2

System Models : Transformations on Vector Spaces

The fundamental purpose in modeling a system is to develop a mechanism for predicting the condition or change in condition of the system. In the abstract model $\mathbf{Tx} = \mathbf{y}$ of (1.1), $\mathbf{T}$ represents (or is a model of) the system, whereas $\mathbf{x}$ and $\mathbf{y}$ have to do with the condition of the system. We explore first some familiar models for the condition or changes in condition of systems. These examples lead us to use a generalization of the usual notion of a vector as a model for the condition of a system. We then develop the concept of a transformation of vectors as a model of the system itself. The rest of the chapter is devoted to examination of the most commonly used models-linear models-and their matrix representations.

2.1 The Condition of a System

The physical condition (or change in condition) of many simple systems has been found to possess a magnitude and a direction in our physical three-dimensional space. It is natural, therefore, that a mathematical concept of condition (or change in condition) has developed over time which has these two properties; this concept is the vector. Probably the most obvious example of the use of this concept is the use of arrows in a two-dimensional plane to represent changes in the position of an object on the two-dimensional surface of the earth (see Figure 2.1). Using the usual techniques of analytic geometry, we can represent each such arrow by a pair of numbers that indicates the components of that arrow along each of a pair of coordinate axes. Thus pairs of numbers serve as an equivalent model for changes in position.

An ordinary road map is another model for the two-dimensional surface of the earth. It is equivalent to the arrow diagram; points on the map are

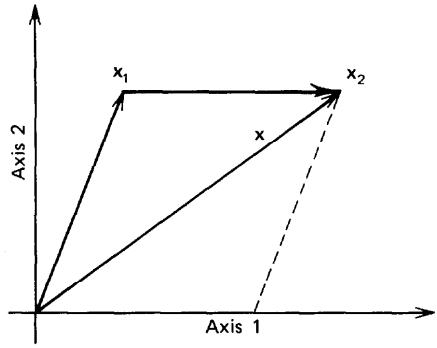

Figure 2.1. An "arrow vector" diagram.

equivalent to the arrow tips of Figure 2.1. The only significant difference between these two models is that the map emphasizes the position (or condition) of an object on the earth, whereas the arrow diagram stresses the changes in position and the manner in which intermediate changes in position add to yield a total change in position. We can also interpret a position on the map as a change from some reference position. The manner in which we combine arrows or changes in position (the parallelogram rule) is the most significant characteristic of either model. Consequently we focus on the arrow model which emphasizes the combination process.

Reference arrows (coordinate axes) are used to tie the arrow model to the physical world. By means of a reference position and a pair of reference "position changes" on the surface of the earth, we relate the positions and changes in position on the earth to positions and arrows in the arrow diagram. However, there are no inherent reference axes on either the physical earth or the two-dimensional plane of arrows.

The same vector model that we use to represent changes in position can be used to represent the forces acting at a point on a physical object. The reason we can use the same model is that the magnitudes and directions of forces also combine according to the parallelogram rule. The physical natures of the reference vectors are different in these three situations: in one case they are changes in position on the earth, in another they are arrows, in the third, forces. Yet once reference vectors are chosen in each, all three situations become in some sense equivalent; corresponding to each vector in one situation is a vector in the other two; corresponding to each sum of vectors in one is a corresponding sum in the other two. We use the set of arrows as a model for the other two situations because it is the most convenient of the three to work with.

The set of complex numbers is one more example of a set of objects which is equivalent to the set of arrows. We usually choose as references in

the set of complex numbers the two numbers 1 and i. Based on these reference numbers and two reference arrows, we interpret every arrow as a complex number. Here we have one set of mathematical (or geometrical) objects serving as a model for another set of mathematical objects.

Consider now a physical system which is more complicated than the two physical systems discussed above. Imagine a flat metal sheet exposed to the sun and partly submerged in a stream. (The sheet is representative of any object subject to heat sources and coolants.) The thermal condition of the sheet is described by the temperature distribution over the surface of the sheet. A change in the cloud cover in the sky will change the pattern in which the sun falls on the sheet. As a result, the temperature distribution will change. Assuming the temperature distribution reaches a new steady state, the new distribution equals the old distribution plus the change in the distribution. We model this situation as follows. Let (s,t) denote a position in some two-dimensional coordinate system on the surface of the sheet. Let $\mathbf{f}(s,$ be the temperature at the point $(s,$ measured in degrees centigrade, for all points (s,t) on the sheet. We model a change in the thermal condition of the sheet by

$$\mathbf{f}_{\text{new}}(s,t) = \mathbf{f}_{\text{old}}(s,t) + \mathbf{f}_{\text{change}}(s,t) \tag{2.1}$$

for all $(s,$ on the sheet. In effect, (2.1) defines $\mathbf{f}_{\text{change}}$. However, we hope to use a model of the system to *predict* $\mathbf{f}_{\text{change}}$. Then (2.1) will determine $\mathbf{f}_{\text{new}}$. Equation (2.1) is a "distributed" equivalent of the arrow diagram in Figure 2.1; each of these models illustrates the manner in which changes in condition combine to yield a net condition of the system in question. Once again, references have been chosen in both the physical system and the model (mathematical system) in order to equate the two systems; choosing physical units of measurement (degrees centigrade) amounts to fixing the relationship between the physical and mathematical systems.

The most significant difference between a system' modeled by Figure 2.1 and a system modeled by (2.1) consists in the nature of the conditions in each system. In one case we have a quantity with magnitude and direction (e.g., force); in the other, a quantity without magnitude and direction—a quantity that is distributed over a two-dimensional region. Yet there are important similarities between the two systems. The changes in condition of the system are under scrutiny; also, several changes in condition combine by simple rules to yield a total or net condition.

Vector Spaces

By expressing various types of problems in a common framework, we learn to use concepts derived from one type of problem in understanding other types of problems. In particular, we are able to draw useful analogies

between algebraic equations and differential equations by expressing both types of equations as "vector" equations. Therefore, we now generalize the common notion of a vector to include all the examples discussed in the previous section.

Definition. A **linear space** (or **vector space)** $\mathcal{V}$ is a set of elements $\mathbf{x}, \mathbf{y}, \mathbf{z}, \ldots,$ called vectors, together with definitions of *vector addition* and *scalar multiplication.*

a. The definition of vector addition is such that:
   1. To every pair, $\mathbf{x}$ and $\mathbf{y}$, of vectors in $\mathcal{V}$ there corresponds a unique vector $\mathbf{x}+\mathbf{y}$ in $\mathcal{V}$, called the **sum** of $\mathbf{x}$ and $\mathbf{y}$.
   2. $\mathbf{x}+\mathbf{y}=\mathbf{y}+\mathbf{x}$.
   3. $(\mathbf{x}+\mathbf{y})+\mathbf{z}=\mathbf{x}+(\mathbf{y}+\mathbf{z})$.
   4. There is a unique vector $\boldsymbol{\theta}$ in $\mathcal{V}$, called the **zero vector** (or origin), such that $\mathbf{x}+\boldsymbol{\theta}=\mathbf{x}$ for all $\mathbf{x}$ in $\mathcal{V}$.
   5. Corresponding to each $\mathbf{x}$ in $\mathcal{V}$ there is a unique vector "$-\mathbf{x}$" in $\mathcal{V}$ such that $\mathbf{x}+(-\mathbf{x})=\boldsymbol{\theta}$.

b. The definition of scalar multiplication is such that:
   1. To every vector $\mathbf{x}$ in $\mathcal{V}$ and every scalar a there corresponds a unique vector $a\mathbf{x}$ in $\mathcal{V}$, called the **scalar multiple** of $\mathbf{x}$.*
   2. $a(b\mathbf{x})=(ab)\mathbf{x}$.
   3. $1(\mathbf{x})=\mathbf{x}$ (where 1 is the unit scalar).
   4. $a(\mathbf{x}+\mathbf{y})=a\mathbf{x}+a\mathbf{y}$.
   5. $(a+b)\mathbf{x}=a\mathbf{x}+b\mathbf{x}$.

Notice that a vector space includes not only a set of elements (vectors) but also "valid" definitions of vector addition and scalar multiplication. Also inherent in the definition is the fact that the vector space $\mathcal{V}$ contains all "combinations" of its own vectors: if $\mathbf{x}$ and $\mathbf{y}$ are in $\mathcal{V}$, then $a\mathbf{x}+b\mathbf{y}$ is also in $\mathcal{V}$. The rules of algebra are so much a part of us that some of the requirements may at first appear above definition; however, they are necessary. A few more vector space properties which may be deduced from the above definition are as follows:

1. $0\mathbf{x}=\boldsymbol{\theta}$ (where "0" is the zero scalar).
2. $a\boldsymbol{\theta}=\boldsymbol{\theta}$.
3. $(-1)\mathbf{x}=-\mathbf{x}$.

Example 1. The Real 3-tuple Space $\mathcal{R}^3$. The space $\mathcal{R}^3$ consists in the set of all

*The scalars are any set of elements which obey the usual rules of algebra. A set of elements which obeys these rules constitutes a field (see Hoffman and Kunze [2.6]). We usually use as scalars either the real numbers or the complex numbers. There are other useful fields, however (P&C 2.4).

real 3-tuples (all ordered sequences of three real numbers), $\mathbf{x} = (\xi_1, \mathbf{y} = (\eta_1, \eta_2, \eta_3)$, with the following definitions of addition and scalar multiplication:

$$\begin{aligned} \mathbf{x} + \mathbf{y} &\triangleq (\xi_1 + \eta_1, \xi_2 + \eta_2, \xi_3 + \eta_3) \\ a\mathbf{x} &\triangleq (a\xi_1, a\xi_2, a\xi_3) \end{aligned} \tag{2.2}$$

It is clear that the zero vector for this 3-tuple space, $\boldsymbol{\theta} = (0,0,0)$, satisfies $\mathbf{x} + \boldsymbol{\theta} = \mathbf{x}$. We show that $\boldsymbol{\theta}$ is unique by assuming another vector $\mathbf{y}$ also satisfies $\mathbf{x} + \mathbf{y} = \mathbf{x}$; that is,

$$(\xi_1 + \eta_1, \xi_2 + \eta_2, \xi_3 + \eta_3) = (\xi_1, \xi_2, \xi_3)$$

or $\xi_i + \eta_i = \xi_i$. The properties of scalars then require $\eta_i = 0$ (or $\mathbf{y} = \boldsymbol{\theta}$). It is easy to prove that $\mathfrak{R}^3$, as defined above, satisfies the other requirements for a linear space. In each instance, questions about vectors are reduced to questions about scalars.

We emphasize that the definition of $\mathfrak{R}^3$ says nothing about coordinates. Coordinates are multipliers for reference vectors (reference arrows, for instance). The 3-tuples are vectors in their own right. However, there is a commonly used correspondence between $\mathfrak{R}^3$ and the set of vectors (arrows) in the usual three-dimensional space which makes it difficult not to think of the 3-tuples as coordinates. The two sets of vectors are certainly equivalent. We will, in fact, use this natural correspondence to help illustrate vector concepts graphically.

Example 2. The Two-Dimensional Space of Points (or Arrows). This space consists in the set of all points in a plane. Addition is defined by the parallelogram rule using a fixed reference point (see Figure 2.2). Scalar multiplication is defined as "length" multiplication using the reference point. The zero vector is obviously the

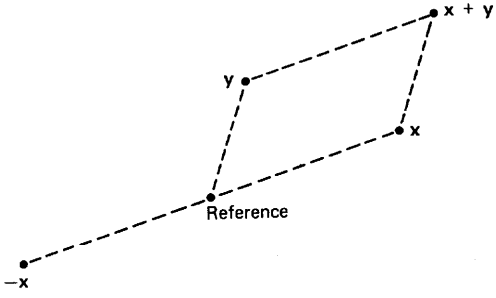

Figure 2.2. The two-dimensional space of points.

reference point. Each of the requirements can be verified by geometrical arguments.

An equivalent (but not identical) space is one where the vectors are not the points, but rather, arrows to the points from the reference point. We distinguish only the magnitude and direction of each arrow; *two parallel arrows of the same length are considered identical.*

Both the arrow space and the point space are easily visualized: we often use the arrow space in two or three dimensions to demonstrate concepts graphically. Although the arrow space contains no *inherent* reference arrows, we sometimes *specify* reference arrows in order to equate the arrows to vectors in $\mathcal{R}^3$. Because of the equivalence between vectors in $\mathcal{R}^3$ and vectors in the three-dimensional space of points, we occasionally refer to vectors in $\mathcal{R}^3$ and in other spaces as *points.*

Example 3. The Space of Column Vectors $\mathcal{M}^{3\times 1}$**.** The space $\mathcal{M}^{3\times 1}$ consists in the set of all real 3x1 column matrices (or column vectors), denoted by

$$\mathbf{x}=\begin{pmatrix}\xi_1\\ \xi_2\\ \xi_3\end{pmatrix} \qquad \mathbf{y}=\begin{pmatrix}\eta_1\\ \eta_2\\ \eta_3\end{pmatrix}$$

with the following definitions of addition and scalar multiplication:

$$\mathbf{x}+\mathbf{y}\triangleq\begin{pmatrix}\xi_1+\eta_1\\ \xi_2+\eta_2\\ \xi_3+\eta_3\end{pmatrix} \qquad a\mathbf{x}\triangleq\begin{pmatrix}a\xi_1\\ a\xi_2\\ a\xi_3\end{pmatrix} \tag{2.3}$$

In order to save space in writing, we occasionally write vectors from $\mathcal{M}^{3\times 1}$ in the transposed form $\mathbf{x} = (\xi_1\ \xi_2\ \xi_3)^T$. The equivalence between $\mathcal{M}^{3\times 1}$ and $\mathcal{R}^3$ is obvious. The only difference between the two vector spaces is in the nature of their vectors. Vectors in $\mathcal{M}^{3\times 1}$ can be multiplied by $m \times 3$ matrices (as in Section 1.5), whereas vectors in CR^3 cannot.

Example 4. The Space of Real Square-Summable Sequences, l_2**.** The space l_2 consists in the set of all infinite sequences of real numbers, $\mathbf{x} = (\xi_1, \xi_2, \xi_3, \ldots)$, $\mathbf{y} = (\eta_1, \eta_2, \eta_3, \ldots)$ which are square summable; that is, for which $\sum_{i=1}^{\infty}\xi_i^2 < \infty$. Addition and scalar multiplication in l_2 are defined by

$$\begin{aligned}\mathbf{x}+\mathbf{y}&\triangleq(\xi_1+\eta_1, \xi_2+\eta_2, \xi_3+\eta_3, \ldots)\\ a\mathbf{x}&\triangleq(a\xi_1, a\xi_2, a\xi_3, \ldots)\end{aligned} \tag{2.4}$$

Most of the properties required by the definition of a linear space are easily verified for l_2; for instance, the zero vector is obviously $\boldsymbol{\theta}$ = (0,0, . . .). However, there is one subtle difference between l_2 and the space $\mathcal{R}^3$ of Example 1. Because

the sequences in l_2 are infinite, it is not obvious that if $\mathbf{x}$ and $\mathbf{y}$ are in l_2, $\mathbf{x}+\mathbf{y}$ is also in l_2. It can be shown that

$$\sqrt{\Sigma_{i=1}^{\infty}(\xi_i+\eta_i)^2} \leqslant \sqrt{\Sigma_{i=1}^{\infty}\xi_i^2} + \sqrt{\Sigma_{i=1}^{\infty}\eta_i^2}$$

[This fact is known as the triangle inequality (P&C 5.4)]. Therefore,

$$\sum_{i=1}^{\infty}(\xi_i+\eta_i)^2 < \infty$$

and $\mathbf{x}+\mathbf{y}$ is square-summable. The requirement of square summability is a definite restriction on the elements of l_2; the simple sequence $(1, 1, 1, \ldots)$, for instance, is not in l_2.

The definition of $\mathcal{R}^3$ extends easily to $\mathcal{R}^n$, the space of n-tuples of real numbers (where n is a positive integer). The space $\mathcal{M}^{n\times 1}$ is a similar extension of $\mathcal{M}^{3\times 1}$. Mathematically these "n-dimensional" spaces are no more complicated than their three-dimensional counterparts. Yet we are not able to draw arrow-space equivalents because our physical world is three-dimensional. Visualization of an abstract vector space is most easily accomplished by thinking in terms of its three-dimensional counterpart.

The spaces $\mathcal{R}^n$, $\mathcal{M}^{n\times 1}$, and l_2 can also be redefined using complex numbers, rather than real numbers, for scalars. We denote by $\mathcal{R}_c^n$ the complex n-tuple space. We use the symbol $\mathcal{M}_c^{n\times 1}$ for the space of complex $n \times 1$ column vectors. Let l_2^c represent the space of complex square-summable sequences. (We need a slightly different definition of square summability for the space l_2^c: $\Sigma_{i=1}^{\infty}|\xi_i|^2 < \infty$). In most vector space definitions, either set of scalars can be used. A notable exception to interchangeability of scalars is the arrow space in two or three dimensions. The primary value of the arrow space is in graphical illustration. We have already discussed the equivalence of the set of complex scalars to the two-dimensional space of arrows. Therefore, substituting complex scalars in the real two-dimensional arrow space would require four-dimensional graphical illustration.

We eventually find it useful to combine simple vector spaces to form more complicated spaces.

Definition. Suppose $\mathcal{V}$ and $\mathcal{W}$ are vector spaces. We define the **Cartesian product** $\mathcal{V}\times\mathcal{W}$ of the spaces $\mathcal{V}$ and $\mathcal{W}$ to be the set of pairs of vectors $\mathbf{z} \triangleq (\mathbf{x},\mathbf{y})$, with $\mathbf{x}$ in $\mathcal{V}$ and $\mathbf{y}$ in $\mathcal{W}$. We define addition and scalar multiplication of vectors in $\mathcal{V}\times\mathcal{W}$ in terms of the corresponding operations in $\mathcal{V}$ and in $\mathcal{W}$: if $\mathbf{z}_1 = (\mathbf{x}_1,\mathbf{y}_1)$ and $\mathbf{z}_2 = (\mathbf{x}_2,\mathbf{y}_2)$, then

$$\mathbf{z}_1+\mathbf{z}_2 \triangleq (\mathbf{x}_1+\mathbf{x}_2, \mathbf{y}_1+\mathbf{y}_2)$$

$$a\mathbf{z}_1 \triangleq (a\mathbf{x}_1, a\mathbf{y}_1)$$

Example 5. A Cartesian Product. Let $x = (\xi_1, \xi_2)$, a vector in $\mathcal{R}^2$. Let $\mathbf{y} = (\eta_1)$, a vector in $\mathcal{R}^1$. Then $\mathbf{z} \triangleq ((\xi_1, \xi_2), (\eta_1))$ is a typical vector in $\mathcal{R}^2 \times \mathcal{R}^1$. This Cartesian product space is clearly equivalent to $\mathcal{R}^3$. Strictly speaking, however, $\mathbf{z}$ is not in $\mathcal{R}^3$. It is not a 3-tuple, but rather a 2-tuple followed by a 1-tuple. Yet we have no need to distinguish between $\mathcal{R}^3$ and $\mathcal{R}^2 \times \mathcal{R}^1$.

Function Spaces

Each vector in the above examples has discrete elements. It is a small conceptual step from the notion of an infinite sequence of discrete numbers (a vector in l_2) to the usual notion of a function—a "continuum" of numbers. Yet vectors and functions are seldom related in the thinking of engineers. We will find that vectors and functions can be viewed as essentially equivalent objects; functions can be treated as vectors, and vectors can be treated as functions. A **function space** is a linear space whose elements are functions. We usually think of a function as a rule or graph which associates with each scalar in its domain a single scalar value. We do not confuse the graph with particular values of the function. Our notation should also keep this distinction. Let $\mathbf{f}$ denote a **function;** that is, the symbol $\mathbf{f}$ recalls to mind a particular rule or graph. Let $\mathbf{f}(t)$ denote the **value of the function at** t. By $\mathbf{f} = \mathbf{g}$, we mean that the scalars $\mathbf{f}(t)$ and $\mathbf{g}(t)$ are equal for each t of interest.

Example 6. $\mathcal{P}^n$, The Polynomials of Degree Less Than n. The space $\mathcal{P}^n$ consists in all real-valued polynomial functions of degree less than n: $\mathbf{f}(t) = \xi_1 + \xi_2 t + \cdots + \xi_n t^{n-1}$ for all real t. Addition and scalar multiplication of vectors (functions) in $\mathcal{P}^n$ are defined by

$$(\mathbf{f}+\mathbf{g})(t) \triangleq \mathbf{f}(t)+\mathbf{g}(t)$$
$$(a\mathbf{f})(t) \triangleq a(\mathbf{f}(t)) \tag{2.5}$$

for all t. The zero function is $\boldsymbol{\theta}(t) = 0$ for all t. This zero function is unique; if the function $\mathbf{g}$ also satisfied $\mathbf{f} + \mathbf{g} = \mathbf{f}$, then the values of $\mathbf{f}$ and $\mathbf{g}$ would satisfy

$$(\mathbf{f}+\mathbf{g})(t) = \mathbf{f}(t)+\mathbf{g}(t) = \mathbf{f}(t)$$

It would follow that $\mathbf{g}(t) = 0$ for all t, or $\mathbf{g} = \boldsymbol{\theta}$. The other requirements for a vector space are easily verified for $\mathcal{P}^n$.

We emphasize that the vector $\mathbf{f}$ in Example 6 is the entire portrait of the function $\mathbf{f}$. The scalar variable t is a "dummy" variable. The only purpose of this variable is to order the values of the function in precisely the same way that the subscript i orders the elements in the following vector from l_2:

$$\mathbf{x} = (\xi_1, \xi_2, \ldots \xi_i, \ldots)$$

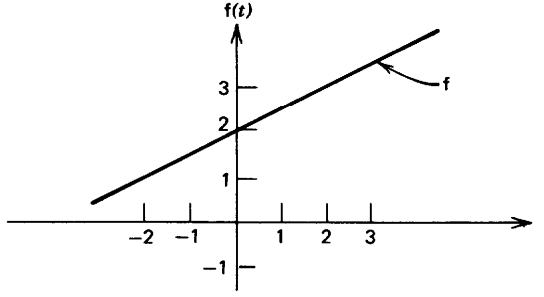

Figure 2.3. A function **f** and its values **f**(*t*).

Figure 2.3 distinguishes graphically between the vector **f** and its value at t for the specific function **f** defined by $\mathbf{f}(t) = 2 + 0.5\,t$. Figure 2.4 distinguishes in a similar manner between an infinite sequence **x** and its ith element.

It is evident that the vector **x** from l_2 is just as much a function as is the polynomial **f** from $\mathcal{P}^n$. In the space of polynomials, the index t is continuous; in the space of infinite sequences the index i is discrete—it takes on only positive integral values. In the latter case, we could as well refer to the ith element ξ_i as the value of **x** at i. In point of fact, most vector spaces can be interpreted as spaces of functions; the terms vector space and function space are somewhat interchangeable. However, it is common practice to use the term function space only for a space in which the index t varies continuously over an interval.

It is unfortunate that the symbol $\mathbf{f}(t)$ is commonly used to represent both a function and the value of that function at t. This blurring of the meaning of symbols is particularly true of the sinusoidal and exponential functions. We will try to be explicit in our distinction between the two concepts. As discussed in the preface, boldface type is used to *emphasize* the interpretation of a function as a vector. However, to avoid overuse of boldface type, it is not used where emphasis on the vector interpretation appears un-

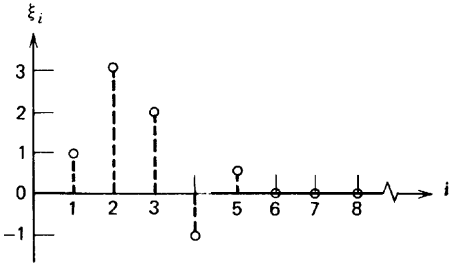

Figure 2.4. The elements ξ_i of an infinite sequence **x**.

necessary; thus the value of a function $\mathbf{f}$ at t may appear either as $\mathbf{f}(t)$ or as $f(t)$. Furthermore, where confusion is unlikely, we sometimes use standard mathematical shorthand; for example, we use $\int_a^b \mathbf{fg}\,dt$ to mean $\int_a^b \mathbf{f}(t)\mathbf{g}(t)\,dt$.

It is difficult to describe or discuss functions in any detail except in terms of their scalar values. In Example 6, for instance, the definitions of addition and scalar multiplication were given in terms of function values. Furthermore, we resorted again to function values to verify that the vector space requirements were met. We will find ourselves continually reducing questions about functions to questions about the scalar values of those functions. Why then do we emphasize the function $\mathbf{f}$ rather than the value f(t)? Because system models act on the whole vector $\mathbf{f}$ rather than on its individual values. As an example, we turn to the one system model we have explored thus far-the matrix equation $\mathbf{Ax} = \mathbf{y}$ which was introduced in Section 1.5. If $\mathbf{A}$ is an $m \times n$ matrix, the vector $\mathbf{x}$ is a column matrix in $\mathfrak{M}^{n\times 1}$; $\mathbf{y}$ is in $\mathfrak{M}^{m\times 1}$. Even though the matrix multiplication requires manipulation of the individual elements (or values) of $\mathbf{x}$, it is impossible to determine *any* element of $\mathbf{y}$ without operating on *all* elements of $\mathbf{x}$. Thus it is natural to think in terms of $\mathbf{A}$ operating on the whole vector $\mathbf{x}$. Similarly, equations involving functions require operations on the whole function (e.g., integration), as we shall see in Section 2.3.

Example 7. The Space $\mathcal{C}(a, b)$ of Continuous Functions. The vectors in $\mathcal{C}$ are those real functions which are defined and continuous on the interval $[a,b]$. Addition and scalar multiplication of functions in $\mathcal{C}(a, b)$ are defined by the standard function space definitions (2.5) for all t in $[a,b]$. It is clear that the sums and scalar multiples of continuous functions are also continuous functions.

Example 8. $\mathcal{L}_2(a,b)$ The Real Square-integrable Functions. The space $\mathcal{L}_2(a,b)$ consists in all real functions which are defined and square integrable on the interval $[a,b]$; that is, functions $\mathbf{f}$ for which*

$$\int_a^b \mathbf{f}^2(t)\,dt < \infty$$

Addition and scalar multiplication of functions in $\mathcal{L}_2(a,b)$ are defined by (2.5) for all t in $[a,$ The space $\mathcal{L}_2(a,b)$ is analogous to l_2. It is not clear that the sum of two square-integrable functions is itself square integrable. As in Example 4, we must rely on P&C 5.4 and the concepts of Chapter 5 to find that

$$\sqrt{\int_a^b [\mathbf{f}(t)+\mathbf{g}(t)]^2\,dt} \leq \sqrt{\int_a^b \mathbf{f}^2(t)\,dt} + \sqrt{\int_a^b \mathbf{g}^2(t)\,dt}$$

*The integral used in the definition of $\mathcal{L}_2(a,b)$ is the Lebesgue integral. For all practical purposes, Lebesgue integration can be considered the same as the usual Riemann integration. Whenever the Riemann integral exists, it yields the same result as the Lebesgue integral. (See Royden [2.l].)

It follows that if $\mathbf{f}$ and $\mathbf{g}$ are square integrable, then $\mathbf{f} + \mathbf{g}$ is square integrable.

Example 9. A Set of Functions. The set of positive real functions [together with the definitions of addition and scalar multiplication in (2.5)] does *not* form a vector space. This set contains a positive valued function $\mathbf{f}$, but not the negative valued function $-\mathbf{f}$; therefore, this set does not include all sums and multiples of its members.

Example 10. ***Functions of a Complex Variable.*** Let $\mathcal{V}$ be the space of all complex functions $\mathbf{w}$ of the complex variable z which are defined and analytic on some region Ω of the complex z plane.* For instance, Ω might be the circle $|z| \leqslant 1$. We define addition and scalar multiplication of functions in $\mathcal{V}$ by

$$\begin{aligned}(\mathbf{w}_1 + \mathbf{w}_2)(z) &\triangleq \mathbf{w}_1(z) + \mathbf{w}_2(z)\\ (a\mathbf{w})(z) &\triangleq a(\mathbf{w}(z))\end{aligned} \tag{2.6}$$

for all z in Ω. In this example, the zero vector $\boldsymbol{\theta}$ is defined by $\boldsymbol{\theta}(z) = 0$ for all z in Ω. (We do not care about the values of the functions $\boldsymbol{\theta}$ and $\mathbf{w}$ outside of Ω.)

Exercise 1. Show that if $\mathbf{w}_1$ and $\mathbf{w}_2$ are in the space $\mathcal{V}$ of Example 10, then $\mathbf{w}_1 + \mathbf{w}_2$ is also in $\mathcal{V}$.

Example 11. A Vector Space of Random Variables † A **random variable** $\mathbf{x}$ is a numerical-valued function whose domain consists in the possible outcomes of an experiment or phenomenon. Associated with the experiment is a probability distribution. Therefore, there is a probability distribution associated with the values of the random variable. For example, the throwing of a single die is an experiment. We define the random variable $\mathbf{x}$ in terms of the possible outcomes σ by

$$\begin{aligned}\mathbf{x}(\sigma) &\triangleq 0 \quad \text{for } \sigma = 2,4,6 \text{ (the die is even)}\\ &\triangleq 1 \quad \text{for } \sigma = 1,3,5 \text{ (the die is odd)}\end{aligned}$$

The probability mass function ω associated with the outcome σ of the experiment is given by

$$\omega(\sigma) = \tfrac{1}{6} \quad \text{for } \sigma = 1,2,3,4,5,6$$

*Express the complex variable z in the form $s + it$, where s and t are real. Let the complex function $\mathbf{w}$ be written as $\mathbf{u} + i\mathbf{v}$, where $\mathbf{u}(z)$ and $\mathbf{v}(z)$ are real. Then $\mathbf{w}$ is analytic in Ω if and only if the partial derivatives of $\mathbf{u}$ and $\mathbf{v}$ are continuous and satisfy the Cauchy-Riemann conditions in Ω:

$$\frac{\partial \mathbf{u}(z)}{\partial s} = \frac{\partial \mathbf{v}(z)}{\partial t}, \quad \frac{\partial \mathbf{v}(z)}{\partial s} = -\frac{\partial \mathbf{u}(z)}{\partial t}$$

For instance, $\mathbf{w}(z) \triangleq z^2$ is analytic in the whole z plane. See Wylie [2.11].

† See Papoulis [2.7], or Cramér and Leadbetter [2.2].

Then the probability mass function $\omega_{\mathbf{x}}$ associated with the values of the random variable $\mathbf{x}$ is

$$\omega_{\mathbf{x}}(x) = \tfrac{1}{2} \quad \text{for } x = 0, 1$$

We can define many other random variables (functions) for the same die-throwing experiment. One other random variable is

$$\mathbf{y}(\sigma) \triangleq 1 \quad \text{for } \sigma = 1 \text{ (the die is 1)}$$

$$\triangleq 0 \quad \text{for } \sigma = 2,3,4,5,6 \text{ (the die is not 1)}$$

where

$$\omega_{\mathbf{y}}(y) = \tfrac{5}{6} \quad \text{for } y = 0$$

$$= \tfrac{1}{6} \quad \text{for } y = 1$$

Two random variables $\mathbf{x}_1$ and $\mathbf{x}_2$ are equal if and only if their values $\mathbf{x}_1(\sigma)$ and $\mathbf{x}_2(\sigma)$ are identical for all possible outcomes σ of the experiment.

A vector **space of random variables** defined on a given experiment consists in a set of functions defined on the possible outcomes of the experiment, together with the following definitions of addition and scalar multiplication*:

$$(a\mathbf{x})(\sigma) \triangleq a(\mathbf{x}(\sigma)) \qquad (\mathbf{x}+\mathbf{y})(\sigma) \triangleq \mathbf{x}(\sigma)+\mathbf{y}(\sigma)$$

for all possible outcomes σ of the experiment. Let $\mathcal{V}$ be the space of all possible random variables defined on the above die-throwing experiment. If $\mathbf{x}$ and $\mathbf{y}$ are the particular vectors described above, then $\mathbf{x}+\mathbf{y}$ is given by

$$(\mathbf{x}+\mathbf{y})(\sigma) \triangleq 2 \quad \text{for } \sigma = 1$$

$$\triangleq 1 \quad \text{for } \sigma = 3,5$$

$$\triangleq 0 \quad \text{for } \sigma = 2,4,6$$

and

$$\omega_{\mathbf{x}+\mathbf{y}}(z) = \tfrac{1}{2} \quad \text{for } z = 0$$

$$= \tfrac{1}{3} \quad \text{for } z = 1$$

$$= \tfrac{1}{6} \quad \text{for } z = 2$$

What is the zero random variable for the vector space $\mathcal{V}$? It is $\boldsymbol{\theta}(\sigma) = 0$ for $\sigma = 1,\ldots,6$.

*We note that the set of functions must be such that it includes all sums and scalar multiples of its members.

2.2 Relations Among Vectors

Combining Vectors

Assuming a vector represents the condition or change in condition of a system, we can use the definitions of addition and scalar multiplication of vectors to find the net result of several successive changes in condition of the system.

Definition. A vector $\mathbf{x}$ is said to be a **linear combination** of the vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ if it can be expressed as

$$\mathbf{x} = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_n\mathbf{x}_n \tag{2.7}$$

for some set of scalars ci, . . . , cn. This concept is illustrated in Figure 2.5 where x= fxi +x2-xX,.

A vector space 'Y is simply a set of elements and a definition of linear combination (addition and scalar multiplication); the space V includes all linear combinations of its own elements. If $\boldsymbol{S}$ is a subset of ?r, the set of all linear combinations of vectors from $\boldsymbol{S}$, using the same definition of linear combination, is also a vector space. We call it a subspace of $\mathcal{V}$. A line or plane through the origin of the three-dimensional arrow space is an example of a subspace.

Definition. A subset $\mathcal{W}$ of a linear space $\mathcal{V}$ is a **linear subspace** (or **linear manifold)** of $\mathcal{V}$ if along with every pair, $\mathbf{x}_1$ and $\mathbf{x}_2$, of vectors in $\mathcal{W}$, every linear combination $c_1\mathbf{x}_1 + c_2\mathbf{x}_2$ is also in $\mathcal{W}$.* We call $\mathcal{W}$ a *proper subspace* if it is smaller than $\mathcal{V}$; that is if $\mathcal{W}$ is not $\mathcal{V}$ itself.

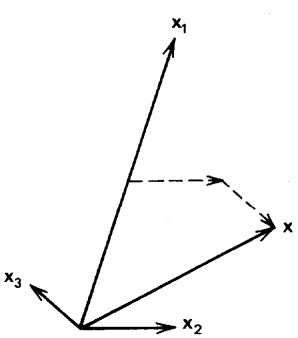

Figure 2.5. A linear combination of arrows.

*In the discussion of infinite-dimensional Hilbert spaces (Section 5.3), we distinguish between a linear subspace and a linear manifold. Linear manifold is the correct term to use in this definition. Yet because a finite-dimensional linear manifold is a linear subspace as well, we emphasize the physically motivated term subspace.

Example 1. A Linear Subspace. The set of vectors from $\mathcal{R}^3$ which are of the form $(c_1, c_2, c_1 + c_2)$ forms a subspace of $\mathcal{R}^3$. It is, in fact, the set of all linear combinations of the two vectors (1, 0, 1) and (0, 1, 1).

Example 2. A Solution Space. The set $\mathcal{W}$ of all solutions to the matrix equation

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 2 & 3 & 3 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

is a subspace of $\mathcal{M}^{3\times 1}$. By elimination (Section 1.5), we find that $\mathcal{W}$ contains all vectors of the form $(0\ \xi_2\ -\xi_2)^T$. Clearly, $\mathcal{W}$ consists in all linear combinations of the single vector $(0\ 1\ -1)^T$. This example extends to general matrices. Let A be an $m \times n$ matrix. Let $\mathbf{x}$ be in $\mathcal{M}^{n\times 1}$. Using the rules of matrix multiplication (Appendix 1) it can be shown that if $\mathbf{x}_1$ and $\mathbf{x}_2$ are solutions to $\mathbf{Ax} = \boldsymbol{\theta}$, then an arbitrary linear combination $c_1\mathbf{x}_1 + c_2\mathbf{x}_2$ is also a solution. Thus the space of solutions is a subspace of $\mathcal{M}^{n\times 1}$.

Example 3. Subspaces (Linear Manifolds) of Functions. Let $\mathcal{C}^2(\Omega)$ be the space of all real-valued functions which are defined and have continuous second partial derivatives in the two-dimensional region Ω. (This region could be the square $0 \leqslant s \leqslant 1, 0 \leqslant t \leqslant 1$, for instance.) Let Γ denote the boundary of the region Ω. Linear combination in $\mathcal{C}^2(\Omega)$ is defined by

$$\begin{aligned} (\mathbf{f}+\mathbf{g})(s,t) &\triangleq \mathbf{f}(s,t)+\mathbf{g}(s,t) \\ (a\mathbf{f})(s,t) &\triangleq a(\mathbf{f}(s,t)) \end{aligned} \tag{2.8}$$

for all (s, t) in Ω. The functions $\mathbf{f}$ in $\mathcal{C}^2(\Omega)$ which satisfy the homogeneous boundary condition $\mathbf{f}(s, t) = 0$ for (s, t) on Γ constitute a linear manifold of $\mathcal{C}^2(\Omega)$. For if $\mathbf{f}_1$ and $\mathbf{f}_2$ satisfy the boundary condition, then $(c_1\mathbf{f}_1 + c_2\mathbf{f}_2)(s, t) = c_1\mathbf{f}_1(s, t) + c_2\mathbf{f}_2(s, t) = 0$, and the arbitrary linear combination $c_1\mathbf{f}_1 + c_2\mathbf{f}_2$ also satisfies the boundary condition.

The set of solutions to Laplace's equation,

$$\frac{\partial^2 \mathbf{f}(s,t)}{\partial s^2} + \frac{\partial^2 \mathbf{f}(s,t)}{\partial t^2} = 0 \tag{2.9}$$

for all (s, t) in Ω, also forms a linear manifold of $\mathcal{C}^2(\Omega)$. For if $\mathbf{f}_1$ and $\mathbf{f}_2$ both satisfy (2.9), then

$$\frac{\partial^2 [c_1\mathbf{f}_1(s,t) + c_2\mathbf{f}_2(s,t)]}{\partial s^2} + \frac{\partial^2 [c_1\mathbf{f}_1(s,t) + c_2\mathbf{f}_2(s,t)]}{\partial t^2} = 0$$

and the arbitrary linear combination $c_1\mathbf{f}_1 + c_2\mathbf{f}_2$ also satisfies (2.9). Equation (2.9) is phrased in terms of the values of $\mathbf{f}$. Laplace's equation can also be expressed in the

vector notation

$$\nabla^2 \mathbf{f} = \boldsymbol{\theta} \tag{2.10}$$

The domain of definition Ω is implicit in (2.10). The vector $\boldsymbol{\theta}$ is defined by $\boldsymbol{\theta}(s, t) = 0$ for all (s, t) in Ω.

In using vector diagrams to analyze physical problems, we often resolve a vector into a linear combination of component vectors. We usually do this in a unique manner. In Figure 2.5, $\mathbf{x}$ is not a unique linear combination of $\mathbf{x}_1$, $\mathbf{x}_2$, and $\mathbf{x}_3$; $\mathbf{x} = 0\mathbf{x}_1 + 3\mathbf{x}_2 + 2\mathbf{x}_3$ is a second resolution of $\mathbf{x}$; the number of possible resolutions is infinite. In point of fact, $\mathbf{x}$ can be represented as a linear combination of any two of the other vectors; the three vectors $\mathbf{x}_1$, $\mathbf{x}_2$, and $\mathbf{x}_3$ are redundant as far as representation of $\mathbf{x}$ is concerned.

Definition. The vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ are **linearly dependent** (or coplanar) if at least one of them can be written as a linear combination of the others. Otherwise they are **linearly independent.** (We often refer to sets of vectors as simply "dependent" or "independent.")

In Figure 2.5 the set $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ is dependent. Any two of the vectors form an independent set. In any vector space, a set which contains the $\boldsymbol{\theta}$ vector is dependent, for $\boldsymbol{\theta}$ can be written as zero times any other vector in the set. We define the $\boldsymbol{\theta}$ vector by itself as a dependent set.

The following statement is equivalent to the above definition of independence: the vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ are linearly independent if and only if

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_n\mathbf{x}_n = \boldsymbol{\theta} \implies c_1 = \cdots = c_n = 0 \tag{2.11}$$

Equation (2.11) says the "zero combination" is the only combination that equals $\boldsymbol{\theta}$. For if c_i were not 0, we could simply divide by c_i to find $\mathbf{x}_i$ as a linear combination of the other vectors, and the set $\{\mathbf{x}_j\}$ would be dependent. If $c_i = 0$, $\mathbf{x}_i$ cannot be a linear combination of the other vectors. Equation (2.11) is a practical tool for determining independence of vectors.

Exercise 1. Explore graphically and by means of (2.11) the following set of vectors from $\mathcal{R}^3$: $\{\mathbf{x}_1 = (1, 0, 0), \mathbf{x}_2 = (0, 1, 0), \mathbf{x}_3 = (1, 1, 0), \mathbf{x}_4 = (0, 0, 1)\}$.

Example 4. Determining Independence In the space $\mathcal{R}^3$ let $\mathbf{x}_1 = (1, 2, 1)$, $\mathbf{x}_2 = (2, 3, 1)$, and $\mathbf{x}_3 = (4, 7, 3)$. Equation (2.11) becomes

$$\begin{aligned} c_1(1,2,1) + c_2(2,3,1) + c_3(4,7,3) \\ &= (c_1 + 2c_2 + 4c_3,\quad 2c_1 + 3c_2 + 7c_3,\quad c_1 + c_2 + 3c_3) \\ &= (0,0,0) \end{aligned}$$

Each component of this vector equation is a scalar-valued linear algebraic equation. We write the three equations in the matrix form:

$$\begin{pmatrix} 1 & 2 & 4 \\ 2 & 3 & 7 \\ 1 & 1 & 3 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

We solve this equation by elimination (Section 1.5) to find $c_1 = -2c_3$ and $c_2 = -c_3$. Any choice for c_3 will yield a particular nonzero linear combination of the vectors $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ which equals $\boldsymbol{\theta}$. The set is linearly dependent.

Definition. Let $\mathcal{S} \triangleq \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ be a set of vectors from a linear space $\mathcal{V}$. The set of all linear combinations of vectors from $\mathcal{S}$ is called the subspace of $\mathcal{V}$ **spanned** (or generated) by $\mathcal{S}$.* We often refer to this subspace as span($\mathcal{S}$) or span $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$.

Bases and Coordinates

We have introduced the vector space concept in order to provide a common mathematical framework for different types of systems. We can make the similarities between systems more apparent by converting their vector space representations to a standard form. We perform this standardization by introducing coordinate systems. In the example of Figure 2.5, the vectors $\{\mathbf{x}, \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ span a plane; yet any two of them will span the same plane. Two of them are redundant as far as generation of the plane is concerned.

Definition. A **basis** (or coordinate system) for a linear space $\mathcal{V}$ is a linearly independent set of vectors from $\mathcal{V}$ which spans $\mathcal{V}$.

Example 5. The Standard Bases for $\mathcal{R}^n$, $\mathcal{M}^{n \times 1}$, and $\mathcal{P}^n$. It is evident that any three linearly independent vectors in $\mathcal{R}^3$ form a basis for $\mathcal{R}^3$. The n-tuples

$$\begin{aligned} \boldsymbol{\varepsilon}_1 &= (1, 0, \ldots, 0) \\ \boldsymbol{\varepsilon}_2 &= (0, 1, 0, \ldots, 0) \\ &\vdots \\ \boldsymbol{\varepsilon}_n &= (0, \ldots, 0, 1) \end{aligned} \tag{2.12}$$

form a basis for $\mathcal{R}^n$. The set $\mathcal{E} \triangleq \{\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_n\}$ is called the **standard basis for $\mathcal{R}^n$.** We use the same notation to represent the standard basis for $\mathcal{M}^{n \times 1}$: $\mathcal{E} \triangleq \{\boldsymbol{\varepsilon}_i\}$, where $\boldsymbol{\varepsilon}_i$ is a column vector of zeros except for a 1 in the ith place. The set $\mathcal{N} \triangleq \{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$ defined by $\mathbf{f}_k(t) = t^{k-1}$ forms a basis for $\mathcal{P}^n$; it is analogous to the standard bases for $\mathcal{R}^n$ and $\mathcal{M}^{n \times 1}$.

*The definition of the space spanned by an infinite set of vectors depends on limiting concepts. We delay the definition until Section 5.3.

Example 6. The Zero Vector Space. The set $\{\boldsymbol{\theta}\}$ together with the obvious definitions of addition and scalar multiplication forms a vector space which we denote $\mathcal{O}$. However, the vector $\boldsymbol{\theta}$, by itself, is a dependent set. Therefore $\mathcal{O}$ has no basis.

If $\mathcal{X} \triangleq \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ is a basis for the space $\mathcal{V}$, any vector $\mathbf{x}$ in $\mathcal{V}$ can be written uniquely as some linear combination

$$\mathbf{x} = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_n\mathbf{x}_n \tag{2.13}$$

of vectors in $\mathcal{X}$. The multipliers c_i are called the **coordinates of x relative to the ordered basis** $\mathcal{X}$. It is easy to show that the coordinates relative to a particular ordered basis are unique: just expand $\mathbf{x}$ as in (2.13) for a second set $\{d_i\}$ of coordinates; then independence of the basis vectors implies $d_i = c_i$.

It is common to write the coordinates of a vector relative to a particular basis as a column matrix. We will denote by $[\mathbf{x}]_{\mathcal{X}}$ the **coordinate matrix** of the vector $\mathbf{x}$ relative to the (ordered) basis $\mathcal{X}$; thus corresponding to (2.13) we have

$$[\mathbf{x}]_{\mathcal{X}} \triangleq \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \tag{2.14}$$

Some bases are more natural or convenient than others. We use the term **natural basis** to mean a basis relative to which we can find coordinates by inspection. The bases of Example 5 are natural bases for $\mathcal{R}^n$, $\mathcal{M}^{n\times 1}$, and $\mathcal{P}^n$. Thus if $\mathbf{f}(t) = \xi_1 + \xi_2 t + \cdots + \xi_n t^{n-1}$, then $[\mathbf{f}]_{\mathcal{R}} = (\xi_1\ \xi_2 \cdots \xi_n)^T$.

Example 7. Coordinates for Vectors in $\mathcal{R}^3$. Let $\mathcal{X} \triangleq \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ be an ordered basis for $\mathcal{R}^3$, where $\mathbf{x}_1 = (1, 2, 3)$, $\mathbf{x}_2 = (2, 3, 2)$, and $\mathbf{x}_3 = (2, 5, 5)$. Let $\mathbf{x} = (1, 1, 1)$. To find $[\mathbf{x}]_{\mathcal{X}}$, we must solve (2.13):

$$\begin{aligned}(1,1,1) &= c_1(1,2,3) + c_2(2,3,2) + c_3(2,5,5). \\ &= (c_1 + 2c_2 + 2c_3,\quad 2c_1 + 3c_2 + 5c_3,\quad 3c_1 + 2c_2 + 5c_3)\end{aligned}$$

We rewrite the vector (3-tuple) equation in the matrix notation:

$$\begin{pmatrix} 1 & 2 & 2 \\ 2 & 3 & 5 \\ 3 & 2 & 5 \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \tag{2.15}$$

We solved this equation in Example 1 of Section 1.5. The result is

$$[\mathbf{x}]_{\mathfrak{X}} \triangleq [(1,1,1)]_{\mathfrak{X}} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} \frac{3}{5} \\ \frac{3}{5} \\ -\frac{2}{5} \end{pmatrix}$$

The coordinate matrix of Example 7 is merely a simple way of stating that $\mathbf{x} = \frac{3}{5}\mathbf{x}_1 + \frac{3}{5}\mathbf{x}_2 - \frac{2}{5}\mathbf{x}_3$. We choose to write the coordinates of a vector $\mathbf{x}$ as a column matrix because it allows us to carry out in a standard matrix format all manipulations involving the coordinates of $\mathbf{x}$.

In Example 4 of Section 1.5 we solved (2.15) with a general right-hand side; that is, for $\mathbf{x} = (\eta_1, \eta_2, \eta_3)$. That solution allows us to determine quickly the coordinate matrix, relative to the basis $\mathfrak{X}$ of Example 7, for *any* vector $\mathbf{x}$ in $\mathcal{R}^3$, including the case $\mathbf{x} = (0, 0, 0)$. In general, (2.13) includes (2.11); inherent in the process of finding coordinates for an arbitrary vector $\mathbf{x}$ is the process of determining whether $\mathfrak{X}$ is a basis. If $\mathfrak{X}$ is not independent, there will exist nonzero coordinates for $\mathbf{x} = \boldsymbol{\theta}$. If $\mathfrak{X}$ does not span the space, there will be some vector $\mathbf{x}$ for which no coordinates exist (P&C 2.7).

Example 8. Coordinates for Vectors in $\mathcal{P}^3$. Let $\mathcal{F} \triangleq \{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$ be an ordered basis for $\mathcal{P}^3$, where $\mathbf{f}_1(t) = 1 + 2t + 3t^2$, $\mathbf{f}_2(t) = 2 + 3t + 2t^2$, and $\mathbf{f}_3(t) = 2 + 5t + 5t^2$. Let $\mathbf{f}$ be defined by $\mathbf{f}(t) = 1 + t + t^2$. To find $[\mathbf{f}]_{\mathcal{F}}$, we solve (2.13), $\mathbf{f} = c_1\mathbf{f}_1 + c_2\mathbf{f}_2 + c_3\mathbf{f}_3$. To solve this equation, we evaluate both sides at t:

$$\begin{aligned} \mathbf{f}(t) &= (c_1\mathbf{f}_1 + c_2\mathbf{f}_2 + c_3\mathbf{f}_3)(t) \\ &= c_1\mathbf{f}_1(t) + c_2\mathbf{f}_2(t) + c_3\mathbf{f}_3(t) \end{aligned} \tag{2.16}$$

or

$$\begin{aligned} 1 + t + t^2 &= c_1(1 + 2t + 3t^2) + c_2(2 + 3t + 2t^2) + c_3(2 + 5t + 5t^2) \\ &= (c_1 + 2c_2 + 2c_3) + (2c_1 + 3c_2 + 5c_3)t + (3c_1 + 2c_2 + 5c_3)t^2 \end{aligned}$$

Equating coefficients on like powers of t we again obtain (2.15). The coordinate matrix of f is

$$[\mathbf{f}]_{\mathcal{F}} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} \frac{3}{5} \\ \frac{3}{5} \\ -\frac{2}{5} \end{pmatrix}$$

In order to solve the vector (function) equation (2.16) we converted it to a set of scalar equations expressed in matrix form. A second method for