

If  $U$  is another linear transformation from  $V$  into  $W$  and  $B = [B_1, \dots, B_n]$  is the matrix of  $U$  relative to the ordered bases  $\mathfrak{B}, \mathfrak{B}'$  then  $cA + B$  is the matrix of  $cT + U$  relative to  $\mathfrak{B}, \mathfrak{B}'$ . That is clear because

$$\begin{aligned} cA_j + B_j &= c[T\alpha_j]_{\mathfrak{B}'} + [U\alpha_j]_{\mathfrak{B}'} \\ &= [cT\alpha_j + U\alpha_j]_{\mathfrak{B}'} \\ &= [(cT + U)\alpha_j]_{\mathfrak{B}'} \end{aligned}$$

**Theorem 12.** *Let  $V$  be an  $n$ -dimensional vector space over the field  $F$  and let  $W$  be an  $m$ -dimensional vector space over  $F$ . For each pair of ordered bases  $\mathfrak{B}, \mathfrak{B}'$  for  $V$  and  $W$  respectively, the function which assigns to a linear transformation  $T$  its matrix relative to  $\mathfrak{B}, \mathfrak{B}'$  is an isomorphism between the space  $L(V, W)$  and the space of all  $m \times n$  matrices over the field  $F$ .*

*Proof.* We observed above that the function in question is linear, and as stated in Theorem 11, this function is one-one and maps  $L(V, W)$  onto the set of  $m \times n$  matrices. ■

We shall be particularly interested in the representation by matrices of linear transformations of a space into itself, i.e., linear operators on a space  $V$ . In this case it is most convenient to use the same ordered basis in each case, that is, to take  $\mathfrak{B} = \mathfrak{B}'$ . We shall then call the representing matrix simply the **matrix of  $T$  relative to the ordered basis  $\mathfrak{B}$** . Since this concept will be so important to us, we shall review its definition. If  $T$  is a linear operator on the finite-dimensional vector space  $V$  and  $\mathfrak{B} = \{\alpha_1, \dots, \alpha_n\}$  is an ordered basis for  $V$ , the matrix of  $T$  relative to  $\mathfrak{B}$  (or, the matrix of  $T$  in the ordered basis  $\mathfrak{B}$ ) is the  $n \times n$  matrix  $A$  whose entries  $A_{ij}$  are defined by the equations

$$(3-5) \quad T\alpha_j = \sum_{i=1}^n A_{ij}\alpha_i, \quad j = 1, \dots, n.$$

One must always remember that this matrix representing  $T$  depends upon the ordered basis  $\mathfrak{B}$ , and that there is a representing matrix for  $T$  in each ordered basis for  $V$ . (For transformations of one space into another the matrix depends upon two ordered bases, one for  $V$  and one for  $W$ .) In order that we shall not forget this dependence, we shall use the notation

$$[T]_{\mathfrak{B}}$$

for the matrix of the linear operator  $T$  in the ordered basis  $\mathfrak{B}$ . The manner in which this matrix and the ordered basis describe  $T$  is that for each  $\alpha$  in  $V$

$$[T\alpha]_{\mathfrak{B}} = [T]_{\mathfrak{B}}[\alpha]_{\mathfrak{B}}.$$

**EXAMPLE 13.** Let  $V$  be the space of  $n \times 1$  column matrices over the field  $F$ ; let  $W$  be the space of  $m \times 1$  matrices over  $F$ ; and let  $A$  be a fixed  $m \times n$  matrix over  $F$ . Let  $T$  be the linear transformation of  $V$  into  $W$  defined by  $T(X) = AX$ . Let  $\mathfrak{B}$  be the ordered basis for  $V$  analogous to the

standard basis in  $F^n$ , i.e., the  $i$ th vector in  $\mathfrak{B}$  in the  $n \times 1$  matrix  $X_i$  with a 1 in row  $i$  and all other entries 0. Let  $\mathfrak{B}'$  be the corresponding ordered basis for  $W$ , i.e., the  $j$ th vector in  $\mathfrak{B}'$  is the  $m \times 1$  matrix  $Y_j$  with a 1 in row  $j$  and all other entries 0. Then the matrix of  $T$  relative to the pair  $\mathfrak{B}, \mathfrak{B}'$  is the matrix  $A$  itself. This is clear because the matrix  $AX_j$  is the  $j$ th column of  $A$ .

EXAMPLE 14. Let  $F$  be a field and let  $T$  be the operator on  $F^2$  defined by

$$T(x_1, x_2) = (x_1, 0).$$

It is easy to see that  $T$  is a linear operator on  $F^2$ . Let  $\mathfrak{B}$  be the standard ordered basis for  $F^2$ ,  $\mathfrak{B} = \{\epsilon_1, \epsilon_2\}$ . Now

$$T\epsilon_1 = T(1, 0) = (1, 0) = 1\epsilon_1 + 0\epsilon_2$$

$$T\epsilon_2 = T(0, 1) = (0, 0) = 0\epsilon_1 + 0\epsilon_2$$

so the matrix of  $T$  in the ordered basis  $\mathfrak{B}$  is

$$[T]_{\mathfrak{B}} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

EXAMPLE 15. Let  $V$  be the space of all polynomial functions from  $R$  into  $R$  of the form

$$f(x) = c_0 + c_1x + c_2x^2 + c_3x^3$$

that is, the space of polynomial functions of degree three or less. The differentiation operator  $D$  of Example 2 maps  $V$  into  $V$ , since  $D$  is 'degree decreasing.' Let  $\mathfrak{B}$  be the ordered basis for  $V$  consisting of the four functions  $f_1, f_2, f_3, f_4$  defined by  $f_i(x) = x^{i-1}$ . Then

$$(Df_1)(x) = 0, \quad Df_1 = 0f_1 + 0f_2 + 0f_3 + 0f_4$$

$$(Df_2)(x) = 1, \quad Df_2 = 1f_1 + 0f_2 + 0f_3 + 0f_4$$

$$(Df_3)(x) = 2x, \quad Df_3 = 0f_1 + 2f_2 + 0f_3 + 0f_4$$

$$(Df_4)(x) = 3x^2, \quad Df_4 = 0f_1 + 0f_2 + 3f_3 + 0f_4$$

so that the matrix of  $D$  in the ordered basis  $\mathfrak{B}$  is

$$[D]_{\mathfrak{B}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

We have seen what happens to representing matrices when transformations are added, namely, that the matrices add. We should now like to ask what happens when we compose transformations. More specifically, let  $V$ ,  $W$ , and  $Z$  be vector spaces over the field  $F$  of respective dimensions  $n$ ,  $m$ , and  $p$ . Let  $T$  be a linear transformation from  $V$  into  $W$  and  $U$  a linear transformation from  $W$  into  $Z$ . Suppose we have ordered bases

$$\mathfrak{B} = \{\alpha_1, \dots, \alpha_n\}, \quad \mathfrak{B}' = \{\beta_1, \dots, \beta_m\}, \quad \mathfrak{B}'' = \{\gamma_1, \dots, \gamma_p\}$$

for the respective spaces  $V$ ,  $W$ , and  $Z$ . Let  $A$  be the matrix of  $T$  relative to the pair  $\mathfrak{B}$ ,  $\mathfrak{B}'$  and let  $B$  be the matrix of  $U$  relative to the pair  $\mathfrak{B}'$ ,  $\mathfrak{B}''$ . It is then easy to see that the matrix  $C$  of the transformation  $UT$  relative to the pair  $\mathfrak{B}$ ,  $\mathfrak{B}''$  is the product of  $B$  and  $A$ ; for, if  $\alpha$  is any vector in  $V$

$$\begin{aligned} [T\alpha]_{\mathfrak{B}'} &= A[\alpha]_{\mathfrak{B}} \\ [U(T\alpha)]_{\mathfrak{B}''} &= B[T\alpha]_{\mathfrak{B}'} \end{aligned}$$

and so

$$[(UT)(\alpha)]_{\mathfrak{B}''} = BA[\alpha]_{\mathfrak{B}}$$

and hence, by the definition and uniqueness of the representing matrix, we must have  $C = BA$ . One can also see this by carrying out the computation

$$\begin{aligned} (UT)(\alpha_j) &= U(T\alpha_j) \\ &= U\left(\sum_{k=1}^m A_{kj}\beta_k\right) \\ &= \sum_{k=1}^m A_{kj}(U\beta_k) \\ &= \sum_{k=1}^m A_{kj} \sum_{i=1}^p B_{ik}\gamma_i \\ &= \sum_{i=1}^p \left(\sum_{k=1}^m B_{ik}A_{kj}\right)\gamma_i \end{aligned}$$

so that we must have

$$(3-6) \quad C_{ij} = \sum_{k=1}^m B_{ik}A_{kj}.$$

We motivated the definition (3-6) of matrix multiplication via operations on the rows of a matrix. One sees here that a very strong motivation for the definition is to be found in composing linear transformations. Let us summarize formally.

**Theorem 13.** *Let  $V$ ,  $W$ , and  $Z$  be finite-dimensional vector spaces over the field  $F$ ; let  $T$  be a linear transformation from  $V$  into  $W$  and  $U$  a linear transformation from  $W$  into  $Z$ . If  $\mathfrak{B}$ ,  $\mathfrak{B}'$ , and  $\mathfrak{B}''$  are ordered bases for the spaces  $V$ ,  $W$ , and  $Z$ , respectively, if  $A$  is the matrix of  $T$  relative to the pair  $\mathfrak{B}$ ,  $\mathfrak{B}'$ , and  $B$  is the matrix of  $U$  relative to the pair  $\mathfrak{B}'$ ,  $\mathfrak{B}''$ , then the matrix of the composition  $UT$  relative to the pair  $\mathfrak{B}$ ,  $\mathfrak{B}''$  is the product matrix  $C = BA$ .*

We remark that Theorem 13 gives a proof that matrix multiplication is associative—a proof which requires no calculations and is independent of the proof we gave in Chapter 1. We should also point out that we proved a special case of Theorem 13 in Example 12.

It is important to note that if  $T$  and  $U$  are linear operators on a space  $V$  and we are representing by a single ordered basis  $\mathfrak{B}$ , then Theorem 13 assumes the simple form  $[UT]_{\mathfrak{B}} = [U]_{\mathfrak{B}}[T]_{\mathfrak{B}}$ . Thus in this case, the

correspondence which  $\mathfrak{B}$  determines between operators and matrices is not only a vector space isomorphism but also preserves products. A simple consequence of this is that the linear operator  $T$  is invertible if and only if  $[T]_{\mathfrak{B}}$  is an invertible matrix. For, the identity operator  $I$  is represented by the identity matrix in any ordered basis, and thus

$$UT = TU = I$$

is equivalent to

$$[U]_{\mathfrak{B}}[T]_{\mathfrak{B}} = [T]_{\mathfrak{B}}[U]_{\mathfrak{B}} = I.$$

Of course, when  $T$  is invertible

$$[T^{-1}]_{\mathfrak{B}} = [T]_{\mathfrak{B}}^{-1}.$$

Now we should like to inquire what happens to representing matrices when the ordered basis is changed. For the sake of simplicity, we shall consider this question only for linear operators on a space  $V$ , so that we can use a single ordered basis. The specific question is this. Let  $T$  be a linear operator on the finite-dimensional space  $V$ , and let

$$\mathfrak{B} = \{\alpha_1, \dots, \alpha_n\} \quad \text{and} \quad \mathfrak{B}' = \{\alpha'_1, \dots, \alpha'_n\}$$

be two ordered bases for  $V$ . How are the matrices  $[T]_{\mathfrak{B}}$  and  $[T]_{\mathfrak{B}'}$  related? As we observed in Chapter 2, there is a unique (invertible)  $n \times n$  matrix  $P$  such that

$$(3-7) \quad [\alpha]_{\mathfrak{B}} = P[\alpha]_{\mathfrak{B}'}$$

for every vector  $\alpha$  in  $V$ . It is the matrix  $P = [P_1, \dots, P_n]$  where  $P_j = [\alpha'_j]_{\mathfrak{B}}$ . By definition

$$(3-8) \quad [T\alpha]_{\mathfrak{B}} = [T]_{\mathfrak{B}}[\alpha]_{\mathfrak{B}}.$$

Applying (3-7) to the vector  $T\alpha$ , we have

$$(3-9) \quad [T\alpha]_{\mathfrak{B}} = P[T\alpha]_{\mathfrak{B}'}$$

Combining (3-7), (3-8), and (3-9), we obtain

$$[T]_{\mathfrak{B}}P[\alpha]_{\mathfrak{B}'} = P[T\alpha]_{\mathfrak{B}'}$$

or

$$P^{-1}[T]_{\mathfrak{B}}P[\alpha]_{\mathfrak{B}'} = [T\alpha]_{\mathfrak{B}'}$$

and so it must be that

$$(3-10) \quad [T]_{\mathfrak{B}'} = P^{-1}[T]_{\mathfrak{B}}P.$$

This answers our question.

Before stating this result formally, let us observe the following. There is a unique linear operator  $U$  which carries  $\mathfrak{B}$  onto  $\mathfrak{B}'$ , defined by

$$U\alpha_j = \alpha'_j, \quad j = 1, \dots, n.$$

This operator  $U$  is invertible since it carries a basis for  $V$  onto a basis for

V. The matrix  $P$  (above) is precisely the matrix of the operator  $U$  in the ordered basis  $\mathfrak{B}$ . For,  $P$  is defined by

$$\alpha'_j = \sum_{i=1}^n P_{ij} \alpha_i$$

and since  $U\alpha_j = \alpha'_j$ , this equation can be written

$$U\alpha_j = \sum_{i=1}^n P_{ij} \alpha_i.$$

So  $P = [U]_{\mathfrak{B}}$ , by definition.

**Theorem 14.** Let  $V$  be a finite-dimensional vector space over the field  $F$ , and let

$$\mathfrak{B} = \{\alpha_1, \dots, \alpha_n\} \quad \text{and} \quad \mathfrak{B}' = \{\alpha'_1, \dots, \alpha'_n\}$$

be ordered bases for  $V$ . Suppose  $T$  is a linear operator on  $V$ . If  $P = [P_1, \dots, P_n]$  is the  $n \times n$  matrix with columns  $P_j = [\alpha'_j]_{\mathfrak{B}}$ , then

$$[T]_{\mathfrak{B}'} = P^{-1}[T]_{\mathfrak{B}}P.$$

Alternatively, if  $U$  is the invertible operator on  $V$  defined by  $U\alpha_j = \alpha'_j$ ,  $j = 1, \dots, n$ , then

$$[T]_{\mathfrak{B}'} = [U]_{\mathfrak{B}}^{-1}[T]_{\mathfrak{B}}[U]_{\mathfrak{B}}.$$

EXAMPLE 16. Let  $T$  be the linear operator on  $R^2$  defined by  $T(x_1, x_2) = (x_1, 0)$ . In Example 14 we showed that the matrix of  $T$  in the standard ordered basis  $\mathfrak{B} = \{\epsilon_1, \epsilon_2\}$  is

$$[T]_{\mathfrak{B}} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Suppose  $\mathfrak{B}'$  is the ordered basis for  $R^2$  consisting of the vectors  $\epsilon'_1 = (1, 1)$ ,  $\epsilon'_2 = (2, 1)$ . Then

$$\begin{aligned} \epsilon'_1 &= \epsilon_1 + \epsilon_2 \\ \epsilon'_2 &= 2\epsilon_1 + \epsilon_2 \end{aligned}$$

so that  $P$  is the matrix

$$P = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}.$$

By a short computation

$$P^{-1} = \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}.$$

Thus

$$\begin{aligned} [T]_{\mathfrak{B}'} &= P^{-1}[T]_{\mathfrak{B}}P \\ &= \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} -1 & -2 \\ 1 & 2 \end{bmatrix}. \end{aligned}$$

We can easily check that this is correct because

$$\begin{aligned}T\epsilon'_1 &= (1, 0) = -\epsilon'_1 + \epsilon'_2 \\T\epsilon'_2 &= (2, 0) = -2\epsilon'_1 + 2\epsilon'_2.\end{aligned}$$

EXAMPLE 17. Let  $V$  be the space of polynomial functions from  $R$  into  $R$  which have 'degree' less than or equal to 3. As in Example 15, let  $D$  be the differentiation operator on  $V$ , and let

$$\mathfrak{B} = \{f_1, f_2, f_3, f_4\}$$

be the ordered basis for  $V$  defined by  $f_i(x) = x^{i-1}$ . Let  $t$  be a real number and define  $g_i(x) = (x + t)^{i-1}$ , that is

$$\begin{aligned}g_1 &= f_1 \\g_2 &= tf_1 + f_2 \\g_3 &= t^2f_1 + 2tf_2 + f_3 \\g_4 &= t^3f_1 + 3t^2f_2 + 3tf_3 + f_4.\end{aligned}$$

Since the matrix

$$P = \begin{bmatrix} 1 & t & t^2 & t^3 \\ 0 & 1 & 2t & 3t^2 \\ 0 & 0 & 1 & 3t \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

is easily seen to be invertible with

$$P^{-1} = \begin{bmatrix} 1 & -t & t^2 & -t^3 \\ 0 & 1 & -2t & 3t^2 \\ 0 & 0 & 1 & -3t \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

it follows that  $\mathfrak{B}' = \{g_1, g_2, g_3, g_4\}$  is an ordered basis for  $V$ . In Example 15, we found that the matrix of  $D$  in the ordered basis  $\mathfrak{B}$  is

$$[D]_{\mathfrak{B}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The matrix of  $D$  in the ordered basis  $\mathfrak{B}'$  is thus

$$\begin{aligned}P^{-1}[D]_{\mathfrak{B}}P &= \begin{bmatrix} 1 & -t & t^2 & t^3 \\ 0 & 1 & -2t & 3t^2 \\ 0 & 0 & 1 & -3t \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & t & t^2 & t^3 \\ 0 & 1 & 2t & 3t^2 \\ 0 & 0 & 1 & 3t \\ 0 & 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & -t & t^2 & t^3 \\ 0 & 1 & -2t & 3t^2 \\ 0 & 0 & 1 & -3t \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2t & 3t^2 \\ 0 & 0 & 2 & 6t \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.\end{aligned}$$