It is quite difficult in this case to obtain a general expression for the rth number in the sequence by observation. On the other hand, the sequence can be described by the relation

$$a_r = a_{r-1} + a_{r-2}$$

together with the conditions that $a_0 = 1$ and $a_1 = 1$.

For a numeric function $(a_0, a_1, a_2, \ldots, a_r, \ldots)$, an equation relating a_r, for any r, to one or more of the a_i's, $i < r$, is called a *recurrence relation*. A recurrence relation is also called a *difference equation*, and these two terms will be used interchangeably. In many discrete computation problems, it is sometimes easier to obtain a specification of a numeric function in terms of a recurrence relation than to obtain a general expression for the value of the numeric function at r or a closed form expression for its generating function. It is clear that according to the recurrence relation, we can carry out a step-by-step computation to determine a_r from $a_{r-1}, a_{r-2}, \ldots$, to determine a_{r+1} from a_r, $a_{r-1}, \ldots$, and so on, provided that the value of the function at one or more points is given so that the computation can be initiated. These given values of the function are called *boundary conditions*. In the first example above, the boundary condition is $a_0 = 1$, and in the second example above, the boundary conditions are $a_0 = 1$ and $a_1 = 1$. We thus conclude that a numeric function can be described by a recurrence relation together with an appropriate set of boundary conditions.

One step beyond determining the values of a numeric function in a step-by-step computation according to a given recurrence relation is to obtain from the recurrence relation either a general expression for the value of the function at r or a closed form expression for its generating function. This is known as *solving* a recurrence relation. Unfortunately, no general method is known for solving all recurrence relations. In the following, we shall study the solution of a class of recurrence relations known as *linear recurrence relations with constant coefficients*.

7.2 LINEAR RECURRENCE RELATIONS WITH CONSTANT COEFFICIENTS

A recurrence relation of the form

$$C_0 a_r + C_1 a_{r-1} + C_2 a_{r-2} + \cdots + C_k a_{r-k} = f(r) \tag{7.2}$$

where the C_i's are constants, is called a *linear recurrence relation with constant coefficients*. The recurrence relation in (7.2) is known as a *kth-order recurrence relation*, provided that both C_0 and C_k are nonzero. For example,

$$3a_r - 5a_{r-1} + 2a_{r-2} = r^2 + 5 \tag{7.3}$$

is a second-order linear recurrence relation with constant coefficients.

If k consecutive values of the numeric function $\mathbf{a}$, $a_{m-k}, a_{m-k+1}, \ldots, a_{m-1}$, are known for some m, the value of a_m can be calculated according to (7.2), namely,

$$a_m = -\frac{1}{C_0}[C_1 a_{m-1} + C_2 a_{m-2} + \cdots + C_k a_{m-k} - f(m)]$$

Furthermore, the value of a_{m+1} can be computed as

$$a_{m+1} = -\frac{1}{C_0}[C_1 a_m + C_2 a_{m-1} + \cdots + C_k a_{m-k+1} - f(m+1)]$$

and the values of $a_{m+2}, a_{m+3}, \ldots$ can be computed in a similar manner. Also, the value of a_{m-k-1} can be computed as

$$a_{m-k-1} = -\frac{1}{C_k}[C_0 a_{m-1} + C_1 a_{m-2} + \cdots + C_{k-1} a_{m-k} - f(m-1)]$$

and the values of $a_{m-k-2}, a_{m-k-3}, \ldots$ can be computed in a similar manner. For example, for the recurrence relation in (7.3), given that $a_3 = 3$ and $a_4 = 6$, we can compute a_5 as

$$a_5 = -\tfrac{1}{3}[-5 \times 6 + 2 \times 3 - (5^2 + 5)] = 18$$

we can then compute a_6 as

$$a_6 = -\tfrac{1}{3}[-5 \times 18 + 2 \times 6 - (6^2 + 5)] = \tfrac{119}{3}$$

and so on. Also, we can compute a_2 as

$$a_2 = -\tfrac{1}{2}[3 \times 6 - 5 \times 3 - (4^2 + 5)] = 9$$

and so on. We thus conclude that the solution to (7.2) is determined uniquely by the values of k consecutive a_i's. Indeed, for a kth-order linear recurrence relation, the values of k consecutive a_i's constitute an appropriate set of boundary conditions. In other words, the values of k consecutive a_i's will always be sufficient to determine the numeric function **a** uniquely. However, the values of k nonconsecutive a_i's might or might not constitute an appropriate set of boundary conditions, depending on the specific recurrence relation we have. We shall not study the problem of what constitutes an appropriate set of boundary conditions, since it is not a significantly important one. See, however, Prob. 7.5.

The (total) solution of a linear difference equation with constant coefficients is the sum of two parts, the *homogeneous solution*, which satisfies the difference equation when the right-hand side of the equation is set to 0, and the *particular solution*, which satisfies the difference equation with $f(r)$ on the right-hand side.† Let $\mathbf{a}^{(h)} = (a_0^{(h)}, a_1^{(h)}, \ldots, a_r^{(h)}, \ldots)$ denote the homogeneous solution and $\mathbf{a}^{(p)} = (a_0^{(p)}, a_1^{(p)}, \ldots, a_r^{(p)}, \ldots)$ denote the particular solution to the difference equation. Since

$$C_0 a_r^{(h)} + C_1 a_{r-1}^{(h)} + \cdots + C_k a_{r-k}^{(h)} = 0$$

† For a reader who has prior exposure to the topic of differential equations, the analogy between the solution of linear differential equations with constant coefficients and that of linear difference equations with constant coefficients should be quite obvious.

and

$$C_0 a_r^{(p)} + C_1 a_{r-1}^{(p)} + \cdots + C_k a_{r-k}^{(p)} = f(r)$$

we have

$$C_0(a_r^{(h)} + a_r^{(p)}) + C_1(a_{r-1}^{(h)} + a_{r-1}^{(p)}) + \cdots + C_k(a_{r-k}^{(h)} + a_{r-k}^{(p)}) = f(r)$$

Clearly, the total solution, $\mathbf{a} = \mathbf{a}^{(h)} + \mathbf{a}^{(p)}$, satisfies the difference equation.

A homogeneous solution of a linear difference equation with constant coefficients is of the form $A\alpha_1^r$, where α_1 is called a *characteristic root* and A is a constant determined by the boundary conditions. Substituting $A\alpha^r$ for a_r in the difference equation with the right-hand side of the equation set to 0, we obtain

$$C_0 A\alpha^r + C_1 A\alpha^{r-1} + C_2 A\alpha^{r-2} + \cdots + C_k A\alpha^{r-k} = 0$$

This equation can be simplified to

$$C_0\alpha^k + C_1\alpha^{k-1} + C_2\alpha^{k-2} + \cdots + C_k = 0$$

which is called the *characteristic equation* of the difference equation. Therefore, if α_1 is one of the roots of the characteristic equation (it is for this reason that α_1 is called a characteristic root), $A\alpha_1^r$ is a homogeneous solution to the difference equation.

A characteristic equation of kth degree has k characteristic roots. Suppose the roots of the characteristic equation are distinct. In this case it is easy to verify that

$$a_r^{(h)} = A_1\alpha_1^r + A_2\alpha_2^r + \cdots + A_k\alpha_k^r$$

is also a homogeneous solution to the difference equation, where $\alpha_1, \alpha_2, \ldots, \alpha_k$ are the distinct characteristic roots and $A_1, A_2, \ldots, A_k$ are constants which are to be determined by the boundary conditions.†

Example 7.1 Let us revisit the example of the Fibonacci sequence of numbers discussed in Sec. 7.1. The recurrence relation for the Fibonacci sequence of numbers is

$$a_r = a_{r-1} + a_{r-2}$$

The corresponding characteristic equation is

$$\alpha^2 - \alpha - 1 = 0$$

which has the two distinct roots

$$\alpha_1 = \frac{1+\sqrt{5}}{2} \qquad \alpha_2 = \frac{1-\sqrt{5}}{2}$$

It follows that

$$a_r^{(h)} = A_1\left(\frac{1+\sqrt{5}}{2}\right)^r + A_2\left(\frac{1-\sqrt{5}}{2}\right)^r$$

is a homogeneous solution where the two constants A_1 and A_2 are to be determined from the boundary conditions $a_0 = 1$ and $a_1 = 1$. □

† The question of the uniqueness of the solution will be discussed later.

Now suppose that some of the roots of the characteristic equation are multiple roots. Let α_1 be a root of multiplicity m. We shall show that the corresponding homogeneous solution is

$$(A_1 r^{m-1} + A_2 r^{m-2} + \cdots + A_{m-2} r^2 + A_{m-1} r + A_m)\alpha_1^r$$

where the A_i's are constants to be determined by the boundary conditions. It is clear that $A_m \alpha_1^r$ is a homogeneous solution of the difference equation in (7.2). To show that $A_{m-1} r \alpha_1^r$ is also a homogeneous solution, we recall that α_1 not only is a root of the equation

$$C_0 \alpha^r + C_1 \alpha^{r-1} + C_2 \alpha^{r-2} + \cdots + C_k \alpha^{r-k} = 0 \tag{7.4}$$

but also is a root of the derivative equation of (7.4),

$$C_0 r \alpha^{r-1} + C_1 (r-1)\alpha^{r-2} + C_2 (r-2)\alpha^{r-3} + \cdots + C_k (r-k)\alpha^{r-k-1} = 0 \tag{7.5}$$

because α_1 is a multiple root of (7.4). Multiplying (7.5) by $A_{m-1}\alpha$ and replacing α by α_1, we obtain

$$\begin{aligned} C_0 A_{m-1} r \alpha_1^r &+ C_1 A_{m-1}(r-1)\alpha_1^{r-1} \\ &+ C_2 A_{m-1}(r-2)\alpha_1^{r-2} + \cdots + C_k A_{m-1}(r-k)\alpha_1^{r-k} = 0 \end{aligned}$$

which shows that $A_{m-1} r \alpha_1^r$ is indeed a homogeneous solution.

The fact that α_1 satisfies the second, third, ..., $(m-1)$st derivative equations of (7.4) enables us to prove that $A_{m-2} r^2 \alpha_1^r$, $A_{m-3} r^3 \alpha_1^r$, ..., $A_1 r^{m-1} \alpha_1^r$ are also homogeneous solutions in a similar manner.

Example 7.2 Consider the difference equation:

$$a_r + 6a_{r-1} + 12a_{r-2} + 8a_{r-3} = 0$$

The characteristic equation is

$$\alpha^3 + 6\alpha^2 + 12\alpha + 8 = 0$$

Thus,

$$a_r = (A_1 r^2 + A_2 r + A_3)(-2)^r$$

is a homogeneous solution since -2 is a triple characteristic root. □

There is no general way of finding the particular solution. However, in simple cases, it can be obtained by the method of inspection which we shall discuss briefly here. In the next section, we shall also present a method for determining both the homogeneous solution and the particular solution at the same time. Consider the difference equation

$$a_r + 5a_{r-1} + 6a_{r-2} = 3r^2 \tag{7.6}$$

We assume that the general form of the particular solution is†

$$P_1r^2 + P_2r + P_3 \tag{7.7}$$

where P_1, P_2, and P_3 are constants to be determined. Substituting the expression in (7.7) into the left-hand side of (7.6), we obtain

$$P_1r^2 + P_2r + P_3 + 5P_1(r-1)^2 + 5P_2(r-1) + 5P_3 + 6P_1(r-2)^2 + 6P_2(r-2) + 6P_3$$

which simplifies to

$$12P_1r^2 - (34P_1 - 12P_2)r + (29P_1 - 17P_2 + 12P_3) \tag{7.8}$$

Comparing (7.8) with the right-hand side of (7.6), we obtain the equations

$$12P_1 = 3$$

$$34P_1 - 12P_2 = 0$$

$$29P_1 - 17P_2 + 12P_3 = 0$$

which yield

$$P_1 = \tfrac{1}{4} \qquad P_2 = \tfrac{17}{24} \qquad P_3 = \tfrac{115}{288}$$

Therefore, the particular solution is

$$a_r^{(p)} = \tfrac{1}{4}r^2 + \tfrac{17}{24}r + \tfrac{115}{288}$$

In general, when the right-hand side of the difference equation is of the form r^t, the corresponding particular solution will be of the form

$$P_1r^t + P_2r^{t-1} + \cdots + P_tr + P_{t+1}$$

As another example, consider the difference equation

$$a_r + 5a_{r-1} + 6a_{r-2} = 42 \cdot 4^r \tag{7.9}$$

We assume that the general form of the particular solution is

$$a_r = P4^r \tag{7.10}$$

Substituting the expression in (7.10) into the left-hand side of (7.9), we obtain

$$P4^r + 5P4^{r-1} + 6P4^{r-2}$$

which simplifies to

$$\tfrac{21}{8}P4^r \tag{7.11}$$

Comparing (7.11) with the right-hand side of (7.9), we obtain

$$P = 16$$

Therefore, the particular solution is

$$a_r = 16 \cdot 4^r$$

† There is no systematic way to determine a general form of the particular solution of a difference equation. In the present case, it is not difficult to see that if the right-hand side of the difference equation is a polynomial in r, the particular solution will also be a polynomial in r.

In general, when the right-hand side of the difference equation is of the form β^r, the corresponding particular solution is of the form $P\beta^r$ if β is not a characteristic root of the difference equation; it is of the form $P_r\beta^r$ if β is a distinct characteristic root of the difference equation; and it is of the form

$$P_r^{m-1}\beta^r$$

if β is a characteristic root of multiplicity $m-1$ of the difference equation. For further details, see Prob. 7.6.

Finally, we must determine the undetermined coefficients in the homogeneous solution. For a kth-order difference equation, the k undetermined coefficients $A_1, A_2, \ldots, A_k$ in the homogeneous solution can be determined by the boundary conditions, $a_{r_0}, a_{r_0+1}, \ldots, a_{r_0+k-1}$, for any r_0. Suppose the characteristic roots of the difference equation are all distinct. The total solution is of the form

$$a_r = A_1\alpha_1^r + A_2\alpha_2^r + \cdots + A_k\alpha_k^r + p(r)$$

where $p(r)$ is the particular solution. Thus, for $r = r_0, r_0+1, \ldots, r_0+k-1$, we have the system of linear equations:

$$\begin{aligned}
a_{r_0} &= A_1\alpha_1^{r_0} + A_2\alpha_2^{r_0} + \cdots + A_k\alpha_k^{r_0} + p(r_0) \\
a_{r_0+1} &= A_1\alpha_1^{r_0+1} + A_2\alpha_2^{r_0+1} + \cdots + A_k\alpha_k^{r_0+1} + p(r_0+1) \qquad (7.12)\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
a_{r_0+k-1} &= A_1\alpha_1^{r_0+k-1} + A_2\alpha_2^{r_0+k-1} + \cdots + A_k\alpha_k^{r_0+k-1} + p(r_0+k-1)
\end{aligned}$$

These k linear equations can be solved for $A_1, A_2, \ldots, A_k$. For example, for the difference equation in (7.9), the total solution is

$$a_r = A_1(-2)^r + A_2(-3)^r + 16 \cdot 4^r$$

Suppose we are given the boundary conditions $a_2 = 278$ and $a_3 = 962$. Solving the equations

$$278 = 4A_1 + 9A_2 + 256$$

$$962 = -8A_1 - 27A_2 + 1024$$

we obtain

$$A_1 = 1 \qquad A_2 = 2$$

Thus,

$$a_r = (-2)^r + 2(-3)^r + 16 \cdot 4^r$$

is the total solution of the difference equation.

One might question how we can be sure that solutions of the k equations in (7.12) are always unique. It can be shown† that this is indeed the case. On the other hand, if we are given the value of the numeric function at k not necessarily successive points, although we can set up k equations for the undetermined coefficients $A_1, A_2, \ldots, A_k$ similar to that in (7.12), it is not always the case that these equations can be solved uniquely.

† See, for example, chap. 3 of Liu [5].

When the characteristic roots of the difference equation are not all distinct, a derivation similar to the foregoing can be carried out. Again, the undetermined coefficients in the homogeneous solution can be determined uniquely by the value of the numeric function at k successive points.

Although we have shown that the total solution obtained according to the procedure presented above satisfies the difference equation and the given boundary conditions, one might wonder whether there are other solutions that will also satisfy the difference equation and the boundary conditions not covered by our procedure. The answer to this question is a negative one. It can be shown that the solution obtained according to the procedure just presented is indeed unique.

7.3 SOLUTION BY THE METHOD OF GENERATING FUNCTIONS

Instead of solving a difference equation for an expression for the value of a numeric function as we did in Sec. 7.2, we can also determine the generating function of the numeric function from the difference equation. In many cases, once the generating function is determined, an expression for the value of the numeric function can easily be obtained.

For a given kth-order difference equation that specifies a numeric function, we should know for what values of r the equation is valid. First of all, we note that the equation is valid only if $r \geqq k$ because for $r < k$ the equation will involve a_{-i}'s which are not defined. Furthermore, in many cases a difference equation arises from a physical problem in which the value of a_i is meaningful only for $i \geqq t$ for some t larger than or equal to 0. In that case, the difference equation is valid only for $r - k \geqq t$.

Suppose we toss a coin r times. There are 2^r possible sequences of outcomes. We want to know the number of sequences of outcomes in which heads never appear on successive tosses. Let a_r denote the number of such sequences. To each sequence of $r-1$ heads and tails in which there are no consecutive heads, we can append a tail to obtain a sequence of r heads and tails in which there are no consecutive heads. To each sequence of $r-2$ heads and tails in which there are no consecutive heads, we can append a tail and then a head to obtain a sequence of r heads and tails in which there are no consecutive heads. Moreover, these exhaust all sequences of r heads and tails in which there are no consecutive heads. We thus have the recurrence relation

$$a_r = a_{r-1} + a_{r-2} \tag{7.13}$$

Note that $a_1 = 2$ and $a_2 = 3$. The value of a_0 has no physical significance, and a reasonable choice of its value is 0. In this case, the difference equation is *not* valid for $r = 2$, and is only valid for $r \geqq 3$.

We now proceed to show how we can obtain $A(z)$ from (7.13). Multiplying