf{M}_{2\times 2}(R) \to \mathsf{P}_2(R) \quad \text{by} \quad \mathsf{T}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = (a + b) + (2d)x + bx^2.$$
   Let
   $$\beta = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\} \quad \text{and} \quad \gamma = \{1, x, x^2\}.$$
   Compute $[\mathsf{T}]_\beta^\gamma$.

5. Let
   $$\alpha = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\},$$
   $$\beta = \{1, x, x^2\},$$
   and
   $$\gamma = \{1\}.$$

(a) Define $\mathsf{T}\colon \mathsf{M}_{2\times 2}(F) \to \mathsf{M}_{2\times 2}(F)$ by $\mathsf{T}(A) = A^t$. Compute $[\mathsf{T}]_\alpha$.
(b) Define

$$\mathsf{T}\colon \mathsf{P}_2(R) \to \mathsf{M}_{2\times 2}(R) \quad \text{by} \quad \mathsf{T}(f(x)) = \begin{pmatrix} f'(0) & 2f(1) \\ 0 & f''(3) \end{pmatrix},$$

where $'$ denotes differentiation. Compute $[\mathsf{T}]_\beta^\alpha$.
(c) Define $\mathsf{T}\colon \mathsf{M}_{2\times 2}(F) \to F$ by $\mathsf{T}(A) = \operatorname{tr}(A)$. Compute $[\mathsf{T}]_\alpha^\gamma$.
(d) Define $\mathsf{T}\colon \mathsf{P}_2(R) \to R$ by $\mathsf{T}(f(x)) = f(2)$. Compute $[\mathsf{T}]_\beta^\gamma$.
(e) If

$$A = \begin{pmatrix} 1 & -2 \\ 0 & 4 \end{pmatrix},$$

compute $[A]_\alpha$.
(f) If $f(x) = 3 - 6x + x^2$, compute $[f(x)]_\beta$.
(g) For $a \in F$, compute $[a]_\gamma$.

6. Complete the proof of part (b) of Theorem 2.7.

7. Prove part (b) of Theorem 2.8.

8.† Let V be an n-dimensional vector space with an ordered basis β. Define $\mathsf{T}\colon \mathsf{V} \to \mathsf{F}^n$ by $\mathsf{T}(x) = [x]_\beta$. Prove that T is linear.

9. Let V be the vector space of complex numbers over the field R. Define $\mathsf{T}\colon \mathsf{V} \to \mathsf{V}$ by $\mathsf{T}(z) = \overline{z}$, where $\overline{z}$ is the complex conjugate of z. Prove that T is linear, and compute $[\mathsf{T}]_\beta$, where $\beta = \{1, i\}$. (Recall by Exercise 38 of Section 2.1 that T is not linear if V is regarded as a vector space over the field C.)

10. Let V be a vector space with the ordered basis $\beta = \{v_1, v_2, \ldots, v_n\}$. Define $v_0 = 0$. By Theorem 2.6 (p. 72), there exists a linear transformation $\mathsf{T}\colon \mathsf{V} \to \mathsf{V}$ such that $\mathsf{T}(v_j) = v_j + v_{j-1}$ for $j = 1, 2, \ldots, n$. Compute $[\mathsf{T}]_\beta$.

11. Let V be an n-dimensional vector space, and let $\mathsf{T}\colon \mathsf{V} \to \mathsf{V}$ be a linear transformation. Suppose that W is a T-invariant subspace of V (see the exercises of Section 2.1) having dimension k. Show that there is a basis β for V such that $[\mathsf{T}]_\beta$ has the form

$$\begin{pmatrix} A & B \\ O & C \end{pmatrix},$$

where A is a $k \times k$ matrix and O is the $(n-k) \times k$ zero matrix.

12. Let V be a finite-dimensional vector space and T be the projection on W along W', where W and W' are subspaces of V. (See the definition in the exercises of Section 2.1 on page 76.) Find an ordered basis β for V such that $[\mathsf{T}]_\beta$ is a diagonal matrix.

13. Let V and W be vector spaces, and let T and U be nonzero linear transformations from V into W. If $\mathsf{R}(\mathsf{T}) \cap \mathsf{R}(\mathsf{U}) = \{\mathit{0}\}$, prove that $\{\mathsf{T}, \mathsf{U}\}$ is a linearly independent subset of $\mathcal{L}(\mathsf{V}, \mathsf{W})$.

14. Let $\mathsf{V} = \mathsf{P}(R)$, and for $j \geq 1$ define $\mathsf{T}_j(f(x)) = f^{(j)}(x)$, where $f^{(j)}(x)$ is the jth derivative of $f(x)$. Prove that the set $\{\mathsf{T}_1, \mathsf{T}_2, \ldots, \mathsf{T}_n\}$ is a linearly independent subset of $\mathcal{L}(\mathsf{V})$ for any positive integer n.

15. Let V and W be vector spaces, and let S be a subset of V. Define $S^0 = \{\mathsf{T} \in \mathcal{L}(\mathsf{V}, \mathsf{W}) \colon \mathsf{T}(x) = \mathit{0} \text{ for all } x \in S\}$. Prove the following statements.

(a) S^0 is a subspace of $\mathcal{L}(\mathsf{V}, \mathsf{W})$.
(b) If S_1 and S_2 are subsets of V and $S_1 \subseteq S_2$, then $S_2^0 \subseteq S_1^0$.
(c) If V_1 and V_2 are subspaces of V, then $(\mathsf{V}_1 + \mathsf{V}_2)^0 = \mathsf{V}_1^0 \cap \mathsf{V}_2^0$.

16. Let V and W be vector spaces such that $\dim(\mathsf{V}) = \dim(\mathsf{W})$, and let $\mathsf{T} \colon \mathsf{V} \to \mathsf{W}$ be linear. Show that there exist ordered bases β and γ for V and W, respectively, such that $[\mathsf{T}]_\beta^\gamma$ is a diagonal matrix.

2.3 COMPOSITION OF LINEAR TRANSFORMATIONS AND MATRIX MULTIPLICATION

In Section 2.2, we learned how to associate a matrix with a linear transformation in such a way that both sums and scalar multiples of matrices are associated with the corresponding sums and scalar multiples of the transformations. The question now arises as to how the matrix representation of a composite of linear transformations is related to the matrix representation of each of the associated linear transformations. The attempt to answer this question leads to a definition of matrix multiplication. We use the more convenient notation of UT rather than $\mathsf{U} \circ \mathsf{T}$ for the composite of linear transformations U and T. (See Appendix B.)

Our first result shows that the composite of linear transformations is linear.

Theorem 2.9. *Let V, W, and Z be vector spaces over the same field F, and let $\mathsf{T} \colon \mathsf{V} \to \mathsf{W}$ and $\mathsf{U} \colon \mathsf{W} \to \mathsf{Z}$ be linear. Then $\mathsf{UT} \colon \mathsf{V} \to \mathsf{Z}$ is linear.*

Proof. Let $x, y \in \mathsf{V}$ and $a \in F$. Then

$$\begin{aligned}\mathsf{UT}(ax + y) &= \mathsf{U}(\mathsf{T}(ax + y)) = \mathsf{U}(a\mathsf{T}(x) + \mathsf{T}(y)) \\ &= a\mathsf{U}(\mathsf{T}(x)) + \mathsf{U}(\mathsf{T}(y)) = a(\mathsf{UT})(x) + \mathsf{UT}(y).\end{aligned}$$ ∎

The following theorem lists some of the properties of the composition of linear transformations.

Theorem 2.10. *Let* V *be a vector space. Let* $\mathsf{T}, \mathsf{U}_1, \mathsf{U}_2 \in \mathcal{L}(\mathsf{V})$. *Then*
(a) $\mathsf{T}(\mathsf{U}_1 + \mathsf{U}_2) = \mathsf{TU}_1 + \mathsf{TU}_2$ *and* $(\mathsf{U}_1 + \mathsf{U}_2)\mathsf{T} = \mathsf{U}_1\mathsf{T} + \mathsf{U}_2\mathsf{T}$
(b) $\mathsf{T}(\mathsf{U}_1\mathsf{U}_2) = (\mathsf{TU}_1)\mathsf{U}_2$
(c) $\mathsf{TI} = \mathsf{IT} = \mathsf{T}$
(d) $a(\mathsf{U}_1\mathsf{U}_2) = (a\mathsf{U}_1)\mathsf{U}_2 = \mathsf{U}_1(a\mathsf{U}_2)$ *for all scalars* a.

Proof. Exercise. ∎

A more general result holds for linear transformations that have domains unequal to their codomains. (See Exercise 8.)

Let $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ and $\mathsf{U}\colon \mathsf{W} \to \mathsf{Z}$ be linear transformations, and let $A = [\mathsf{U}]_\beta^\gamma$ and $B = [\mathsf{T}]_\alpha^\beta$, where $\alpha = \{v_1, v_2, \ldots, v_n\}$, $\beta = \{w_1, w_2, \ldots, w_m\}$, and $\gamma = \{z_1, z_2, \ldots, z_p\}$ are ordered bases for V, W, and Z, respectively. We would like to define the product AB of two matrices so that $AB = [\mathsf{UT}]_\alpha^\gamma$. Consider the matrix $[\mathsf{UT}]_\alpha^\gamma$. For $1 \le j \le n$, we have

$$\begin{aligned}(\mathsf{UT})(v_j) = \mathsf{U}(\mathsf{T}(v_j)) &= \mathsf{U}\left(\sum_{k=1}^m B_{kj}w_k\right) = \sum_{k=1}^m B_{kj}\mathsf{U}(w_k)\\ &= \sum_{k=1}^m B_{kj}\left(\sum_{i=1}^p A_{ik}z_i\right) = \sum_{i=1}^p\left(\sum_{k=1}^m A_{ik}B_{kj}\right)z_i\\ &= \sum_{i=1}^p C_{ij}z_i,\end{aligned}$$

where

$$C_{ij} = \sum_{k=1}^m A_{ik}B_{kj}.$$

This computation motivates the following definition of matrix multiplication.

Definition. *Let* A *be an* $m \times n$ *matrix and* B *be an* $n \times p$ *matrix. We define the* **product** *of* A *and* B, *denoted* AB, *to be the* $m \times p$ *matrix such that*

$$(AB)_{ij} = \sum_{k=1}^n A_{ik}B_{kj} \quad \text{for } 1 \le i \le m, \quad 1 \le j \le p.$$

Note that $(AB)_{ij}$ is the sum of products of corresponding entries from the ith row of A and the jth column of B. Some interesting applications of this definition are presented at the end of this section.

The reader should observe that in order for the product AB to be defined, there are restrictions regarding the relative sizes of A and B. The following mnemonic device is helpful: "$(m \times n)\cdot(n \times p) = (m \times p)$"; that is, in order for the product AB to be defined, the two "inner" dimensions must be equal, and the two "outer" dimensions yield the size of the product.

Example 1

We have

$$\begin{pmatrix} 1 & 2 & 1 \\ 0 & 4 & -1 \end{pmatrix} \begin{pmatrix} 4 \\ 2 \\ 5 \end{pmatrix} = \begin{pmatrix} 1\cdot 4 + 2\cdot 2 + 1\cdot 5 \\ 0\cdot 4 + 4\cdot 2 + (-1)\cdot 5 \end{pmatrix} = \begin{pmatrix} 13 \\ 3 \end{pmatrix}.$$

Notice again the symbolic relationship $(2 \times 3)\cdot(3 \times 1) = 2 \times 1$. ◆

As in the case with composition of functions, we have that matrix multiplication is not commutative. Consider the following two products:

$$\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}.$$

Hence we see that even if both of the matrix products AB and BA are defined, it need not be true that $AB = BA$.

Recalling the definition of the transpose of a matrix from Section 1.3, we show that if A is an $m\times n$ matrix and B is an $n\times p$ matrix, then $(AB)^t = B^tA^t$. Since

$$(AB)^t_{ij} = (AB)_{ji} = \sum_{k=1}^{n} A_{jk}B_{ki}$$

and

$$(B^tA^t)_{ij} = \sum_{k=1}^{n} (B^t)_{ik}(A^t)_{kj} = \sum_{k=1}^{n} B_{ki}A_{jk},$$

we are finished. Therefore the transpose of a product is the product of the transposes *in the opposite order*.

The next theorem is an immediate consequence of our definition of matrix multiplication.

Theorem 2.11. *Let* V, W, *and* Z *be finite-dimensional vector spaces with ordered bases* α, β, *and* γ, *respectively. Let* $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ *and* $\mathsf{U}\colon \mathsf{W} \to \mathsf{Z}$ *be linear transformations. Then*

$$[\mathsf{UT}]_\alpha^\gamma = [\mathsf{U}]_\beta^\gamma [\mathsf{T}]_\alpha^\beta.$$

Corollary. *Let* V *be a finite-dimensional vector space with an ordered basis* β. *Let* $\mathsf{T}, \mathsf{U} \in \mathcal{L}(\mathsf{V})$. *Then* $[\mathsf{UT}]_\beta = [\mathsf{U}]_\beta[\mathsf{T}]_\beta$.

We illustrate Theorem 2.11 in the next example.

Example 2

Let $\mathsf{U}\colon \mathsf{P}_3(R) \to \mathsf{P}_2(R)$ and $\mathsf{T}\colon \mathsf{P}_2(R) \to \mathsf{P}_3(R)$ be the linear transformations respectively defined by

$$\mathsf{U}(f(x)) = f'(x) \qquad \text{and} \qquad \mathsf{T}(f(x)) = \int_0^x f(t)\,dt.$$

Let α and β be the standard ordered bases of $\mathsf{P}_3(R)$ and $\mathsf{P}_2(R)$, respectively. From calculus, it follows that $\mathsf{UT} = \mathsf{I}$, the identity transformation on $\mathsf{P}_2(R)$. To illustrate Theorem 2.11, observe that

$$[\mathsf{UT}]_\beta = [\mathsf{U}]_\alpha^\beta[\mathsf{T}]_\beta^\alpha = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = [\mathsf{I}]_\beta. \quad \blacklozenge$$

The preceding 3×3 diagonal matrix is called an *identity matrix* and is defined next, along with a very useful notation, the *Kronecker delta*.

Definitions. *We define the* **Kronecker delta** δ_{ij} *by* $\delta_{ij} = 1$ *if* $i = j$ *and* $\delta_{ij} = 0$ *if* $i \neq j$. *The* $n \times n$ **identity matrix** I_n *is defined by* $(I_n)_{ij} = \delta_{ij}$.

Thus, for example,

$$I_1 = (1), \quad I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{and} \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The next theorem provides analogs of (a), (c), and (d) of Theorem 2.10. Theorem 2.10(b) has its analog in Theorem 2.16. Observe also that part (c) of the next theorem illustrates that the identity matrix acts as a multiplicative identity in $\mathsf{M}_{n\times n}(F)$. When the context is clear, we sometimes omit the subscript n from I_n.

Theorem 2.12. *Let* A *be an* $m \times n$ *matrix,* B *and* C *be* $n \times p$ *matrices, and* D *and* E *be* $q \times m$ *matrices. Then*

(a) $A(B+C) = AB + AC$ *and* $(D+E)A = DA + EA$.

(b) $a(AB) = (aA)B = A(aB)$ *for any scalar* a.

(c) $I_mA = A = AI_n$.

(d) *If* V *is an* n*-dimensional vector space with an ordered basis* β, *then* $[\mathsf{I_V}]_\beta = I_n$.

Proof. We prove the first half of (a) and (c) and leave the remaining proofs as an exercise. (See Exercise 5.)

(a) We have

$$\begin{aligned}[A(B+C)]_{ij} &= \sum_{k=1}^{n} A_{ik}(B+C)_{kj} = \sum_{k=1}^{n} A_{ik}(B_{kj}+C_{kj}) \\ &= \sum_{k=1}^{n}(A_{ik}B_{kj}+A_{ik}C_{kj}) = \sum_{k=1}^{n} A_{ik}B_{kj} + \sum_{k=1}^{n} A_{ik}C_{kj} \\ &= (AB)_{ij} + (AC)_{ij} = [AB+AC]_{ij}.\end{aligned}$$

So $A(B+C) = AB + AC$.

(c) We have

$$(I_mA)_{ij} = \sum_{k=1}^{m}(I_m)_{ik}A_{kj} = \sum_{k=1}^{m}\delta_{ik}A_{kj} = A_{ij}. \qquad \blacksquare$$

Corollary. *Let A be an $m \times n$ matrix, $B_1, B_2, \ldots, B_k$ be $n \times p$ matrices, $C_1, C_2, \ldots, C_k$ be $q \times m$ matrices, and $a_1, a_2, \ldots, a_k$ be scalars. Then*

$$A\left(\sum_{i=1}^{k} a_iB_i\right) = \sum_{i=1}^{k} a_iAB_i$$

and

$$\left(\sum_{i=1}^{k} a_iC_i\right)A = \sum_{i=1}^{k} a_iC_iA.$$

Proof. Exercise. $\blacksquare$

For an $n \times n$ matrix A, we define $A^1 = A$, $A^2 = AA$, $A^3 = A^2A$, and, in general, $A^k = A^{k-1}A$ for $k = 2, 3, \ldots$. We define $A^0 = I_n$.

With this notation, we see that if

$$A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

then $A^2 = O$ (the zero matrix) even though $A \neq O$. Thus the cancellation property for multiplication in fields is not valid for matrices. To see why, assume that the cancellation law is valid. Then, from $A \cdot A = A^2 = O = A \cdot O$, we would conclude that $A = O$, which is false.

Theorem 2.13. *Let A be an $m \times n$ matrix and B be an $n \times p$ matrix. For each j $(1 \leq j \leq p)$ let u_j and v_j denote the jth columns of AB and B, respectively. Then*

(a) $u_j = Av_j$
(b) $v_j = Be_j$, *where* e_j *is the* j*th standard vector of* F^p.

Proof. (a) We have

$$u_j = \begin{pmatrix} (AB)_{1j} \\ (AB)_{2j} \\ \vdots \\ (AB)_{mj} \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{n} A_{1k}B_{kj} \\ \sum_{k=1}^{n} A_{2k}B_{kj} \\ \vdots \\ \sum_{k=1}^{n} A_{mk}B_{kj} \end{pmatrix} = A \begin{pmatrix} B_{1j} \\ B_{2j} \\ \vdots \\ B_{nj} \end{pmatrix} = Av_j.$$

Hence (a) is proved. The proof of (b) is left as an exercise. (See Exercise 6.) ∎

It follows (see Exercise 14) from Theorem 2.13 that column j of AB is a linear combination of the columns of A with the coefficients in the linear combination being the entries of column j of B. An analogous result holds for rows; that is, row i of AB is a linear combination of the rows of B with the coefficients in the linear combination being the entries of row i of A.

The next result justifies much of our past work. It utilizes both the matrix representation of a linear transformation and matrix multiplication in order to evaluate the transformation at any given vector.

Theorem 2.14. *Let* V *and* W *be finite-dimensional vector spaces having ordered bases* β *and* γ*, respectively, and let* $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ *be linear. Then, for each* $u \in \mathsf{V}$*, we have*

$$[\mathsf{T}(u)]_\gamma = [\mathsf{T}]_\beta^\gamma [u]_\beta.$$

Proof. Fix $u \in \mathsf{V}$, and define the linear transformations $f\colon F \to \mathsf{V}$ by $f(a) = au$ and $g\colon F \to \mathsf{W}$ by $g(a) = a\mathsf{T}(u)$ for all $a \in F$. Let $\alpha = \{1\}$ be the standard ordered basis for F. Notice that $g = \mathsf{T}f$. Identifying column vectors as matrices and using Theorem 2.11, we obtain

$$[\mathsf{T}(u)]_\gamma = [g(1)]_\gamma = [g]_\alpha^\gamma = [\mathsf{T}f]_\alpha^\gamma = [\mathsf{T}]_\beta^\gamma [f]_\alpha^\beta = [\mathsf{T}]_\beta^\gamma [f(1)]_\beta = [\mathsf{T}]_\beta^\gamma [u]_\beta. \quad ∎$$

Example 3

Let $\mathsf{T}\colon \mathsf{P}_3(R) \to \mathsf{P}_2(R)$ be the linear transformation defined by $\mathsf{T}(f(x)) = f'(x)$, and let β and γ be the standard ordered bases for $\mathsf{P}_3(R)$ and $\mathsf{P}_2(R)$, respectively. If $A = [\mathsf{T}]_\beta^\gamma$, then, from Example 4 of Section 2.2, we have

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

We illustrate Theorem 2.14 by verifying that $[\mathsf{T}(p(x))]_\gamma = [\mathsf{T}]_\beta^\gamma [p(x)]_\beta$, where $p(x) \in \mathsf{P}_3(R)$ is the polynomial $p(x) = 2-4x+x^2+3x^3$. Let $q(x) = \mathsf{T}(p(x))$; then $q(x) = p'(x) = -4+2x+9x^2$. Hence

$$[\mathsf{T}(p(x))]_\gamma = [q(x)]_\gamma = \begin{pmatrix} -4 \\ 2 \\ 9 \end{pmatrix},$$

but also

$$[\mathsf{T}]_\beta^\gamma [p(x)]_\beta = A[p(x)]_\beta = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 2 \\ -4 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} -4 \\ 2 \\ 9 \end{pmatrix}. \qquad \blacklozenge$$

We complete this section with the introduction of the *left-multiplication transformation* L_A, where A is an $m \times n$ matrix. This transformation is probably the most important tool for transferring properties about transformations to analogous properties about matrices and vice versa. For example, we use it to prove that matrix multiplication is associative.

Definition. *Let A be an $m \times n$ matrix with entries from a field F. We denote by L_A the mapping $\mathsf{L}_A \colon \mathsf{F}^n \to \mathsf{F}^m$ defined by $\mathsf{L}_A(x) = Ax$ (the matrix product of A and x) for each column vector $x \in \mathsf{F}^n$. We call L_A a* ***left-multiplication transformation.***

Example 4

Let

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}.$$

Then $A \in \mathsf{M}_{2\times 3}(R)$ and $\mathsf{L}_A \colon \mathsf{R}^3 \to \mathsf{R}^2$. If

$$x = \begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix},$$

then

$$\mathsf{L}_A(x) = Ax = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 6 \\ 1 \end{pmatrix}. \qquad \blacklozenge$$

We see in the next theorem that not only is L_A linear, but, in fact, it has a great many other useful properties. These properties are all quite natural and so are easy to remember.

Theorem 2.15. *Let A be an $m \times n$ matrix with entries from F. Then the left-multiplication transformation $\mathsf{L}_A\colon \mathsf{F}^n \to \mathsf{F}^m$ is linear. Furthermore, if B is any other $m \times n$ matrix (with entries from F) and β and γ are the standard ordered bases for F^n and F^m, respectively, then we have the following properties.*

(a) $[\mathsf{L}_A]_\beta^\gamma = A$.
(b) $\mathsf{L}_A = \mathsf{L}_B$ *if and only if* $A = B$.
(c) $\mathsf{L}_{A+B} = \mathsf{L}_A + \mathsf{L}_B$ *and* $\mathsf{L}_{aA} = a\mathsf{L}_A$ *for all* $a \in F$.
(d) *If* $\mathsf{T}\colon \mathsf{F}^n \to \mathsf{F}^m$ *is linear, then there exists a unique* $m \times n$ *matrix* C *such that* $\mathsf{T} = \mathsf{L}_C$. *In fact,* $C = [\mathsf{T}]_\beta^\gamma$.
(e) *If* E *is an* $n \times p$ *matrix, then* $\mathsf{L}_{AE} = \mathsf{L}_A\mathsf{L}_E$.
(f) *If* $m = n$, *then* $\mathsf{L}_{I_n} = \mathsf{I}_{\mathsf{F}^n}$.

Proof. The fact that L_A is linear follows immediately from Theorem 2.12.

(a) The jth column of $[\mathsf{L}_A]_\beta^\gamma$ is equal to $\mathsf{L}_A(e_j)$. However $\mathsf{L}_A(e_j) = Ae_j$, which is also the jth column of A by Theorem 2.13(b). So $[\mathsf{L}_A]_\beta^\gamma = A$.

(b) If $\mathsf{L}_A = \mathsf{L}_B$, then we may use (a) to write $A = [\mathsf{L}_A]_\beta^\gamma = [\mathsf{L}_B]_\beta^\gamma = B$. Hence $A = B$. The proof of the converse is trivial.

(c) The proof is left as an exercise. (See Exercise 7.)

(d) Let $C = [\mathsf{T}]_\beta^\gamma$. By Theorem 2.14, we have $[\mathsf{T}(x)]_\gamma = [\mathsf{T}]_\beta^\gamma[x]_\beta$, or $\mathsf{T}(x) = Cx = \mathsf{L}_C(x)$ for all $x \in F^n$. So $\mathsf{T} = \mathsf{L}_C$. The uniqueness of C follows from (b).

(e) For any j $(1 \le j \le p)$, we may apply Theorem 2.13 several times to note that $(AE)e_j$ is the jth column of AE and that the jth column of AE is also equal to $A(Ee_j)$. So $(AE)e_j = A(Ee_j)$. Thus

$$\mathsf{L}_{AE}(e_j) = (AE)e_j = A(Ee_j) = \mathsf{L}_A(Ee_j) = \mathsf{L}_A(\mathsf{L}_E(e_j)).$$

Hence $\mathsf{L}_{AE} = \mathsf{L}_A\mathsf{L}_E$ by the corollary to Theorem 2.6 (p. 73).

(f) The proof is left as an exercise. (See Exercise 7.) ∎

We now use left-multiplication transformations to establish the associativity of matrix multiplication.

Theorem 2.16. *Let A, B, and C be matrices such that $A(BC)$ is defined. Then $(AB)C$ is also defined and $A(BC) = (AB)C$; that is, matrix multiplication is associative.*

Proof. It is left to the reader to show that $(AB)C$ is defined. Using (e) of Theorem 2.15 and the associativity of functional composition (see Appendix B), we have

$$\mathsf{L}_{A(BC)} = \mathsf{L}_A\mathsf{L}_{BC} = \mathsf{L}_A(\mathsf{L}_B\mathsf{L}_C) = (\mathsf{L}_A\mathsf{L}_B)\mathsf{L}_C = \mathsf{L}_{AB}\mathsf{L}_C = \mathsf{L}_{(AB)C}.$$

So from (b) of Theorem 2.15, it follows that $A(BC) = (AB)C$. ∎

Needless to say, this theorem could be proved directly from the definition of matrix multiplication (see Exercise 18). The proof above, however, provides a prototype of many of the arguments that utilize the relationships between linear transformations and matrices.

Applications

A large and varied collection of interesting applications arises in connection with special matrices called *incidence matrices*. An **incidence matrix** is a square matrix in which all the entries are either zero or one and, for convenience, all the diagonal entries are zero. If we have a relationship on a set of n objects that we denote by $1, 2, \ldots, n$, then we define the associated incidence matrix A by $A_{ij} = 1$ if i is related to j, and $A_{ij} = 0$ otherwise.

To make things concrete, suppose that we have four people, each of whom owns a communication device. If the relationship on this group is "can transmit to," then $A_{ij} = 1$ if i can send a message to j, and $A_{ij} = 0$ otherwise. Suppose that

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix}.$$

Then since $A_{34} = 1$ and $A_{14} = 0$, we see that person 3 can send to 4 but 1 cannot send to 4.

We obtain an interesting interpretation of the entries of A^2. Consider, for instance,

$$(A^2)_{31} = A_{31}A_{11} + A_{32}A_{21} + A_{33}A_{31} + A_{34}A_{41}.$$

Note that any term $A_{3k}A_{k1}$ equals 1 if and only if both A_{3k} and A_{k1} equal 1, that is, if and only if 3 can send to k and k can send to 1. Thus $(A^2)_{31}$ gives the number of ways in which 3 can send to 1 in two *stages* (or in one *relay*). Since

$$A^2 = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 2 & 0 & 0 \\ 2 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \end{pmatrix},$$

we see that there are two ways 3 can send to 1 in two stages. In general, $(A + A^2 + \cdots + A^m)_{ij}$ is the number of ways in which i can send to j in at most m stages.

A maximal collection of three or more people with the property that any two can send to each other is called a **clique**. The problem of determining cliques is difficult, but there is a simple method for determining if someone belongs to a clique. If we define a new matrix B by $B_{ij} = 1$ if i and j can send to each other, and $B_{ij} = 0$ otherwise, then it can be shown (see Exercise 19) that person i belongs to a clique if and only if $(B^3)_{ii} > 0$. For example, suppose that the incidence matrix associated with some relationship is

$$A = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}.$$

To determine which people belong to cliques, we form the matrix B, described earlier, and compute B^3. In this case,

$$B = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix} \quad \text{and} \quad B^3 = \begin{pmatrix} 0 & 4 & 0 & 4 \\ 4 & 0 & 4 & 0 \\ 0 & 4 & 0 & 4 \\ 4 & 0 & 4 & 0 \end{pmatrix}.$$

Since all the diagonal entries of B^3 are zero, we conclude that there are no cliques in this relationship.

Our final example of the use of incidence matrices is concerned with the concept of *dominance*. A relation among a group of people is called a **dominance relation** if the associated incidence matrix A has the property that for all distinct pairs i and j, $A_{ij} = 1$ if and only if $A_{ji} = 0$, that is, given any two people, exactly one of them *dominates* (or, using the terminology of our first example, can send a message to) the other. Since A is an incidence matrix, $A_{ii} = 0$ for all i. For such a relation, it can be shown (see Exercise 21) that the matrix $A + A^2$ has a row [column] in which each entry is positive except for the diagonal entry. In other words, there is at least one person who dominates [is dominated by] all others in one or two stages. In fact, it can be shown that any person who dominates [is dominated by] the greatest number of people in the first stage has this property. Consider, for example, the matrix

$$A = \begin{pmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 \end{pmatrix}.$$

The reader should verify that this matrix corresponds to a dominance relation. Now

$$A + A^2 = \begin{pmatrix} 0 & 2 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 2 & 0 & 2 & 1 \\ 1 & 2 & 2 & 0 & 1 \\ 2 & 2 & 2 & 2 & 0 \end{pmatrix}.$$

Thus persons 1, 3, 4, and 5 dominate (can send messages to) all the others in at most two stages, while persons 1, 2, 3, and 4 are dominated by (can receive messages from) all the others in at most two stages.

EXERCISES

1. Label the following statements as true or false. In each part, V, W, and Z denote vector spaces with ordered (finite) bases α, β, and γ, respectively; $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ and $\mathsf{U}\colon \mathsf{W} \to \mathsf{Z}$ denote linear transformations; and A and B denote matrices.
   (a) $[\mathsf{UT}]_\alpha^\gamma = [\mathsf{T}]_\alpha^\beta [\mathsf{U}]_\beta^\gamma$.
   (b) $[\mathsf{T}(v)]_\beta = [\mathsf{T}]_\alpha^\beta [v]_\alpha$ for all $v \in \mathsf{V}$.
   (c) $[\mathsf{U}(w)]_\beta = [\mathsf{U}]_\alpha^\beta [w]_\beta$ for all $w \in \mathsf{W}$.
   (d) $[\mathsf{I_V}]_\alpha = I$.
   (e) $[\mathsf{T}^2]_\alpha^\beta = ([\mathsf{T}]_\alpha^\beta)^2$.
   (f) $A^2 = I$ implies that $A = I$ or $A = -I$.
   (g) $\mathsf{T} = \mathsf{L}_A$ for some matrix A.
   (h) $A^2 = O$ implies that $A = O$, where O denotes the zero matrix.
   (i) $\mathsf{L}_{A+B} = \mathsf{L}_A + \mathsf{L}_B$.
   (j) If A is square and $A_{ij} = \delta_{ij}$ for all i and j, then $A = I$.

2. (a) Let

$$A = \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 & -3 \\ 4 & 1 & 2 \end{pmatrix},$$

$$C = \begin{pmatrix} 1 & 1 & 4 \\ -1 & -2 & 0 \end{pmatrix}, \quad \text{and} \quad D = \begin{pmatrix} 2 \\ -2 \\ 3 \end{pmatrix}.$$

   Compute $A(2B + 3C), (AB)D$, and $A(BD)$.
   (b) Let

$$A = \begin{pmatrix} 2 & 5 \\ -3 & 1 \\ 4 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 3 & -2 & 0 \\ 1 & -1 & 4 \\ 5 & 5 & 3 \end{pmatrix}, \quad \text{and} \quad C = \begin{pmatrix} 4 & 0 & 3 \end{pmatrix}.$$

   Compute A^t, A^tB, BC^t, CB, and CA.

3. Let $g(x) = 3 + x$. Let $\mathsf{T}\colon \mathsf{P}_2(R) \to \mathsf{P}_2(R)$ and $\mathsf{U}\colon \mathsf{P}_2(R) \to \mathsf{R}^3$ be the linear transformations respectively defined by

$$\mathsf{T}(f(x)) = f'(x)g(x) + 2f(x) \quad \text{and} \quad U(a + bx + cx^2) = (a + b, c, a - b).$$

   Let β and γ be the standard ordered bases of $\mathsf{P}_2(R)$ and R^3, respectively.

(a) Compute $[\mathsf{U}]_\beta^\gamma$, $[\mathsf{T}]_\beta$, and $[\mathsf{UT}]_\beta^\gamma$ directly. Then use Theorem 2.11 to verify your result.

(b) Let $h(x) = 3 - 2x + x^2$. Compute $[h(x)]_\beta$ and $[\mathsf{U}(h(x))]_\gamma$. Then use $[\mathsf{U}]_\beta^\gamma$ from (a) and Theorem 2.14 to verify your result.

4. For each of the following parts, let T be the linear transformation defined in the corresponding part of Exercise 5 of Section 2.2. Use Theorem 2.14 to compute the following vectors:

(a) $[\mathsf{T}(A)]_\alpha$, where $A = \begin{pmatrix} 1 & 4 \\ -1 & 6 \end{pmatrix}$.

(b) $[\mathsf{T}(f(x))]_\alpha$, where $f(x) = 4 - 6x + 3x^2$.

(c) $[\mathsf{T}(A)]_\gamma$, where $A = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$.

(d) $[\mathsf{T}(f(x))]_\gamma$, where $f(x) = 6 - x + 2x^2$.

5. Complete the proof of Theorem 2.12 and its corollary.

6. Prove (b) of Theorem 2.13.

7. Prove (c) and (f) of Theorem 2.15.

8. Prove Theorem 2.10. Now state and prove a more general result involving linear transformations with domains unequal to their codomains.

9. Find linear transformations $\mathsf{U}, \mathsf{T}\colon \mathsf{F}^2 \to \mathsf{F}^2$ such that $\mathsf{UT} = \mathsf{T}_0$ (the zero transformation) but $\mathsf{TU} \neq \mathsf{T}_0$. Use your answer to find matrices A and B such that $AB = O$ but $BA \neq O$.

10. Let A be an $n \times n$ matrix. Prove that A is a diagonal matrix if and only if $A_{ij} = \delta_{ij} A_{ij}$ for all i and j.

11. Let V be a vector space, and let $\mathsf{T}\colon \mathsf{V} \to \mathsf{V}$ be linear. Prove that $\mathsf{T}^2 = \mathsf{T}_0$ if and only if $\mathsf{R}(\mathsf{T}) \subseteq \mathsf{N}(\mathsf{T})$.

12. Let V, W, and Z be vector spaces, and let $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ and $\mathsf{U}\colon \mathsf{W} \to \mathsf{Z}$ be linear.

(a) Prove that if UT is one-to-one, then T is one-to-one. Must U also be one-to-one?

(b) Prove that if UT is onto, then U is onto. Must T also be onto?

(c) Prove that if U and T are one-to-one and onto, then UT is also.

13. Let A and B be $n \times n$ matrices. Recall that the trace of A is defined by

$$\operatorname{tr}(A) = \sum_{i=1}^{n} A_{ii}.$$

Prove that $\operatorname{tr}(AB) = \operatorname{tr}(BA)$ and $\operatorname{tr}(A) = \operatorname{tr}(A^t)$.

14. Assume the notation in Theorem 2.13.

(a) Suppose that z is a (column) vector in F^p. Use Theorem 2.13(b) to prove that Bz is a linear combination of the columns of B. In particular, if $z = (a_1, a_2, \ldots, a_p)^t$, then show that

$$Bz = \sum_{j=1}^{p} a_j v_j.$$

(b) Extend (a) to prove that column j of AB is a linear combination of the columns of A with the coefficients in the linear combination being the entries of column j of B.

(c) For any row vector $w \in \mathsf{F}^m$, prove that wA is a linear combination of the rows of A with the coefficients in the linear combination being the coordinates of w. *Hint:* Use properties of the transpose operation applied to (a).

(d) Prove the analogous result to (b) about rows: Row i of AB is a linear combination of the rows of B with the coefficients in the linear combination being the entries of row i of A.

15.† Let M and A be matrices for which the product matrix MA is defined. If the jth column of A is a linear combination of a set of columns of A, prove that the jth column of MA is a linear combination of the corresponding columns of MA with the same corresponding coefficients.

16. Let V be a finite-dimensional vector space, and let $\mathsf{T}\colon \mathsf{V} \to \mathsf{V}$ be linear.

(a) If $\mathrm{rank}(\mathsf{T}) = \mathrm{rank}(\mathsf{T}^2)$, prove that $\mathsf{R}(\mathsf{T}) \cap \mathsf{N}(\mathsf{T}) = \{0\}$. Deduce that $\mathsf{V} = \mathsf{R}(\mathsf{T}) \oplus \mathsf{N}(\mathsf{T})$ (see the exercises of Section 1.3).

(b) Prove that $\mathsf{V} = \mathsf{R}(\mathsf{T}^k) \oplus \mathsf{N}(\mathsf{T}^k)$ for some positive integer k.

17. Let V be a vector space. Determine all linear transformations $\mathsf{T}\colon \mathsf{V} \to \mathsf{V}$ such that $\mathsf{T} = \mathsf{T}^2$. *Hint:* Note that $x = \mathsf{T}(x) + (x - \mathsf{T}(x))$ for every x in V, and show that $\mathsf{V} = \{y \colon \mathsf{T}(y) = y\} \oplus \mathsf{N}(\mathsf{T})$ (see the exercises of Section 1.3).

18. Using only the definition of matrix multiplication, prove that multiplication of matrices is associative.

19. For an incidence matrix A with related matrix B defined by $B_{ij} = 1$ if i is related to j and j is related to i, and $B_{ij} = 0$ otherwise, prove that i belongs to a clique if and only if $(B^3)_{ii} > 0$.

20. Use Exercise 19 to determine the cliques in the relations corresponding to the following incidence matrices.

(a) $\begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$ (b) $\begin{pmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$

21. Let A be an incidence matrix that is associated with a dominance relation. Prove that the matrix $A + A^2$ has a row [column] in which each entry is positive except for the diagonal entry.

22. Prove that the matrix

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

corresponds to a dominance relation. Use Exercise 21 to determine which persons dominate [are dominated by] each of the others within two stages.

23. Let A be an $n \times n$ incidence matrix that corresponds to a dominance relation. Determine the number of nonzero entries of A.

2.4 INVERTIBILITY AND ISOMORPHISMS

The concept of invertibility is introduced quite early in the study of functions. Fortunately, many of the intrinsic properties of functions are shared by their inverses. For example, in calculus we learn that the properties of being continuous or differentiable are generally retained by the inverse functions. We see in this section (Theorem 2.17) that the inverse of a linear transformation is also linear. This result greatly aids us in the study of *inverses* of matrices. As one might expect from Section 2.3, the inverse of the left-multiplication transformation L_A (when it exists) can be used to determine properties of the inverse of the matrix A.

In the remainder of this section, we apply many of the results about invertibility to the concept of *isomorphism.* We will see that finite-dimensional vector spaces (over F) of equal dimension may be identified. These ideas will be made precise shortly.

The facts about inverse functions presented in Appendix B are, of course, true for linear transformations. Nevertheless, we repeat some of the definitions for use in this section.

Definition. *Let V and W be vector spaces, and let $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ be linear. A function $\mathsf{U}\colon \mathsf{W} \to \mathsf{V}$ is said to be an* **inverse** *of T if $\mathsf{TU} = \mathsf{I_W}$ and $\mathsf{UT} = \mathsf{I_V}$. If T has an inverse, then T is said to be* **invertible**. *As noted in Appendix B, if T is invertible, then the inverse of T is unique and is denoted by T^{-1}.*

The following facts hold for invertible functions T and U.

1. $(\mathsf{TU})^{-1} = \mathsf{U}^{-1}\mathsf{T}^{-1}$.
2. $(\mathsf{T}^{-1})^{-1} = \mathsf{T}$; in particular, T^{-1} is invertible.

We often use the fact that a function is invertible if and only if it is both one-to-one and onto. We can therefore restate Theorem 2.5 as follows.

3. Let $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ be a linear transformation, where V and W are finite-dimensional spaces of equal dimension. Then T is invertible if and only if $\operatorname{rank}(\mathsf{T}) = \dim(\mathsf{V})$.

Example 1

Let $\mathsf{T}\colon \mathsf{P}_1(R) \to \mathsf{R}^2$ be the linear transformation defined by $\mathsf{T}(a + bx) = (a, a+b)$. The reader can verify directly that $\mathsf{T}^{-1}\colon \mathsf{R}^2 \to \mathsf{P}_1(R)$ is defined by $\mathsf{T}^{-1}(c,d) = c + (d-c)x$. Observe that T^{-1} is also linear. As Theorem 2.17 demonstrates, this is true in general. ◆

Theorem 2.17. *Let V and W be vector spaces, and let $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ be linear and invertible. Then $\mathsf{T}^{-1}\colon \mathsf{W} \to \mathsf{V}$ is linear.*

Proof. Let $y_1, y_2 \in \mathsf{W}$ and $c \in F$. Since T is onto and one-to-one, there exist unique vectors x_1 and x_2 such that $\mathsf{T}(x_1) = y_1$ and $\mathsf{T}(x_2) = y_2$. Thus $x_1 = \mathsf{T}^{-1}(y_1)$ and $x_2 = \mathsf{T}^{-1}(y_2)$; so

$$\begin{aligned}\mathsf{T}^{-1}(cy_1 + y_2) &= \mathsf{T}^{-1}[c\mathsf{T}(x_1) + \mathsf{T}(x_2)] = \mathsf{T}^{-1}[\mathsf{T}(cx_1 + x_2)]\\ &= cx_1 + x_2 = c\mathsf{T}^{-1}(y_1) + \mathsf{T}^{-1}(y_2).\end{aligned}$$ ∎

It now follows immediately from Theorem 2.5 (p. 71) that if T is a linear transformation between vector spaces of equal (finite) dimension, then the conditions of being invertible, one-to-one, and onto are all equivalent.

We are now ready to define the inverse of a matrix. The reader should note the analogy with the inverse of a linear transformation.

Definition. *Let A be an $n \times n$ matrix. Then A is* **invertible** *if there exists an $n \times n$ matrix B such that $AB = BA = I$.*

If A is invertible, then the matrix B such that $AB = BA = I$ is unique. (If C were another such matrix, then $C = CI = C(AB) = (CA)B = IB = B$.) The matrix B is called the **inverse** of A and is denoted by A^{-1}.

Example 2

The reader should verify that the inverse of

$$\begin{pmatrix} 5 & 7 \\ 2 & 3 \end{pmatrix} \quad \text{is} \quad \begin{pmatrix} 3 & -7 \\ -2 & 5 \end{pmatrix}.$$ ◆

In Section 3.2, we learn a technique for computing the inverse of a matrix. At this point, we develop a number of results that relate the inverses of matrices to the inverses of linear transformations.

Lemma. *Let* T *be an invertible linear transformation from* V *to* W. *Then* V *is finite-dimensional if and only if* W *is finite-dimensional. In this case,* $\dim(\mathsf{V}) = \dim(\mathsf{W})$.

Proof. Suppose that V is finite-dimensional. Let $\beta = \{x_1, x_2, \ldots, x_n\}$ be a basis for V. By Theorem 2.2 (p. 68), $\mathsf{T}(\beta)$ spans $\mathsf{R}(\mathsf{T}) = \mathsf{W}$; hence W is finite-dimensional by Theorem 1.9 (p. 44). Conversely, if W is finite-dimensional, then so is V by a similar argument, using T^{-1}.

Now suppose that V and W are finite-dimensional. Because T is one-to-one and onto, we have

$$\text{nullity}(\mathsf{T}) = 0 \quad \text{and} \quad \text{rank}(\mathsf{T}) = \dim(\mathsf{R}(\mathsf{T})) = \dim(\mathsf{W}).$$

So by the dimension theorem (p. 70), it follows that $\dim(\mathsf{V}) = \dim(\mathsf{W})$. ∎

Theorem 2.18. *Let* V *and* W *be finite-dimensional vector spaces with ordered bases* β *and* γ, *respectively. Let* $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ *be linear. Then* T *is invertible if and only if* $[\mathsf{T}]_\beta^\gamma$ *is invertible. Furthermore,* $[\mathsf{T}^{-1}]_\gamma^\beta = ([\mathsf{T}]_\beta^\gamma)^{-1}$.

Proof. Suppose that T is invertible. By the lemma, we have $\dim(\mathsf{V}) = \dim(\mathsf{W})$. Let $n = \dim(\mathsf{V})$. So $[\mathsf{T}]_\beta^\gamma$ is an $n \times n$ matrix. Now $\mathsf{T}^{-1}\colon \mathsf{W} \to \mathsf{V}$ satisfies $\mathsf{T}\mathsf{T}^{-1} = \mathsf{I}_\mathsf{W}$ and $\mathsf{T}^{-1}\mathsf{T} = \mathsf{I}_\mathsf{V}$. Thus

$$I_n = [\mathsf{I}_\mathsf{V}]_\beta = [\mathsf{T}^{-1}\mathsf{T}]_\beta = [\mathsf{T}^{-1}]_\gamma^\beta [\mathsf{T}]_\beta^\gamma.$$

Similarly, $[\mathsf{T}]_\beta^\gamma [\mathsf{T}^{-1}]_\gamma^\beta = I_n$. So $[\mathsf{T}]_\beta^\gamma$ is invertible and $\left([\mathsf{T}]_\beta^\gamma\right)^{-1} = [\mathsf{T}^{-1}]_\gamma^\beta$.

Now suppose that $A = [\mathsf{T}]_\beta^\gamma$ is invertible. Then there exists an $n \times n$ matrix B such that $AB = BA = I_n$. By Theorem 2.6 (p. 72), there exists $\mathsf{U} \in \mathcal{L}(\mathsf{W}, \mathsf{V})$ such that

$$\mathsf{U}(w_j) = \sum_{i=1}^{n} B_{ij} v_i \quad \text{for } j = 1, 2, \ldots, n,$$

where $\gamma = \{w_1, w_2, \ldots, w_n\}$ and $\beta = \{v_1, v_2, \ldots, v_n\}$. It follows that $[\mathsf{U}]_\gamma^\beta = B$. To show that $\mathsf{U} = \mathsf{T}^{-1}$, observe that

$$[\mathsf{U}\mathsf{T}]_\beta = [\mathsf{U}]_\gamma^\beta [\mathsf{T}]_\beta^\gamma = BA = I_n = [\mathsf{I}_\mathsf{V}]_\beta$$

by Theorem 2.11 (p. 88). So $\mathsf{U}\mathsf{T} = \mathsf{I}_\mathsf{V}$, and similarly, $\mathsf{T}\mathsf{U} = \mathsf{I}_\mathsf{W}$. ∎

Example 3

Let β and γ be the standard ordered bases of $\mathsf{P}_1(R)$ and R^2, respectively. For T as in Example 1, we have

$$[\mathsf{T}]_\beta^\gamma = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad [\mathsf{T}^{-1}]_\gamma^\beta = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

It can be verified by matrix multiplication that each matrix is the inverse of the other. ♦

Corollary 1. *Let V be a finite-dimensional vector space with an ordered basis β, and let $\mathsf{T}\colon \mathsf{V} \to \mathsf{V}$ be linear. Then T is invertible if and only if $[\mathsf{T}]_\beta$ is invertible. Furthermore, $[\mathsf{T}^{-1}]_\beta = ([\mathsf{T}]_\beta)^{-1}$.*

Proof. Exercise. ∎

Corollary 2. *Let A be an $n \times n$ matrix. Then A is invertible if and only if L_A is invertible. Furthermore, $(\mathsf{L}_A)^{-1} = \mathsf{L}_{A^{-1}}$.*

Proof. Exercise. ∎

The notion of invertibility may be used to formalize what may already have been observed by the reader, that is, that certain vector spaces strongly resemble one another except for the form of their vectors. For example, in the case of $\mathsf{M}_{2\times 2}(F)$ and F^4, if we associate to each matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

the 4-tuple (a, b, c, d), we see that sums and scalar products associate in a similar manner; that is, in terms of the vector space structure, these two vector spaces may be considered identical or *isomorphic.*

Definitions. *Let V and W be vector spaces. We say that V is* ***isomorphic*** *to W if there exists a linear transformation $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ that is invertible. Such a linear transformation is called an* ***isomorphism*** *from V onto W.*

We leave as an exercise (see Exercise 13) the proof that "is isomorphic to" is an equivalence relation. (See Appendix A.) So we need only say that V and W are isomorphic.

Example 4

Define $\mathsf{T}\colon \mathsf{F}^2 \to \mathsf{P}_1(F)$ by $\mathsf{T}(a_1, a_2) = a_1 + a_2 x$. It is easily checked that T is an isomorphism; so F^2 is isomorphic to $\mathsf{P}_1(F)$. ♦

Example 5

Define

$$\mathsf{T}\colon \mathsf{P}_3(R) \to \mathsf{M}_{2\times 2}(R) \quad \text{by } \mathsf{T}(f) = \begin{pmatrix} f(1) & f(2) \\ f(3) & f(4) \end{pmatrix}.$$

It is easily verified that T is linear. By use of the Lagrange interpolation formula in Section 1.6, it can be shown (compare with Exercise 22) that $\mathsf{T}(f) = O$ only when f is the zero polynomial. Thus T is one-to-one (see Exercise 11). Moreover, because $\dim(\mathsf{P}_3(R)) = \dim(\mathsf{M}_{2\times 2}(R))$, it follows that T is invertible by Theorem 2.5 (p. 71). We conclude that $\mathsf{P}_3(R)$ is isomorphic to $\mathsf{M}_{2\times 2}(R)$. ♦

In each of Examples 4 and 5, the reader may have observed that isomorphic vector spaces have equal dimensions. As the next theorem shows, this is no coincidence.

Theorem 2.19. *Let V and W be finite-dimensional vector spaces (over the same field). Then V is isomorphic to W if and only if* $\dim(\mathsf{V}) = \dim(\mathsf{W})$.

Proof. Suppose that V is isomorphic to W and that $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ is an isomorphism from V to W. By the lemma preceding Theorem 2.18, we have that $\dim(\mathsf{V}) = \dim(\mathsf{W})$.

Now suppose that $\dim(\mathsf{V}) = \dim(\mathsf{W})$, and let $\beta = \{v_1, v_2, \ldots, v_n\}$ and $\gamma = \{w_1, w_2, \ldots, w_n\}$ be bases for V and W, respectively. By Theorem 2.6 (p. 72), there exists $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ such that T is linear and $\mathsf{T}(v_i) = w_i$ for $i = 1, 2, \ldots, n$. Using Theorem 2.2 (p. 68), we have

$$\mathsf{R}(\mathsf{T}) = \operatorname{span}(\mathsf{T}(\beta)) = \operatorname{span}(\gamma) = \mathsf{W}.$$

So T is onto. From Theorem 2.5 (p. 71), we have that T is also one-to-one. Hence T is an isomorphism. ∎

By the lemma to Theorem 2.18, if V and W are isomorphic, then either both of V and W are finite-dimensional or both are infinite-dimensional.

Corollary. *Let V be a vector space over F. Then V is isomorphic to F^n if and only if* $\dim(\mathsf{V}) = n$.

Up to this point, we have associated linear transformations with their matrix representations. We are now in a position to prove that, as a vector space, the collection of all linear transformations between two given vector spaces may be identified with the appropriate vector space of $m \times n$ matrices.

Theorem 2.20. *Let V and W be finite-dimensional vector spaces over F of dimensions n and m, respectively, and let β and γ be ordered bases for V and W, respectively. Then the function $\Phi\colon \mathcal{L}(\mathsf{V}, \mathsf{W}) \to \mathsf{M}_{m\times n}(F)$, defined by $\Phi(\mathsf{T}) = [\mathsf{T}]_\beta^\gamma$ for $\mathsf{T} \in \mathcal{L}(\mathsf{V}, \mathsf{W})$, is an isomorphism.*

Proof. By Theorem 2.8 (p. 82), Φ is linear. Hence we must show that Φ is one-to-one and onto. This is accomplished if we show that for every $m \times n$ matrix A, there exists a unique linear transformation $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ such that $\Phi(\mathsf{T}) = A$. Let $\beta = \{v_1, v_2, \ldots, v_n\}$, $\gamma = \{w_1, w_2, \ldots, w_m\}$, and let A be a given $m \times n$ matrix. By Theorem 2.6 (p. 72), there exists a unique linear transformation $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ such that

$$\mathsf{T}(v_j) = \sum_{i=1}^{m} A_{ij} w_i \quad \text{for } 1 \le j \le n.$$

But this means that $[\mathsf{T}]_\beta^\gamma = A$, or $\Phi(\mathsf{T}) = A$. Thus Φ is an isomorphism. ∎

Corollary. *Let* V *and* W *be finite-dimensional vector spaces of dimensions* n *and* m, *respectively. Then* $\mathcal{L}(\mathsf{V}, \mathsf{W})$ *is finite-dimensional of dimension* mn.

Proof. The proof follows from Theorems 2.20 and 2.19 and the fact that $\dim(\mathsf{M}_{m\times n}(F)) = mn$. ∎

We conclude this section with a result that allows us to see more clearly the relationship between linear transformations defined on abstract finite-dimensional vector spaces and linear transformations from F^n to F^m.

We begin by naming the transformation $x \to [x]_\beta$ introduced in Section 2.2.

Definition. *Let* β *be an ordered basis for an* n*-dimensional vector space* V *over the field* F. *The* ***standard representation of*** V ***with respect to*** β *is the function* $\phi_\beta\colon \mathsf{V} \to \mathsf{F}^n$ *defined by* $\phi_\beta(x) = [x]_\beta$ *for each* $x \in \mathsf{V}$.

Example 6

Let $\beta = \{(1,0),(0,1)\}$ and $\gamma = \{(1,2),(3,4)\}$. It is easily observed that β and γ are ordered bases for R^2. For $x = (1,-2)$, we have

$$\phi_\beta(x) = [x]_\beta = \begin{pmatrix} 1 \\ -2 \end{pmatrix} \quad \text{and} \quad \phi_\gamma(x) = [x]_\gamma = \begin{pmatrix} -5 \\ 2 \end{pmatrix}. \quad \blacklozenge$$

We observed earlier that ϕ_β is a linear transformation. The next theorem tells us much more.

Theorem 2.21. *For any finite-dimensional vector space* V *with ordered basis* β, ϕ_β *is an isomorphism.*

Proof. Exercise. ∎

This theorem provides us with an alternate proof that an n-dimensional vector space is isomorphic to F^n (see the corollary to Theorem 2.19).

$$
\begin{tikzcd}
\mathsf{V} \arrow[rr, "\mathsf{T}"] \arrow[d, "\phi_\beta"'] \arrow[drr, phantom, "(1)\quad(2)"] & & \mathsf{W} \arrow[d, "\phi_\gamma"] \\
\mathsf{F}^n \arrow[rr, "\mathsf{L}_A"'] & & \mathsf{F}^m
\end{tikzcd}
$$

Figure 2.2

Let V and W be vector spaces of dimension n and m, respectively, and let $\mathsf{T}: \mathsf{V} \to \mathsf{W}$ be a linear transformation. Define $A = [\mathsf{T}]_\beta^\gamma$, where β and γ are arbitrary ordered bases of V and W, respectively. We are now able to use ϕ_β and ϕ_γ to study the relationship between the linear transformations T and $\mathsf{L}_A: \mathsf{F}^n \to \mathsf{F}^m$.

Let us first consider Figure 2.2. Notice that there are two composites of linear transformations that map V into F^m:

1. Map V into F^n with ϕ_β and follow this transformation with L_A; this yields the composite $\mathsf{L}_A\phi_\beta$.
2. Map V into W with T and follow it by ϕ_γ to obtain the composite $\phi_\gamma\mathsf{T}$.

These two composites are depicted by the dashed arrows in the diagram. By a simple reformulation of Theorem 2.14 (p. 91), we may conclude that

$$\mathsf{L}_A\phi_\beta = \phi_\gamma\mathsf{T};$$

that is, the diagram "commutes." Heuristically, this relationship indicates that after V and W are identified with F^n and F^m via ϕ_β and ϕ_γ, respectively, we may "identify" T with L_A. This diagram allows us to transfer operations on abstract vector spaces to ones on F^n and F^m.

Example 7

Recall the linear transformation $\mathsf{T}: \mathsf{P}_3(R) \to \mathsf{P}_2(R)$ defined in Example 4 of Section 2.2 ($\mathsf{T}(f(x)) = f'(x)$). Let β and γ be the standard ordered bases for $\mathsf{P}_3(R)$ and $\mathsf{P}_2(R)$, respectively, and let $\phi_\beta: \mathsf{P}_3(R) \to \mathsf{R}^4$ and $\phi_\gamma: \mathsf{P}_2(R) \to \mathsf{R}^3$ be the corresponding standard representations of $\mathsf{P}_3(R)$ and $\mathsf{P}_2(R)$. If $A = [\mathsf{T}]_\beta^\gamma$, then

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Consider the polynomial $p(x) = 2+x-3x^2+5x^3$. We show that $\mathsf{L}_A\phi_\beta(p(x)) = \phi_\gamma \mathsf{T}(p(x))$. Now

$$\mathsf{L}_A\phi_\beta(p(x)) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ -3 \\ 5 \end{pmatrix} = \begin{pmatrix} 1 \\ -6 \\ 15 \end{pmatrix}.$$

But since $\mathsf{T}(p(x)) = p'(x) = 1 - 6x + 15x^2$, we have

$$\phi_\gamma \mathsf{T}(p(x)) = \begin{pmatrix} 1 \\ -6 \\ 15 \end{pmatrix}.$$

So $\mathsf{L}_A\phi_\beta(p(x)) = \phi_\gamma \mathsf{T}(p(x))$. ♦

Try repeating Example 7 with different polynomials $p(x)$.

EXERCISES

1. Label the following statements as true or false. In each part, V and W are vector spaces with ordered (finite) bases α and β, respectively, $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ is linear, and A and B are matrices.

- **(a)** $([\mathsf{T}]_\alpha^\beta)^{-1} = [\mathsf{T}^{-1}]_\alpha^\beta$.
- **(b)** T is invertible if and only if T is one-to-one and onto.
- **(c)** $\mathsf{T} = \mathsf{L}_A$, where $A = [\mathsf{T}]_\alpha^\beta$.
- **(d)** $\mathsf{M}_{2\times 3}(F)$ is isomorphic to F^5.
- **(e)** $\mathsf{P}_n(F)$ is isomorphic to $\mathsf{P}_m(F)$ if and only if $n = m$.
- **(f)** $AB = I$ implies that A and B are invertible.
- **(g)** If A is invertible, then $(A^{-1})^{-1} = A$.
- **(h)** A is invertible if and only if L_A is invertible.
- **(i)** A must be square in order to possess an inverse.

2. For each of the following linear transformations T, determine whether T is invertible and justify your answer.

- **(a)** $\mathsf{T}\colon \mathsf{R}^2 \to \mathsf{R}^3$ defined by $\mathsf{T}(a_1, a_2) = (a_1 - 2a_2, a_2, 3a_1 + 4a_2)$.
- **(b)** $\mathsf{T}\colon \mathsf{R}^2 \to \mathsf{R}^3$ defined by $\mathsf{T}(a_1, a_2) = (3a_1 - a_2, a_2, 4a_1)$.
- **(c)** $\mathsf{T}\colon \mathsf{R}^3 \to \mathsf{R}^3$ defined by $\mathsf{T}(a_1, a_2, a_3) = (3a_1 - 2a_3, a_2, 3a_1 + 4a_2)$.
- **(d)** $\mathsf{T}\colon \mathsf{P}_3(R) \to \mathsf{P}_2(R)$ defined by $\mathsf{T}(p(x)) = p'(x)$.
- **(e)** $\mathsf{T}\colon \mathsf{M}_{2\times 2}(R) \to \mathsf{P}_2(R)$ defined by $\mathsf{T}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = a + 2bx + (c+d)x^2$.
- **(f)** $\mathsf{T}\colon \mathsf{M}_{2\times 2}(R) \to \mathsf{M}_{2\times 2}(R)$ defined by $\mathsf{T}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+b & a \\ c & c+d \end{pmatrix}$.

3. Which of the following pairs of vector spaces are isomorphic? Justify your answers.

(a) F^3 and $\mathsf{P}_3(F)$.
(b) F^4 and $\mathsf{P}_3(F)$.
(c) $\mathsf{M}_{2\times 2}(R)$ and $\mathsf{P}_3(R)$.
(d) $\mathsf{V} = \{A \in \mathsf{M}_{2\times 2}(R)\colon \operatorname{tr}(A) = 0\}$ and R^4.

4.† Let A and B be $n \times n$ invertible matrices. Prove that AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$.

5.† Let A be invertible. Prove that A^t is invertible and $(A^t)^{-1} = (A^{-1})^t$.

6. Prove that if A is invertible and $AB = O$, then $B = O$.

7. Let A be an $n \times n$ matrix.

(a) Suppose that $A^2 = O$. Prove that A is not invertible.
(b) Suppose that $AB = O$ for some nonzero $n \times n$ matrix B. Could A be invertible? Explain.

8. Prove Corollaries 1 and 2 of Theorem 2.18.

9. Let A and B be $n \times n$ matrices such that AB is invertible. Prove that A and B are invertible. Give an example to show that arbitrary matrices A and B need not be invertible if AB is invertible.

10.† Let A and B be $n \times n$ matrices such that $AB = I_n$.

(a) Use Exercise 9 to conclude that A and B are invertible.
(b) Prove $A = B^{-1}$ (and hence $B = A^{-1}$). (We are, in effect, saying that for square matrices, a "one-sided" inverse is a "two-sided" inverse.)
(c) State and prove analogous results for linear transformations defined on finite-dimensional vector spaces.

11. Verify that the transformation in Example 5 is one-to-one.

12. Prove Theorem 2.21.

13. Let $\sim$ mean "is isomorphic to." Prove that $\sim$ is an equivalence relation on the class of vector spaces over F.

14. Let

$$\mathsf{V} = \left\{ \begin{pmatrix} a & a+b \\ 0 & c \end{pmatrix} : a, b, c \in F \right\}.$$

Construct an isomorphism from V to F^3.

15. Let V and W be n-dimensional vector spaces, and let $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ be a linear transformation. Suppose that β is a basis for V. Prove that T is an isomorphism if and only if $\mathsf{T}(\beta)$ is a basis for W.

16. Let B be an $n \times n$ invertible matrix. Define $\Phi\colon \mathsf{M}_{n\times n}(F) \to \mathsf{M}_{n\times n}(F)$ by $\Phi(A) = B^{-1}AB$. Prove that Φ is an isomorphism.

17.† Let V and W be finite-dimensional vector spaces and $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ be an isomorphism. Let V_0 be a subspace of V.

(a) Prove that $\mathsf{T}(\mathsf{V}_0)$ is a subspace of W.
(b) Prove that $\dim(\mathsf{V}_0) = \dim(\mathsf{T}(\mathsf{V}_0))$.

18. Repeat Example 7 with the polynomial $p(x) = 1 + x + 2x^2 + x^3$.

19. In Example 5 of Section 2.1, the mapping $\mathsf{T}\colon \mathsf{M}_{2\times 2}(R) \to \mathsf{M}_{2\times 2}(R)$ defined by $\mathsf{T}(M) = M^t$ for each $M \in \mathsf{M}_{2\times 2}(R)$ is a linear transformation. Let $\beta = \{E^{11}, E^{12}, E^{21}, E^{22}\}$, which is a basis for $\mathsf{M}_{2\times 2}(R)$, as noted in Example 3 of Section 1.6.

(a) Compute $[\mathsf{T}]_\beta$.
(b) Verify that $\mathsf{L}_A\phi_\beta(M) = \phi_\beta\mathsf{T}(M)$ for $A = [\mathsf{T}]_\beta$ and

$$M = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

20.† Let $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ be a linear transformation from an n-dimensional vector space V to an m-dimensional vector space W. Let β and γ be ordered bases for V and W, respectively. Prove that $\operatorname{rank}(\mathsf{T}) = \operatorname{rank}(\mathsf{L}_A)$ and that $\operatorname{nullity}(\mathsf{T}) = \operatorname{nullity}(\mathsf{L}_A)$, where $A = [\mathsf{T}]_\beta^\gamma$. *Hint:* Apply Exercise 17 to Figure 2.2.

21. Let V and W be finite-dimensional vector spaces with ordered bases $\beta = \{v_1, v_2, \ldots, v_n\}$ and $\gamma = \{w_1, w_2, \ldots, w_m\}$, respectively. By Theorem 2.6 (p. 72), there exist linear transformations $\mathsf{T}_{ij}\colon \mathsf{V} \to \mathsf{W}$ such that

$$\mathsf{T}_{ij}(v_k) = \begin{cases} w_i & \text{if } k = j \\ 0 & \text{if } k \neq j. \end{cases}$$

First prove that $\{\mathsf{T}_{ij}\colon 1 \le i \le m,\ 1 \le j \le n\}$ is a basis for $\mathcal{L}(\mathsf{V},\mathsf{W})$. Then let M^{ij} be the $m \times n$ matrix with 1 in the ith row and jth column and 0 elsewhere, and prove that $[\mathsf{T}_{ij}]_\beta^\gamma = M^{ij}$. Again by Theorem 2.6, there exists a linear transformation $\Phi\colon \mathcal{L}(\mathsf{V},\mathsf{W}) \to \mathsf{M}_{m\times n}(F)$ such that $\Phi(\mathsf{T}_{ij}) = M^{ij}$. Prove that Φ is an isomorphism.

22. Let $c_0, c_1, \ldots, c_n$ be distinct scalars from an infinite field F. Define $\mathsf{T}\colon \mathsf{P}_n(F) \to \mathsf{F}^{n+1}$ by $\mathsf{T}(f) = (f(c_0), f(c_1), \ldots, f(c_n))$. Prove that T is an isomorphism. *Hint:* Use the Lagrange polynomials associated with $c_0, c_1, \ldots, c_n$.

23. Let V denote the vector space defined in Example 5 of Section 1.2, and let $\mathsf{W} = \mathsf{P}(F)$. Define

$$\mathsf{T}\colon \mathsf{V} \to \mathsf{W} \quad \text{by} \quad \mathsf{T}(\sigma) = \sum_{i=0}^{n} \sigma(i)x^i,$$

where n is the largest integer such that $\sigma(n) \neq 0$. Prove that T is an isomorphism.

The following exercise requires familiarity with the concept of *quotient space* defined in Exercise 31 of Section 1.3 and with Exercise 40 of Section 2.1.

24. Let $\mathsf{T}\colon \mathsf{V} \to \mathsf{Z}$ be a linear transformation of a vector space V onto a vector space Z. Define the mapping

$$\overline{\mathsf{T}}\colon \mathsf{V}/\mathsf{N}(\mathsf{T}) \to \mathsf{Z} \quad \text{by} \quad \overline{\mathsf{T}}(v + \mathsf{N}(\mathsf{T})) = \mathsf{T}(v)$$

for any coset $v + \mathsf{N}(\mathsf{T})$ in $\mathsf{V}/\mathsf{N}(\mathsf{T})$.

(a) Prove that $\overline{\mathsf{T}}$ is well-defined; that is, prove that if $v + \mathsf{N}(\mathsf{T}) = v' + \mathsf{N}(\mathsf{T})$, then $\mathsf{T}(v) = \mathsf{T}(v')$.
(b) Prove that $\overline{\mathsf{T}}$ is linear.
(c) Prove that $\overline{\mathsf{T}}$ is an isomorphism.
(d) Prove that the diagram shown in Figure 2.3 commutes; that is, prove that $\mathsf{T} = \overline{\mathsf{T}}\eta$.

$$\begin{tikzcd}
\mathsf{V} \arrow[rr, dashed, "\mathsf{T}"] \arrow[dr, dashed, "\eta"'] & & \mathsf{Z} \\
& \mathsf{V}/\mathsf{N}(\mathsf{T}) \arrow[ur, dashed, "\overline{\mathsf{T}}"'] &
\end{tikzcd}$$

Figure 2.3

25. Let V be a nonzero vector space over a field F, and suppose that S is a basis for V. (By the corollary to Theorem 1.13 (p. 60) in Section 1.7, every vector space has a basis). Let $\mathcal{C}(S, F)$ denote the vector space of all functions $f \in \mathcal{F}(S, F)$ such that $f(s) = 0$ for all but a finite number of vectors in S. (See Exercise 14 of Section 1.3.) Let $\Psi\colon \mathcal{C}(S,F) \to \mathsf{V}$ be defined by $\Psi(f) = \mathit{0}$ if f is the zero function, and

$$\Psi(f) = \sum_{s\in S, f(s)\neq 0} f(s)s,$$

otherwise. Prove that Ψ is an isomorphism. Thus every nonzero vector space can be viewed as a space of functions.

2.5 THE CHANGE OF COORDINATE MATRIX

In many areas of mathematics, a change of variable is used to simplify the appearance of an expression. For example, in calculus an antiderivative of $2xe^{x^2}$ can be found by making the change of variable $u = x^2$. The resulting expression is of such a simple form that an antiderivative is easily recognized:

$$\int 2xe^{x^2}\,dx = \int e^u\,du = e^u + c = e^{x^2} + c.$$

Similarly, in geometry the change of variable

$$x = \frac{2}{\sqrt{5}}x' - \frac{1}{\sqrt{5}}y'$$

$$y = \frac{1}{\sqrt{5}}x' + \frac{2}{\sqrt{5}}y'$$

can be used to transform the equation $2x^2 - 4xy + 5y^2 = 1$ into the simpler equation $(x')^2+6(y')^2 = 1$, in which form it is easily seen to be the equation of an ellipse. (See Figure 2.4.) We see how this change of variable is determined in Section 6.5. Geometrically, the change of variable

$$\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} x' \\ y' \end{pmatrix}$$

is a change in the way that the position of a point P in the plane is described. This is done by introducing a new frame of reference, an $x'y'$-coordinate system with coordinate axes rotated from the original xy-coordinate axes. In this case, the new coordinate axes are chosen to lie in the direction of the axes of the ellipse. The unit vectors along the x'-axis and the y'-axis form an ordered basis

$$\beta' = \left\{\frac{1}{\sqrt{5}}\begin{pmatrix} 2 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{5}}\begin{pmatrix} -1 \\ 2 \end{pmatrix}\right\}$$

for R^2, and the change of variable is actually a change from $[P]_\beta = \begin{pmatrix} x \\ y \end{pmatrix}$, the coordinate vector of P relative to the standard ordered basis $\beta = \{e_1, e_2\}$, to

$[P]_{\beta'} = \begin{pmatrix} x' \\ y' \end{pmatrix}$, the coordinate vector of P relative to the new rotated basis β'.

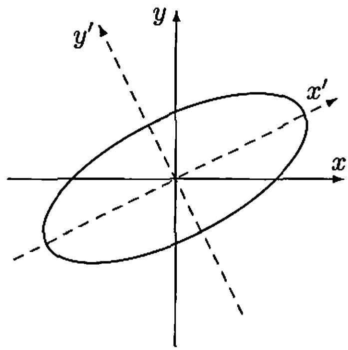


Figure 2.4

A natural question arises: How can a coordinate vector relative to one basis be changed into a coordinate vector relative to the other? Notice that the system of equations relating the new and old coordinates can be represented by the matrix equation

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}.$$

Notice also that the matrix

$$Q = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$$

equals $[\mathsf{I}]_{\beta'}^{\beta}$, where I denotes the identity transformation on R^2. Thus $[v]_\beta = Q[v]_{\beta'}$ for all $v \in \mathsf{R}^2$. A similar result is true in general.

Theorem 2.22. *Let β and β' be two ordered bases for a finite-dimensional vector space V, and let $Q = [\mathsf{I_V}]_{\beta'}^{\beta}$. Then*
(a) *Q is invertible.*
(b) *For any $v \in \mathsf{V}$, $[v]_\beta = Q[v]_{\beta'}$.*

Proof. (a) Since $\mathsf{I_V}$ is invertible, Q is invertible by Theorem 2.18 (p. 101).
(b) For any $v \in \mathsf{V}$,

$$[v]_\beta = [\mathsf{I_V}(v)]_\beta = [\mathsf{I_V}]_{\beta'}^{\beta}[v]_{\beta'} = Q[v]_{\beta'}$$

by Theorem 2.14 (p. 91). ∎

The matrix $Q = [\mathsf{I_V}]_{\beta'}^{\beta}$ defined in Theorem 2.22 is called a **change of coordinate matrix.** Because of part (b) of the theorem, we say that Q **changes β'-coordinates into β-coordinates.** Observe that if $\beta = \{x_1, x_2, \ldots, x_n\}$ and $\beta' = \{x'_1, x'_2, \ldots, x'_n\}$, then

$$x'_j = \sum_{i=1}^{n} Q_{ij} x_i$$

for $j = 1, 2, \ldots, n$; that is, the jth column of Q is $[x'_j]_\beta$.

Notice that if Q changes β'-coordinates into β-coordinates, then Q^{-1} changes β-coordinates into β'-coordinates. (See Exercise 11.)

Example 1

In R^2, let $\beta = \{(1,1), (1,-1)\}$ and $\beta' = \{(2,4), (3,1)\}$. Since

$$(2,4) = 3(1,1) - 1(1,-1) \qquad \text{and} \qquad (3,1) = 2(1,1) + 1(1,-1),$$

the matrix that changes β'-coordinates into β-coordinates is

$$Q = \begin{pmatrix} 3 & 2 \\ -1 & 1 \end{pmatrix}.$$

Thus, for instance,

$$[(2,4)]_\beta = Q[(2,4)]_{\beta'} = Q\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}. \quad \blacklozenge$$

For the remainder of this section, we consider only linear transformations that map a vector space V into itself. Such a linear transformation is called a **linear operator** on V. Suppose now that T is a linear operator on a finite-dimensional vector space V and that β and β' are ordered bases for V. Then V can be represented by the matrices $[\mathsf{T}]_\beta$ and $[\mathsf{T}]_{\beta'}$. What is the relationship between these matrices? The next theorem provides a simple answer using a change of coordinate matrix.

Theorem 2.23. *Let T be a linear operator on a finite-dimensional vector space V, and let β and β' be ordered bases for V. Suppose that Q is the change of coordinate matrix that changes β'-coordinates into β-coordinates. Then*

$$[\mathsf{T}]_{\beta'} = Q^{-1}[\mathsf{T}]_\beta Q.$$

Proof. Let I be the identity transformation on V. Then $\mathsf{T} = \mathsf{IT} = \mathsf{TI}$; hence, by Theorem 2.11 (p. 88),

$$Q[\mathsf{T}]_{\beta'} = [\mathsf{I}]_{\beta'}^{\beta}[\mathsf{T}]_{\beta'}^{\beta'} = [\mathsf{IT}]_{\beta'}^{\beta} = [\mathsf{TI}]_{\beta'}^{\beta} = [\mathsf{T}]_{\beta}^{\beta}[\mathsf{I}]_{\beta'}^{\beta} = [\mathsf{T}]_\beta Q.$$

Therefore $[\mathsf{T}]_{\beta'} = Q^{-1}[\mathsf{T}]_\beta Q$. ∎

Example 2

Let T be the linear operator on R^2 defined by

$$\mathsf{T}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 3a - b \\ a + 3b \end{pmatrix},$$

and let β and β' be the ordered bases in Example 1. The reader should verify that

$$[\mathsf{T}]_\beta = \begin{pmatrix} 3 & 1 \\ -1 & 3 \end{pmatrix}.$$

In Example 1, we saw that the change of coordinate matrix that changes β'-coordinates into β-coordinates is

$$Q = \begin{pmatrix} 3 & 2 \\ -1 & 1 \end{pmatrix},$$

and it is easily verified that

$$Q^{-1} = \frac{1}{5}\begin{pmatrix} 1 & -2 \\ 1 & 3 \end{pmatrix}.$$

Hence, by Theorem 2.23,

$$[\mathsf{T}]_{\beta'} = Q^{-1}[\mathsf{T}]_\beta Q = \begin{pmatrix} 4 & 1 \\ -2 & 2 \end{pmatrix}.$$

To show that this is the correct matrix, we can verify that the image under T of each vector of β' is the linear combination of the vectors of β' with the entries of the corresponding column as its coefficients. For example, the image of the second vector in β' is

$$\mathsf{T}\begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 8 \\ 6 \end{pmatrix} = 1\begin{pmatrix} 2 \\ 4 \end{pmatrix} + 2\begin{pmatrix} 3 \\ 1 \end{pmatrix}.$$

Notice that the coefficients of the linear combination are the entries of the second column of $[\mathsf{T}]_{\beta'}$. ♦

It is often useful to apply Theorem 2.23 to compute $[\mathsf{T}]_\beta$, as the next example shows.

Example 3

Recall the reflection about the x-axis in Example 3 of Section 2.1. The rule $(x, y) \to (x, -y)$ is easy to obtain. We now derive the less obvious rule for the reflection T about the line $y = 2x$. (See Figure 2.5.) We wish to find an expression for $\mathsf{T}(a, b)$ for any (a, b) in R^2. Since T is linear, it is completely

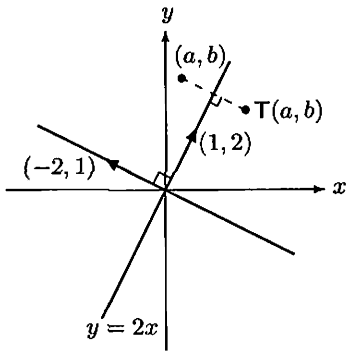


Figure 2.5

determined by its values on a basis for R^2. Clearly, $\mathsf{T}(1,2) = (1,2)$ and $\mathsf{T}(-2,1) = -(-2,1) = (2,-1)$. Therefore if we let

$$\beta' = \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \end{pmatrix} \right\},$$

then β' is an ordered basis for R^2 and

$$[\mathsf{T}]_{\beta'} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Let β be the standard ordered basis for R^2, and let Q be the matrix that changes β'-coordinates into β-coordinates. Then

$$Q = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}$$

and $Q^{-1}[\mathsf{T}]_\beta Q = [\mathsf{T}]_{\beta'}$. We can solve this equation for $[\mathsf{T}]_\beta$ to obtain that $[\mathsf{T}]_\beta = Q[\mathsf{T}]_{\beta'}Q^{-1}$. Because

$$Q^{-1} = \frac{1}{5}\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix},$$

the reader can verify that

$$[\mathsf{T}]_\beta = \frac{1}{5}\begin{pmatrix} -3 & 4 \\ 4 & 3 \end{pmatrix}.$$

Since β is the standard ordered basis, it follows that T is left-multiplication by $[\mathsf{T}]_\beta$. Thus for any (a,b) in R^2, we have

$$\mathsf{T}\begin{pmatrix} a \\ b \end{pmatrix} = \frac{1}{5}\begin{pmatrix} -3 & 4 \\ 4 & 3 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \frac{1}{5}\begin{pmatrix} -3a + 4b \\ 4a + 3b \end{pmatrix}. \quad \blacklozenge$$

A useful special case of Theorem 2.23 is contained in the next corollary, whose proof is left as an exercise.

Corollary. *Let $A \in \mathsf{M}_{n\times n}(F)$, and let γ be an ordered basis for F^n. Then $[\mathsf{L}_A]_\gamma = Q^{-1}AQ$, where Q is the $n \times n$ matrix whose jth column is the jth vector of γ.*

Example 4

Let

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 3 \\ 0 & -1 & 0 \end{pmatrix},$$

and let

$$\gamma = \left\{ \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\},$$

which is an ordered basis for R^3. Let Q be the 3×3 matrix whose jth column is the jth vector of γ. Then

$$Q = \begin{pmatrix} -1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad Q^{-1} = \begin{pmatrix} -1 & 2 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

So by the preceding corollary,

$$[\mathsf{L}_A]_\gamma = Q^{-1}AQ = \begin{pmatrix} 0 & 2 & 8 \\ -1 & 4 & 6 \\ 0 & -1 & -1 \end{pmatrix}. \quad \blacklozenge$$

The relationship between the matrices $[\mathsf{T}]_{\beta'}$ and $[\mathsf{T}]_\beta$ in Theorem 2.23 will be the subject of further study in Chapters 5, 6, and 7. At this time, however, we introduce the name for this relationship.

Definition. *Let A and B be matrices in $\mathsf{M}_{n\times n}(F)$. We say that B is* **similar** *to A if there exists an invertible matrix Q such that $B = Q^{-1}AQ$.*

Observe that the relation of similarity is an equivalence relation (see Exercise 9). So we need only say that A and B are similar.

Notice also that in this terminology Theorem 2.23 can be stated as follows: If T is a linear operator on a finite-dimensional vector space V, and if β and β' are any ordered bases for V, then $[\mathsf{T}]_{\beta'}$ is similar to $[\mathsf{T}]_\beta$.

Theorem 2.23 can be generalized to allow $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$, where V is distinct from W. In this case, we can change bases in V as well as in W (see Exercise 8).

EXERCISES

1. Label the following statements as true or false.

(a) Suppose that $\beta = \{x_1, x_2, \ldots, x_n\}$ and $\beta' = \{x'_1, x'_2, \ldots, x'_n\}$ are ordered bases for a vector space and Q is the change of coordinate matrix that changes β'-coordinates into β-coordinates. Then the jth column of Q is $[x_j]_{\beta'}$.

(b) Every change of coordinate matrix is invertible.

(c) Let T be a linear operator on a finite-dimensional vector space V, let β and β' be ordered bases for V, and let Q be the change of coordinate matrix that changes β'-coordinates into β-coordinates. Then $[\mathsf{T}]_\beta = Q[\mathsf{T}]_{\beta'}Q^{-1}$.

(d) The matrices $A, B \in \mathsf{M}_{n\times n}(F)$ are called similar if $B = Q^tAQ$ for some $Q \in \mathsf{M}_{n\times n}(F)$.

(e) Let T be a linear operator on a finite-dimensional vector space V. Then for any ordered bases β and γ for V, $[\mathsf{T}]_\beta$ is similar to $[\mathsf{T}]_\gamma$.

2. For each of the following pairs of ordered bases β and β' for R^2, find the change of coordinate matrix that changes β'-coordinates into β-coordinates.

(a) $\beta = \{e_1, e_2\}$ and $\beta' = \{(a_1, a_2), (b_1, b_2)\}$

(b) $\beta = \{(-1, 3), (2, -1)\}$ and $\beta' = \{(0, 10), (5, 0)\}$

(c) $\beta = \{(2, 5), (-1, -3)\}$ and $\beta' = \{e_1, e_2\}$

(d) $\beta = \{(-4, 3), (2, -1)\}$ and $\beta' = \{(2, 1), (-4, 1)\}$

3. For each of the following pairs of ordered bases β and β' for $\mathsf{P}_2(R)$, find the change of coordinate matrix that changes β'-coordinates into β-coordinates.

(a) $\beta = \{x^2, x, 1\}$ and
$\beta' = \{a_2x^2 + a_1x + a_0, b_2x^2 + b_1x + b_0, c_2x^2 + c_1x + c_0\}$

(b) $\beta = \{1, x, x^2\}$ and
$\beta' = \{a_2x^2 + a_1x + a_0, b_2x^2 + b_1x + b_0, c_2x^2 + c_1x + c_0\}$

(c) $\beta = \{2x^2 - x, 3x^2 + 1, x^2\}$ and $\beta' = \{1, x, x^2\}$

(d) $\beta = \{x^2 - x + 1, x + 1, x^2 + 1\}$ and
$\beta' = \{x^2 + x + 4, 4x^2 - 3x + 2, 2x^2 + 3\}$

(e) $\beta = \{x^2 - x, x^2 + 1, x - 1\}$ and
$\beta' = \{5x^2 - 2x - 3, -2x^2 + 5x + 5, 2x^2 - x - 3\}$

(f) $\beta = \{2x^2 - x + 1, x^2 + 3x - 2, -x^2 + 2x + 1\}$ and
$\beta' = \{9x - 9, x^2 + 21x - 2, 3x^2 + 5x + 2\}$

4. Let T be the linear operator on R^2 defined by

$$\mathsf{T}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 2a + b \\ a - 3b \end{pmatrix},$$

let β be the standard ordered basis for R^2, and let

$$\beta' = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\}.$$

Use Theorem 2.23 and the fact that

$$\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$$

to find $[\mathsf{T}]_{\beta'}$.

5. Let T be the linear operator on $\mathsf{P}_1(R)$ defined by $\mathsf{T}(p(x)) = p'(x)$, the derivative of $p(x)$. Let $\beta = \{1, x\}$ and $\beta' = \{1 + x, 1 - x\}$. Use Theorem 2.23 and the fact that

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}$$

to find $[\mathsf{T}]_{\beta'}$.

6. For each matrix A and ordered basis β, find $[\mathsf{L}_A]_\beta$. Also, find an invertible matrix Q such that $[\mathsf{L}_A]_\beta = Q^{-1}AQ$.

(a) $A = \begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix}$ and $\beta = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\}$

(b) $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ and $\beta = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$

(c) $A = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$ and $\beta = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \right\}$

(d) $A = \begin{pmatrix} 13 & 1 & 4 \\ 1 & 13 & 4 \\ 4 & 4 & 10 \end{pmatrix}$ and $\beta = \left\{ \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$

7. In R^2, let L be the line $y = mx$, where $m \neq 0$. Find an expression for $\mathsf{T}(x, y)$, where

(a) T is the reflection of R^2 about L.

(b) T is the projection on L along the line perpendicular to L. (See the definition of *projection* in the exercises of Section 2.1.)

8. Prove the following generalization of Theorem 2.23. Let $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ be a linear transformation from a finite-dimensional vector space V to a finite-dimensional vector space W. Let β and β' be ordered bases for

V, and let γ and γ' be ordered bases for W. Then $[\mathsf{T}]_{\beta'}^{\gamma'} = P^{-1}[\mathsf{T}]_{\beta}^{\gamma}Q$, where Q is the matrix that changes β'-coordinates into β-coordinates and P is the matrix that changes γ'-coordinates into γ-coordinates.

9. Prove that "is similar to" is an equivalence relation on $\mathsf{M}_{n\times n}(F)$.

10. Prove that if A and B are similar $n \times n$ matrices, then $\operatorname{tr}(A) = \operatorname{tr}(B)$. *Hint:* Use Exercise 13 of Section 2.3.

11. Let V be a finite-dimensional vector space with ordered bases α, β, and γ.

(a) Prove that if Q and R are the change of coordinate matrices that change α-coordinates into β-coordinates and β-coordinates into γ-coordinates, respectively, then RQ is the change of coordinate matrix that changes α-coordinates into γ-coordinates.

(b) Prove that if Q changes α-coordinates into β-coordinates, then Q^{-1} changes β-coordinates into α-coordinates.

12. Prove the corollary to Theorem 2.23.

13.† Let V be a finite-dimensional vector space over a field F, and let $\beta = \{x_1, x_2, \ldots, x_n\}$ be an ordered basis for V. Let Q be an $n \times n$ invertible matrix with entries from F. Define

$$x'_j = \sum_{i=1}^{n} Q_{ij}x_i \quad \text{for } 1 \le j \le n,$$

and set $\beta' = \{x'_1, x'_2, \ldots, x'_n\}$. Prove that β' is a basis for V and hence that Q is the change of coordinate matrix changing β'-coordinates into β-coordinates.

14. Prove the converse of Exercise 8: If A and B are each $m \times n$ matrices with entries from a field F, and if there exist invertible $m \times m$ and $n \times n$ matrices P and Q, respectively, such that $B = P^{-1}AQ$, then there exist an n-dimensional vector space V and an m-dimensional vector space W (both over F), ordered bases β and β' for V and γ and γ' for W, and a linear transformation $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ such that

$$A = [\mathsf{T}]_{\beta}^{\gamma} \quad \text{and} \quad B = [\mathsf{T}]_{\beta'}^{\gamma'}.$$

Hints: Let $\mathsf{V} = \mathsf{F}^n$, $\mathsf{W} = \mathsf{F}^m$, $\mathsf{T} = \mathsf{L}_A$, and β and γ be the standard ordered bases for F^n and F^m, respectively. Now apply the results of Exercise 13 to obtain ordered bases β' and γ' from β and γ via Q and P, respectively.

2.6* DUAL SPACES

In this section, we are concerned exclusively with linear transformations from a vector space V into its field of scalars F, which is itself a vector space of dimension 1 over F. Such a linear transformation is called a **linear functional** on V. We generally use the letters $\mathsf{f}, \mathsf{g}, \mathsf{h}, \ldots$ to denote linear functionals. As we see in Example 1, the definite integral provides us with one of the most important examples of a linear functional in mathematics.

Example 1

Let V be the vector space of continuous real-valued functions on the interval $[0, 2\pi]$. Fix a function $g \in \mathsf{V}$. The function $\mathsf{h}\colon \mathsf{V} \to R$ defined by

$$\mathsf{h}(x) = \frac{1}{2\pi}\int_0^{2\pi} x(t)g(t)\,dt$$

is a linear functional on V. In the cases that $g(t)$ equals $\sin nt$ or $\cos nt$, $h(x)$ is often called the **nth Fourier coefficient of x.** ♦

Example 2

Let $\mathsf{V} = \mathsf{M}_{n\times n}(F)$, and define $\mathsf{f}\colon \mathsf{V} \to F$ by $\mathsf{f}(A) = \mathrm{tr}(A)$, the trace of A. By Exercise 6 of Section 1.3, we have that f is a linear functional. ♦

Example 3

Let V be a finite-dimensional vector space, and let $\beta = \{x_1, x_2, \ldots, x_n\}$ be an ordered basis for V. For each $i = 1, 2, \ldots, n$, define $\mathsf{f}_i(x) = a_i$, where

$$[x]_\beta = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

is the coordinate vector of x relative to β. Then f_i is a linear functional on V called the **ith coordinate function with respect to the basis β.** Note that $\mathsf{f}_i(x_j) = \delta_{ij}$, where δ_{ij} is the Kronecker delta. These linear functionals play an important role in the theory of dual spaces (see Theorem 2.24). ♦

Definition. *For a vector space V over F, we define the* **dual space** *of V to be the vector space $\mathcal{L}(\mathsf{V}, F)$, denoted by V^*.*

Thus V^* is the vector space consisting of all linear functionals on V with the operations of addition and scalar multiplication as defined in Section 2.2. Note that if V is finite-dimensional, then by the corollary to Theorem 2.20 (p. 104)

$$\dim(\mathsf{V}^*) = \dim(\mathcal{L}(\mathsf{V}, F)) = \dim(\mathsf{V})\cdot\dim(F) = \dim(\mathsf{V}).$$

Hence by Theorem 2.19 (p. 103), V and V^* are isomorphic. We also define the **double dual** V^{**} of V to be the dual of V^*. In Theorem 2.26, we show, in fact, that there is a natural identification of V and V^{**} in the case that V is finite-dimensional.

Theorem 2.24. *Suppose that V is a finite-dimensional vector space with the ordered basis $\beta = \{x_1, x_2, \ldots, x_n\}$. Let f_i $(1 \le i \le n)$ be the ith coordinate function with respect to β as just defined, and let $\beta^* = \{\mathsf{f}_1, \mathsf{f}_2, \ldots, \mathsf{f}_n\}$. Then β^* is an ordered basis for V^*, and, for any $\mathsf{f} \in \mathsf{V}^*$, we have*

$$\mathsf{f} = \sum_{i=1}^{n} \mathsf{f}(x_i)\mathsf{f}_i.$$

Proof. Let $\mathsf{f} \in \mathsf{V}^*$. Since $\dim(\mathsf{V}^*) = n$, we need only show that

$$\mathsf{f} = \sum_{i=1}^{n} \mathsf{f}(x_i)\mathsf{f}_i,$$

from which it follows that β^* generates V^*, and hence is a basis by Corollary 2(a) to the replacement theorem (p. 47). Let

$$\mathsf{g} = \sum_{i=1}^{n} \mathsf{f}(x_i)\mathsf{f}_i.$$

For $1 \le j \le n$, we have

$$\begin{aligned} \mathsf{g}(x_j) &= \left(\sum_{i=1}^{n} \mathsf{f}(x_i)\mathsf{f}_i\right)(x_j) = \sum_{i=1}^{n} \mathsf{f}(x_i)\mathsf{f}_i(x_j) \\ &= \sum_{i=1}^{n} \mathsf{f}(x_i)\delta_{ij} = \mathsf{f}(x_j). \end{aligned}$$

Therefore $\mathsf{f} = \mathsf{g}$ by the corollary to Theorem 2.6 (p. 72). ∎

Definition. *Using the notation of Theorem 2.24, we call the ordered basis $\beta^* = \{\mathsf{f}_1, \mathsf{f}_2, \ldots, \mathsf{f}_n\}$ of V^* that satisfies $\mathsf{f}_i(x_j) = \delta_{ij}$ $(1 \le i, j \le n)$ the **dual basis** of β.*

Example 4

Let $\beta = \{(2,1), (3,1)\}$ be an ordered basis for R^2. Suppose that the dual basis of β is given by $\beta^* = \{\mathsf{f}_1, \mathsf{f}_2\}$. To explicitly determine a formula for f_1, we need to consider the equations

$$\begin{aligned} 1 &= \mathsf{f}_1(2,1) = \mathsf{f}_1(2e_1 + e_2) = 2\mathsf{f}_1(e_1) + \mathsf{f}_1(e_2) \\ 0 &= \mathsf{f}_1(3,1) = \mathsf{f}_1(3e_1 + e_2) = 3\mathsf{f}_1(e_1) + \mathsf{f}_1(e_2). \end{aligned}$$

Solving these equations, we obtain $\mathsf{f}_1(e_1) = -1$ and $\mathsf{f}_1(e_2) = 3$; that is, $\mathsf{f}_1(x, y) = -x + 3y$. Similarly, it can be shown that $\mathsf{f}_2(x, y) = x - 2y$. ♦

We now assume that V and W are finite-dimensional vector spaces over F with ordered bases β and γ, respectively. In Section 2.4, we proved that there is a one-to-one correspondence between linear transformations $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ and $m \times n$ matrices (over F) via the correspondence $\mathsf{T} \leftrightarrow [\mathsf{T}]_\beta^\gamma$. For a matrix of the form $A = [\mathsf{T}]_\beta^\gamma$, the question arises as to whether or not there exists a linear transformation U associated with T in some natural way such that U may be represented in some basis as A^t. Of course, if $m \neq n$, it would be impossible for U to be a linear transformation from V into W. We now answer this question by applying what we have already learned about dual spaces.

Theorem 2.25. *Let* V *and* W *be finite-dimensional vector spaces over* F *with ordered bases* β *and* γ*, respectively. For any linear transformation* $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$*, the mapping* $\mathsf{T}^t\colon \mathsf{W}^* \to \mathsf{V}^*$ *defined by* $\mathsf{T}^t(\mathsf{g}) = \mathsf{gT}$ *for all* $\mathsf{g} \in \mathsf{W}^*$ *is a linear transformation with the property that* $[\mathsf{T}^t]_{\gamma^*}^{\beta^*} = ([\mathsf{T}]_\beta^\gamma)^t$.

Proof. For $\mathsf{g} \in \mathsf{W}^*$, it is clear that $\mathsf{T}^t(\mathsf{g}) = \mathsf{gT}$ is a linear functional on V and hence is in V^*. Thus T^t maps W^* into V^*. We leave the proof that T^t is linear to the reader.

To complete the proof, let $\beta = \{x_1, x_2, \ldots, x_n\}$ and $\gamma = \{y_1, y_2, \ldots, y_m\}$ with dual bases $\beta^* = \{\mathsf{f}_1, \mathsf{f}_2, \ldots, \mathsf{f}_n\}$ and $\gamma^* = \{\mathsf{g}_1, \mathsf{g}_2, \ldots, \mathsf{g}_m\}$, respectively. For convenience, let $A = [\mathsf{T}]_\beta^\gamma$. To find the jth column of $[\mathsf{T}^t]_{\gamma^*}^{\beta^*}$, we begin by expressing $\mathsf{T}^t(\mathsf{g}_j)$ as a linear combination of the vectors of β^*. By Theorem 2.24, we have

$$\mathsf{T}^t(\mathsf{g}_j) = \mathsf{g}_j\mathsf{T} = \sum_{s=1}^{n} (\mathsf{g}_j\mathsf{T})(x_s)\mathsf{f}_s.$$

So the row i, column j entry of $[\mathsf{T}^t]_{\gamma^*}^{\beta^*}$ is

$$\begin{aligned}(\mathsf{g}_j\mathsf{T})(x_i) = \mathsf{g}_j(\mathsf{T}(x_i)) &= \mathsf{g}_j\left(\sum_{k=1}^{m} A_{ki}y_k\right) \\ &= \sum_{k=1}^{m} A_{ki}\mathsf{g}_j(y_k) = \sum_{k=1}^{m} A_{ki}\delta_{jk} = A_{ji}.\end{aligned}$$

Hence $[\mathsf{T}^t]_{\gamma^*}^{\beta^*} = A^t$. ∎

The linear transformation T^t defined in Theorem 2.25 is called the **transpose** of T. It is clear that T^t is the unique linear transformation U such that $[\mathsf{U}]_{\gamma^*}^{\beta^*} = ([\mathsf{T}]_\beta^\gamma)^t$.

We illustrate Theorem 2.25 with the next example.

Example 5

Define $\mathsf{T}\colon \mathsf{P}_1(R) \to \mathsf{R}^2$ by $\mathsf{T}(p(x)) = (p(0), p(2))$. Let β and γ be the standard ordered bases for $\mathsf{P}_1(R)$ and R^2, respectively. Clearly,

$$[\mathsf{T}]_\beta^\gamma = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}.$$

We compute $[\mathsf{T}^t]_{\gamma^*}^{\beta^*}$ directly from the definition. Let $\beta^* = \{\mathsf{f}_1, \mathsf{f}_2\}$ and $\gamma^* = \{\mathsf{g}_1, \mathsf{g}_2\}$. Suppose that $[\mathsf{T}^t]_{\gamma^*}^{\beta^*} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then $\mathsf{T}^t(\mathsf{g}_1) = a\mathsf{f}_1 + c\mathsf{f}_2$. So

$$\mathsf{T}^t(\mathsf{g}_1)(1) = (a\mathsf{f}_1 + c\mathsf{f}_2)(1) = a\mathsf{f}_1(1) + c\mathsf{f}_2(1) = a(1) + c(0) = a.$$

But also

$$(\mathsf{T}^t(\mathsf{g}_1))(1) = \mathsf{g}_1(\mathsf{T}(1)) = \mathsf{g}_1(1,1) = 1.$$

So $a = 1$. Using similar computations, we obtain that $c = 0$, $b = 1$, and $d = 2$. Hence a direct computation yields

$$[\mathsf{T}^t]_{\gamma^*}^{\beta^*} = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} = \left([\mathsf{T}]_\beta^\gamma\right)^t,$$

as predicted by Theorem 2.25. ♦

We now concern ourselves with demonstrating that any finite-dimensional vector space V can be identified in a natural way with its double dual V^{**}. There is, in fact, an isomorphism between V and V^{**} that does not depend on any choice of bases for the two vector spaces.

For a vector $x \in \mathsf{V}$, we define $\widehat{x}\colon \mathsf{V}^* \to F$ by $\widehat{x}(\mathsf{f}) = \mathsf{f}(x)$ for every $\mathsf{f} \in \mathsf{V}^*$. It is easy to verify that $\widehat{x}$ is a linear functional on V^*, so $\widehat{x} \in \mathsf{V}^{**}$. The correspondence $x \leftrightarrow \widehat{x}$ allows us to define the desired isomorphism between V and V^{**}.

Lemma. *Let V be a finite-dimensional vector space, and let $x \in \mathsf{V}$. If $\widehat{x}(\mathsf{f}) = 0$ for all $\mathsf{f} \in \mathsf{V}^*$, then $x = 0$.*

Proof. Let $x \neq 0$. We show that there exists $\mathsf{f} \in \mathsf{V}^*$ such that $\widehat{x}(\mathsf{f}) \neq 0$. Choose an ordered basis $\beta = \{x_1, x_2, \ldots, x_n\}$ for V such that $x_1 = x$. Let $\{\mathsf{f}_1, \mathsf{f}_2, \ldots, \mathsf{f}_n\}$ be the dual basis of β. Then $\mathsf{f}_1(x_1) = 1 \neq 0$. Let $\mathsf{f} = \mathsf{f}_1$. ∎

Theorem 2.26. *Let V be a finite-dimensional vector space, and define $\psi\colon \mathsf{V} \to \mathsf{V}^{**}$ by $\psi(x) = \widehat{x}$. Then ψ is an isomorphism.*

Proof. (a) ψ is linear: Let $x, y \in \mathsf{V}$ and $c \in F$. For $\mathsf{f} \in \mathsf{V}^*$, we have

$$\psi(cx+y)(\mathsf{f}) = \mathsf{f}(cx+y) = c\mathsf{f}(x) + \mathsf{f}(y) = c\widehat{x}(\mathsf{f}) + \widehat{y}(\mathsf{f})$$
$$= (c\widehat{x} + \widehat{y})(\mathsf{f}).$$

Therefore

$$\psi(cx+y) = c\widehat{x} + \widehat{y} = c\psi(x) + \psi(y).$$

(b) ψ is one-to-one: Suppose that $\psi(x)$ is the zero functional on V^* for some $x \in \mathsf{V}$. Then $\widehat{x}(\mathsf{f}) = 0$ for every $\mathsf{f} \in \mathsf{V}^*$. By the previous lemma, we conclude that $x = \mathit{0}$.

(c) ψ is an isomorphism: This follows from (b) and the fact that $\dim(\mathsf{V}) = \dim(\mathsf{V}^{**})$. ∎

Corollary. *Let* V *be a finite-dimensional vector space with dual space* V^**. Then every ordered basis for* V^* *is the dual basis for some basis for* V.

Proof. Let $\{\mathsf{f}_1, \mathsf{f}_2, \ldots, \mathsf{f}_n\}$ be an ordered basis for V^*. We may combine Theorems 2.24 and 2.26 to conclude that for this basis for V^* there exists a dual basis $\{\widehat{x}_1, \widehat{x}_2, \ldots, \widehat{x}_n\}$ in V^{**}, that is, $\delta_{ij} = \widehat{x}_i(\mathsf{f}_j) = \mathsf{f}_j(x_i)$ for all i and j. Thus $\{\mathsf{f}_1, \mathsf{f}_2, \ldots, \mathsf{f}_n\}$ is the dual basis of $\{x_1, x_2, \ldots, x_n\}$. ∎

Although many of the ideas of this section, (e.g., the existence of a dual space), can be extended to the case where V is not finite-dimensional, only a finite-dimensional vector space is isomorphic to its double dual via the map $x \to \widehat{x}$. In fact, for infinite-dimensional vector spaces, no two of V, V^*, and V^{**} are isomorphic.

EXERCISES

1. Label the following statements as true or false. Assume that all vector spaces are finite-dimensional.
   - **(a)** Every linear transformation is a linear functional.
   - **(b)** A linear functional defined on a field may be represented as a 1×1 matrix.
   - **(c)** Every vector space is isomorphic to its dual space.
   - **(d)** Every vector space is the dual of some other vector space.
   - **(e)** If T is an isomorphism from V onto V^* and β is a finite ordered basis for V, then $\mathsf{T}(\beta) = \beta^*$.
   - **(f)** If T is a linear transformation from V to W, then the domain of $(\mathsf{T}^t)^t$ is V^{**}.
   - **(g)** If V is isomorphic to W, then V^* is isomorphic to W^*.

(h) The derivative of a function may be considered as a linear functional on the vector space of differentiable functions.

2. For the following functions f on a vector space V, determine which are linear functionals.

(a) $\mathsf{V} = \mathsf{P}(R)$; $\mathsf{f}(p(x)) = 2p'(0) + p''(1)$, where $'$ denotes differentiation
(b) $\mathsf{V} = \mathsf{R}^2$; $\mathsf{f}(x, y) = (2x, 4y)$
(c) $\mathsf{V} = \mathsf{M}_{2\times 2}(F)$; $\mathsf{f}(A) = \mathrm{tr}(A)$
(d) $\mathsf{V} = \mathsf{R}^3$; $\mathsf{f}(x, y, z) = x^2 + y^2 + z^2$
(e) $\mathsf{V} = \mathsf{P}(R)$; $\mathsf{f}(p(x)) = \int_0^1 p(t)\,dt$
(f) $\mathsf{V} = \mathsf{M}_{2\times 2}(F)$; $\mathsf{f}(A) = A_{11}$

3. For each of the following vector spaces V and bases β, find explicit formulas for vectors of the dual basis β^* for V^*, as in Example 4.

(a) $\mathsf{V} = \mathsf{R}^3$; $\beta = \{(1, 0, 1), (1, 2, 1), (0, 0, 1)\}$
(b) $\mathsf{V} = \mathsf{P}_2(R)$; $\beta = \{1, x, x^2\}$

4. Let $\mathsf{V} = \mathsf{R}^3$, and define $\mathsf{f}_1, \mathsf{f}_2, \mathsf{f}_3 \in \mathsf{V}^*$ as follows:

$$\mathsf{f}_1(x, y, z) = x - 2y, \quad \mathsf{f}_2(x, y, z) = x + y + z, \quad \mathsf{f}_3(x, y, z) = y - 3z.$$

Prove that $\{\mathsf{f}_1, \mathsf{f}_2, \mathsf{f}_3\}$ is a basis for V^*, and then find a basis for V for which it is the dual basis.

5. Let $\mathsf{V} = \mathsf{P}_1(R)$, and, for $p(x) \in \mathsf{V}$, define $\mathsf{f}_1, \mathsf{f}_2 \in \mathsf{V}^*$ by

$$\mathsf{f}_1(p(x)) = \int_0^1 p(t)\,dt \quad \text{and} \quad \mathsf{f}_2(p(x)) = \int_0^2 p(t)\,dt.$$

Prove that $\{\mathsf{f}_1, \mathsf{f}_2\}$ is a basis for V^*, and find a basis for V for which it is the dual basis.

6. Define $\mathsf{f} \in (\mathsf{R}^2)^*$ by $\mathsf{f}(x, y) = 2x + y$ and $\mathsf{T}\colon \mathsf{R}^2 \to \mathsf{R}^2$ by $\mathsf{T}(x, y) = (3x + 2y, x)$.

(a) Compute $\mathsf{T}^t(\mathsf{f})$.
(b) Compute $[\mathsf{T}^t]_{\beta^*}$, where β is the standard ordered basis for R^2 and $\beta^* = \{\mathsf{f}_1, \mathsf{f}_2\}$ is the dual basis, by finding scalars a, b, c, and d such that $\mathsf{T}^t(\mathsf{f}_1) = a\mathsf{f}_1 + c\mathsf{f}_2$ and $\mathsf{T}^t(\mathsf{f}_2) = b\mathsf{f}_1 + d\mathsf{f}_2$.
(c) Compute $[\mathsf{T}]_\beta$ and $([\mathsf{T}]_\beta)^t$, and compare your results with (b).

7. Let $\mathsf{V} = \mathsf{P}_1(R)$ and $\mathsf{W} = \mathsf{R}^2$ with respective standard ordered bases β and γ. Define $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ by

$$\mathsf{T}(p(x)) = (p(0) - 2p(1), p(0) + p'(0)),$$

where $p'(x)$ is the derivative of $p(x)$.

(a) For $\mathsf{f} \in \mathsf{W}^*$ defined by $\mathsf{f}(a, b) = a - 2b$, compute $\mathsf{T}^t(\mathsf{f})$.
(b) Compute $[\mathsf{T}^t]_{\gamma^*}^{\beta^*}$ without appealing to Theorem 2.25.
(c) Compute $[\mathsf{T}]_\beta^\gamma$ and its transpose, and compare your results with (b).

8. Show that every plane through the origin in R^3 may be identified with the null space of a vector in $(\mathsf{R}^3)^*$. State an analogous result for R^2.

9. Prove that a function $\mathsf{T}\colon \mathsf{F}^n \to \mathsf{F}^m$ is linear if and only if there exist $\mathsf{f}_1, \mathsf{f}_2, \ldots, \mathsf{f}_m \in (\mathsf{F}^n)^*$ such that $\mathsf{T}(x) = (\mathsf{f}_1(x), \mathsf{f}_2(x), \ldots, \mathsf{f}_m(x))$ for all $x \in \mathsf{F}^n$. *Hint:* If T is linear, define $\mathsf{f}_i(x) = (\mathsf{g}_i\mathsf{T})(x)$ for $x \in \mathsf{F}^n$; that is, $\mathsf{f}_i = \mathsf{T}^t(\mathsf{g}_i)$ for $1 \le i \le m$, where $\{\mathsf{g}_1, \mathsf{g}_2, \ldots, \mathsf{g}_m\}$ is the dual basis of the standard ordered basis for F^m.

10. Let $\mathsf{V} = \mathsf{P}_n(F)$, and let $c_0, c_1, \ldots, c_n$ be distinct scalars in F.
(a) For $0 \le i \le n$, define $\mathsf{f}_i \in \mathsf{V}^*$ by $\mathsf{f}_i(p(x)) = p(c_i)$. Prove that $\{\mathsf{f}_0, \mathsf{f}_1, \ldots, \mathsf{f}_n\}$ is a basis for V^*. *Hint:* Apply any linear combination of this set that equals the zero transformation to $p(x) = (x - c_1)(x - c_2)\cdots(x - c_n)$, and deduce that the first coefficient is zero.
(b) Use the corollary to Theorem 2.26 and (a) to show that there exist unique polynomials $p_0(x), p_1(x), \ldots, p_n(x)$ such that $p_i(c_j) = \delta_{ij}$ for $0 \le i \le n$. These polynomials are the Lagrange polynomials defined in Section 1.6.
(c) For any scalars $a_0, a_1, \ldots, a_n$ (not necessarily distinct), deduce that there exists a unique polynomial $q(x)$ of degree at most n such that $q(c_i) = a_i$ for $0 \le i \le n$. In fact,

$$q(x) = \sum_{i=0}^{n} a_i p_i(x).$$

(d) Deduce the Lagrange interpolation formula:

$$p(x) = \sum_{i=0}^{n} p(c_i) p_i(x)$$

for any $p(x) \in \mathsf{V}$.
(e) Prove that

$$\int_a^b p(t)\, dt = \sum_{i=0}^{n} p(c_i) d_i,$$

where

$$d_i = \int_a^b p_i(t)\, dt.$$

Suppose now that

$$c_i = a + \frac{i(b-a)}{n} \quad \text{for } i = 0, 1, \ldots, n.$$

For $n = 1$, the preceding result yields the trapezoidal rule for evaluating the definite integral of a polynomial. For $n = 2$, this result yields Simpson's rule for evaluating the definite integral of a polynomial.

11. Let V and W be finite-dimensional vector spaces over F, and let ψ_1 and ψ_2 be the isomorphisms between V and V** and W and W**, respectively, as defined in Theorem 2.26. Let $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ be linear, and define $\mathsf{T}^{tt} = (\mathsf{T}^t)^t$. Prove that the diagram depicted in Figure 2.6 commutes (i.e., prove that $\psi_2\mathsf{T} = \mathsf{T}^{tt}\psi_1$).

$$\begin{tikzcd} \mathsf{V} \arrow[r, "\mathsf{T}"] \arrow[d, "\psi_1"'] & \mathsf{W} \arrow[d, "\psi_2"] \\ \mathsf{V}^{**} \arrow[r, "\mathsf{T}^{tt}"] & \mathsf{W}^{**} \end{tikzcd}$$

Figure 2.6

12. Let V be a finite-dimensional vector space with the ordered basis β. Prove that $\psi(\beta) = \beta^{**}$, where ψ is defined in Theorem 2.26.

In Exercises 13 through 17, V denotes a finite-dimensional vector space over F. For every subset S of V, define the **annihilator** S^0 of S as

$$S^0 = \{\mathsf{f} \in \mathsf{V}^* \colon \mathsf{f}(x) = 0 \text{ for all } x \in S\}.$$

13. (a) Prove that S^0 is a subspace of V*.
(b) If W is a subspace of V and $x \notin \mathsf{W}$, prove that there exists $\mathsf{f} \in \mathsf{W}^0$ such that $\mathsf{f}(x) \neq 0$.
(c) Prove that $(S^0)^0 = \operatorname{span}(\psi(S))$, where ψ is defined as in Theorem 2.26.
(d) For subspaces W_1 and W_2, prove that $\mathsf{W}_1 = \mathsf{W}_2$ if and only if $\mathsf{W}_1^0 = \mathsf{W}_2^0$.
(e) For subspaces W_1 and W_2, show that $(\mathsf{W}_1 + \mathsf{W}_2)^0 = \mathsf{W}_1^0 \cap \mathsf{W}_2^0$.

14. Prove that if W is a subspace of V, then $\dim(\mathsf{W}) + \dim(\mathsf{W}^0) = \dim(\mathsf{V})$.
Hint: Extend an ordered basis $\{x_1, x_2, \ldots, x_k\}$ of W to an ordered basis $\beta = \{x_1, x_2, \ldots, x_n\}$ of V. Let $\beta^* = \{\mathsf{f}_1, \mathsf{f}_2, \ldots, \mathsf{f}_n\}$. Prove that $\{\mathsf{f}_{k+1}, \mathsf{f}_{k+2}, \ldots, \mathsf{f}_n\}$ is a basis for W^0.

15. Suppose that W is a finite-dimensional vector space and that $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ is linear. Prove that $\mathsf{N}(\mathsf{T}^t) = (\mathsf{R}(\mathsf{T}))^0$.

16. Use Exercises 14 and 15 to deduce that $\operatorname{rank}(\mathsf{L}_{A^t}) = \operatorname{rank}(\mathsf{L}_A)$ for any $A \in \mathsf{M}_{m\times n}(F)$.

17. Let T be a linear operator on V, and let W be a subspace of V. Prove that W is T-invariant (as defined in the exercises of Section 2.1) if and only if W^0 is T^t-invariant.

18. Let V be a nonzero vector space over a field F, and let S be a basis for V. (By the corollary to Theorem 1.13 (p. 60) in Section 1.7, every vector space has a basis.) Let $\Phi\colon \mathsf{V}^* \to \mathcal{F}(S, F)$ be the mapping defined by $\Phi(\mathsf{f}) = \mathsf{f}_S$, the restriction of f to S. Prove that Φ is an isomorphism. *Hint:* Apply Exercise 34 of Section 2.1.

19. Let V be a nonzero vector space, and let W be a proper subspace of V (i.e., $\mathsf{W} \neq \mathsf{V}$). Prove that there exists a nonzero linear functional $\mathsf{f} \in \mathsf{V}^*$ such that $\mathsf{f}(x) = 0$ for all $x \in \mathsf{W}$. *Hint:* For the infinite-dimensional case, use Exercise 34 of Section 2.1 as well as results about extending linearly independent sets to bases in Section 1.7.

20. Let V and W be nonzero vector spaces over the same field, and let $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ be a linear transformation.

(a) Prove that T is onto if and only if T^t is one-to-one.
(b) Prove that T^t is onto if and only if T is one-to-one.

Hint: Parts of the proof require the result of Exercise 19 for the infinite-dimensional case.

2.7* HOMOGENEOUS LINEAR DIFFERENTIAL EQUATIONS WITH CONSTANT COEFFICIENTS

As an introduction to this section, consider the following physical problem. A weight of mass m is attached to a vertically suspended spring that is allowed to stretch until the forces acting on the weight are in equilibrium. Suppose that the weight is now motionless and impose an xy-coordinate system with the weight at the origin and the spring lying on the positive y-axis (see Figure 2.7).

Suppose that at a certain time, say $t = 0$, the weight is lowered a distance s along the y-axis and released. The spring then begins to oscillate.

We describe the motion of the spring. At any time $t \geq 0$, let $F(t)$ denote the force acting on the weight and $y(t)$ denote the position of the weight along the y-axis. For example, $y(0) = -s$. The second derivative of y with respect

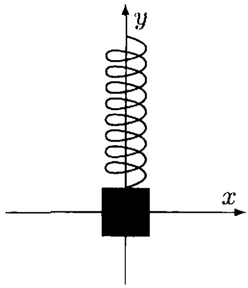


Figure 2.7

to time, $y''(t)$, is the acceleration of the weight at time t; hence, by Newton's second law of motion,

$$F(t) = my''(t). \tag{1}$$

It is reasonable to assume that the force acting on the weight is due totally to the tension of the spring, and that this force satisfies Hooke's law: *The force acting on the weight is proportional to its displacement from the equilibrium position, but acts in the opposite direction.* If $k > 0$ is the proportionality constant, then Hooke's law states that

$$F(t) = -ky(t). \tag{2}$$

Combining (1) and (2), we obtain $my'' = -ky$ or

$$y'' + \frac{k}{m}y = 0. \tag{3}$$

The expression (3) is an example of a *differential equation.* A **differential equation** in an unknown function $y = y(t)$ is an equation involving y, t, and derivatives of y. If the differential equation is of the form

$$a_n y^{(n)} + a_{n-1} y^{(n-1)} + \cdots + a_1 y^{(1)} + a_0 y = f, \tag{4}$$

where $a_0, a_1, \ldots, a_n$ and f are functions of t and $y^{(k)}$ denotes the kth derivative of y, then the equation is said to be **linear**. The functions a_i are called the **coefficients** of the differential equation (4). Thus (3) is an example of a linear differential equation in which the coefficients are constants and the function f is identically zero. When f is identically zero, (4) is called **homogeneous**.

In this section, we apply the linear algebra we have studied to solve homogeneous linear differential equations with constant coefficients. If $a_n \neq 0$, we say that differential equation (4) is of **order** n. In this case, we divide both sides by a_n to obtain a new, but equivalent, equation

$$y^{(n)} + b_{n-1}y^{(n-1)} + \cdots + b_1 y^{(1)} + b_0 y = 0,$$

where $b_i = a_i/a_n$ for $i = 0, 1, \ldots, n-1$. Because of this observation, we always assume that the coefficient a_n in (4) is 1.

A solution to (4) is a function that when substituted for y reduces (4) to an identity.

Example 1

The function $y(t) = \sin\sqrt{k/m}\,t$ is a solution to (3) since

$$y''(t) + \frac{k}{m}y(t) = -\frac{k}{m}\sin\sqrt{\frac{k}{m}}\,t + \frac{k}{m}\sin\sqrt{\frac{k}{m}}\,t = 0$$

for all t. Notice, however, that substituting $y(t) = t$ into (3) yields

$$y''(t) + \frac{k}{m}y(t) = \frac{k}{m}t,$$

which is not identically zero. Thus $y(t) = t$ is not a solution to (3). ◆

In our study of differential equations, it is useful to regard solutions as complex-valued functions of a real variable even though the solutions that are meaningful to us in a physical sense are real-valued. The convenience of this viewpoint will become clear later. Thus we are concerned with the vector space $\mathcal{F}(R, C)$ (as defined in Example 3 of Section 1.2). In order to consider complex-valued functions of a real variable as solutions to differential equations, we must define what it means to differentiate such functions. Given a complex-valued function $x \in \mathcal{F}(R, C)$ of a real variable t, there exist unique real-valued functions x_1 and x_2 of t, such that

$$x(t) = x_1(t) + ix_2(t) \quad \text{for} \quad t \in R,$$

where i is the imaginary number such that $i^2 = -1$. We call x_1 the **real part** and x_2 the **imaginary part** of x.

Definitions. *Given a function $x \in \mathcal{F}(R, C)$ with real part x_1 and imaginary part x_2, we say that x is* ***differentiable*** *if x_1 and x_2 are differentiable. If x is differentiable, we define the* ***derivative*** *x' of x by*

$$x' = x_1' + ix_2'.$$

We illustrate some computations with complex-valued functions in the following example.

Example 2

Suppose that $x(t) = \cos 2t + i \sin 2t$. Then

$$x'(t) = -2\sin 2t + 2i\cos 2t.$$

We next find the real and imaginary parts of x^2. Since

$$\begin{aligned} x^2(t) &= (\cos 2t + i\sin 2t)^2 = (\cos^2 2t - \sin^2 2t) + i(2\sin 2t\cos 2t) \\ &= \cos 4t + i\sin 4t, \end{aligned}$$

the real part of $x^2(t)$ is $\cos 4t$, and the imaginary part is $\sin 4t$. ◆

The next theorem indicates that we may limit our investigations to a vector space considerably smaller than $\mathcal{F}(R, C)$. Its proof, which is illustrated in Example 3, involves a simple induction argument, which we omit.

Theorem 2.27. *Any solution to a homogeneous linear differential equation with constant coefficients has derivatives of all orders; that is, if x is a solution to such an equation, then $x^{(k)}$ exists for every positive integer k.*

Example 3

To illustrate Theorem 2.27, consider the equation

$$y^{(2)} + 4y = 0.$$

Clearly, to qualify as a solution, a function y must have two derivatives. If y is a solution, however, then

$$y^{(2)} = -4y.$$

Thus since $y^{(2)}$ is a constant multiple of a function y that has two derivatives, $y^{(2)}$ must have two derivatives. Hence $y^{(4)}$ exists; in fact,

$$y^{(4)} = -4y^{(2)}.$$

Since $y^{(4)}$ is a constant multiple of a function that we have shown has at least two derivatives, it also has at least two derivatives; hence $y^{(6)}$ exists. Continuing in this manner, we can show that any solution has derivatives of all orders. ◆

Definition. *We use C^∞ to denote the set of all functions in $\mathcal{F}(R, C)$ that have derivatives of all orders.*

It is a simple exercise to show that C^∞ is a subspace of $\mathcal{F}(R, C)$ and hence a vector space over C. In view of Theorem 2.27, it is this vector space that is of interest to us. For $x \in \mathsf{C}^\infty$, the derivative x' of x also lies in C^∞. We can use the derivative operation to define a mapping $\mathsf{D}\colon \mathsf{C}^\infty \to \mathsf{C}^\infty$ by

$$\mathsf{D}(x) = x' \quad \text{for } x \in \mathsf{C}^\infty.$$

It is easy to show that D is a linear operator. More generally, consider any polynomial over C of the form

$$p(t) = a_n t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0.$$

If we define

$$p(\mathsf{D}) = a_n \mathsf{D}^n + a_{n-1}\mathsf{D}^{n-1} + \cdots + a_1 D + a_0 \mathsf{I},$$

then $p(\mathsf{D})$ is a linear operator on C^∞. (See Appendix E.)

Definitions. *For any polynomial $p(t)$ over C of positive degree, $p(\mathsf{D})$ is called a* **differential operator.** *The* **order** *of the differential operator $p(\mathsf{D})$ is the degree of the polynomial $p(t)$.*

Differential operators are useful since they provide us with a means of reformulating a differential equation in the context of linear algebra. Any homogeneous linear differential equation with constant coefficients,

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1 y^{(1)} + a_0 y = 0,$$

can be rewritten using differential operators as

$$(\mathsf{D}^n + a_{n-1}\mathsf{D}^{n-1} + \cdots + a_1\mathsf{D} + a_0\mathsf{I})(y) = 0.$$

Definition. *Given the differential equation above, the complex polynomial*

$$p(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$$

is called the **auxiliary polynomial** *associated with the equation.*

For example, (3) has the auxiliary polynomial

$$p(t) = t^2 + \frac{k}{m}.$$

Any homogeneous linear differential equation with constant coefficients can be rewritten as

$$p(\mathsf{D})(y) = 0,$$

where $p(t)$ is the auxiliary polynomial associated with the equation. Clearly, this equation implies the following theorem.

Theorem 2.28. *The set of all solutions to a homogeneous linear differential equation with constant coefficients coincides with the null space of $p(\mathsf{D})$, where $p(t)$ is the auxiliary polynomial associated with the equation.*

Proof. Exercise. ∎

Corollary. *The set of all solutions to a homogeneous linear differential equation with constant coefficients is a subspace of C^∞.*

In view of the preceding corollary, we call the set of solutions to a homogeneous linear differential equation with constant coefficients the **solution space** of the equation. A practical way of describing such a space is in terms of a basis. We now examine a certain class of functions that is of use in finding bases for these solution spaces.

For a real number s, we are familiar with the real number e^s, where e is the unique number whose natural logarithm is 1 (i.e., $\ln e = 1$). We know, for instance, certain properties of exponentiation, namely,

$$e^{s+t} = e^s e^t \quad \text{and} \quad e^{-t} = \frac{1}{e^t}$$

for any real numbers s and t. We now extend the definition of powers of e to include complex numbers in such a way that these properties are preserved.

Definition. *Let $c = a + ib$ be a complex number with real part a and imaginary part b. Define*

$$e^c = e^a(\cos b + i \sin b).$$

The special case

$$e^{ib} = \cos b + i \sin b$$

is called ***Euler's formula.***

For example, for $c = 2 + i(\pi/3)$,

$$e^c = e^2\left(\cos\frac{\pi}{3} + i\sin\frac{\pi}{3}\right) = e^2\left(\frac{1}{2} + i\frac{\sqrt{3}}{2}\right).$$

Clearly, if c is real ($b = 0$), then we obtain the usual result: $e^c = e^a$. Using the approach of Example 2, we can show by the use of trigonometric identities that

$$e^{c+d} = e^c e^d \quad \text{and} \quad e^{-c} = \frac{1}{e^c}$$

for any complex numbers c and d.

Definition. *A function $f\colon R \to C$ defined by $f(t) = e^{ct}$ for a fixed complex number c is called an* **exponential function.**

The derivative of an exponential function, as described in the next theorem, is consistent with the real version. The proof involves a straightforward computation, which we leave as an exercise.

Theorem 2.29. *For any exponential function $f(t) = e^{ct}$, $f'(t) = ce^{ct}$.*

Proof. Exercise. ∎

We can use exponential functions to describe all solutions to a homogeneous linear differential equation of order 1. Recall that the **order** of such an equation is the degree of its auxiliary polynomial. Thus an equation of order 1 is of the form

$$y' + a_0 y = 0. \tag{5}$$

Theorem 2.30. *The solution space for (5) is of dimension 1 and has $\{e^{-a_0 t}\}$ as a basis.*

Proof. Clearly (5) has $e^{-a_0 t}$ as a solution. Suppose that $x(t)$ is any solution to (5). Then

$$x'(t) = -a_0 x(t) \quad \text{for all } t \in R.$$

Define

$$z(t) = e^{a_0 t} x(t).$$

Differentiating z yields

$$z'(t) = (e^{a_0 t})' x(t) + e^{a_0 t} x'(t) = a_0 e^{a_0 t} x(t) - a_0 e^{a_0 t} x(t) = 0.$$

(Notice that the familiar product rule for differentiation holds for complex-valued functions of a real variable. A justification of this involves a lengthy, although direct, computation.)

Since z' is identically zero, z is a constant function. (Again, this fact, well known for real-valued functions, is also true for complex-valued functions. The proof, which relies on the real case, involves looking separately at the real and imaginary parts of z.) Thus there exists a complex number k such that

$$z(t) = e^{a_0 t} x(t) = k \quad \text{for all } t \in R.$$

So

$$x(t) = ke^{-a_0 t}.$$

We conclude that any solution to (5) is a linear combination of $e^{-a_0 t}$. ∎

Another way of stating Theorem 2.30 is as follows.

Corollary. *For any complex number c, the null space of the differential operator* $\mathsf{D} - c\mathsf{I}$ *has* $\{e^{ct}\}$ *as a basis.*

We next concern ourselves with differential equations of order greater than one. Given an nth order homogeneous linear differential equation with constant coefficients,

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1 y^{(1)} + a_0 y = 0,$$

its auxiliary polynomial

$$p(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$$

factors into a product of polynomials of degree 1, that is,

$$p(t) = (t - c_1)(t - c_2)\cdots(t - c_n),$$

where $c_1, c_2, \ldots, c_n$ are (not necessarily distinct) complex numbers. (This follows from the fundamental theorem of algebra in Appendix D.) Thus

$$p(\mathsf{D}) = (\mathsf{D} - c_1\mathsf{I})(\mathsf{D} - c_2\mathsf{I})\cdots(\mathsf{D} - c_n\mathsf{I}).$$

The operators $\mathsf{D} - c_i\mathsf{I}$ commute, and so, by Exercise 9, we have that

$$\mathsf{N}(\mathsf{D} - c_i\mathsf{I}) \subseteq \mathsf{N}(p(\mathsf{D})) \quad \text{for all } i.$$

Since $\mathsf{N}(p(\mathsf{D}))$ coincides with the solution space of the given differential equation, we can deduce the following result from the preceding corollary.

Theorem 2.31. *Let $p(t)$ be the auxiliary polynomial for a homogeneous linear differential equation with constant coefficients. For any complex number c, if c is a zero of $p(t)$, then e^{ct} is a solution to the differential equation.*

Example 4

Given the differential equation

$$y'' - 3y' + 2y = 0,$$

its auxiliary polynomial is

$$p(t) = t^2 - 3t + 2 = (t - 1)(t - 2).$$

Hence, by Theorem 2.31, e^t and e^{2t} are solutions to the differential equation because $c = 1$ and $c = 2$ are zeros of $p(t)$. Since the solution space of the differential equation is a subspace of C^∞, span($\{e^t, e^{2t}\}$) lies in the solution space. It is a simple matter to show that $\{e^t, e^{2t}\}$ is linearly independent. Thus if we can show that the solution space is two-dimensional, we can conclude that $\{e^t, e^{2t}\}$ is a basis for the solution space. This result is a consequence of the next theorem. ♦

Theorem 2.32. *For any differential operator* $p(\mathsf{D})$ *of order* n*, the null space of* $p(\mathsf{D})$ *is an* n*-dimensional subspace of* C^∞.

As a preliminary to the proof of Theorem 2.32, we establish two lemmas.

Lemma 1. *The differential operator* $\mathsf{D} - c\mathsf{I}\colon \mathsf{C}^\infty \to \mathsf{C}^\infty$ *is onto for any complex number* c.

Proof. Let $v \in \mathsf{C}^\infty$. We wish to find a $u \in \mathsf{C}^\infty$ such that $(\mathsf{D} - c\mathsf{I})u = v$. Let $w(t) = v(t)e^{-ct}$ for $t \in R$. Clearly, $w \in \mathsf{C}^\infty$ because both v and e^{-ct} lie in C^∞. Let w_1 and w_2 be the real and imaginary parts of w. Then w_1 and w_2 are continuous because they are differentiable. Hence they have antiderivatives, say, W_1 and W_2, respectively. Let $W\colon R \to C$ be defined by

$$W(t) = W_1(t) + iW_2(t) \quad \text{for } t \in R.$$

Then $W \in \mathsf{C}^\infty$, and the real and imaginary parts of W are W_1 and W_2, respectively. Furthermore, $W' = w$. Finally, let $u\colon R \to C$ be defined by $u(t) = W(t)e^{ct}$ for $t \in R$. Clearly $u \in \mathsf{C}^\infty$, and since

$$\begin{aligned}(\mathsf{D} - c\mathsf{I})u(t) &= u'(t) - cu(t)\\ &= W'(t)e^{ct} + W(t)ce^{ct} - cW(t)e^{ct}\\ &= w(t)e^{ct}\\ &= v(t)e^{-ct}e^{ct}\\ &= v(t),\end{aligned}$$

we have $(\mathsf{D} - c\mathsf{I})u = v$. ∎

Lemma 2. *Let* V *be a vector space, and suppose that* T *and* U *are linear operators on* V *such that* U *is onto and the null spaces of* T *and* U *are finite-dimensional. Then the null space of* TU *is finite-dimensional, and*

$$\dim(\mathsf{N}(\mathsf{TU})) = \dim(\mathsf{N}(\mathsf{T})) + \dim(\mathsf{N}(\mathsf{U})).$$

Proof. Let $p = \dim(\mathsf{N}(\mathsf{T}))$, $q = \dim(\mathsf{N}(\mathsf{U}))$, and $\{u_1, u_2, \ldots, u_p\}$ and $\{v_1, v_2, \ldots, v_q\}$ be bases for $\mathsf{N}(\mathsf{T})$ and $\mathsf{N}(\mathsf{U})$, respectively. Since U is onto, we can choose for each i $(1 \le i \le p)$ a vector $w_i \in \mathsf{V}$ such that $\mathsf{U}(w_i) = u_i$. Note that the w_i's are distinct. Furthermore, for any i and j, $w_i \neq v_j$, for otherwise $u_i = \mathsf{U}(w_i) = \mathsf{U}(v_j) = 0$—a contradiction. Hence the set

$$\beta = \{w_1, w_2, \ldots, w_p, v_1, v_2, \ldots, v_q\}$$

contains $p+q$ distinct vectors. To complete the proof of the lemma, it suffices to show that β is a basis for $\mathsf{N}(\mathsf{TU})$.

We first show that β generates $\mathsf{N}(\mathsf{TU})$. Since for any w_i and v_j in β, $\mathsf{TU}(w_i) = \mathsf{T}(u_i) = \mathit{0}$ and $\mathsf{TU}(v_j) = \mathsf{T}(\mathit{0}) = \mathit{0}$, it follows that $\beta \subseteq \mathsf{N}(\mathsf{TU})$. Now suppose that $v \in \mathsf{N}(\mathsf{TU})$. Then $\mathit{0} = \mathsf{TU}(v) = \mathsf{T}(\mathsf{U}(v))$. Thus $\mathsf{U}(v) \in \mathsf{N}(\mathsf{T})$. So there exist scalars $a_1, a_2, \ldots, a_p$ such that

$$\begin{aligned}\mathsf{U}(v) &= a_1u_1 + a_2u_2 + \cdots + a_pu_p \\ &= a_1U(w_1) + a_2U(w_2) + \cdots + a_pU(w_p) \\ &= \mathsf{U}(a_1w_1 + a_2w_2 + \cdots + a_pw_p).\end{aligned}$$

Hence

$$\mathsf{U}(v - (a_1w_1 + a_2w_2 + \cdots + a_pw_p)) = \mathit{0}.$$

Consequently, $v - (a_1w_1 + a_2w_2 + \cdots + a_pw_p)$ lies in $\mathsf{N}(\mathsf{U})$. It follows that there exist scalars $b_1, b_2, \ldots, b_q$ such that

$$v - (a_1w_1 + a_2w_2 + \cdots + a_pw_p) = b_1v_1 + b_2v_2 + \cdots + b_qv_q$$

or

$$v = a_1w_1 + a_2w_2 + \cdots + a_pw_p + b_1v_1 + b_2v_2 + \cdots + b_qv_q.$$

Therefore β spans $\mathsf{N}(\mathsf{TU})$.

To prove that β is linearly independent, let $a_1, a_2, \ldots, a_p, b_1, b_2, \ldots, b_q$ be any scalars such that

$$a_1w_1 + a_2w_2 + \cdots + a_pw_p + b_1v_1 + b_2v_2 + \cdots + b_qv_q = \mathit{0}. \tag{6}$$

Applying U to both sides of (6), we obtain

$$a_1u_1 + a_2u_2 + \cdots + a_pu_p = \mathit{0}.$$

Since $\{u_1, u_2, \ldots, u_p\}$ is linearly independent, the a_i's are all zero. Thus (6) reduces to

$$b_1v_1 + b_2v_2 + \cdots + b_qv_q = \mathit{0}.$$

Again, the linear independence of $\{v_1, v_2, \ldots, v_q\}$ implies that the b_i's are all zero. We conclude that β is a basis for $\mathsf{N}(\mathsf{TU})$. Hence $\mathsf{N}(\mathsf{TU})$ is finite-dimensional, and $\dim(\mathsf{N}(\mathsf{TU})) = p + q = \dim(\mathsf{N}(\mathsf{T})) + \dim(\mathsf{N}(\mathsf{U}))$. ∎

Proof of Theorem 2.32. The proof is by mathematical induction on the order of the differential operator $p(\mathsf{D})$. The first-order case coincides with Theorem 2.30. For some integer $n > 1$, suppose that Theorem 2.32 holds for any differential operator of order less than n, and consider a differential operator $p(\mathsf{D})$ of order n. The polynomial $p(t)$ can be factored into a product of two polynomials as follows:

$$p(t) = q(t)(t - c),$$

where $q(t)$ is a polynomial of degree $n - 1$ and c is a complex number. Thus the given differential operator may be rewritten as

$$p(\mathsf{D}) = q(\mathsf{D})(\mathsf{D} - c\mathsf{I}).$$

Now, by Lemma 1, $\mathsf{D} - c\mathsf{I}$ is onto, and by the corollary to Theorem 2.30, $\dim(\mathsf{N}(\mathsf{D} - c\mathsf{I})) = 1$. Also, by the induction hypothesis, $\dim(\mathsf{N}(q(\mathsf{D})) = n - 1$. Thus, by Lemma 2, we conclude that

$$\begin{aligned}\dim(\mathsf{N}(p(\mathsf{D}))) &= \dim(\mathsf{N}(q(\mathsf{D}))) + \dim(\mathsf{N}(\mathsf{D} - c\mathsf{I})) \\ &= (n - 1) + 1 = n.\end{aligned}$$ ∎

Corollary. *The solution space of any nth-order homogeneous linear differential equation with constant coefficients is an n-dimensional subspace of* C^∞.

The corollary to Theorem 2.32 reduces the problem of finding all solutions to an nth-order homogeneous linear differential equation with constant coefficients to finding a set of n linearly independent solutions to the equation. By the results of Chapter 1, any such set must be a basis for the solution space. The next theorem enables us to find a basis quickly for many such equations. Hints for its proof are provided in the exercises.

Theorem 2.33. *Given n distinct complex numbers $c_1, c_2, \ldots, c_n$, the set of exponential functions $\{e^{c_1 t}, e^{c_2 t}, \ldots, e^{c_n t}\}$ is linearly independent.*

Proof. Exercise. (See Exercise 10.) ∎

Corollary. *For any nth-order homogeneous linear differential equation with constant coefficients, if the auxiliary polynomial has n distinct zeros $c_1, c_2, \ldots, c_n$, then $\{e^{c_1 t}, e^{c_2 t}, \ldots, e^{c_n t}\}$ is a basis for the solution space of the differential equation.*

Proof. Exercise. (See Exercise 10.) ∎

Example 5

We find all solutions to the differential equation

$$y'' + 5y' + 4y = 0.$$

Since the auxiliary polynomial factors as $(t+4)(t+1)$, it has two distinct zeros, -1 and -4. Thus $\{e^{-t}, e^{-4t}\}$ is a basis for the solution space. So any solution to the given equation is of the form

$$y(t) = b_1 e^{-t} + b_2 e^{-4t}$$

for unique scalars b_1 and b_2. ♦

Example 6

We find all solutions to the differential equation

$$y'' + 9y = 0.$$

The auxiliary polynomial $t^2 + 9$ factors as $(t-3i)(t+3i)$ and hence has distinct zeros $c_1 = 3i$ and $c_2 = -3i$. Thus $\{e^{3it}, e^{-3it}\}$ is a basis for the solution space. Since

$$\cos 3t = \frac{1}{2}(e^{3it} + e^{-3it}) \quad \text{and} \quad \sin 3t = \frac{1}{2i}(e^{3it} - e^{-3it}),$$

it follows from Exercise 7 that $\{\cos 3t, \sin 3t\}$ is also a basis for this solution space. This basis has an advantage over the original one because it consists of the familiar sine and cosine functions and makes no reference to the imaginary number i. Using this latter basis, we see that any solution to the given equation is of the form

$$y(t) = b_1 \cos 3t + b_2 \sin 3t$$

for unique scalars b_1 and b_2. ♦

Next consider the differential equation

$$y'' + 2y' + y = 0,$$

for which the auxiliary polynomial is $(t+1)^2$. By Theorem 2.31, e^{-t} is a solution to this equation. By the corollary to Theorem 2.32, its solution space is two-dimensional. In order to obtain a basis for the solution space, we need a solution that is linearly independent of e^{-t}. The reader can verify that te^{-t} is a such a solution. The following lemma extends this result.

Lemma. *For a given complex number c and positive integer n, suppose that $(t-c)^n$ is the auxiliary polynomial of a homogeneous linear differential equation with constant coefficients. Then the set*

$$\beta = \{e^{ct}, te^{ct}, \ldots, t^{n-1}e^{ct}\}$$

is a basis for the solution space of the equation.

Proof. Since the solution space is n-dimensional, we need only show that β is linearly independent and lies in the solution space. First, observe that for any positive integer k,

$$\begin{aligned}(\mathsf{D} - c\mathsf{I})(t^k e^{ct}) &= kt^{k-1}e^{ct} + ct^k e^{ct} - ct^k e^{ct}\\ &= kt^{k-1}e^{ct}.\end{aligned}$$

Hence for $k < n$,

$$(\mathsf{D} - c\mathsf{I})^n(t^k e^{ct}) = \mathit{0}.$$

It follows that β is a subset of the solution space.

We next show that β is linearly independent. Consider any linear combination of vectors in β such that

$$b_0 e^{ct} + b_1 t e^{ct} + \cdots + b_{n-1} t^{n-1} e^{ct} = \mathit{0} \tag{7}$$

for some scalars $b_0, b_1, \ldots, b_{n-1}$. Dividing by e^{ct} in (7), we obtain

$$b_0 + b_1 t + \cdots + b_{n-1} t^{n-1} = \mathit{0}. \tag{8}$$

Thus the left side of (8) must be the zero polynomial function. We conclude that the coefficients $b_0, b_1, \ldots, b_{n-1}$ are all zero. So β is linearly independent and hence is a basis for the solution space. ∎

Example 7

We find all solutions to the differential equation

$$y^{(4)} - 4y^{(3)} + 6y^{(2)} - 4y^{(1)} + y = \mathit{0}.$$

Since the auxiliary polynomial is

$$t^4 - 4t^3 + 6t^2 - 4t + 1 = (t-1)^4,$$

we can immediately conclude by the preceding lemma that $\{e^t, te^t, t^2e^t, t^3e^t\}$ is a basis for the solution space. So any solution y to the given differential equation is of the form

$$y(t) = b_1 e^t + b_2 te^t + b_3 t^2 e^t + b_4 t^3 e^t$$

for unique scalars b_1, b_2, b_3, and b_4. ◆

The most general situation is stated in the following theorem.

Theorem 2.34. *Given a homogeneous linear differential equation with constant coefficients and auxiliary polynomial*

$$(t - c_1)^{n_1}(t - c_2)^{n_2} \cdots (t - c_k)^{n_k},$$

where $n_1, n_2, \ldots, n_k$ are positive integers and $c_1, c_2, \ldots, c_k$ are distinct complex numbers, the following set is a basis for the solution space of the equation:

$$\{e^{c_1 t}, te^{c_1 t}, \ldots, t^{n_1 - 1}e^{c_1 t}, \ldots, e^{c_k t}, te^{c_k t}, \ldots, t^{n_k - 1}e^{c_k t}\}.$$

Proof. Exercise. ∎

Example 8

The differential equation

$$y^{(3)} - 4y^{(2)} + 5y^{(1)} - 2y = 0$$

has the auxiliary polynomial

$$t^3 - 4t^2 + 5t - 2 = (t-1)^2(t-2).$$

By Theorem 2.34, $\{e^t, te^t, e^{2t}\}$ is a basis for the solution space of the differential equation. Thus any solution y has the form

$$y(t) = b_1e^t + b_2te^t + b_3e^{2t}$$

for unique scalars b_1, b_2, and b_3. ♦

EXERCISES

1. Label the following statements as true or false.
   (a) The set of solutions to an nth-order homogeneous linear differential equation with constant coefficients is an n-dimensional subspace of C^∞.
   (b) The solution space of a homogeneous linear differential equation with constant coefficients is the null space of a differential operator.
   (c) The auxiliary polynomial of a homogeneous linear differential equation with constant coefficients is a solution to the differential equation.
   (d) Any solution to a homogeneous linear differential equation with constant coefficients is of the form ae^{ct} or at^ke^{ct}, where a and c are complex numbers and k is a positive integer.
   (e) Any linear combination of solutions to a given homogeneous linear differential equation with constant coefficients is also a solution to the given equation.
   (f) For any homogeneous linear differential equation with constant coefficients having auxiliary polynomial $p(t)$, if $c_1, c_2, \ldots, c_k$ are the distinct zeros of $p(t)$, then $\{e^{c_1t}, e^{c_2t}, \ldots, e^{c_kt}\}$ is a basis for the solution space of the given differential equation.
   (g) Given any polynomial $p(t) \in \mathsf{P}(C)$, there exists a homogeneous linear differential equation with constant coefficients whose auxiliary polynomial is $p(t)$.

2. For each of the following parts, determine whether the statement is true or false. Justify your claim with either a proof or a counterexample, whichever is appropriate.

(a) Any finite-dimensional subspace of C^∞ is the solution space of a homogeneous linear differential equation with constant coefficients.
(b) There exists a homogeneous linear differential equation with constant coefficients whose solution space has the basis $\{t, t^2\}$.
(c) For any homogeneous linear differential equation with constant coefficients, if x is a solution to the equation, so is its derivative x'.

Given two polynomials $p(t)$ and $q(t)$ in $\mathsf{P}(C)$, if $x \in \mathsf{N}(p(\mathsf{D}))$ and $y \in \mathsf{N}(q(\mathsf{D}))$, then

(d) $x + y \in \mathsf{N}(p(\mathsf{D})q(\mathsf{D}))$.
(e) $xy \in \mathsf{N}(p(\mathsf{D})q(\mathsf{D}))$.

3. Find a basis for the solution space of each of the following differential equations.

(a) $y'' + 2y' + y = 0$
(b) $y''' = y'$
(c) $y^{(4)} - 2y^{(2)} + y = 0$
(d) $y'' + 2y' + y = 0$
(e) $y^{(3)} - y^{(2)} + 3y^{(1)} + 5y = 0$

4. Find a basis for each of the following subspaces of C^∞.

(a) $\mathsf{N}(\mathsf{D}^2 - \mathsf{D} - \mathsf{I})$
(b) $\mathsf{N}(\mathsf{D}^3 - 3\mathsf{D}^2 + 3\mathsf{D} - \mathsf{I})$
(c) $\mathsf{N}(\mathsf{D}^3 + 6\mathsf{D}^2 + 8\mathsf{D})$

5. Show that C^∞ is a subspace of $\mathcal{F}(R, C)$.

6. (a) Show that $\mathsf{D}\colon \mathsf{C}^\infty \to \mathsf{C}^\infty$ is a linear operator.
(b) Show that any differential operator is a linear operator on C^∞.

7. Prove that if $\{x, y\}$ is a basis for a vector space over C, then so is

$$\left\{\frac{1}{2}(x+y), \frac{1}{2i}(x-y)\right\}.$$

8. Consider a second-order homogeneous linear differential equation with constant coefficients in which the auxiliary polynomial has distinct conjugate complex roots $a + ib$ and $a - ib$, where $a, b \in R$. Show that $\{e^{at}\cos bt, e^{at}\sin bt\}$ is a basis for the solution space.

9. Suppose that $\{\mathsf{U}_1, \mathsf{U}_2, \ldots, \mathsf{U}_n\}$ is a collection of pairwise commutative linear operators on a vector space V (i.e., operators such that $\mathsf{U}_i\mathsf{U}_j = \mathsf{U}_j\mathsf{U}_i$ for all i, j). Prove that, for any i $(1 \le i \le n)$,

$$\mathsf{N}(\mathsf{U}_i) \subseteq \mathsf{N}(\mathsf{U}_1\mathsf{U}_2\cdots\mathsf{U}_n).$$

10. Prove Theorem 2.33 and its corollary. *Hint:* Suppose that

$$b_1e^{c_1t} + b_2e^{c_2t} + \cdots + b_ne^{c_nt} = 0 \quad \text{(where the } c_i\text{'s are distinct).}$$

To show the b_i's are zero, apply mathematical induction on n as follows. Verify the theorem for $n = 1$. Assuming that the theorem is true for $n - 1$ functions, apply the operator $\mathsf{D} - c_n\mathsf{I}$ to both sides of the given equation to establish the theorem for n distinct exponential functions.

11. Prove Theorem 2.34. *Hint:* First verify that the alleged basis lies in the solution space. Then verify that this set is linearly independent by mathematical induction on k as follows. The case $k = 1$ is the lemma to Theorem 2.34. Assuming that the theorem holds for $k - 1$ distinct c_i's, apply the operator $(\mathsf{D} - c_k\mathsf{I})^{n_k}$ to any linear combination of the alleged basis that equals 0.

12. Let V be the solution space of an nth-order homogeneous linear differential equation with constant coefficients having auxiliary polynomial $p(t)$. Prove that if $p(t) = g(t)h(t)$, where $g(t)$ and $h(t)$ are polynomials of positive degree, then

$$\mathsf{N}(h(\mathsf{D})) = \mathsf{R}(g(\mathsf{D}_\mathsf{V})) = g(\mathsf{D})(\mathsf{V}),$$

where $\mathsf{D}_\mathsf{V}\colon \mathsf{V} \to \mathsf{V}$ is defined by $\mathsf{D}_\mathsf{V}(x) = x'$ for $x \in \mathsf{V}$. *Hint:* First prove $g(\mathsf{D})(\mathsf{V}) \subseteq \mathsf{N}(h(\mathsf{D}))$. Then prove that the two spaces have the same finite dimension.

13. A differential equation

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1y^{(1)} + a_0y = x$$

is called a **nonhomogeneous** linear differential equation with constant coefficients if the a_i's are constant and x is a function that is not identically zero.

(a) Prove that for any $x \in \mathsf{C}^\infty$ there exists $y \in \mathsf{C}^\infty$ such that y is a solution to the differential equation. *Hint:* Use Lemma 1 to Theorem 2.32 to show that for any polynomial $p(t)$, the linear operator $p(\mathsf{D})\colon \mathsf{C}^\infty \to \mathsf{C}^\infty$ is onto.

(b) Let V be the solution space for the homogeneous linear equation

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1 y^{(1)} + a_0 y = 0.$$

Prove that if z is any solution to the associated nonhomogeneous linear differential equation, then the set of all solutions to the nonhomogeneous linear differential equation is

$$\{z + y \colon y \in \mathsf{V}\}.$$

14. Given any nth-order homogeneous linear differential equation with constant coefficients, prove that, for any solution x and any $t_0 \in R$, if $x(t_0) = x'(t_0) = \cdots = x^{(n-1)}(t_0) = 0$, then $x = 0$ (the zero function). *Hint:* Use mathematical induction on n as follows. First prove the conclusion for the case $n = 1$. Next suppose that it is true for equations of order $n - 1$, and consider an nth-order differential equation with auxiliary polynomial $p(t)$. Factor $p(t) = q(t)(t - c)$, and let $z = q((D))x$. Show that $z(t_0) = 0$ and $z' - cz = 0$ to conclude that $z = 0$. Now apply the induction hypothesis.

15. Let V be the solution space of an nth-order homogeneous linear differential equation with constant coefficients. Fix $t_0 \in R$, and define a mapping $\Phi\colon \mathsf{V} \to \mathsf{C}^n$ by

$$\Phi(x) = \begin{pmatrix} x(t_0) \\ x'(t_0) \\ \vdots \\ x^{(n-1)}(t_0) \end{pmatrix} \quad \text{for each } x \text{ in } \mathsf{V}.$$

(a) Prove that Φ is linear and its null space is the zero subspace of V. Deduce that Φ is an isomorphism. *Hint:* Use Exercise 14.

(b) Prove the following: For any nth-order homogeneous linear differential equation with constant coefficients, any $t_0 \in R$, and any complex numbers $c_0, c_1, \ldots, c_{n-1}$ (not necessarily distinct), there exists exactly one solution, x, to the given differential equation such that $x(t_0) = c_0$ and $x^{(k)}(t_0) = c_k$ for $k = 1, 2, \ldots n - 1$.

16. *Pendular Motion.* It is well known that the motion of a pendulum is approximated by the differential equation

$$\theta'' + \frac{g}{l}\theta = 0,$$

where $\theta(t)$ is the angle in radians that the pendulum makes with a vertical line at time t (see Figure 2.8), interpreted so that θ is positive if the pendulum is to the right and negative if the pendulum is to the

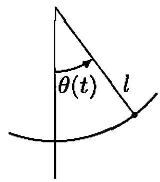


Figure 2.8

left of the vertical line as viewed by the reader. Here l is the length of the pendulum and g is the magnitude of acceleration due to gravity. The variable t and constants l and g must be in compatible units (e.g., t in seconds, l in meters, and g in meters per second per second).

(a) Express an arbitrary solution to this equation as a linear combination of two real-valued solutions.

(b) Find the unique solution to the equation that satisfies the conditions

$$\theta(0) = \theta_0 > 0 \quad \text{and} \quad \theta'(0) = 0.$$

(The significance of these conditions is that at time $t = 0$ the pendulum is released from a position displaced from the vertical by θ_0.)

(c) Prove that it takes $2\pi\sqrt{l/g}$ units of time for the pendulum to make one circuit back and forth. (This time is called the **period** of the pendulum.)

17. *Periodic Motion of a Spring without Damping.* Find the general solution to (3), which describes the periodic motion of a spring, ignoring frictional forces.

18. *Periodic Motion of a Spring with Damping.* The ideal periodic motion described by solutions to (3) is due to the ignoring of frictional forces. In reality, however, there is a frictional force acting on the motion that is proportional to the speed of motion, but that acts in the opposite direction. The modification of (3) to account for the frictional force, called the *damping force*, is given by

$$my'' + ry' + ky = 0,$$

where $r > 0$ is the proportionality constant.

(a) Find the general solution to this equation.

(b) Find the unique solution in (a) that satisfies the initial conditions $y(0) = 0$ and $y'(0) = v_0$, the initial velocity.

(c) For $y(t)$ as in (b), show that the amplitude of the oscillation decreases to zero; that is, prove that $\lim_{t \to \infty} y(t) = 0$.

19. In our study of differential equations, we have regarded solutions as complex-valued functions even though functions that are useful in describing physical motion are real-valued. Justify this approach.

20. The following parts, which do not involve linear algebra, are included for the sake of completeness.

(a) Prove Theorem 2.27. *Hint:* Use mathematical induction on the number of derivatives possessed by a solution.

(b) For any $c, d \in C$, prove that

$$e^{c+d} = c^c e^d \quad \text{and} \quad e^{-c} = \frac{1}{e^c}.$$

(c) Prove Theorem 2.28.

(d) Prove Theorem 2.29.

(e) Prove the product rule for differentiating complex-valued functions of a real variable: For any differentiable functions x and y in $\mathcal{F}(R, C)$, the product xy is differentiable and

$$(xy)' = x'y + xy'.$$

Hint: Apply the rules of differentiation to the real and imaginary parts of xy.

(f) Prove that if $x \in \mathcal{F}(R, C)$ and $x' = 0$, then x is a constant function.

INDEX OF DEFINITIONS FOR CHAPTER 2