If ϵ is small, say $\epsilon = 0.01$, then

$$\mathbf{x}_2 = 50\begin{pmatrix} 0.505 \\ -1 \end{pmatrix} \quad \text{and} \quad \mathbf{x}_3 = (50)^2\begin{pmatrix} 1.51 \\ -3.01 \end{pmatrix}$$

After only three iterations, the sequence $\mathbf{x}_k$ has settled; the vector $\mathbf{x}_3$ provides a good description of the near nullspace of $\mathbf{T}$. If $\epsilon = 0$, $\mathbf{T}$ is singular; $\mathbf{x}_3$ lies almost in the nullspace of this singular operator (Figure 2.7). Were we to try other starting vectors $\mathbf{x}_0$, we would obtain other vectors $\mathbf{x}_k$ nearly parallel to $(-1 \;\; 2)^T$. This near nullspace of $\mathbf{T}$ should be considered one-dimensional.

We note from Example 7 that the vector $\mathbf{x}_k$ in the inverse iteration grows drastically in size. Practical computer implementations of inverse iteration include normalization of $\mathbf{x}_k$ at each step in order to avoid numbers too large for the computer. A description for a two-dimensional near nullspace is sought in P&C 2.26. In Section 4.2 we analyze the inverse iteration more precisely in terms of eigenvalues and eigenvectors. Forsythe [2.3] gives some interesting examples of the treatment of nearly singular operators.

The Role of Linear Transformations

The purpose of modeling a system is to develop insight concerning the system, to develop an intuitive feel for the input-output relationship. In order to decide whether or not a particular model, linear or nonlinear, is a good model, we must compare the input-output relationship of the model with the corresponding, but measurable, input-output relationship of the system being modeled. If the model and the system are sufficiently in agreement for our purposes, we need not distinguish between the system and the model.

Almost all physical systems are to some degree nonlinear. Yet most systems act in a nearly linear manner if the range of variation of the variables is restricted. For example, the current through a resistor is essentially proportional to the applied voltage if the current is not large enough to heat the resistor significantly. We are able to develop adequate models for a wide variety of static and dynamic physical systems using only linear transformations. For linear models there is available a vast array of mathematical results; most mathematical analysis is linear analysis. Furthermore, the analysis or optimization of a *nonlinear* system is usually based on linearization (Chapters 7 and 8). Even in solving a nonlinear equation for a given input, we typically must resort to repetitive linearization.

The examples and exercises of this section have demonstrated the variety of familiar transformations which are linear: matrix multiplication, differentiation, integration, etc. We introduce other linear transformations as we need them. The next few chapters pertain only to linear transformations. In Chapter 3 we focus on the peculiarities of linear differential systems. In Chapter 4 we develop the concepts of spectral decomposition of linear systems. The discussion of infinite-dimensional systems in Chapter 5 is also directed toward linear systems. Because we use the symbols **T** and **U** so much in reference to linear transformations, hereinafter we employ the symbols **F** and **G** to emphasize concepts which apply as well to nonlinear transformations. We begin to examine nonlinear concepts in Chapter 6. We do not return fully to the subject of nonlinear systems, however, until we introduce the concepts of linearization and repetitive linearization in Chapters 7 and 8.

2.5 Matrices of Linear Transformations

By the process of picking an ordered basis for an n-dimensional vector space $\mathcal{V}$, we associate with each vector in $\mathcal{V}$ a unique $n \times 1$ column matrix. In effect, we convert the vectors in $\mathcal{V}$ into an equivalent set of vectors which are suitable for matrix manipulation and, therefore, automatic computation by computer. By taking coordinates, we can also convert a linear equation, $\mathbf{Tx} = \mathbf{y}$, into a matrix equation. Suppose $\mathbf{T}$: $\mathcal{V} \to \mathcal{W}$ is a linear transformation, $\dim(\mathcal{V}) = n$, and $\dim(\mathcal{W}) = m$. Pick as bases for $\mathcal{V}$ and $\mathcal{W}$ the sets $\mathcal{X} \triangleq \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ and $\mathcal{Y} \triangleq \{\mathbf{y}_1, \ldots, \mathbf{y}_m\}$, respectively. The vectors $\mathbf{x}$ in $\mathcal{V}$ and $\mathbf{Tx}$ in $\mathcal{W}$ can be represented by their coordinate matrices $[\mathbf{x}]_{\mathcal{X}}$ and $[\mathbf{Tx}]_{\mathcal{Y}}$. The vectors $\mathbf{x}$ and $\mathbf{Tx}$ are linearly related (by the linear transformation $\mathbf{T}$). By (2.41), we know that a vector and its coordinates are also linearly related. Therefore, we expect $[\mathbf{x}]_{\mathcal{X}}$ and $[\mathbf{Tx}]_{\mathcal{Y}}$ to be linearly related as well. Furthermore, we intuitively expect the linear relation between the $n \times 1$ matrix $[\mathbf{x}]_{\mathcal{X}}$ and the $m \times 1$ matrix $[\mathbf{Tx}]_{\mathcal{Y}}$ to be multiplication by an $m \times n$ matrix. We denote this matrix by $[\mathbf{T}]_{\mathcal{X}\mathcal{Y}}$ and refer to it as the **matrix of T relative to the ordered bases $\mathcal{X}$ and $\mathcal{Y}$**; it must satisfy

$$[\mathbf{T}]_{\mathcal{X}\mathcal{Y}}[\mathbf{x}]_{\mathcal{X}} \triangleq [\mathbf{Tx}]_{\mathcal{Y}} \tag{2.45}$$

for all x in $\mathcal{V}$. Assume we can find such a matrix. Then by taking coordinates (with respect to $\mathcal{Y}$) of each side of the linear equation $\mathbf{Tx} = \mathbf{y}$, we convert the equation to the equivalent matrix equation.

$$[\mathbf{T}]_{\mathcal{X}\mathcal{Y}}[\mathbf{x}]_{\mathcal{X}} = [\mathbf{y}]_{\mathcal{Y}} \tag{2.46}$$

We will show that we can represent any linear transformation of $\mathcal{V}$ into $\mathcal{W}$ by a matrix multiplication by selecting bases for $\mathcal{V}$ and $\mathcal{W}$—we can convert any linear equation involving finite-dimensional vector spaces into a matrix equation. We first show how to determine the matrix of $\mathbf{T}$, then we show that it satisfies the defining equation (2.45) for all vectors $\mathbf{x}$ in $\mathcal{V}$.

Example 1. Determining the Matrix of a Linear Transformation Let $\mathbf{x}=(\xi_1, \xi_2, \xi_3)$, an arbitrary vector in $\mathfrak{R}^3$. Define $\mathbf{T}: \mathfrak{R}^3 \to \mathfrak{R}^2$ by

$$\mathbf{T}(\xi_1,\xi_2,\xi_3) \triangleq (2\xi_2-\xi_1, \xi_1+\xi_2+\xi_3)$$

We now find $[\mathbf{T}]_{\mathcal{E}_3\mathcal{E}_2}$, where $\mathcal{E}_3$ and $\mathcal{E}_2$ are the standard bases for $\mathfrak{R}^3$ and $\mathfrak{R}^2$, respectively. By (2.45), we have

$$[\mathbf{T}]_{\mathcal{E}_3\mathcal{E}_2}[(\xi_1,\xi_2,\xi_3)]_{\mathcal{E}_3}=[(2\xi_2-\xi_1,\xi_1+\xi_2+\xi_3)]_{\mathcal{E}_2}$$

for all vectors (ξ_1, ξ_2, ξ_3), or

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} = \begin{pmatrix} 2\xi_2-\xi_1 \\ \xi_1+\xi_2+\xi_3 \end{pmatrix} \tag{2.47}$$

where we have used $\{a_{ij}\}$ to represent the elements of $[\mathbf{T}]_{\mathcal{E}_3\mathcal{E}_2}$. By making three independent choices of the scalars ξ_1, ξ_2, and ξ_3, we could convert this matrix equation into six equations in the six unknowns $\{a_{ij}\}$. However, by using a little ingenuity, we reduce this effort. Think of the matrix multiplication in terms of the columns of $[\mathbf{T}]_{\mathcal{E}_3\mathcal{E}_2}$. The ith element of $[\mathbf{x}]_{\mathcal{E}_3}$ multiplies the ith column of $[\mathbf{T}]_{\mathcal{E}_3\mathcal{E}_2}$. If we choose $\mathbf{x}=(1, 0, 0)$, then $[(1, 0, 0)]_{\mathcal{E}_3}=\begin{pmatrix}1\\0\\0\end{pmatrix}$, and (2.47) becomes

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

We have found the first column of $[\mathbf{T}]_{\mathcal{E}_3\mathcal{E}_2}$ directly. We obtain the other two columns of $[\mathbf{T}]_{\mathcal{E}_3\mathcal{E}_2}$ from (2.47) by successive substitution of $\mathbf{x}=(0, 1, 0)$ and $\mathbf{x}=(0, 0, 1)$. The result is

$$[\mathbf{T}]_{\mathcal{E}_3\mathcal{E}_2}=\begin{pmatrix} -1 & 2 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

In Example 1 we avoided the need for simultaneous equations by substituting the basis vectors $\boldsymbol{\varepsilon}_1$, $\boldsymbol{\varepsilon}_2$, and $\boldsymbol{\varepsilon}_3$ into (2.47) to pick out the columns of $[\mathbf{T}]_{\mathcal{E}_3\mathcal{E}_2}$. This same technique can be used to find the matrix of any linear transformation acting on a finite-dimensional space. We refer again to $\mathbf{T}: \mathcal{V} \to \mathcal{W}$, with $\dim(\mathcal{V}) = n$, $\dim(\mathcal{W}) = m$, $\mathcal{X}$ a basis for $\mathcal{V}$, and $\mathcal{Y}$ a basis for $\mathcal{W}$. If we substitute into (2.45) the vector $\mathbf{x}_i$, the ith vector of the basis $\mathcal{X}$, we pick out the ith column of $[\mathbf{T}]_{\mathcal{X}\mathcal{Y}}$:

$$[\mathbf{T}]_{\mathcal{X}\mathcal{Y}}[\mathbf{x}_i]_{\mathcal{X}}=[\mathbf{T}]_{\mathcal{X}\mathcal{Y}}\begin{pmatrix}0\\ \vdots\\ 0\\ 1_i\\ 0\\ \vdots\\ 0\end{pmatrix}=i\text{th column of }[\mathbf{T}]_{\mathcal{X}\mathcal{Y}}=[\mathbf{T}\mathbf{x}_i]_{\mathcal{Y}}$$

We can find each column of $[\mathbf{T}]_{\mathcal{X}\mathcal{Y}}$ independently. The only computational effort is that in determining the coordinate matrices $[\mathbf{T}\mathbf{x}_i]_{\mathcal{Y}}$. Therefore,

$$[\mathbf{T}]_{\mathcal{X}\mathcal{Y}}=([\mathbf{T}\mathbf{x}_1]_{\mathcal{Y}}\ \vdots\ [\mathbf{T}\mathbf{x}_2]_{\mathcal{Y}}\ \vdots\ \cdots\ \vdots\ [\mathbf{T}\mathbf{x}_n]_{\mathcal{Y}}) \tag{2.48}$$

Example 2. The Matrix of a Linear Operator. Define the differential operator $\mathbf{D}: \mathcal{P}^3\to\mathcal{P}^3$ as in (2.36). The set $\mathcal{N}\triangleq\{\mathbf{f}_1,\mathbf{f}_2,\mathbf{f}_3\}$, where $\mathbf{f}_1(t)=1$, $\mathbf{f}_2(t)=t$, $\mathbf{f}_3(t)=t^2$, is a natural basis for $\mathcal{P}^3$. We use (2.48) to find

$$\begin{aligned}[\mathbf{D}]_{\mathcal{N}\mathcal{N}}&=([\mathbf{D}\mathbf{f}_1]_{\mathcal{N}}\ \vdots\ [\mathbf{D}\mathbf{f}_2]_{\mathcal{N}}\ \vdots\ [\mathbf{D}\mathbf{f}_3]_{\mathcal{N}})\\ &=([\boldsymbol{\theta}]_{\mathcal{N}}\ \vdots\ [\mathbf{f}_1]_{\mathcal{N}}\ \vdots\ [2\mathbf{f}_2]_{\mathcal{N}})\\ &=\begin{pmatrix}0 & \vdots & 1 & \vdots & 0\\ 0 & \vdots & 0 & \vdots & 2\\ 0 & \vdots & 0 & \vdots & 0\end{pmatrix}\end{aligned}$$

From the method used to determine $[\mathbf{T}]_{\mathcal{X}\mathcal{Y}}$ in (2.48), we know that this matrix correctly represents the action of $\mathbf{T}$ on the basis vectors $\{\mathbf{x}_i\}$. We now show that the matrix (2.48) also represents correctly the action of $\mathbf{T}$ on all other vectors in $\mathcal{V}$. An arbitrary vector $\mathbf{x}$ in $\mathcal{V}$ may be written in terms of the basis vectors for $\mathcal{V}$:

$$\mathbf{x}=\sum_{i=1}^{n}\xi_i\mathbf{x}_i$$

Since the transformation $\mathbf{T}$ is linear,

$$\mathbf{T}\mathbf{x}=\sum_{i=1}^{n}\xi_i\mathbf{T}\mathbf{x}_i$$

Because the process of taking coordinates is linear [see (2.41)],

$$\begin{aligned}[\mathbf{Tx}]_{\mathcal{Y}} &= \sum_{i=1}^{n} \xi_i [\mathbf{Tx}_i]_{\mathcal{Y}} \\ &= \left([\mathbf{Tx}_1]_{\mathcal{Y}} \;\vdots\; \cdots \;\vdots\; [\mathbf{Tx}_n]_{\mathcal{Y}}\right)\begin{pmatrix}\xi_1\\ \vdots \\ \xi_n\end{pmatrix} \\ &= [\mathbf{T}]_{\mathcal{X}\mathcal{Y}}[\mathbf{x}]_{\mathcal{X}}\end{aligned}$$

Thus, continuing Example 2 above, if $\mathbf{f}$ is the arbitrary vector defined by $\mathbf{f}(t) \stackrel{\Delta}{=} \xi_1 + \xi_2 t + \xi_3 t^2$, then

$$(\mathbf{Df})(t) = \xi_2 + 2\xi_3 t,\ [\mathbf{f}]_{\mathcal{N}} = \begin{pmatrix}\xi_1\\ \xi_2\\ \xi_3\end{pmatrix},\ [\mathbf{Df}]_{\mathcal{N}} = \begin{pmatrix}\xi_2\\ 2\xi_3\\ 0\end{pmatrix},\ \text{and } [\mathbf{D}]_{\mathcal{N}\mathcal{N}}[\mathbf{f}]_{\mathcal{N}} = [\mathbf{Df}]_{\mathcal{N}}$$

When the domain and range space of $\mathbf{T}$ are identical, and the same basis is used for both spaces (as it is in Example 2), we sometimes refer to the matrix $[\mathbf{T}]_{\mathcal{X}\mathcal{X}}$ as the **matrix of the operator T relative to the basis** $\mathcal{X}$**.**

We expect the matrix of a linear transformation to possess the basic characteristics of that transformation. The only basic characteristics of a linear transformation that we have discussed thus far are its rank and nullity. The picking of coordinate systems $\mathcal{X}$ and $\mathcal{Y}$ converts the transformation equation $\mathbf{Tx} = \mathbf{y}$ to a precisely equivalent matrix equation, $[\mathbf{Tx}]_{\mathcal{Y}} = [\mathbf{T}]_{\mathcal{X}\mathcal{Y}}[\mathbf{x}]_{\mathcal{X}} = [\mathbf{y}]_{\mathcal{Y}}$; for every $\mathbf{x}$ and $\mathbf{y}$ in the one equation, there is a unique $[\mathbf{x}]_{\mathcal{X}}$ and $[\mathbf{y}]_{\mathcal{Y}}$ in the other. The dimensions of the nullspace and range of the transformation "multiplication by $[\mathbf{T}]_{\mathcal{X}\mathcal{Y}}$" must be the same, therefore, as the dimensions of the nullspace and range of $\mathbf{T}$. We speak loosely of the rank and nullity of $[\mathbf{T}]_{\mathcal{X}\mathcal{Y}}$ when we actually mean the rank and nullity of the transformation "multiplication by $[\mathbf{T}]_{\mathcal{X}\mathcal{Y}}$." We refer to the nullity and rank of a matrix as if it were the matrix of a linear transformation. The nullspace and range of matrix multiplications are explored in P&C 2.19; the problem demonstrates that for an $m \times n$ matrix $\mathbf{A}$,

$$\begin{aligned}\operatorname{rank}(\mathbf{A}) &= \text{the number of independent columns of } \mathbf{A} \\ &= \text{the number of independent rows of } \mathbf{A} \\ \operatorname{nullity}(\mathbf{A}) &= n - \operatorname{rank}(\mathbf{A}) \\ \operatorname{nullity}(\mathbf{A}^{\mathrm{T}}) &= m - \operatorname{rank}(\mathbf{A})\end{aligned}$$

Once again referring to Example 2, we see that the nullity of $\mathbf{D}$ is 1 [the vector $\mathbf{f}_1$ is a basis for nullspace($\mathbf{D}$)]. The nullity of $[\mathbf{D}]_{\mathcal{N}\mathcal{N}}$ is also 1 ($[\mathbf{D}]_{\mathcal{N}\mathcal{N}}$ contains one dependent column). The matrix $[\mathbf{D}]_{\mathcal{N}\mathcal{N}}$ does possess the same nullity and rank as the operator $\mathbf{D}$.

It is apparent that determination of the matrix of a transformation reduces to the determination of coordinate matrices for the set of vectors $\{\mathbf{Tx}_i\}$ of (2.48). We found in Section 2.2 that determination of the coordinate matrix of a vector $\mathbf{x}$ with respect to a basis $\mathcal{X} = \{\mathbf{x}_i\}$ can be reduced to performing elimination on the matrix equation (2.17):

$$[\mathbf{x}]_{\mathcal{N}} = ([\mathbf{x}_1]_{\mathcal{N}} \vdots \cdots \vdots [\mathbf{x}_n]_{\mathcal{N}})[\mathbf{x}]_{\mathcal{X}}$$

where $\mathcal{N}$ is a natural basis for the space $\mathcal{V}$ of which $\mathbf{x}$ is a member (i.e., a basis with respect to which coordinates can be determined by inspection).

Exercise 1. Show that $[\mathbf{T}]_{\mathcal{X}\mathcal{Y}}$ of (2.48) can be obtained by the row reduction

$$\left([\mathbf{y}_1]_{\mathcal{N}} \vdots \cdots \vdots [\mathbf{y}_n]_{\mathcal{N}} \vdots [\mathbf{Tx}_1]_{\mathcal{N}} \vdots \cdots \vdots [\mathbf{Tx}_n]_{\mathcal{N}}\right) \rightarrow \left(\mathbf{I} \vdots [\mathbf{T}]_{\mathcal{X}\mathcal{Y}}\right) \tag{2.49}$$

where $\mathcal{N}$ is a natural basis for the range of definition $\mathcal{W}$. (Hint: if the elements of $[\mathbf{Tx}_i]_{\mathcal{Y}}$ are denoted by $[\mathbf{Tx}_i]_{\mathcal{Y}} = (c_{1i} \cdots c_{ni})^{\mathrm{T}}$, then $\mathbf{Tx}_i = \sum_j c_{ji}\mathbf{y}_i$, and $[\mathbf{Tx}_i]_{\mathcal{N}} = \sum_j c_{ji}[\mathbf{y}_i]_{\mathcal{N}}$.) Use this approach to find $[\mathbf{T}]_{\mathcal{E}_3\mathcal{E}_2}$ of Example 1.

Example 3. The Matrix of a Matrix Transformation. Let $\mathbf{T}: \mathcal{M}^{n\times 1} \rightarrow \mathcal{M}^{m\times 1}$ be defined by $\mathbf{Tx} \triangleq \mathbf{Ax}$, where $\mathbf{A}$ is an $m \times n$ matrix. Denoting the standard bases for $\mathcal{M}^{n\times 1}$ and $\mathcal{M}^{m\times 1}$ by $\mathcal{E}_n$ and $\mathcal{E}_m$, respectively, we find $[\mathbf{T}]_{\mathcal{E}_n\mathcal{E}} = \mathbf{A}$. Although $[\mathbf{x}]_{\mathcal{X}}$ and $\mathbf{x}$ are identical in this example, we should distinguish between them, for it is certainly incorrect to equate the matrix $[\mathbf{T}]_{\mathcal{E}_n\mathcal{E}_m}$ to the transformation $\mathbf{T}$.

Suppose $\mathbf{T}: \mathcal{V} \rightarrow \mathcal{W}$ is invertible and linear; $\mathcal{V}$ and $\mathcal{W}$ are finite-dimensional with bases $\mathcal{X}$ and $\mathcal{Y}$, respectively. It follows from (2.45) that

$$[\mathbf{T}^{-1}]_{\mathcal{Y}\mathcal{X}}[\mathbf{y}]_{\mathcal{Y}} = [\mathbf{T}^{-1}\mathbf{y}]_{\mathcal{X}} \tag{2.50}$$

for all y in $\mathcal{W}$. Then, for each x in $\mathcal{V}$,

$$[\mathbf{x}]_{\mathcal{X}} = [\mathbf{T}^{-1}\mathbf{Tx}]_{\mathcal{X}} = [\mathbf{T}^{-1}]_{\mathcal{Y}\mathcal{X}}[\mathbf{Tx}]_{\mathcal{Y}} = [\mathbf{T}^{-1}]_{\mathcal{Y}\mathcal{X}}[\mathbf{T}]_{\mathcal{X}\mathcal{Y}}[\mathbf{x}]_{\mathcal{X}}$$

A similar relationship can be established with $\mathbf{T}$ and $\mathbf{T}^{-1}$ reversed. Then as a consequence of (2.29),

$$[\mathbf{T}^{-1}]_{\mathcal{Y}\mathcal{X}} = [\mathbf{T}]_{\mathcal{X}\mathcal{Y}}^{-1} \tag{2.51}$$

Exercise 2. Suppose $\mathcal{V}$, $\mathcal{W}$, and $\mathcal{U}$ are finite-dimensional vector spaces with bases $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$, respectively. Show that

a. If $\mathbf{T}: \mathcal{V} \to \mathcal{W}$ and $\mathbf{U}: \mathcal{V} \to \mathcal{W}$ are linear, then

$$[a\mathbf{T}+b\mathbf{U}]_{\mathcal{X}\mathcal{Y}} = a[\mathbf{T}]_{\mathcal{X}\mathcal{Y}} + b[\mathbf{U}]_{\mathcal{X}\mathcal{Y}} \tag{2.52}$$

b. If $\mathbf{T}: \mathcal{V} \to \mathcal{W}$ and $\mathbf{U}: \mathcal{W} \to \mathcal{U}$ are linear, then

$$[\mathbf{UT}]_{\mathcal{X}\mathcal{Z}} = [\mathbf{U}]_{\mathcal{Y}\mathcal{Z}}[\mathbf{T}]_{\mathcal{X}\mathcal{Y}} \tag{2.53}$$

Changes in Coordinate System

In Chapter 4 we discuss coordinate systems which are particularly suitable for analysis of a given linear transformation—coordinate systems for which the matrix of the transformation is diagonal. In preparation for that discussion we now explore the effect of a change of coordinate system on a coordinate matrix $[\mathbf{x}]$ and on the matrix of a transformation $[\mathbf{T}]$.

Suppose $\mathcal{X}$ and $\mathcal{Z}$ are two different bases for an n-dimensional vector space $\mathcal{V}$. We know by (2.41) that the transformations

$$\mathbf{x} \to [\mathbf{x}]_{\mathcal{X}} \quad \text{and} \quad \mathbf{x} \to [\mathbf{x}]_{\mathcal{Z}}$$

are linear and invertible. Thus we expect $[\mathbf{x}]_{\mathcal{X}}$ and $[\mathbf{x}]_{\mathcal{Z}}$ to be related by

$$\mathbf{S}[\mathbf{x}]_{\mathcal{X}} = [\mathbf{x}]_{\mathcal{Z}} \tag{2.54}$$

where $\mathbf{S}$ is an $n \times n$ invertible matrix. In fact, multiplication of $[\mathbf{x}]_{\mathcal{X}}$ by any invertible matrix represents a change from the coordinate system $\mathcal{X}$ to some new coordinate system. We sometimes denote the matrix $\mathbf{S}$ of (2.54) by the symbol $\mathbf{S}_{\mathcal{X}\mathcal{Z}}$, thereby making explicit the fact that $\mathbf{S}$ converts coordinates relative to $\mathcal{X}$ into coordinates relative to $\mathcal{Z}$. Then $(\mathbf{S}_{\mathcal{X}\mathcal{Z}})^{-1} = S_{\mathcal{Z}\mathcal{X}}$.

Determination of the specific change-of-coordinates matrix $\mathbf{S}$ defined in (2.54) follows the same line of thought as that used to determine $[\mathbf{T}]$ in (2.48). By successively substituting into (2.54) the vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ from the basis $\mathcal{X}$, we isolate the columns of $\mathbf{S}$: the ith column of $\mathbf{S}$ is $[\mathbf{x}_i]_{\mathcal{Z}}$. Thus the unique invertible matrix $\mathbf{S}$ which transforms coordinate matrices relative to $\mathcal{X}$ into coordinate matrices relative to $\mathcal{Z}$ is

$$\mathbf{S} = \mathbf{S}_{\mathcal{X}\mathcal{Z}} = ([\mathbf{x}_1]_{\mathcal{Z}} \vdots \cdots \vdots [\mathbf{x}_n]_{\mathcal{Z}}) \tag{2.55}$$

where the $\mathbf{x}_i$ are the vectors in the basis $\mathcal{X}$.

Since a change-of-coordinates matrix is always invertible, we determine from (2.54) that

$$\mathbf{S}^{-1}[\mathbf{x}]_{\mathfrak{Z}} = [\mathbf{x}]_{\mathfrak{X}}$$

and

$$\mathbf{S}^{-1} = \mathbf{S}_{\mathfrak{X}\mathfrak{Z}}^{-1} = \mathbf{S}_{\mathfrak{Z}\mathfrak{X}} = ([\mathbf{z}_1]_{\mathfrak{X}} \vdots \cdots \vdots [\mathbf{z}_n]_{\mathfrak{X}}) \tag{2.56}$$

where the $\mathbf{z}_i$ are the vectors in the basis $\mathfrak{Z}$. If $\mathfrak{Z}$ is a natural basis for the space, then $\mathbf{S}$ can be found by inspection. On the other hand, if $\mathfrak{X}$ is a natural basis, we find $\mathbf{S}^{-1}$ by inspection. It is appropriate to use either (2.55) or (2.56) in determining $\mathbf{S}$. We need both $\mathbf{S}$ and $\mathbf{S}^{-1}$ to allow conversion back and forth between the two coordinate systems. Besides, the placing of $\mathbf{S}$ on the left side of (2.54) was arbitrary.

Example 4. A Change-of-Coordinates Matrix, Let $\mathcal{E}$ be the standard basis for $\mathfrak{R}^3$. Another basis for $\mathfrak{R}^3$ is $\mathfrak{Z} = \{\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3\}$, where $\mathbf{z}_1 = (1, 1, 1)$, $\mathbf{z}_2 = (1, 1, 0)$, and $\mathbf{z}_3 = (1, 0, 0)$. Since $\mathcal{E}$ is a natural basis for $\mathfrak{R}^3$, we use (2.56) to find

$$\mathbf{S}^{-1} = ([\mathbf{z}_1]_{\mathcal{E}} \vdots [\mathbf{z}_2]_{\mathcal{E}} \vdots [\mathbf{z}_3]_{\mathcal{E}}) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \tag{2.57}$$

A straightforward elimination (Section 1.5) yields

$$\mathbf{S} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix} \tag{2.58}$$

We note that for an arbitrary vector $\mathbf{x} = (\xi_1, \xi_2, \xi_3)$ in $\mathfrak{R}^3$, $[\mathbf{x}]_{\mathcal{E}} = (\xi_1 \ \xi_2 \ \xi_3)^{\mathrm{T}}$. By (2.54),

$$[\mathbf{x}]_{\mathfrak{Z}} = \mathbf{S}[\mathbf{x}]_{\mathcal{E}} = (\xi_3 \quad \xi_2 - \xi_3 \quad \xi_1 - \xi_2)^{\mathrm{T}} \tag{2.59}$$

But then,

$$\begin{aligned} \mathbf{x} &= (\xi_3)\mathbf{z}_1 + (\xi_2 - \xi_3)\mathbf{z}_2 + (\xi_1 - \xi_2)\mathbf{z}_3 \\ &= (\xi_3)(1,1,1) + (\xi_2 - \xi_3)(1,1,0) + (\xi_1 - \xi_2)(1,0,0) \\ &= (\xi_1, \xi_2, \xi_3) \end{aligned} \tag{2.60}$$

and the validity of the change of coordinates matrix $\mathbf{S}$ is verified.

If neither $\mathfrak{X}$ nor $\mathfrak{Z}$ is a natural basis, the determination of $\mathbf{S}$ can still be systematized by the introduction of an intermediate step which does involve a natural basis.

Exercise 3. Suppose we need the change-of-coordinates matrix $\mathbf{S}$ such that $\mathbf{S}[\mathbf{x}]_{\mathcal{X}} = [\mathbf{x}]_{\mathcal{Z}}$, where neither $\mathcal{X}$ nor $\mathcal{Z}$ is a natural basis for $\mathcal{V}$. Suppose $\mathcal{N}$ is a natural basis. Show, by introducing an intermediate change to the coordinates $[\mathbf{x}]_{\mathcal{N}}$, that

$$\mathbf{S}=\left([\mathbf{z}_1]_{\mathcal{N}} \vdots \cdots \vdots [\mathbf{z}_n]_{\mathcal{N}}\right)^{-1}\left([\mathbf{x}_1]_{\mathcal{N}} \vdots \cdots \vdots [\mathbf{x}_n]_{\mathcal{N}}\right) \tag{2.61}$$

Example 5. Change of Coordinates via an Intermediate Natural Basis. Two bases for $\mathcal{P}^3$ are $\mathcal{F} \triangleq \{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$ and $\mathcal{G} \triangleq \{\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3\}$, where

$$\mathbf{f}_1(t)=1, \quad \mathbf{f}_2(t)=1+t, \quad \mathbf{f}_3(t)=1+t^2$$

$$\mathbf{g}_1(t)=1+t, \quad \mathbf{g}_2(t)=t, \quad \mathbf{g}_3(t)=t+t^2$$

To find $\mathbf{S}$ such that $\mathbf{S}[\mathbf{f}]_{\mathcal{F}} = [\mathbf{f}]_{\mathcal{G}}$, we introduce the natural basis $\mathcal{N} \triangleq \{\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3\}$, where $\mathbf{h}_i(t) = t^{i-1}$. Then, by (2.61),

$$\mathbf{S}= ([\mathbf{g}_1]_{\mathcal{N}} \vdots [\mathbf{g}_2]_{\mathcal{N}} \vdots [\mathbf{g}_3]_{\mathcal{N}})^{-1}([\mathbf{f}_1]_{\mathcal{N}} \vdots [\mathbf{f}_2]_{\mathcal{N}} \vdots [\mathbf{f}_3]_{\mathcal{N}})$$

$$=\begin{pmatrix}1&0&0\\1&1&1\\0&0&1\end{pmatrix}^{-1}\begin{pmatrix}1&1&1\\0&1&0\\0&0&1\end{pmatrix}$$

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

Similarity and Equivalence Transformations

Now that we have a process for changing coordinate systems, we explore the effect of such a change on the matrix of a transformation. Suppose $\mathbf{T}$ is a linear operator on $\mathcal{V}$, and that $\mathcal{X}$ and $\mathcal{Z}$ are two different bases for $\mathcal{V}$. Then $[\mathbf{T}]_{\mathcal{X}\mathcal{X}}$ is defined by

$$[\mathbf{T}]_{\mathcal{X}\mathcal{X}}[\mathbf{x}]_{\mathcal{X}} = [\mathbf{T}\mathbf{x}]_{\mathcal{X}}$$

The change from the $\mathcal{X}$ to the $\mathcal{Z}$ coordinate system is described by

$$\mathbf{S}[\mathbf{x}]_{\mathcal{X}} = [\mathbf{x}]_{\mathcal{Z}}$$

The change-of-coordinates matrix $\mathbf{S}$ also applies to the vector $\mathbf{T}\mathbf{x}$ in $\mathcal{V}$:

$$\mathbf{S}[\mathbf{T}\mathbf{x}]_{\mathcal{X}} = [\mathbf{T}\mathbf{x}]_{\mathcal{Z}}$$

By substituting $[\mathbf{x}]_{\mathfrak{X}}$ and $[\mathbf{Tx}]_{\mathfrak{X}}$ from these last two equations into the defining equation for $[\mathbf{Tx}]_{\mathfrak{Z}}$, we find

$$[\mathbf{T}]_{\mathfrak{X}\mathfrak{X}}\mathbf{S}^{-1}[\mathbf{x}]_{\mathfrak{Z}} = \mathbf{S}^{-1}[\mathbf{Tx}]_{\mathfrak{Z}}$$

or

$$\left(\mathbf{S}[\mathbf{T}]_{\mathfrak{X}\mathfrak{X}}\mathbf{S}^{-1}\right)[\mathbf{x}]_{\mathfrak{Z}} = [\mathbf{Tx}]_{\mathfrak{Z}}$$

But this is the defining equation for $[\mathbf{T}]_{\mathfrak{Z}\mathfrak{Z}}$. It is apparent that

$$[\mathbf{T}]_{\mathfrak{Z}\mathfrak{Z}} = \mathbf{S}[\mathbf{T}]_{\mathfrak{X}\mathfrak{X}}\mathbf{S}^{-1} \tag{2.62}$$

where $\mathbf{S}$ converts from the $\mathfrak{X}$ coordinate system to the $\mathfrak{Z}$ coordinate system. Equation (2.62) describes an invertible linear transformation on $[\mathbf{T}]_{\mathfrak{X}\mathfrak{X}}$ known as a **similarity transformation.** In Section 4.2, we find that a similarity transformation preserves the basic spectral properties of the matrix. It is comforting to know that any two matrix representations of a linear system have the same properties-these properties are inherent in the model, $\mathbf{T}$, and should not be affected by the coordinate system we select.

Example 6. A Similarity Transformation. In Example 2 we found the matrix of the differential operator on $\mathcal{P}^3$ relative to the natural basis for $\mathcal{P}^3$:

$$[\mathbf{D}]_{\mathfrak{N}\mathfrak{N}} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

Another basis for $\mathcal{P}^3$ is $\mathcal{G} = \{\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3\}$, where $\mathbf{g}_1(t) = 1 + t$, $\mathbf{g}_2(t) = t$, and $\mathbf{g}_3(t) = t + t^2$. The change-of-coordinates matrix which relates the two bases $\mathfrak{N}$ and $\mathcal{G}$ is defined by $\mathbf{S}[\mathbf{f}]_{\mathfrak{N}} = [\mathbf{f}]_{\mathcal{G}}$; we find it using (2.56):

$$\mathbf{S}^{-1} = \left([\mathbf{g}_1]_{\mathfrak{N}} \vdots [\mathbf{g}_2]_{\mathfrak{N}} \vdots [\mathbf{g}_3]_{\mathfrak{N}}\right)$$
$$= \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

The inverse matrix is

$$\mathbf{S} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$$