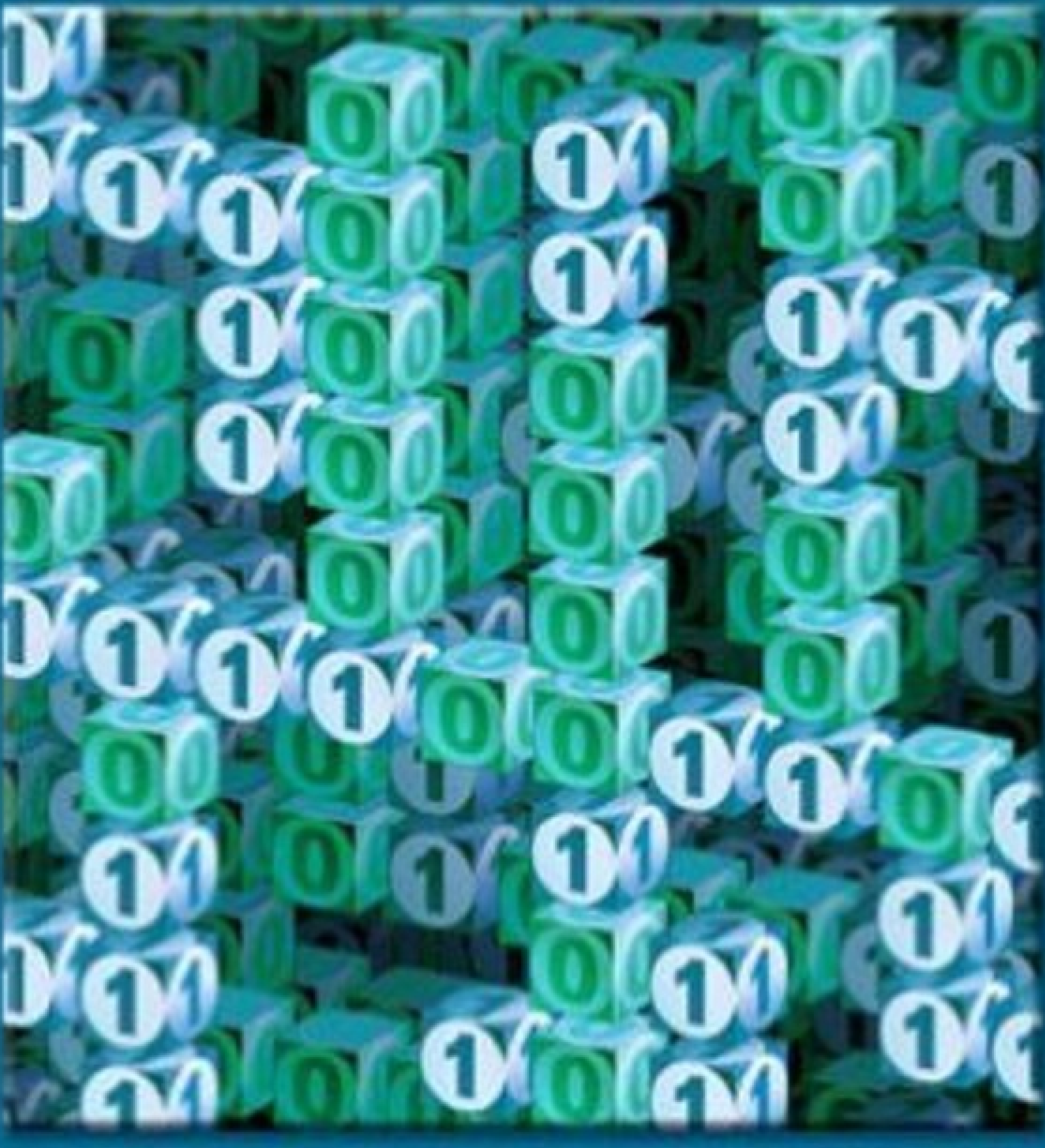
DISCRETE and COMBINATORIAL MATHEMATICS

An Applied Introduction



Ralph P. Grimaldi

Fifth Edition

DISCRETE AND COMBINATORIAL MATHEMATICS

An Applied Introduction

FIFTH EDITION

RALPH P. GRIMALDI
Rose-Hulman Institute of Technology

PEARSON
Addison Wesley

Boston San Francisco New York
London Toronto Sydney Tokyo Singapore Madrid
Mexico City Munich Paris Cape Town Hong Kong Montreal

NOTATION

LOGIC	p, q	statements (or propositions)
	$\neg p$	the negation of (statement) p: *not p*
	$p \wedge q$	the conjunction of p, q: *p and q*
	$p \vee q$	the disjunction of p, q: *p or q*
	$p \rightarrow q$	the implication of q by p: *p implies q*
	$p \leftrightarrow q$	the biconditional of p and q: *p if and only if q*
	iff	if and only if
	$p \Rightarrow q$	logical implication: *p logically implies q*
	$p \Longleftrightarrow q$	logical equivalence: p is *logically equivalent* to q
	T_0	tautology
	F_0	contradiction
	$\forall x$	For *all* x (the universal quantifier)
	$\exists x$	For *some* x (the existential quantifier)
SET THEORY	$x \in A$	element x is a member of set A
	$x \notin A$	element x is not a member of set A
	$\mathcal{U}$	the universal set
	$A \subseteq B, B \supseteq A$	A is a subset of B
	$A \subset B, B \supset A$	A is a proper subset of B
	$A \nsubseteq B$	A is not a subset of B
	$A \not\subset B$	A is not a proper subset of B
	$\|A\|$	the cardinality, or size, of set A — that is, the number of elements in A
	$\emptyset = \{ \}$	the empty, or null, set
	$\mathcal{P}(A)$	the power set of A — that is, the collection of all subsets of A
	$A \cap B$	the intersection of sets A, B: $\{x \mid x \in A \text{ and } x \in B\}$
	$A \cup B$	the union of sets A, B: $\{x \mid x \in A \text{ or } x \in B\}$
	$A \triangle B$	the symmetric difference of sets A, B: $\{x \mid x \in A \text{ or } x \in B, \text{ but } x \notin A \cap B\}$
	$\overline{A}$	the complement of set A: $\{x \mid x \in \mathcal{U} \text{ and } x \notin A\}$
	$A - B$	the (relative) complement of set B in set A: $\{x \mid x \in A \text{ and } x \notin B\}$
	$\bigcup_{i \in I} A_i$	$\{x \mid x \in A_i, \text{ for at least one } i \in I\}$, where I is an index set
	$\bigcap_{i \in I} A_i$	$\{x \mid x \in A_i, \text{ for every } i \in I\}$, where I is an index set
PROBABILITY	S	the sample space for an experiment $\mathcal{E}$
	$A \subseteq S$	A is an event
	$Pr(A)$	the probability of event A
	$Pr(A\|B)$	the probability of A given B; conditional probability
	X	random variable
	$E(X)$	the expected value of X, a random variable
	$\text{Var}(X) = \sigma_X^2$	the variance of X, a random variable
	σ_X	the standard deviation of X, a random variable
NUMBERS	$a\|b$	a divides b, for $a, b \in \mathbf{Z}, a \neq 0$
	$a \nmid b$	a does not divide b, for $a, b \in \mathbf{Z}, a \neq 0$
	$\gcd(a, b)$	the greatest common divisor of the integers a, b
	$\text{lcm}(a, b)$	the least common multiple of the integers a, b
	$\phi(n)$	Euler's phi function for $n \in \mathbf{Z}^+$
	$\lfloor x \rfloor$	the greatest integer less than or equal to the real number x: the greatest integer in x: the *floor* of x

NOTATION

	$\lceil x \rceil$	the smallest integer greater than or equal to the real number x: the *ceiling* of x
	$a \equiv b \pmod{n}$	a is congruent to b modulo n
RELATIONS	$A \times B$	the Cartesian, or cross, product of sets A, B: $\{(a, b) \mid a \in A, b \in B\}$
	$\mathcal{R} \subseteq A \times B$	$\mathcal{R}$ is a relation from A to B
	$a \mathcal{R} b$; $(a, b) \in \mathcal{R}$	a is related to b
	$a \not\mathcal{R} b$; $(a, b) \notin \mathcal{R}$	a is not related to b
	$\mathcal{R}^c$	the converse of relation $\mathcal{R}$: $(a, b) \in \mathcal{R}$ iff $(b, a) \in \mathcal{R}^c$
	$\mathcal{R} \circ \mathcal{S}$	the composite relation for $\mathcal{R} \subseteq A \times B$, $\mathcal{S} \subseteq B \times C$: $(a, c) \in \mathcal{R} \circ \mathcal{S}$ if $(a, b) \in \mathcal{R}$, $(b, c) \in \mathcal{S}$ for some $b \in B$
	$\text{lub}\{a, b\}$	the least upper bound of a and b
	$\text{glb}\{a, b\}$	the greatest lower bound of a and b
	$[a]$	the equivalence class of element a (relative to an equivalence relation $\mathcal{R}$ on a set A): $\{x \in A \mid x \mathcal{R} a\}$
FUNCTIONS	$f: A \to B$	f is a function from A to B
	$f(A_1)$	for $f: A \to B$ and $A_1 \subseteq A$, $f(A_1)$ is the image of A_1 under f — that is, $\{f(a) \mid a \in A_1\}$
	$f(A)$	for $f: A \to B$, $f(A)$ is the range of f
	$f: A \times A \to B$	f is a binary operation on A
	$f: A \times A \to B\ (\subseteq A)$	f is a closed binary operation on A
	$1_A: A \to A$	the identity function on A: $1_A(a) = a$ for each $a \in A$
	$f\|_{A_1}$	the restriction of $f: A \to B$ to $A_1 \subseteq A$
	$g \circ f$	the composite function for $f: A \to B$, $g: B \to C$: $(g \circ f)a = g(f(a))$, for $a \in A$
	f^{-1}	the inverse of function f
	$f^{-1}(B_1)$	the preimage of $B_1 \subseteq B$ for $f: A \to B$
	$f \in O(g)$	f is "big Oh" of g; f is of order g
THE ALGEBRA OF STRINGS	$\mathbf{\Sigma}$	a finite set of symbols called an alphabet
	λ	the empty string
	$\\|x\\|$	the length of string x
	$\mathbf{\Sigma}^n$	$\{x_1 x_2 \cdots x_n \mid x_i \in \mathbf{\Sigma}\}$, $n \in \mathbf{Z}^+$
	$\mathbf{\Sigma}^0$	$\{\lambda\}$
	$\mathbf{\Sigma}^+$	$\bigcup_{n \in \mathbf{Z}^+} \mathbf{\Sigma}^n$: the set of all strings of positive length
	$\mathbf{\Sigma}^*$	$\bigcup_{n \geq 0} \mathbf{\Sigma}^n$: the set of all finite strings
	$A \subseteq \mathbf{\Sigma}^*$	A is a language
	AB	the concatenation of languages $A, B \subseteq \mathbf{\Sigma}^*$: $\{ab \mid a \in A, b \in B\}$
	A^n	$\{a_1 a_2 \cdots a_n \mid a_i \in A \subseteq \mathbf{\Sigma}^*\}$, $n \in \mathbf{Z}^+$
	A^0	$\{\lambda\}$
	A^+	$\bigcup_{n \in \mathbf{Z}^+} A^n$
	A^*	$\bigcup_{n \geq 0} A^n$: the Kleene closure of language A
	$M = (S, \mathcal{I}, \mathcal{O}, \nu, \omega)$	a finite state machine M with internal states S, input alphabet $\mathcal{I}$, output alphabet $\mathcal{O}$, next state function $\nu: S \times \mathcal{I} \to S$ and output function $\omega: S \times \mathcal{I} \to \mathcal{O}$

PART 2

FURTHER TOPICS IN ENUMERATION

8

The Principle of Inclusion and Exclusion

We now return to the topic of enumeration as we investigate the *Principle of Inclusion and Exclusion*. Extending the ideas in the counting problems on Venn diagrams in Chapter 3, this principle will assist us in establishing the formula we conjectured in Section 5.3 for the number of onto functions $f: A \to B$, where A, B are finite (nonempty) sets. Other applications of this principle will demonstrate its versatile nature in combinatorial mathematics.

8.1 The Principle of Inclusion and Exclusion

In this section we develop some notation for stating this new counting principle. Then we establish the principle by a combinatorial argument. Following this, a wide range of examples demonstrate how this principle may be applied.

We shall motivate the Principle of Inclusion and Exclusion with a series of three examples, the first two of which will be reminiscent of the work we did with counting and Venn diagrams in Section 3.3.

EXAMPLE 8.1

Let S represent the set of 100 students enrolled in the freshman engineering program at Central College. Then $|S| = 100$. Now let c_1, c_2 denote the following conditions (or properties) satisfied by some of the elements of S:

c_1: A student at Central College is among the 100 students in the freshman engineering program and is enrolled in Freshman Composition.

c_2: A student at Central College is among the 100 students in the freshman engineering program and is enrolled in Introduction to Economics.

Suppose that 35 of these 100 students are enrolled in Freshman Composition and that 30 of them are enrolled in Introduction to Economics. We shall denote this by

$$N(c_1) = 35 \quad \text{and} \quad N(c_2) = 30.$$

If nine of these 100 students are enrolled in both Freshman Composition and Introduction to Economics then we write $N(c_1c_2) = 9$.

Further, of these 100 students, there are 100 − 35 = 65 who are *not* taking Freshman Composition. Denoting $|S|$ by N, we can designate this by writing $N(\overline{c}_1) = N - N(c_1)$. In a similar way we designate that there are $N(\overline{c}_2) = N - N(c_2) = 100 - 30 = 70$ of these students who are not taking Introduction to Economics. The number who *are* taking Freshman Composition and who are *not* taking Introduction to Economics is $N(c_1\overline{c}_2) = N(c_1) - N(c_1c_2) = 35 - 9 = 26$. Likewise, of these 100 students, there are $N(\overline{c}_1c_2) = N(c_2) - N(c_1c_2) = 30 - 9 = 21$ who are enrolled in Introduction to Economics but not in Freshman Composition. Of particular interest are those students (from among these 100 freshmen) who are taking neither Freshman Composition nor Introduction to Economics — that is, they are *not* taking Freshman Composition and they are also *not* taking Introduction to Economics. Their number is $N(\overline{c}_1\overline{c}_2)$. And since $N(\overline{c}_1) = N(\overline{c}_1c_2) + N(\overline{c}_1\overline{c}_2)$, we learn that $N(\overline{c}_1\overline{c}_2) = N(\overline{c}_1) - N(\overline{c}_1c_2) = 65 - 21 = 44$.

The preceding observations also demonstrate that

$$\begin{aligned} N(\overline{c}_1\overline{c}_2) &= N(\overline{c}_1) - N(\overline{c}_1c_2) = [N - N(c_1)] - [N(c_2) - N(c_1c_2)] \\ &= N - N(c_1) - N(c_2) + N(c_1c_2) = N - [N(c_1) + N(c_2)] + N(c_1c_2) \\ &= 100 - [35 + 30] + 9 = 44, \text{ as we saw above.} \end{aligned}$$

From the Venn diagram in Fig. 8.1, we see that if $N(c_1)$ denotes the number of elements of S in the left-hand circle and $N(c_2)$ denotes the number in the right-hand circle, then $N(c_1c_2)$ is the number of these elements from S in the overlap, while $N(\overline{c}_1\overline{c}_2)$ counts those elements of S that are outside the union of these two circles. Consequently, we see once again — this time from the figure — that

$$N(\overline{c}_1\overline{c}_2) = N - [N(c_1) + N(c_2)] + N(c_1c_2),$$

where the last term is added on because it was eliminated twice in the term $[N(c_1) + N(c_2)]$. (Also, at this point, the reader may wish to look back at the second formula following Example 3.25 to find the same result presented with a different notation.)

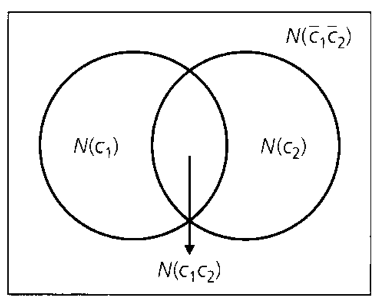


Figure 8.1

[Before we advance to our next example where we will introduce a third condition, let us note that $N(\overline{c}_1\overline{c}_2)$ is *not* the same as $N(\overline{c_1c_2})$. For $N(\overline{c_1c_2}) = N - N(c_1c_2) = 100 - 9 = 91$, in this example, while $N(\overline{c}_1\overline{c}_2) = 44$, as we learned earlier. However, $N(\overline{c}_1 \text{ or } \overline{c}_2) = N(\overline{c_1c_2}) = 91 = 65 + 70 - 44 = N(\overline{c}_1) + N(\overline{c}_2) - N(\overline{c}_1\overline{c}_2)$.]

EXAMPLE 8.2

We start with the same 100 students as in Example 8.1 and the same conditions c_1, c_2, but now we consider a third condition, given as follows:

c_3: A student at Central College is among the 100 students in the freshman engineering program and is enrolled in Fundamentals of Computer Programming.

It is still the case that $N(c_1) = 35$, $N(c_2) = 30$, and $N(c_1c_2) = 9$, but now we are also given that $N(c_3) = 30$, $N(c_1c_3) = 11$, $N(c_2c_3) = 10$, and $N(c_1c_2c_3) = 5$ (that is, there are five of these 100 freshmen who are taking Freshman Composition, Introduction to Economics, and Fundamentals of Computer Programming). Looking to Fig. 8.2, we learn that

$$N(\overline{c}_1\overline{c}_2\overline{c}_3) = N - [N(c_1) + N(c_2) + N(c_3)] + [N(c_1c_2) + N(c_1c_3) + N(c_2c_3)] - N(c_1c_2c_3).$$

So here we have $N(\overline{c}_1\overline{c}_2\overline{c}_3) = 100 - [35 + 30 + 30] + [9 + 11 + 10] - 5 = 30$. That is, out of these 100 students there are 30 who are *not* enrolled in any of the courses: (i) Freshman Composition; (ii) Introduction to Economics; or (iii) Fundamentals of Computer Programming.

[We also learn here that $N(\overline{c}_3) = 70 = 100 - 30 = N - N(c_3)$, $N(\overline{c}_1\overline{c}_3) = 46 = 100 - [35 + 30] + 11 = N - [N(c_1) + N(c_3)] + N(c_1c_3)$, and $N(\overline{c}_2\overline{c}_3) = 50 = 100 - [30 + 30] + 10 = N - [N(c_2) + N(c_3)] + N(c_2c_3)$. Furthermore, we note the similarity here with the result for $|\overline{A} \cap \overline{B} \cap \overline{C}|$ given in the second formula following Example 3.26.]

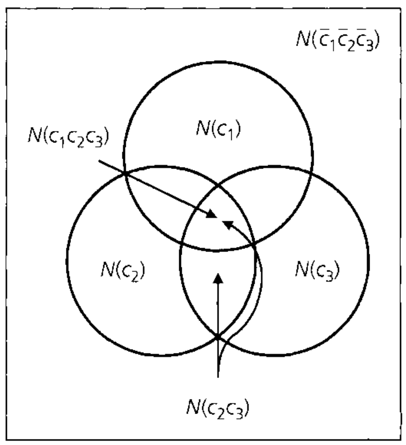


Figure 8.2

EXAMPLE 8.3

Based on the results in the previous two examples we may now feel that for a given finite set S (with $|S| = N$) and four conditions c_1, c_2, c_3, c_4 we should have

$$\begin{aligned} N(\overline{c}_1\overline{c}_2\overline{c}_3\overline{c}_4) = N &- [N(c_1) + N(c_2) + N(c_3) + N(c_4)] \qquad (*) \\ &+ [N(c_1c_2) + N(c_1c_3) + N(c_1c_4) + N(c_2c_3) + N(c_2c_4) + N(c_3c_4)] \\ &- [N(c_1c_2c_3) + N(c_1c_2c_4) + N(c_1c_3c_4) + N(c_2c_3c_4)] \\ &+ N(c_1c_2c_3c_4). \end{aligned}$$

To show that this is the case we consider an arbitrary element x from S and show that it is counted the same number of times on both sides of the above equation.

0) If x satisfies none of the four conditions, then it is counted once on the left side of Eq. (*) [in $N(\overline{c}_1\overline{c}_2\overline{c}_3\overline{c}_4)$], and once on the right side of Eq. (*) [in N].

1) If x satisfies only one of the conditions, say c_1, then it is not counted at all on the left side of Eq. (*). But on the right side of Eq. (*), x is counted once in N and once in $N(c_1)$, for a total of $1 - 1 = 0$ times.

2) Now suppose that x satisfies conditions c_2, c_4 but does *not* satisfy conditions c_1, c_3. Once again x is not counted on the left side of Eq. (*). For the right side of Eq. (*), x is counted once in N, once in each of $N(c_2)$ and $N(c_4)$, and then once in $N(c_2c_4)$, totaling $1 - [1 + 1] + 1 = 1 - \binom{2}{1} + \binom{2}{2} = 0$ times.

3) Continuing with the case for three conditions, we'll suppose here that x satisfies conditions c_1, c_2, and c_4, but *not* c_3. As in the previous two cases, x is not counted on the left side of Eq. (*). On the right side of Eq. (*), x is counted once in N, once in each of $N(c_1)$, $N(c_2)$, and $N(c_4)$, once in each of $N(c_1c_2)$, $N(c_1c_4)$, and $N(c_2c_4)$, and, finally, once in $N(c_1c_2c_4)$. So on the right side of Eq. (*), x is counted $1 - [1 + 1 + 1] + [1 + 1 + 1] - 1 = 1 - \binom{3}{1} + \binom{3}{2} - \binom{3}{3} = 0$ times, in total.

4) Finally, if x satisfies all four of the conditions c_1, c_2, c_3, c_4, then once again it is not counted on the left side of Eq. (*). On the right side of Eq. (*), x is counted once for each of the 16 terms on the right side of this equation — for a total of $1 - [1 + 1 + 1 + 1] + [1 + 1 + 1 + 1 + 1 + 1] - [1 + 1 + 1 + 1] + 1 = 1 - \binom{4}{1} + \binom{4}{2} - \binom{4}{3} + \binom{4}{4} = 0$ times.

Consequently, from these preceding five cases we have shown that the two sides of Eq. (*) count the same elements from S, and this provides a combinatorial proof for the formula for $N(\overline{c}_1\overline{c}_2\overline{c}_3\overline{c}_4)$.

So now we shall reconsider the situation in Example 8.2 and introduce a fourth condition as follows:

c_4: A student at Central College is among the 100 students in the freshman engineering program and is enrolled in Introduction to Design.

We already know that $N(c_1) = 35$, $N(c_2) = 30$, $N(c_3) = 30$, $N(c_1c_2) = 9$, $N(c_1c_3) = 11$, $N(c_2c_3) = 10$, and $N(c_1c_2c_3) = 5$. If $N(c_4) = 41$, $N(c_1c_4) = 13$, $N(c_2c_4) = 14$, $N(c_3c_4) = 10$, $N(c_1c_2c_4) = 6$, $N(c_1c_3c_4) = 6$, $N(c_2c_3c_4) = 6$, and $N(c_1c_2c_3c_4) = 4$, then, using the equation we derived above, it follows that $N(\overline{c}_1\overline{c}_2\overline{c}_3\overline{c}_4) = 100 - [35 + 30 + 30 + 41] + [9 + 11 + 13 + 10 + 14 + 10] - [5 + 6 + 6 + 6] + 4 = 100 - 136 + 67 - 23 + 4 = 12$. Thus, of the 100 students in the freshman engineering program at Central College, there are 12 who are not taking any of the four courses: Freshman Composition, Introduction to Economics, Fundamentals of Computer Programming, or Introduction to Design.

If we are interested in the number (from these 100 students) who are taking Freshman Composition, but none of the other three courses, then we should want to compute $N(c_1\overline{c}_2\overline{c}_3\overline{c}_4)$. To do so we start by observing that

$$N(\overline{c}_2\overline{c}_3\overline{c}_4) = N(c_1\overline{c}_2\overline{c}_3\overline{c}_4) + N(\overline{c}_1\overline{c}_2\overline{c}_3\overline{c}_4),$$

which can be established by an argument similar to the one above for $N(\overline{c}_1\overline{c}_2\overline{c}_3\overline{c}_4)$. This then leads us to

$$N(c_1\overline{c}_2\overline{c}_3\overline{c}_4) = N(\overline{c}_2\overline{c}_3\overline{c}_4) - N(\overline{c}_1\overline{c}_2\overline{c}_3\overline{c}_4).$$

Using the result in Example 8.2 we find that

$$\begin{aligned} N(\overline{c}_2\overline{c}_3\overline{c}_4) &= N - [N(c_2) + N(c_3) + N(c_4)] + [N(c_2c_3) + N(c_2c_4) + N(c_3c_4)] \\ &\quad - N(c_2c_3c_4) \\ &= 100 - [30 + 30 + 41] + [10 + 14 + 10] - 6 = 27, \text{ and} \\ N(c_1\overline{c}_2\overline{c}_3\overline{c}_4) &= N(\overline{c}_2\overline{c}_3\overline{c}_4) - N(\overline{c}_1\overline{c}_2\overline{c}_3\overline{c}_4) = 27 - 12 = 15. \end{aligned}$$

So there are 15 students in this set of 100 who are taking Freshman Composition, but none of the other courses: Introduction to Economics, Fundamentals of Computer Programming, or Introduction to Design.

Further, we also observe that

$$\begin{aligned}
N(c_1\overline{c}_2\overline{c}_3\overline{c}_4) &= N(\overline{c}_2\overline{c}_3\overline{c}_4) - N(\overline{c}_1\overline{c}_2\overline{c}_3\overline{c}_4) \\
&= \{N - [N(c_2) + N(c_3) + N(c_4)] + [N(c_2c_3) + N(c_2c_4) + N(c_3c_4)] \\
&\quad - N(c_2c_3c_4)\} - \{N - [N(c_1) + N(c_2) + N(c_3) + N(c_4)] \\
&\quad + [N(c_1c_2) + N(c_1c_3) + N(c_1c_4) + N(c_2c_3) + N(c_2c_4) + N(c_3c_4)] \\
&\quad - [N(c_1c_2c_3) + N(c_1c_2c_4) + N(c_1c_3c_4) + N(c_2c_3c_4)] + N(c_1c_2c_3c_4)\}, \text{ or}
\end{aligned}$$

$$\begin{aligned}
N(c_1\overline{c}_2\overline{c}_3\overline{c}_4) &= N(c_1) - [N(c_1c_2) + N(c_1c_3) + N(c_1c_4)] \\
&\quad + [N(c_1c_2c_3) + N(c_1c_2c_4) + N(c_1c_3c_4)] - N(c_1c_2c_3c_4).
\end{aligned}$$

So here $N(c_1\overline{c}_2\overline{c}_3\overline{c}_4) = 35 - [9 + 11 + 13] + [5 + 6 + 6] - 4 = 35 - 33 + 17 - 4 = 15$, as we found above.

Having seen the results in Examples 8.1, 8.2, and 8.3, now it is time for us to generalize these results and establish the Principle of Inclusion and Exclusion. To do so we once again let S be a set with $|S| = N$, and we let $c_1, c_2, \ldots, c_t$ be a collection of t conditions or properties — each of which may be satisfied by some of the elements of S. Some elements of S may satisfy more than one of the conditions, whereas others may not satisfy any of them. For all $1 \leq i \leq t$, $N(c_i)$ will denote the number of elements in S that satisfy condition c_i. (Elements of S are counted here when they satisfy only condition c_i, as well as when they satisfy c_i and other conditions c_j, for $j \neq i$.) For all $i, j \in \{1, 2, 3, \ldots, t\}$ where $i \neq j$, $N(c_ic_j)$ will denote the number of elements in S that satisfy both of the conditions c_i, c_j, and perhaps some others. [$N(c_ic_j)$ does *not* count the elements of S that satisfy *only* c_i, c_j.] Continuing, if $1 \leq i, j, k \leq t$ are three distinct integers, then $N(c_ic_jc_k)$ denotes the number of elements in S satisfying, perhaps among others, each of the conditions c_i, c_j, and c_k.

For each $1 \leq i \leq t$, $N(\overline{c}_i) = N - N(c_i)$ denotes the number of elements in S that do not satisfy condition c_i. If $1 \leq i, j \leq t$ with $i \neq j$, $N(\overline{c}_i\overline{c}_j) =$ the number of elements in S that do not satisfy either of the conditions c_i or c_j. [This is *not* the same as $N(\overline{c_ic_j})$, as we observed at the end of Example 8.1.]

With the necessary preliminaries now in hand we state the following theorem.

THEOREM 8.1 *The Principle of Inclusion and Exclusion.* Consider a set S, with $|S| = N$, and conditions c_i, $1 \leq i \leq t$, each of which may be satisfied by some of the elements of S. The number of elements of S that satisfy *none* of the conditions c_i, $1 \leq i \leq t$, is denoted by $\overline{N} = N(\overline{c}_1\overline{c}_2\overline{c}_3 \cdots \overline{c}_t)$ where

$$\begin{aligned}
\overline{N} &= N - [N(c_1) + N(c_2) + N(c_3) + \cdots + N(c_t)] \qquad (1) \\
&\quad + [N(c_1c_2) + N(c_1c_3) + \cdots + N(c_1c_t) + N(c_2c_3) + \cdots + N(c_{t-1}c_t)] \\
&\quad - [N(c_1c_2c_3) + N(c_1c_2c_4) + \cdots + N(c_1c_2c_t) + N(c_1c_3c_4) + \cdots \\
&\quad + N(c_1c_3c_t) + \cdots + N(c_{t-2}c_{t-1}c_t)] + \cdots + (-1)^t N(c_1c_2c_3 \cdots c_t),
\end{aligned}$$

or

$$\overline{N} = N - \sum_{1 \le i \le t} N(c_i) + \sum_{1 \le i < j \le t} N(c_i c_j) - \sum_{1 \le i < j < k \le t} N(c_i c_j c_k) + \cdots \tag{2}$$
$$+ (-1)^t N(c_1 c_2 c_3 \cdots c_t).$$

Proof: Although this result can be established by applying the Principle of Mathematical Induction to the number t of conditions, we shall give a combinatorial proof. The argument will be reminiscent of the ideas we saw in Example 8.3 in establishing the formula for $N(\overline{c}_1\overline{c}_2\overline{c}_3\overline{c}_4)$.

For each $x \in S$ we show that x contributes the same count, either 0 or 1, to each side of Eq. (2).

If x satisfies none of the conditions, then x is counted once in $\overline{N}$ and once in N, but not in any of the other terms in Eq. (2). Consequently, x contributes a count of 1 to each side of the equation.

The other possibility is that x satisfies *exactly* r of the conditions where $1 \le r \le t$. In this case x contributes nothing to $\overline{N}$. But on the right-hand side of Eq. (2), x is counted

(1) One time in N.

(2) r times in $\sum_{1 \le i \le t} N(c_i)$. (Once for each of the r conditions.)

(3) $\binom{r}{2}$ times in $\sum_{1 \le i < j \le t} N(c_i c_j)$. (Once for each pair of conditions selected from the r conditions it satisfies.)

(4) $\binom{r}{3}$ times in $\sum_{1 \le i < j < k \le t} N(c_i c_j c_k)$. (Why?)

. . . .

$(r+1)$ $\binom{r}{r} = 1$ time in $\sum N(c_{i_1} c_{i_2} \cdots c_{i_r})$, where the summation is taken over all selections of size r from the t conditions.

Consequently, on the right-hand side of Eq. (2), x is counted

$$1 - r + \binom{r}{2} - \binom{r}{3} + \cdots + (-1)^r \binom{r}{r} = [1 + (-1)]^r = 0^r = 0 \text{ times},$$

by the binomial theorem. Therefore, the two sides of Eq. (2) count the same elements from S, and the equality is verified.

An immediate corollary of this principle is given as follows:

COROLLARY 8.1 Under the hypotheses of Theorem 8.1, the number of elements in S that satisfy at least one of the conditions c_i, where $1 \le i \le t$, is given by $N(c_1 \text{ or } c_2 \text{ or } \ldots \text{ or } c_t) = N - \overline{N}$.

Before solving some examples, we examine some further notation for simplifying the statement of Theorem 8.1.

We write

$$S_0 = N,$$
$$S_1 = [N(c_1) + N(c_2) + \cdots + N(c_t)],$$
$$S_2 = [N(c_1c_2) + N(c_1c_3) + \cdots + N(c_1c_t) + N(c_2c_3) + \cdots + N(c_{t-1}c_t)],$$

and, in general,

$$S_k = \sum N(c_{i_1}c_{i_2} \cdots c_{i_k}),\ 1 \le k \le t,$$

where the summation is taken over all selections of size k from the collection of t conditions. Hence S_k has $\binom{t}{k}$ summands in it.

Using this notation we can rewrite the result in Eq. (2) as

$$\overline{N} = S_0 - S_1 + S_2 - S_3 + \cdots + (-1)^t S_t.$$

Now let us look at how this principle is used to solve certain enumeration problems.

EXAMPLE 8.4

Determine the number of positive integers n where $1 \le n \le 100$ and n is *not* divisible by 2, 3, or 5.

Here $S = \{1, 2, 3, \ldots, 100\}$ and $N = 100$. For $n \in S$, n satisfies

a) condition c_1 if n is divisible by 2,

b) condition c_2 if n is divisible by 3, and

c) condition c_3 if n is divisible by 5.

Then the answer to this problem is $N(\overline{c}_1\overline{c}_2\overline{c}_3)$.

As in Section 5.2 we use the notation $\lfloor r \rfloor$ to denote the greatest integer less than or equal to r, for any real number r. This function proves to be helpful in this problem as we find that

$N(c_1) = \lfloor 100/2 \rfloor = 50$ [since the 50 $(= \lfloor 100/2 \rfloor)$ positive integers 2, 4, 6, 8, ..., 96, 98 $(= 2 \cdot 49)$, 100 $(= 2 \cdot 50)$ are divisible by 2];

$N(c_2) = \lfloor 100/3 \rfloor = \lfloor 33\ 1/3 \rfloor = 33$ [since the 33 $(= \lfloor 100/3 \rfloor)$ positive integers 3, 6, 9, 12, ..., 96 $(= 3 \cdot 32)$, 99 $(= 3 \cdot 33)$ are divisible by 3];

$N(c_3) = \lfloor 100/5 \rfloor = 20$;

$N(c_1c_2) = \lfloor 100/6 \rfloor = 16$ [since there are 16 $(= \lfloor 100/6 \rfloor)$ elements in S that are divisible by both 2 and 3—hence divisible by lcm(2, 3) $= 2 \cdot 3 = 6$];

$N(c_1c_3) = \lfloor 100/10 \rfloor = 10$;

$N(c_2c_3) = \lfloor 100/15 \rfloor = 6$; and

$N(c_1c_2c_3) = \lfloor 100/30 \rfloor = 3$.

Applying the Principle of Inclusion and Exclusion, we find that

$$\begin{aligned} N(\overline{c}_1\overline{c}_2\overline{c}_3) &= S_0 - S_1 + S_2 - S_3 = N - [N(c_1) + N(c_2) + N(c_3)] \\ &\quad + [N(c_1c_2) + N(c_1c_3) + N(c_2c_3)] - N(c_1c_2c_3) \\ &= 100 - [50 + 33 + 20] + [16 + 10 + 6] - 3 = 26. \end{aligned}$$

(These 26 numbers are 1, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 49, 53, 59, 61, 67, 71, 73, 77, 79, 83, 89, 91, and 97.)

EXAMPLE 8.5

In Chapter 1 we found the number of nonnegative integer solutions to the equation $x_1 + x_2 + x_3 + x_4 = 18$. We now answer the same question with the extra restriction that $x_i \le 7$, for all $1 \le i \le 4$.

Here S is the set of solutions of $x_1 + x_2 + x_3 + x_4 = 18$, with $0 \le x_i$ for all $1 \le i \le 4$. So $|S| = N = S_0 = \binom{4+18-1}{18} = \binom{21}{18}$.

We say that a solution x_1, x_2, x_3, x_4 satisfies condition c_i, where $1 \le i \le 4$, if $x_i > 7$ (or $x_i \ge 8$). The answer to the problem is then $N(\overline{c}_1\overline{c}_2\overline{c}_3\overline{c}_4)$.

Here by symmetry $N(c_1) = N(c_2) = N(c_3) = N(c_4)$. To compute $N(c_1)$, we consider the integer solutions for $x_1 + x_2 + x_3 + x_4 = 10$, with each $x_i \ge 0$ for all $1 \le i \le 4$. Then we add 8 to the value of x_1 and get the solutions of $x_1 + x_2 + x_3 + x_4 = 18$ that satisfy condition c_1. Hence $N(c_i) = \binom{4+10-1}{10} = \binom{13}{10}$, for each $1 \le i \le 4$, and $S_1 = \binom{4}{1}\binom{13}{10}$.

Likewise, $N(c_1c_2)$ is the number of integer solutions of $x_1 + x_2 + x_3 + x_4 = 2$, where $x_i \ge 0$ for all $1 \le i \le 4$. So $N(c_1c_2) = \binom{4+2-1}{2} = \binom{5}{2}$, and $S_2 = \binom{4}{2}\binom{5}{2}$.

Since $N(c_ic_jc_k) = 0$ for every selection of three conditions, and $N(c_1c_2c_3c_4) = 0$, we have

$$N(\overline{c}_1\overline{c}_2\overline{c}_3\overline{c}_4) = S_0 - S_1 + S_2 - S_3 + S_4 = \binom{21}{18} - \binom{4}{1}\binom{13}{10} + \binom{4}{2}\binom{5}{2} - 0 + 0 = 246.$$

So of the 1330 nonnegative integer solutions of $x_1 + x_2 + x_3 + x_4 = 18$, only 246 of them satisfy $x_i \le 7$ for each $1 \le i \le 4$.

Our next example establishes the formula conjectured in Section 5.3 for counting onto functions.

EXAMPLE 8.6

For finite sets A, B, where $|A| = m \ge n = |B|$, let $A = \{a_1, a_2, \ldots, a_m\}$, $B = \{b_1, b_2, \ldots, b_n\}$, and S = the set of all functions $f\colon A \to B$. Then $N = S_0 = |S| = n^m$.

For all $1 \le i \le n$, let c_i denote the condition on S where a function $f\colon A \to B$ satisfies c_i if b_i is *not* in the range of f. (Note the difference between c_i here and c_i in Examples 8.4 and 8.5.) Then $N(\overline{c}_i)$ is the number of functions in S that have b_i in their range, and $N(\overline{c}_1\overline{c}_2 \cdots \overline{c}_n)$ counts the number of onto functions $f\colon A \to B$.

For all $1 \le i \le n$, $N(c_i) = (n-1)^m$, because each element of B, except b_i, can be used as the second component of an ordered pair for a function $f\colon A \to B$, whose range does not include b_i. Likewise, for all $1 \le i < j \le n$, there are $(n-2)^m$ functions $f\colon A \to B$ whose range contains neither b_i nor b_j. From these observations we have $S_1 = [N(c_1) + N(c_2) + \cdots + N(c_n)] = n(n-1)^m = \binom{n}{1}(n-1)^m$, and $S_2 = [N(c_1c_2) + N(c_1c_3) + \cdots + N(c_1c_n) + N(c_2c_3) + \cdots + N(c_2c_n) + \cdots + N(c_{n-1}c_n)] = \binom{n}{2}(n-2)^m$. In general, for each $1 \le k \le n$,

$$S_k = \sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} N(c_{i_1}c_{i_2} \cdots c_{i_k}) = \binom{n}{k}(n-k)^m.$$

It then follows by the Principle of Inclusion and Exclusion that the number of onto functions from A to B is

$$N(\overline{c}_1\overline{c}_2\overline{c}_3\cdots\overline{c}_n) = S_0 - S_1 + S_2 - S_3 + \cdots + (-1)^n S_n$$

$$= n^m - \binom{n}{1}(n-1)^m + \binom{n}{2}(n-2)^m - \binom{n}{3}(n-3)^m$$

$$+ \cdots + (-1)^n(n-n)^m = \sum_{i=0}^{n}(-1)^i\binom{n}{i}(n-i)^m$$

$$= \sum_{i=0}^{n}(-1)^i\binom{n}{n-i}(n-i)^m.$$

Before we finish discussing this example, let us note that

$$\sum_{i=0}^{n}(-1)^i\binom{n}{n-i}(n-i)^m$$

can also be evaluated even if $m < n$. Furthermore, for $m < n$, the expression

$$N(\overline{c}_1\overline{c}_2\overline{c}_3\cdots\overline{c}_n)$$

still counts the number of functions $f\colon A \to B$, where $|A| = m$, $|B| = n$, and each element of B is in the range of f. But now this number is 0.

For example, suppose that $m = 3 < 7 = n$. Then $N(\overline{c}_1\overline{c}_2\overline{c}_3\cdots\overline{c}_7)$ counts the number of onto functions $f\colon A \to B$ for $|A| = 3$ and $|B| = 7$. We know this number is 0, and we also find that

$$\sum_{i=0}^{7}(-1)^i\binom{7}{7-i}(7-i)^3 = \binom{7}{7}7^3 - \binom{7}{6}6^3 + \binom{7}{5}5^3 - \binom{7}{4}4^3 + \binom{7}{3}3^3 - \binom{7}{2}2^3 + \binom{7}{1}1^3 - \binom{7}{0}0^3$$

$$= 343 - 1512 + 2625 - 2240 + 945 - 168 + 7 - 0 = 0.$$

Hence, for all $m, n \in \mathbf{Z}^+$, if $m < n$, then

$$\sum_{i=0}^{n}(-1)^i\binom{n}{n-i}(n-i)^m = 0.$$

We now solve a problem similar to those in Chapter 3 that dealt with Venn diagrams.

EXAMPLE 8.7

In how many ways can the 26 letters of the alphabet be permuted so that none of the patterns *car*, *dog*, *pun*, or *byte* occurs?

Let S denote the set of all permutations of the 26 letters. Then $|S| = 26!$ For each $1 \le i \le 4$, a permutation in S is said to satisfy condition c_i if the permutation contains the pattern *car*, *dog*, *pun*, or *byte*, respectively.

In order to compute $N(c_1)$, for example, we count the number of ways the 24 symbols *car*, $b, d, e, f, \ldots, p, q, s, t, \ldots, x, y, z$ can be permuted. So $N(c_1) = 24!$, and in a similar way we obtain

$$N(c_2) = N(c_3) = 24!, \quad \text{while } N(c_4) = 23!$$

For $N(c_1c_2)$ we deal with the 22 symbols *car*, *dog*, $b, e, f, h, i, \ldots, m, n, p, q, s, t, \ldots, x, y, z$, which can be permuted in 22! ways. Hence $N(c_1c_2) = 22!$, and comparable calculations give

$$N(c_1c_3) = N(c_2c_3) = 22!, \qquad N(c_ic_4) = 21!, \quad i \neq 4.$$

Furthermore,

$$N(c_1c_2c_3) = 20!, \qquad N(c_ic_jc_4) = 19!, \qquad 1 \le i < j \le 3,$$
$$N(c_1c_2c_3c_4) = 17!$$

So the number of permutations in S that contain none of the given patterns is

$$N(\overline{c}_1\overline{c}_2\overline{c}_3\overline{c}_4) = 26! - [3(24!) + 23!] + [3(22!) + 3(21!)] - [20! + 3(19!)] + 17!$$

Our next example deals with a number theory problem.

EXAMPLE 8.8

For $n \in \mathbf{Z}^+$, $n \ge 2$, let $\phi(n)$ be the number of positive integers m, where $1 \le m < n$ and $\gcd(m, n) = 1$ — that is, m, n are relatively prime. This function is known as *Euler's phi function*, and it arises in several situations in abstract algebra involving enumeration. We find that $\phi(2) = 1$, $\phi(3) = 2$, $\phi(4) = 2$, $\phi(5) = 4$, and $\phi(6) = 2$. For each prime p, $\phi(p) = p - 1$. We would like to derive a formula for $\phi(n)$ that is related to n so that we need not make a case-by-case comparison for each m, $1 \le m < n$, against the integer n.

The derivation of our formula will use the Principle of Inclusion and Exclusion as in Example 8.4. We proceed as follows: For $n \ge 2$, use the Fundamental Theorem of Arithmetic to write $n = p_1^{e_1}p_2^{e_2}\cdots p_t^{e_t}$, where $p_1, p_2, \ldots, p_t$ are distinct primes and $e_i \ge 1$, for all $1 \le i \le t$. We consider the case where $t = 4$. This will be enough to demonstrate the general idea.

With $S = \{1, 2, 3, \ldots, n\}$, we have $N = S_0 = |S| = n$, and for each $1 \le i \le 4$ we say that $k \in S$ satisfies condition c_i if k is divisible by p_i. For $1 \le k < n$, $\gcd(k, n) = 1$ if k is not divisible by any of the primes p_i, where $1 \le i \le 4$. Hence $\phi(n) = N(\overline{c}_1\overline{c}_2\overline{c}_3\overline{c}_4)$.

For each $1 \le i \le 4$, we have $N(c_i) = n/p_i$; $N(c_ic_j) = n/(p_ip_j)$, for all $1 \le i < j \le 4$. Also, $N(c_ic_jc_\ell) = n/(p_ip_jp_\ell)$, for all $1 \le i < j < \ell \le 4$, and $N(c_1c_2c_3c_4) = n/(p_1p_2p_3p_4)$. So

$$\begin{aligned}
\phi(n) &= S_0 - S_1 + S_2 - S_3 + S_4 \\
&= n - \left[\frac{n}{p_1} + \cdots + \frac{n}{p_4}\right] + \left[\frac{n}{p_1p_2} + \frac{n}{p_1p_3} + \cdots + \frac{n}{p_3p_4}\right] \\
&\quad - \left[\frac{n}{p_1p_2p_3} + \cdots + \frac{n}{p_2p_3p_4}\right] + \frac{n}{p_1p_2p_3p_4} \\
&= n\left[1 - \left(\frac{1}{p_1} + \cdots + \frac{1}{p_4}\right) + \left(\frac{1}{p_1p_2} + \frac{1}{p_1p_3} + \cdots + \frac{1}{p_3p_4}\right)\right. \\
&\quad \left. - \left(\frac{1}{p_1p_2p_3} + \cdots + \frac{1}{p_2p_3p_4}\right) + \frac{1}{p_1p_2p_3p_4}\right] \\
&= \frac{n}{p_1p_2p_3p_4}\,[p_1p_2p_3p_4 - (p_2p_3p_4 + p_1p_3p_4 + p_1p_2p_4 + p_1p_2p_3) \\
&\quad + (p_3p_4 + p_2p_4 + p_2p_3 + p_1p_4 + p_1p_3 + p_1p_2) \\
&\quad - (p_4 + p_3 + p_2 + p_1) + 1] \\
&= \frac{n}{p_1p_2p_3p_4}\,[(p_1 - 1)(p_2 - 1)(p_3 - 1)(p_4 - 1)] \\
&= n\left[\frac{p_1 - 1}{p_1} \cdot \frac{p_2 - 1}{p_2} \cdot \frac{p_3 - 1}{p_3} \cdot \frac{p_4 - 1}{p_4}\right] = n\prod_{i=1}^{4}\left(1 - \frac{1}{p_i}\right).
\end{aligned}$$

In general, $\phi(n) = n \prod_{p|n}(1 - (1/p))$, where the product is taken over all primes p dividing n. When $n = p$, a prime, $\phi(n) = \phi(p) = p\,[1 - (1/p)] = p - 1$, as we observed earlier. If $n = 23{,}100$, for example, we find that

$$\begin{aligned}\phi(23{,}100) &= \phi(2^2 \cdot 3 \cdot 5^2 \cdot 7 \cdot 11)\\ &= (23{,}100)(1 - (1/2))(1 - (1/3))(1 - (1/5))(1 - (1/7))(1 - (1/11))\\ &= 4800.\end{aligned}$$

The Euler phi function has many interesting properties. We shall investigate some of them in the exercises for this section and in the Supplementary Exercises.

The next example provides another encounter with the circular arrangements introduced in Chapter 1.

EXAMPLE 8.9

Six married couples are to be seated at a circular table. In how many ways can they arrange themselves so that no wife sits next to her husband? (Here, as in Example 1.16, two seating arrangements are considered the same if one is a rotation of the other.)

For $1 \le i \le 6$, we let c_i denote the condition where a seating arrangement has couple i seated next to each other.

To determine $N(c_1)$, for instance, we consider arranging 11 distinct objects — namely, couple 1 (considered as one object) and the other 10 people. Eleven distinct objects can be arranged around a circular table in $(11 - 1)! = 10!$ ways. However, here $N(c_1) = 2(10!)$, where the 2 takes into account whether the wife in couple 1 is seated to the left or right of her husband. Similarly, $N(c_i) = 2(10!)$, for $2 \le i \le 6$, and $S_1 = \binom{6}{1}2(10!)$.

Continuing, let us now compute $N(c_ic_j)$, for $1 \le i < j \le 6$. Here we are arranging 10 distinct objects — couple i (considered as one object), couple j (likewise considered as one object), and the other eight people. Ten distinct objects can be arranged around a circular table in $(10 - 1)! = 9!$ ways. So here $N(c_ic_j) = 2^2(9!)$ because there are two ways for the wife in couple i to be seated next to her husband, and two ways for the wife in couple j to be seated next to her husband. Consequently, $S_2 = \binom{6}{2}2^2(9!)$.

Similar reasoning shows us that

$$N(c_1c_2c_3) = 2^3(8!),\ S_3 = \binom{6}{3}2^3(8!) \qquad N(c_1c_2c_3c_4) = 2^4(7!),\ S_4 = \binom{6}{4}2^4(7!)$$
$$N(c_1c_2c_3c_4c_5) = 2^5(6!),\ S_5 = \binom{6}{5}2^5(6!) \qquad N(c_1c_2c_3c_4c_5c_6) = 2^6(5!),\ S_6 = \binom{6}{6}2^6(5!).$$

With S_0 (the total number of arrangements of the 12 people) $= (12 - 1)! = 11!$, we find that the number of arrangements where no couple is seated side by side is

$$\begin{aligned}N(\overline{c}_1\overline{c}_2\cdots\overline{c}_6) &= \sum_{i=0}^{6}(-1)^i S_i = \sum_{i=0}^{6}(-1)^i\binom{6}{i}2^i(11-i)!\\ &= 39{,}916{,}800 - 43{,}545{,}600 + 21{,}772{,}800 - 6{,}451{,}200\\ &\quad + 1{,}209{,}600 - 138{,}240 + 7680\\ &= 12{,}771{,}840.\end{aligned}$$

Our final example recalls some of the graph theory we studied in Chapter 7.

EXAMPLE 8.10

In a certain area of the countryside are five villages. An engineer is to devise a system of two-way roads so that after the system is completed, no village will be isolated. In how many ways can he do this?

Calling the villages a, b, c, d, and e, we seek the number of loop-free undirected graphs on these vertices, where no vertex is isolated. Consequently, we want to count situations such as those illustrated in parts (a) and (b) of Fig. 8.3, but not situations such as those shown in parts (c) and (d).

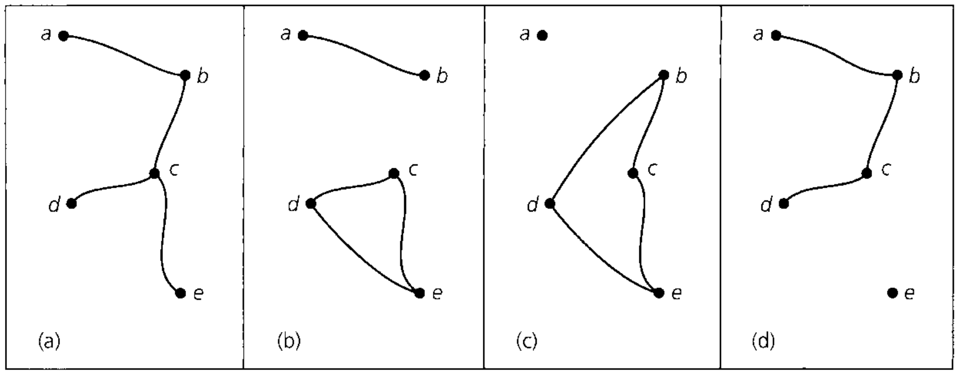


Figure 8.3

Let S be the set of loop-free undirected graphs G on $V = \{a, b, c, d, e\}$. Then $N = S_0 = |S| = 2^{10}$ because there are $\binom{5}{2} = 10$ possible two-way roads for these five villages, and each road can be either included or excluded.

For each $1 \le i \le 5$, let c_i be the condition that a system of these roads isolates village a, b, c, d, and e, respectively. Then the answer to the problem is $N(\overline{c}_1\overline{c}_2\overline{c}_3\overline{c}_4\overline{c}_5)$.

For condition c_1 village a is isolated, so we consider the six edges (roads) $\{b, c\}$, $\{b, d\}$, $\{b, e\}$, $\{c, d\}$, $\{c, e\}$, $\{d, e\}$. With two choices for each edge — namely, put the edge in the graph or leave the edge out — we find that $N(c_1) = 2^6$. Then by symmetry $N(c_i) = 2^6$ for all $2 \le i \le 5$, so $S_1 = \binom{5}{1}2^6$.

When villages a and b are to be isolated, each of the edges $\{c, d\}$, $\{d, e\}$, $\{c, e\}$ may be put in or left out of our graph. This results in 2^3 possibilities, so $N(c_1c_2) = 2^3$, and $S_2 = \binom{5}{2}2^3$.

Similar arguments tell us that $N(c_1c_2c_3) = 2^1$ and $S_3 = \binom{5}{3}2^1$; $N(c_1c_2c_3c_4) = 2^0$ and $S_4 = \binom{5}{4}2^0$; and $N(c_1c_2c_3c_4c_5) = 2^0$ and $S_5 = \binom{5}{5}2^0$.

Consequently,

$$N(\overline{c}_1\overline{c}_2\overline{c}_3\overline{c}_4\overline{c}_5) = 2^{10} - \binom{5}{1}2^6 + \binom{5}{2}2^3 - \binom{5}{3}2^1 + \binom{5}{4}2^0 - \binom{5}{5}2^0 = 768.$$

EXERCISES 8.1

1. Let S be a finite set with $|S| = N$ and let c_1, c_2, c_3, c_4 be four conditions, each of which may be satisfied by one or more of the elements of S. Prove that $N(\overline{c}_2\overline{c}_3\overline{c}_4) = N(c_1\overline{c}_2\overline{c}_3\overline{c}_4) + N(\overline{c}_1\overline{c}_2\overline{c}_3\overline{c}_4)$.

2. Establish the Principle of Inclusion and Exclusion by applying the Principle of Mathematical Induction to the number t of conditions.

3. Of the 100 students in Example 8.3, how many are taking (a) Fundamentals of Computer Programming but none of the other three courses; (b) Fundamentals of Computer Programming and Introduction to Economics but neither of the other two courses?

4. Annually, the 65 members of the maintenance staff sponsor a "Christmas in July" picnic for the 400 summer employees at their company. For these 65 people, 21 bring hot dogs, 35 bring fried chicken, 28 bring salads, 32 bring desserts, 13 bring hot dogs and fried chicken, 10 bring hot dogs and salads, 9 bring hot dogs and desserts, 12 bring fried chicken and salads, 17 bring fried chicken and desserts, 14 bring salads and desserts, 4 bring hot dogs, fried chicken, and salads, 6 bring hot dogs, fried chicken, and desserts, 5 bring hot dogs, salads, and desserts, 7 bring fried chicken, salads, and desserts, and 2 bring all four food items. Those (of the 65) who do not bring any of these four food items are responsible for setting up and cleaning up for the picnic. How many of the 65 maintenance staff will (a) help to set up and clean up for the picnic? (b) bring only hot dogs? (c) bring exactly one food item?

5. Determine the number of positive integers n, $1 \leq n \leq 2000$, that are

a) not divisible by 2, 3, or 5

b) not divisible by 2, 3, 5, or 7

c) not divisible by 2, 3, or 5, but are divisible by 7

6. Determine how many integer solutions there are to $x_1 + x_2 + x_3 + x_4 = 19$, if

a) $0 \leq x_i$ for all $1 \leq i \leq 4$

b) $0 \leq x_i < 8$ for all $1 \leq i \leq 4$

c) $0 \leq x_1 \leq 5$, $0 \leq x_2 \leq 6$, $3 \leq x_3 \leq 7$, $3 \leq x_4 \leq 8$

7. In how many ways can one arrange all of the letters in the word INFORMATION so that no pair of consecutive letters occurs more than once? [Here we want to count arrangements such as IINNOOFRMTA and FORTMAIINON but not INFORINMOTA (where "IN" occurs twice) or NORTFNOIAMI (where "NO" occurs twice).]

8. Determine the number of integer solutions to $x_1 + x_2 + x_3 + x_4 = 19$ where $-5 \leq x_i \leq 10$ for all $1 \leq i \leq 4$.

9. Determine the number of positive integers x where $x \leq 9{,}999{,}999$ and the sum of the digits in x equals 31.

10. Professor Bailey has just completed writing the final examination for his course in advanced engineering mathematics. This examination has 12 questions, whose total value is to be 200 points. In how many ways can Professor Bailey assign the 200 points if each question must count for at least 10, but not more than 25, points and the point value for each question is to be a multiple of 5?

11. At Flo's Flower Shop, Flo wants to arrange 15 different plants on five shelves for a window display. In how many ways can she arrange them so that each shelf has at least one, but no more than four, plants?

12. In how many ways can Troy select nine marbles from a bag of twelve (identical except for color), where three are red, three blue, three white, and three green?

13. Find the number of permutations of $a, b, c, \ldots, x, y, z$, in which none of the patterns *spin*, *game*, *path*, or *net* occurs.

14. Answer the question in Example 8.10 for the case of six villages.

15. If eight distinct dice are rolled, what is the probability that all six numbers appear?

16. How many social security numbers (nine-digit sequences) have each of the digits 1, 3, and 7 appearing at least once?

17. In how many ways can three x's, three y's, and three z's be arranged so that no consecutive triple of the same letter appears?

18. Frostburg township sponsors four Boy Scout troops, each with 20 boys. If the head scoutmaster selects 50 of these boys to represent this township at the state jamboree, what is the probability that his selection will include at least one boy from each of the four troops?

19. If Zachary rolls a fair die five times, what is the probability that the sum of his five rolls is 20?

20. At a 12-week conference in mathematics, Sharon met seven of her friends from college. During the conference she met each friend at lunch 35 times, every pair of them 16 times, every trio eight times, every foursome four times, each set of five twice, and each set of six once, but never all seven at once. If she had lunch every day during the 84 days of the conference, did she ever have lunch alone?

21. Compute $\phi(n)$ for n equal to (a) 51; (b) 420; (c) 12300.

22. Compute $\phi(n)$ for n equal to (a) 5186; (b) 5187; (c) 5188.

23. Let $n \in \mathbf{Z}^+$. (a) Determine $\phi(2^n)$. (b) Determine $\phi(2^n p)$, where p is an odd prime.

24. For which $n \in \mathbf{Z}^+$ is $\phi(n)$ odd?

25. How many positive integers n less than 6000 (a) satisfy $\gcd(n, 6000) = 1$? (b) share a common prime divisor with 6000?

26. If $m, n \in \mathbf{Z}^+$, prove that $\phi(n^m) = n^{m-1}\phi(n)$.

27. Find three values for $n \in \mathbf{Z}^+$ where $\phi(n) = 16$.

28. For which positive integers n is $\phi(n)$ a power of 2?

29. For which positive integers n does 4 divide $\phi(n)$?

30. At an upcoming family reunion, five families — each consisting of a husband, wife, and one child — are to be seated around a circular table. In how many ways can these 15 people be arranged around the table so that no family is seated all together? (Here, as in Example 8.9, two seating arrangements are considered the same if one is a rotation of the other.)

8.2 Generalizations of the Principle

Consider a set S with $|S| = N$, and conditions $c_1, c_2, \ldots, c_t$ satisfied by some of the elements of S. In Section 8.1 we saw how the Principle of Inclusion and Exclusion provides a way to determine $N(\overline{c}_1\overline{c}_2 \cdots \overline{c}_t)$, the number of elements in S that satisfy none of the t conditions. If $m \in \mathbf{Z}^+$ and $1 \leq m \leq t$, we now want to determine E_m, which denotes the number of elements in S that satisfy *exactly m* of the t conditions. (At present we can obtain E_0.)

We can write formulas such as

$$E_1 = N(c_1\overline{c}_2\overline{c}_3 \cdots \overline{c}_t) + N(\overline{c}_1 c_2\overline{c}_3 \cdots \overline{c}_t) + \cdots + N(\overline{c}_1\overline{c}_2\overline{c}_3 \cdots \overline{c}_{t-1}c_t),$$

and

$$E_2 = N(c_1 c_2\overline{c}_3 \cdots \overline{c}_t) + N(c_1\overline{c}_2 c_3 \cdots \overline{c}_t) + \cdots + N(\overline{c}_1\overline{c}_2\overline{c}_3 \cdots \overline{c}_{t-2}c_{t-1}c_t),$$

and although these results do not assist us as much as we should like, they will be a useful starting place as we examine the Venn diagrams for the cases where $t = 3$ and 4.

For Fig. 8.4, where $t = 3$, we place a numbered condition beside the circle representing those elements of S satisfying that particular condition and we also number each of the individual regions shown. Then E_1 equals the number of elements in regions 2, 3, and 4. But we can also write

$$E_1 = N(c_1) + N(c_2) + N(c_3) - 2\,[N(c_1c_2) + N(c_1c_3) + N(c_2c_3)] + 3N(c_1c_2c_3).$$

In $N(c_1) + N(c_2) + N(c_3)$ we count the elements in regions 5, 6, and 7 twice and those in region 8 three times. In the next term, the elements in regions 5, 6, and 7 are deleted twice. We remove the elements in region 8 six times in $2\,[N(c_1c_2) + N(c_1c_3) + N(c_2c_3)]$, so we then add on the term $3N(c_1c_2c_3)$ and end up not counting the elements in region 8 at all. Hence we have $E_1 = S_1 - 2S_2 + 3S_3 = S_1 - \binom{2}{1}S_2 + \binom{3}{2}S_3$.

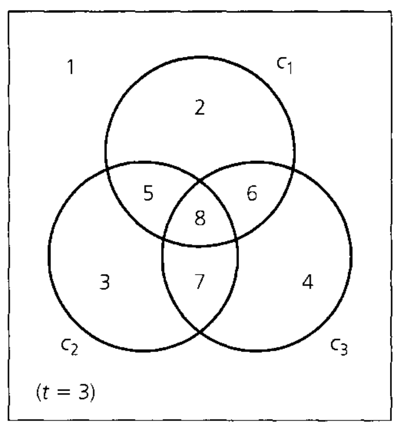


Figure 8.4

When we turn to E_2, our earlier formula indicates that we want to count the elements of S in regions 5, 6, and 7. From the Venn diagram,

$$E_2 = N(c_1c_2) + N(c_1c_3) + N(c_2c_3) - 3N(c_1c_2c_3) = S_2 - 3S_3 = S_2 - \binom{3}{1}S_3,$$

and

$$E_3 = N(c_1c_2c_3) = S_3.$$

In Fig. 8.5, the conditions c_1, c_2, c_3 are associated with circular subsets of S, whereas c_4 is paired with the rather irregularly shaped area made up of regions 4, 8, 9, 11, 12, 13, 14, and 16. For each $1 \le i \le 4$, E_i is determined as follows:

E_1 [regions 2, 3, 4, 5]:

$$\begin{aligned} E_1 &= [N(c_1) + N(c_2) + N(c_3) + N(c_4)] \\ &\quad - 2\,[N(c_1c_2) + N(c_1c_3) + N(c_1c_4) + N(c_2c_3) + N(c_2c_4) + N(c_3c_4)] \\ &\quad + 3\,[N(c_1c_2c_3) + N(c_1c_2c_4) + N(c_1c_3c_4) + N(c_2c_3c_4)] \\ &\quad - 4N(c_1c_2c_3c_4) \\ &= S_1 - 2S_2 + 3S_3 - 4S_4 = S_1 - \binom{2}{1}S_2 + \binom{3}{2}S_3 - \binom{4}{3}S_4. \end{aligned}$$

Note: Taking an element in region 3, we find that it is counted once in E_1 and once in S_1 [in $N(c_3)$]. Taking an element in region 6, we find that it is not counted in E_1; it is counted twice in S_1 [in both $N(c_2)$ and $N(c_3)$] but removed twice in $2S_2$ [for it is counted once in S_2 in $N(c_2c_3)$], so overall it is not counted. The reader should now consider an element from region 12 and one from region 16 and show that each contributes a count of 0 to both sides of the formula for E_1.

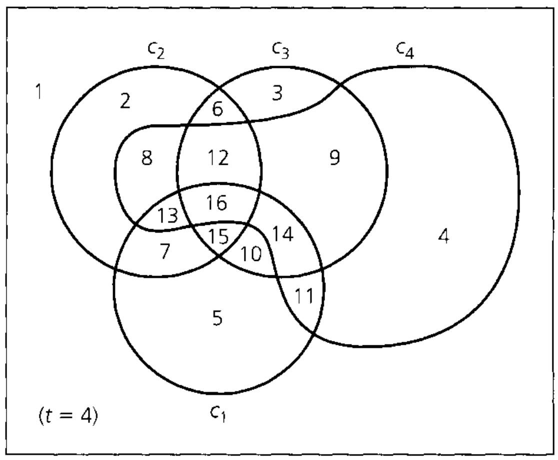


Figure 8.5

E_2 [regions 6–11]:

From Fig. 8.5, $E_2 = S_2 - 3S_3 + 6S_4 = S_2 - \binom{3}{1}S_3 + \binom{4}{2}S_4$. For details on this formula we examine the results in Table 8.1, where next to each summand of S_2, S_3, and S_4 we list the regions whose elements are counted in determining that particular summand. In calculating $S_2 - 3S_3 + 6S_4$ we find the elements in regions 6–11, which are precisely those that are to be counted for E_2.

Table 8.1

S_2	S_3	S_4
$N(c_1c_2)$: 7, 13, 15, 16	$N(c_1c_2c_3)$: 15, 16	$N(c_1c_2c_3c_4)$: 16
$N(c_1c_3)$: 10, 14, 15, 16	$N(c_1c_2c_4)$: 13, 16	
$N(c_1c_4)$: 11, 13, 14, 16	$N(c_1c_3c_4)$: 14, 16	
$N(c_2c_3)$: 6, 12, 15, 16	$N(c_2c_3c_4)$: 12, 16	
$N(c_2c_4)$: 8, 12, 13, 16		
$N(c_3c_4)$: 9, 12, 14, 16		

Finally, the elements for E_3 are found in regions 12–15, and $E_3 = S_3 - 4S_4 = S_3 - \binom{4}{1}S_4$; the elements for E_4 are those in region 16, and $E_4 = S_4$.

These results suggest the following theorem.

THEOREM 8.2 Under the hypotheses of Theorem 8.1, for each $1 \le m \le t$, the number of elements in S that satisfy *exactly* m of the conditions $c_1, c_2, \ldots, c_t$ is given by

$$E_m = S_m - \binom{m+1}{1}S_{m+1} + \binom{m+2}{2}S_{m+2} - \cdots + (-1)^{t-m}\binom{t}{t-m}S_t. \qquad (1)$$

(If $m = 0$, we obtain Theorem 8.1.)

Proof: Arguing as in Theorem 8.1, let $x \in S$ and consider the following three cases.

a) When x satisfies fewer than m conditions, it contributes a count of 0 to each of the terms $E_m, S_m, S_{m+1}, \ldots, S_t$, so it is not counted on either side of the equation.

b) When x satisfies exactly m of the conditions, it is counted once in E_m and once in S_m, but not in $S_{m+1}, \ldots, S_t$. Consequently, it is included once in the count for either side of the equation.

c) Suppose x satisfies r of the conditions, where $m < r \le t$. Then x contributes nothing to E_m. Yet it is counted $\binom{r}{m}$ times in S_m, $\binom{r}{m+1}$ times in $S_{m+1}, \ldots,$ and $\binom{r}{r}$ times in S_r, but 0 times for any term beyond S_r. So on the right-hand side of the equation, x is counted $\binom{r}{m} - \binom{m+1}{1}\binom{r}{m+1} + \binom{m+2}{2}\binom{r}{m+2} - \cdots + (-1)^{r-m}\binom{r}{r-m}\binom{r}{r}$ times.

For $0 \le k \le r - m$,

$$\binom{m+k}{k}\binom{r}{m+k} = \frac{(m+k)!}{k!\,m!}\cdot\frac{r!}{(m+k)!(r-m-k)!}$$
$$= \frac{r!}{m!}\cdot\frac{1}{k!(r-m-k)!} = \frac{r!}{m!(r-m)!}\cdot\frac{(r-m)!}{k!(r-m-k)!}$$
$$= \binom{r}{m}\binom{r-m}{k}.$$

Consequently, on the right-hand side of Eq. (1), x is counted

$$\binom{r}{m}\binom{r-m}{0} - \binom{r}{m}\binom{r-m}{1} + \binom{r}{m}\binom{r-m}{2} - \cdots + (-1)^{r-m}\binom{r}{m}\binom{r-m}{r-m}$$
$$= \binom{r}{m}\left[\binom{r-m}{0} - \binom{r-m}{1} + \binom{r-m}{2} - \cdots + (-1)^{r-m}\binom{r-m}{r-m}\right]$$
$$= \binom{r}{m}[1-1]^{r-m} = \binom{r}{m}\cdot 0 = 0 \text{ times},$$

and the formula is verified.

Based on this result, if L_m denotes the number of elements of S (under the hypotheses of Theorem 8.1) that satisfy *at least* m of the t conditions, then we have the following formula.

COROLLARY 8.2 $L_m = S_m - \binom{m}{m-1}S_{m+1} + \binom{m+1}{m-1}S_{m+2} - \cdots + (-1)^{t-m}\binom{t-1}{m-1}S_t$.

Proof: A proof is outlined in the exercises at the end of this section.

When $m = 1$, the result in Corollary 8.2 becomes

$$L_1 = S_1 - \binom{1}{0}S_2 + \binom{2}{0}S_3 - \cdots + (-1)^{t-1}\binom{t-1}{0}S_t$$
$$= S_1 - S_2 + S_3 - \cdots + (-1)^{t-1}S_t.$$

Comparing this with the result in Theorem 8.1, we find that

$$L_1 = N - \overline{N} = |S| - \overline{N}.$$

This result is not much of a surprise, because an element x of S is counted in L_1 if it satisfies at least one of the conditions $c_1, c_2, c_3, \ldots, c_t$ —that is, if $x \in S$ and x is not counted in $\overline{N} = N(\overline{c}_1\overline{c}_2\overline{c}_3 \cdots \overline{c}_t)$.

EXAMPLE 8.11

Looking back to Example 8.10, we shall find the numbers of systems of two-way roads so that exactly (E_2) and at least (L_2) two of the villages remain isolated.

The previously calculated results for this example show

$$E_2 = S_2 - \binom{3}{1}S_3 + \binom{4}{2}S_4 - \binom{5}{3}S_5 = 80 - 3(20) + 6(5) - 10(1) = 40,$$
$$L_2 = S_2 - \binom{2}{1}S_3 + \binom{3}{1}S_4 - \binom{4}{1}S_5 = 80 - 2(20) + 3(5) - 4(1) = 51.$$

EXERCISES 8.2

1. For the situation in Examples 8.10 and 8.11 compute E_i for $0 \le i \le 5$ and show that $\sum_{i=0}^{5} E_i = N = |S|$.

2. a) In how many ways can the letters in ARRANGEMENT be arranged so that there are exactly two pairs of consecutive identical letters? at least two pairs of consecutive identical letters?

   b) Answer part (a), replacing two with three.

3. In how many ways can one arrange the letters in CORRESPONDENTS so that (a) there is no pair of consecutive identical letters? (b) there are exactly two pairs of consecutive identical letters? (c) there are at least three pairs of consecutive identical letters?

4. Let $A = \{1, 2, 3, \ldots, 10\}$, and $B = \{1, 2, 3, \ldots, 7\}$. How many functions $f: A \to B$ satisfy $|f(A)| = 4$? How many have $|f(A)| \le 4$?

5. In how many ways can one distribute ten distinct prizes among four students with exactly two students getting nothing? How many ways have at least two students getting nothing?

6. Zelma is having a luncheon for herself and nine of the women in her tennis league. On the morning of the luncheon she places name cards at the ten places at her table and then leaves to run a last-minute errand. Her husband, Herbert, comes home from his morning tennis match and unfortunately leaves the back door open. A gust of wind scatters the ten name cards. In how many ways can Herbert replace the ten cards at the places at the table so that exactly four of the ten women will be seated where Zelma had wanted them? In how many ways will at least four of them be seated where they were supposed to be?

7. If 13 cards are dealt from a standard deck of 52, what is the probability that these 13 cards include (a) at least one card from each suit? (b) exactly one void (for example, no clubs)? (c) exactly two voids?

8. The following provides an outline for proving Corollary 8.2. Fill in the needed details.

   a) First note that $E_t = L_t = S_t$.

   b) What is E_{t-1}, and how are L_t and L_{t-1} related?

   c) Show that $L_{t-1} = S_{t-1} - \binom{t-1}{t-2}S_t$.

   d) For all $1 \le m \le t-1$, how are L_m, L_{m+1}, and E_m related?

   e) Using the results in steps (a) through (d), establish the corollary by a backward type of induction.

8.3 Derangements: Nothing Is in Its Right Place

In elementary calculus the Maclaurin series for the exponential function is given by

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!},$$

so

$$e^{-1} = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} = 1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \cdots.$$

To five places, $e^{-1} = 0.36788$ and $1 - 1 + (1/2!) - (1/3!) + \cdots - (1/7!) \doteq 0.36786$. Consequently, for all $k \in \mathbf{Z}^+$, if $k \geq 7$, then $\sum_{n=0}^{k}((-1)^n)/n!$ is a very good approximation to e^{-1}.

We find these ideas helpful in working some of the following examples.

EXAMPLE 8.12

While at the racetrack, Ralph bets on each of the ten horses in a race to come in according to how they are favored. In how many ways can they reach the finish line so that he loses all of his bets?

Removing the words *horses* and *racetrack* from the problem, we really want to know in how many ways we can arrange the numbers $1, 2, 3, \ldots, 10$ so that 1 is not in first place (its natural position), 2 is not in second place (its natural position), ..., and 10 is not in tenth place (its natural position). These arrangements are called the *derangements* of $1, 2, 3, \ldots, 10$.

The Principle of Inclusion and Exclusion provides the key to calculating the number of derangements. For each $1 \leq i \leq 10$, an arrangement of $1, 2, 3, \ldots, 10$ is said to satisfy condition c_i if integer i is in the ith place. We obtain the number of derangements, denoted by d_{10}, as follows:

$$\begin{aligned} d_{10} &= N(\overline{c}_1\overline{c}_2\overline{c}_3 \cdots \overline{c}_{10}) = 10! - \binom{10}{1}9! + \binom{10}{2}8! - \binom{10}{3}7! + \cdots + \binom{10}{10}0! \\ &= 10![1 - \binom{10}{1}(9!/10!) + \binom{10}{2}(8!/10!) - \binom{10}{3}(7!/10!) + \cdots + \binom{10}{10}(0!/10!)] \\ &= 10![1 - 1 + (1/2!) - (1/3!) + \cdots + (1/10!)] \doteq (10!)(e^{-1}). \end{aligned}$$

The sample space here consists of the $10!$ ways the horses can finish. So the *probability* that Ralph will lose every bet is approximately $(10!)(e^{-1})/(10!) = e^{-1}$. This probability remains (more or less) the same if the number of horses in the race is $11, 12, \ldots$. On the other hand, for n horses, where $n \geq 10$, the probability that our gambler wins at least one of his bets is approximately $1 - e^{-1} \doteq 0.63212$.

EXAMPLE 8.13

The number of derangements of 1, 2, 3, 4 is

$$\begin{aligned} d_4 &= 4![1 - 1 + (1/2!) - (1/3!) + (1/4!)] \\ &= 4![(1/2!) - (1/3!) + (1/4!)] = (4)(3) - 4 + 1 = 9. \end{aligned}$$

These nine derangements are

2143	3142	4123
2341	3412	4312
2413	3421	4321.

Among the $24 - 9 = 15$ permutations of 1, 2, 3, 4 that are *not* derangements one finds 1234, 2314, 3241, 1342, 2431, and 2314.

EXAMPLE 8.14

Peggy has seven books to review for the C–H Company, so she hires seven people to review them. She wants two reviews per book, so the first week she gives each person one book to read and then redistributes the books at the start of the second week. In how many ways can she make these two distributions so that she gets two reviews (by different people) of each book?

She can distribute the books in 7! ways the first week. Numbering both the books and the reviewers (for the first week) as $1, 2, \ldots, 7$, for the second distribution she must arrange these numbers so that none of them is in its natural position. This she can do in d_7 ways. By the rule of product, she can make the two distributions in $(7!)d_7 \doteq (7!)^2(e^{-1})$ ways.

EXERCISES 8.3

1. In how many ways can the integers $1, 2, 3, \ldots, 10$ be arranged in a line so that no even integer is in its natural position?

2. **a)** List all the derangements of 1, 2, 3, 4, 5 where the first three numbers are 1, 2, and 3, in some order.

b) List all the derangements of 1, 2, 3, 4, 5, 6 where the first three numbers are 1, 2, and 3, in some order.

3. How many derangements are there for 1, 2, 3, 4, 5?

4. How many permutations of 1, 2, 3, 4, 5, 6, 7 are not derangements?

5. **a)** Let $A = \{1, 2, 3, \ldots, 7\}$. A function $f: A \to A$ is said to have a *fixed point* if for some $x \in A$, $f(x) = x$. How many one-to-one functions $f: A \to A$ have at least one fixed point?

b) In how many ways can we devise a secret code by assigning to each letter of the alphabet a different letter to represent it?

6. How many derangements of 1, 2, 3, 4, 5, 6, 7, 8 start with (a) 1, 2, 3, and 4, in some order? (b) 5, 6, 7, and 8, in some order?

7. For the positive integers $1, 2, 3, \ldots, n-1, n$, there are 11,660 derangements where 1, 2, 3, 4, and 5 appear in the first five positions. What is the value of n?

8. Four applicants for a job are to be interviewed for 30 minutes each: 15 minutes with each of supervisors Nancy and Yolanda. (The interviews are in separate rooms, and interviewing starts at 9:00 A.M.) (a) In how many ways can these interviews be scheduled during a one-hour period? (b) One applicant, named Josephine, arrives at 9:00 A.M. What is the probability that she will have her two interviews one after the other? (c) Regina, another applicant, arrives at 9:00 A.M. and hopes to be finished in time to leave by 9:50 A.M. for another appointment. What is the probability that Regina will be able to leave on time?

9. In how many ways can Mrs. Ford distribute ten distinct books to her ten children (one book to each child) and then collect and redistribute the books so that each child has the opportunity to peruse two different books?

10. **a)** When n balls, numbered $1, 2, 3, \ldots, n$ are taken in succession from a container, a *rencontre* occurs if the mth ball withdrawn is numbered m, for some $1 \le m \le n$. Find the probability of getting (i) no rencontres; (ii) (exactly) one rencontre, (iii) at least one rencontre; and (iv) r rencontres, where $1 \le r \le n$.

b) Approximate the answers to the questions in part (a).

11. Ten women attend a business luncheon. Each woman checks her coat and attaché case. Upon leaving, each woman is given a coat and case at random. (a) In how many ways can the coats and cases be distributed so that no woman gets either of her possessions? (b) In how many ways can they be distributed so that no woman gets back both of her possessions?

12. Ms. Pezzulo teaches geometry and then biology to a class of 12 advanced students in a classroom that has only 12 desks. In how many ways can she assign the students to these desks so that (a) no student is seated at the same desk for both classes? (b) there are exactly six students each of whom occupies the same desk for both classes?

13. Give a combinatorial argument to verify that for all $n \in \mathbf{Z}^+$,

$$n! = \binom{n}{0}d_0 + \binom{n}{1}d_1 + \binom{n}{2}d_2 + \cdots + \binom{n}{n}d_n = \sum_{k=0}^{n}\binom{n}{k}d_k.$$

(For each $1 \le k \le n$, $d_k =$ the number of derangements of $1, 2, 3, \ldots, k$; $d_0 = 1$.)

14. **a)** In how many ways can the integers $1, 2, 3, \ldots, n$ be arranged in a line so that none of the patterns 12, 23, 34, ..., $(n-1)n$ occurs?

b) Show that the result in part (a) equals $d_{n-1} + d_n$. (d_n = the number of derangements of $1, 2, 3, \ldots, n$.)

15. Answer part (a) of Exercise 14 if the numbers are arranged in a circle, and, as we count clockwise about the circle, none of the patterns 12, 23, 34, ..., $(n-1)n$, $n1$ occurs.

16. What is the probability that the gambler in Example 8.12 wins (a) (exactly) five of his bets? (b) at least five of his bets?

8.4 Rook Polynomials

Consider the six-square "chessboard" shown in Fig. 8.6 (*Note*: The shaded squares are not part of the chessboard.). In chess a piece called a *rook* or *castle* is allowed at one turn to be moved horizontally or vertically over as many unoccupied spaces as one wishes. Here a rook in square 3 of the figure could be moved in one turn to squares 1, 2, or 4. A rook at square 5 could be moved to square 6 or square 2 (even though there is no square between squares 5 and 2).

For $k \in \mathbf{Z}^+$ we want to determine the number of ways in which k rooks can be placed on the unshaded squares of this chessboard so that no two of them can take each other—that is, no two of them are in the same row or column of the chessboard. This number is denoted by r_k, or by $r_k(C)$ if we wish to stress that we are working on a particular chessboard C.

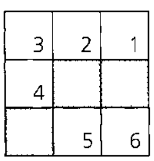


Figure 8.6

For any chessboard, r_1 is the number of squares on the board. Here $r_1 = 6$. Two nontaking rooks can be placed at the following pairs of positions: $\{1, 4\}$, $\{1, 5\}$, $\{2, 4\}$, $\{2, 6\}$, $\{3, 5\}$, $\{3, 6\}$, $\{4, 5\}$, and $\{4, 6\}$, so $r_2 = 8$. Continuing, we find that $r_3 = 2$, using the locations $\{1, 4, 5\}$ and $\{2, 4, 6\}$; $r_k = 0$, for $k \geq 4$.

With $r_0 = 1$, the *rook polynomial*, $r(C, x)$, for the chessboard in Fig. 8.6 is defined as $r(C, x) = 1 + 6x + 8x^2 + 2x^3$. For each $k \geq 0$, the coefficient of x^k is the number of ways we can place k nontaking rooks on chessboard C.

What we have done here (using a case-by-case analysis) soon proves tedious. As the size of the board increases, we have to consider cases wherein numbers such as r_4 and r_5 are nonzero. Consequently, we shall now make some observations that will allow us to make use of small boards and somehow break up a large board into smaller *subboards*.

The chessboard C in Fig. 8.7 is made up of 11 unshaded squares. We note that C consists of a 2×2 subboard C_1 located in the upper left corner and a seven-square subboard C_2 located in the lower right corner. These subboards are *disjoint* because they have no squares in the same row or column of C.

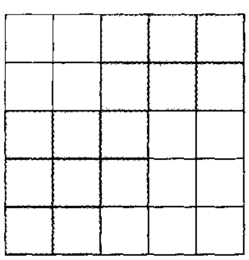


Figure 8.7

Calculating as we did for our first chessboard, here we find

$$r(C_1, x) = 1 + 4x + 2x^2, \qquad r(C_2, x) = 1 + 7x + 10x^2 + 2x^3,$$
$$r(C, x) = 1 + 11x + 40x^2 + 56x^3 + 28x^4 + 4x^5 = r(C_1, x) \cdot r(C_2, x).$$

Hence $r(C, x) = r(C_1, x) \cdot r(C_2, x)$. But did this occur by luck or is something happening here that we should examine more closely? For example, to obtain r_3 for C, we need to know in how many ways three nontaking rooks can be placed on board C. These fall into three cases:

a) All three rooks are on subboard C_2 (and none is on C_1): (2)(1) = 2 ways.

b) Two rooks are on subboard C_2 and one is on C_1: (10)(4) = 40 ways.

c) One rook is on subboard C_2 and two are on C_1: (7)(2) = 14 ways.

Consequently, three nontaking rooks can be placed on board C in $(2)(1) + (10)(4) + (7)(2) = 56$ ways. Here we see that 56 arises just as the coefficient of x^3 does in the product $r(C_1, x) \cdot r(C_2, x)$.

In general, if C is a chessboard made up of *pairwise* disjoint subboards $C_1, C_2, \ldots, C_n$, then $r(C, x) = r(C_1, x)r(C_2, x) \cdots r(C_n, x)$.

The last result for this section demonstrates the type of principle we have seen in other results in combinatorial and discrete mathematics: Given a large chessboard, break it into smaller subboards whose rook polynomials can be determined by inspection.

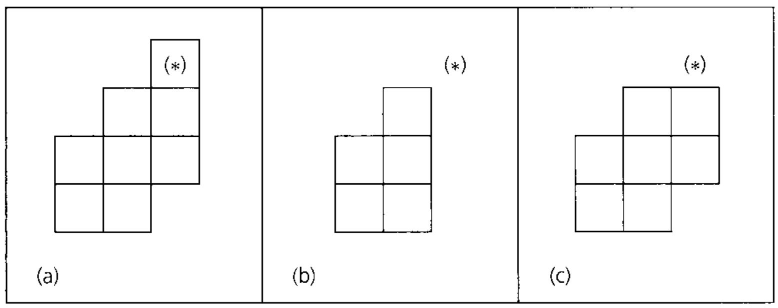


Figure 8.8

Consider chessboard C in Fig. 8.8(a). For $k \geq 1$, suppose we wish to place k nontaking rooks on C. For each square of C, such as the one designated by $(*)$, there are two possibilities to examine.

a) Place one rook on the designated square. Then we remove, as possible locations for the other $k - 1$ rooks, all other squares of C in the same row or column as the designated square. We use C_s to denote the remaining smaller subboard [seen in Fig. 8.8(b)].

b) We do not use the designated square at all. The k rooks are placed on the subboard C_e [C with the one designated square eliminated — as shown in Fig. 8.8(c)].

Since these two cases are all-inclusive and mutually disjoint,

$$r_k(C) = r_{k-1}(C_s) + r_k(C_e).$$

From this we see that

$$r_k(C)x^k = r_{k-1}(C_s)x^k + r_k(C_e)x^k. \tag{1}$$

If n is the number of squares in the chessboard (here n is 8), then Eq. (1) is valid for all $1 \leq k \leq n$, and we write

$$\sum_{k=1}^{n} r_k(C)x^k = \sum_{k=1}^{n} r_{k-1}(C_s)x^k + \sum_{k=1}^{n} r_k(C_e)x^k. \tag{2}$$

For Eq. (2) we realize that the summations may stop before $k = n$. We have seen cases, as in Fig. 8.6, where r_n and some prior r_k's are 0. The summations start at $k = 1$, for otherwise we could find ourselves with the term $r_{-1}(C_s)x^0$ in the first summand on the right-hand side of Eq. (2).

Equation (2) may be rewritten as

$$\sum_{k=1}^{n} r_k(C)x^k = x\sum_{k=1}^{n} r_{k-1}(C_s)x^{k-1} + \sum_{k=1}^{n} r_k(C_e)x^k \tag{3}$$

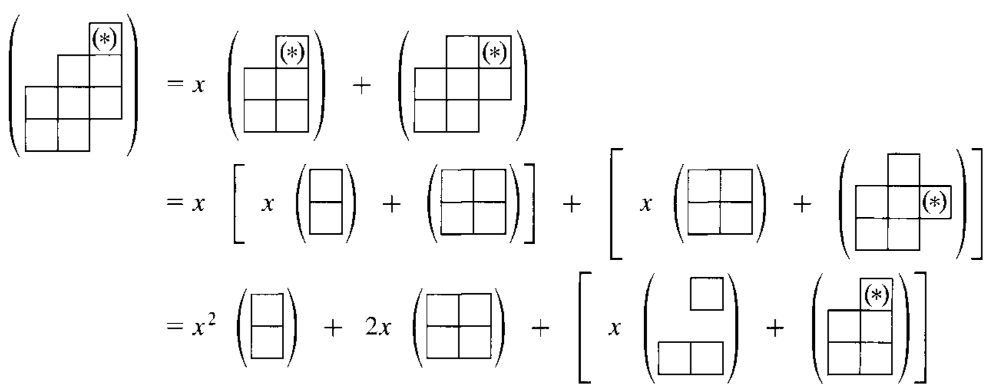
or

$$1 + \sum_{k=1}^{n} r_k(C)x^k = x \cdot r(C_s, x) + \sum_{k=1}^{n} r_k(C_e)x^k + 1,$$

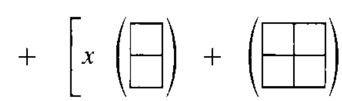
from which it follows that

$$r(C, x) = x \cdot r(C_s, x) + r(C_e, x). \tag{4}$$

We now use this final equation to determine the rook polynomial for the chessboard shown in part (a) of Fig. 8.8. Each time the idea in Eq. (4) is used, we mark the special square we are using with $(*)$. Parentheses are placed about each chessboard to denote the rook polynomial of the board.



$= x^2(1 + 2x) + 2x(1 + 4x + 2x^2) + x(1 + 3x + x^2)$



$= 3x + 12x^2 + 7x^3 + x(1 + 2x) + (1 + 4x + 2x^2) = 1 + 8x + 16x^2 + 7x^3.$

8.5 Arrangements with Forbidden Positions

The rook polynomials of the previous section seem interesting on their own. Now we shall find them useful in solving the following problems.

EXAMPLE 8.15

In making seating arrangements for their son's wedding reception, Grace and Nick are down to four relatives, denoted R_i, for $1 \leq i \leq 4$, who do not get along with one another. There is a single open seat at each of the five tables T_j, where $1 \leq j \leq 5$. Because of family differences,

a) R_1 will not sit at T_1 or T_2.

b) R_2 will not sit at T_2.

c) R_3 will not sit at T_3 or T_4.

d) R_4 will not sit at T_4 or T_5.

This situation is represented in Fig. 8.9. The number of ways we can seat these four people at four different tables, and satisfy conditions (a) through (d), is the number of ways four nontaking rooks can be placed on the chessboard made up of the *unshaded* squares. However, since there are only seven shaded squares, as opposed to thirteen unshaded ones, it would be easier to work with the shaded chessboard.

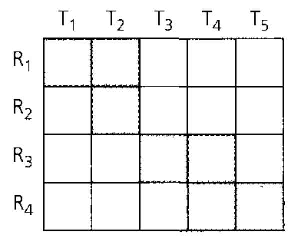


Figure 8.9

We start with the conditions that are required for us to apply the Principle of Inclusion and Exclusion: For each $1 \le i \le 4$, let c_i be the condition where a seating assignment of these four people (at different tables) is made with relative R_i in a forbidden (shaded) position. As usual, $|S|$ denotes the total number of ways we can place the four relatives, one to a table. Then $|S| = N = S_0 = 5!$

To determine S_1 we consider each of the following:

- $N(c_1) = 4! + 4!$, for there are 4! ways to seat R_2, R_3, and R_4 if R_1 is in forbidden position T_1 and another 4! ways if R_1 is at table T_2, his or her other forbidden position.
- $N(c_2) = 4!$, for after placing R_2 at forbidden table T_2, we must place R_1, R_3, and R_4 at T_1, T_3, T_4, and T_5, one person to a table.
- $N(c_3) = 4! + 4!$, one summand for R_3 being in forbidden position T_3, and the other summand for R_3 being in the forbidden position T_4.
- $N(c_4) = 4! + 4!$, each of the two summands arising when R_4 is placed at each of the two forbidden positions T_4 and T_5.

Hence $S_1 = 7(4!)$.

Turning to S_2 we have these considerations:

- $N(c_1c_2) = 3!$, because after we place R_1 at T_1 and R_2 at T_2, there are three tables (T_3, T_4, and T_5) where R_3 and R_4 can be seated.
- $N(c_1c_3) = 3! + 3! + 3! + 3!$, because there are four cases where R_1 and R_3 are located at forbidden positions:

  i) R_1 at T_1; R_3 at T_3 **ii)** R_1 at T_2; R_3 at T_3
  iii) R_1 at T_1; R_3 at T_4 **iv)** R_1 at T_2; R_3 at T_4.

In a similar manner we find that $N(c_1c_4) = 4(3!)$, $N(c_2c_3) = 2(3!)$, $N(c_2c_4) = 2(3!)$, and $N(c_3c_4) = 3(3!)$. Consequently, $S_2 = 16(3!)$.

Before continuing, we make a few observations about S_1 and S_2. For S_1 we have $7(4!) = 7(5-1)!$, where 7 is the number of shaded squares in Fig. 8.9. Also, $S_2 = 16(3!) = 16(5-2)!$, where 16 is the number of ways two nontaking rooks can be placed on the shaded chessboard.

In general, for all $0 \le i \le 4$, $S_i = r_i(5-i)!$, where r_i is the number of ways in which it is possible to place i nontaking rooks on the shaded chessboard shown in Fig. 8.9.

Consequently, to expedite the solution of this problem, we turn to $r(C, x)$, the rook polynomial of this shaded chessboard. Using the decomposition of C into the disjoint subboards in the upper left and lower right corners, we find that

$$r(C, x) = (1 + 3x + x^2)(1 + 4x + 3x^2) = 1 + 7x + 16x^2 + 13x^3 + 3x^4,$$

so

$$N(\overline{c}_1\overline{c}_2\overline{c}_3\overline{c}_4) = S_0 - S_1 + S_2 - S_3 + S_4 = 5! - 7(4!) + 16(3!) - 13(2!) + 3(1!)$$
$$= \sum_{i=0}^{4}(-1)^i r_i(5 - i)! = 25.$$

Grace and Nick can breathe a sigh of relief. There are 25 ways in which they can seat these last four relatives at the reception and avoid any squabbling.

The next example demonstrates how a bit of rearranging of our chessboard can help in our calculations.

EXAMPLE 8.16

We have a pair of dice; one is red, the other green. We roll these dice six times. What is the probability that we obtain all six values on both the red die and the green die if we know that the ordered pairs (1, 2), (2, 1), (2, 5), (3, 4), (4, 1), (4, 5), and (6, 6) did not occur? [Here an ordered pair (a, b) indicates a on the red die and b on the green.]

Recognizing this problem as one dealing with permutations and forbidden positions, we construct the chessboard shown in Fig. 8.10(a), where the row labels represent the outcome on the red die, the column labels the outcome on the green die, and the shaded squares constitute the forbidden positions. In this figure the shaded squares are scattered. Relabeling the rows and columns, we can redraw the chessboard as shown in Fig. 8.10(b), where we have taken shaded squares in the same row (or column) of the board shown in part (a) and made them adjacent. In Fig. 8.10(b), the chessboard C (of seven shaded squares) is the union of four pairwise disjoint subboards, and so

$$r(C, x) = (1 + 4x + 2x^2)(1 + x)^3 = 1 + 7x + 17x^2 + 19x^3 + 10x^4 + 2x^5.$$

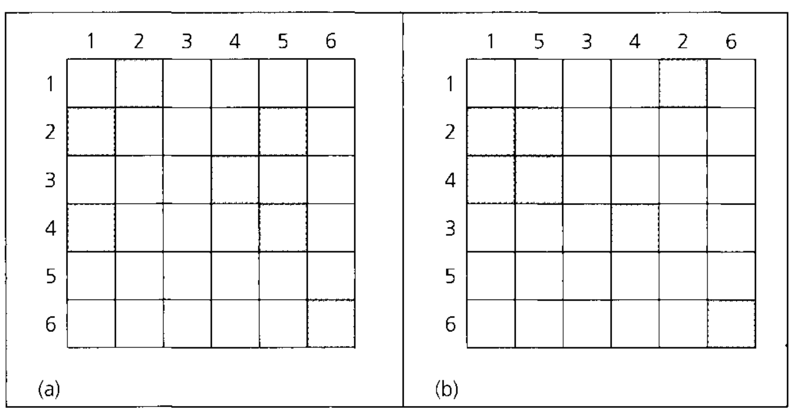


Figure 8.10

For each $1 \leq i \leq 6$, define c_i as the condition where, having rolled the dice six times, we find that all six values occur on both the red die and the green die, but i on the red die is paired with one of the forbidden numbers on the green die. [Note that $N(c_5) = 0$.] Then the number of (ordered) sequences of the six rolls of the dice for the event we are interested in is

$$\begin{aligned}(6!)N(\overline{c}_1\overline{c}_2\overline{c}_3\overline{c}_4\overline{c}_5\overline{c}_6) &= (6!)\sum_{i=0}^{6}(-1)^i S_i = (6!)\sum_{i=0}^{6}(-1)^i r_i (6-i)! \\ &= 6![6! - 7(5!) + 17(4!) - 19(3!) + 10(2!) - 2(1!) + 0(0!)] \\ &= 6![192] = 138{,}240.\end{aligned}$$

Since the sample space consists of all sequences of six ordered pairs selected with repetition from the 29 unshaded squares of the chessboard, the probability of this event is $138{,}240/(29)^6 \doteq 0.00023$.

Our last example provides a unifying idea for what we have done in this section.

EXAMPLE 8.17

Let $A = \{1, 2, 3, 4\}$ and $B = \{u, v, w, x, y, z\}$. How many one-to-one functions $f: A \to B$ satisfy none of the following conditions:

c_1: $f(1) = u$ or v $\quad$ c_2: $f(2) = w$ $\quad$ c_3: $f(3) = w$ or x $\quad$ c_4: $f(4) = x, y,$ or z

As in our two prior examples, we construct a chessboard, as shown in Fig. 8.11. Here we are really interested in the chessboard C made up of the eight shaded squares (which comprise two disjoint subboards). Now

$$r(C, x) = (1 + 2x)(1 + 6x + 9x^2 + 2x^3) = 1 + 8x + 21x^2 + 20x^3 + 4x^4.$$

So

$$\begin{aligned}N(\overline{c}_1\overline{c}_2\overline{c}_3\overline{c}_4) &= S_0 - S_1 + S_2 - S_3 + S_4 \\ &= (6!/2!) - 8(5!/2!) + 21(4!/2!) - 20(3!/2!) + 4(2!/2!) \\ &= \sum_{i=0}^{4}(-1)^i r_i (6-i)!/2! = 76\end{aligned}$$

and there are 76 one-to-one functions $f: A \to B$ where none of the conditions c_1, c_2, c_3, c_4 is satisfied.

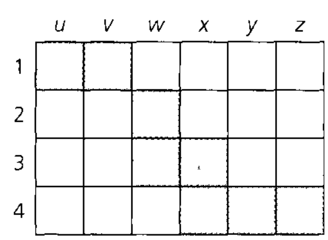


Figure 8.11

Even more so, look back at $N(\overline{c}_1\overline{c}_2\overline{c}_3\overline{c}_4)$ in Example 8.15. Disregarding the vocabulary of the "relatives" and "tables," we realize that we are counting the number of one-to-one functions $g: \{R_1, R_2, R_3, R_4\} \to \{T_1, T_2, T_3, T_4, T_5\}$ where none of the conditions c_1, c_2, c_3, c_4 is satisfied. (The situation is similar for $N(\overline{c}_1\overline{c}_2\overline{c}_3\overline{c}_4\overline{c}_5\overline{c}_6)$ in Example 8.16.)

Finally, for $A = \{1, 2, 3, 4, 5, 6, 7, 8\}$, suppose we want to count the number of one-to-one functions $h: A \to A$ where $h(i) \neq i$ for all $i \in A$. Here the rook polynomial would be

$$r(C, x) = (1 + x)^8 = \sum_{k=0}^{8} \binom{8}{k} x^k$$

and we find that the number of such one-to-one functions h is

$$\binom{8}{0}8! - \binom{8}{1}7! + \binom{8}{2}6! - \binom{8}{3}5! + \cdots + \binom{8}{8}0!$$
$$= 8!\left[1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{1}{8!}\right]$$
$$= d_8, \text{ the number of derangements of } 1, 2, 3, \ldots, 8.$$

EXERCISES 8.4 AND 8.5

1. Verify directly the rook polynomials for (a) the unshaded chessboards in Figs. 8.7 and 8.8(a), and (b) the shaded chessboards in Figs. 8.9 and 8.10(b).

2. Construct or describe a smallest (least number of squares) chessboard for which $r_{10} \neq 0$.

3. a) Find the rook polynomial for the standard 8×8 chessboard.

b) Answer part (a) with 8 replaced by n, for $n \in \mathbf{Z}^+$.

4. Find the rook polynomials for the shaded chessboards in Fig. 8.12.

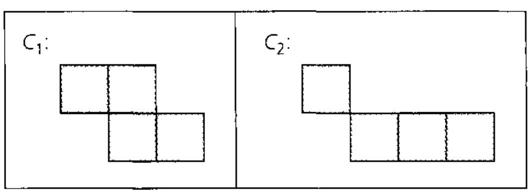


Figure 8.12

5. a) Find the rook polynomials for the shaded chessboards in Fig. 8.13.

b) Generalize the chessboard (and rook polynomial) for Fig. 8.13(i).

6. a) Let C be a chessboard that has m rows and n columns, with $m \le n$ (for a total of mn squares). For $0 \le k \le m$, in how many ways can we arrange k (identical) nontaking rooks on C?

b) For the chessboard C in part (a), determine the rook polynomial $r(C, x)$.

7. Professor Ruth has five graders to correct programs in her courses in Java, C++, SQL, Perl, and VHDL. Graders Jeanne and Charles both dislike SQL, Sandra wants to avoid C++ and VHDL. Paul detests Java and C++, and Todd refuses to work in SQL and Perl. In how many ways can Professor Ruth assign each grader to correct programs in one language, cover all five languages, and keep everyone content?

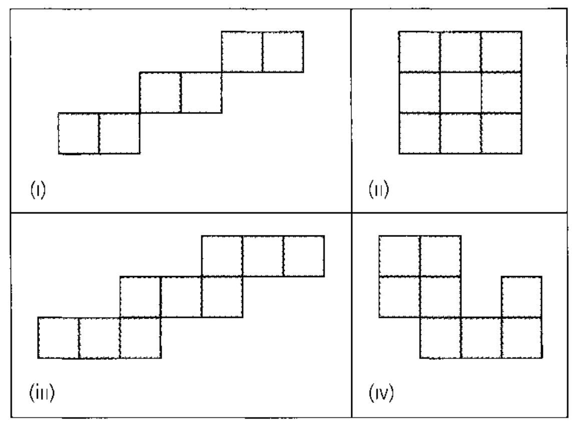


Figure 8.13

8. Why do we have 6! in the term $(6!)N(\overline{c}_1\overline{c}_2 \cdots \overline{c}_6)$ for the solution of Example 8.16?

9. Five professors named Al, Violet, Lynn, Jack, and Mary Lou are to be assigned to teach one class each from among calculus I, calculus II, calculus III, statistics, and combinatorics. Al will not teach calculus II or combinatorics, Lynn cannot stand statistics, Violet and Mary Lou both refuse to teach calculus I or calculus III, and Jack detests calculus II.

a) In how many ways can the head of the mathematics department assign each of these professors one of these five courses and still keep peace in the department?

b) For the assignments in part (a), what is the probability that Violet will get to teach combinatorics?

10. A pair of dice, one red and the other green, is rolled six times. We know that the ordered pairs (1, 1), (1, 5), (2, 4), (3, 6), (4, 2), (4, 4), (5, 1), and (5, 5) did not come up. What is the probability that every value came up on both the red die and the green one?

11. A computer dating service wants to match each of four women with one of six men. According to the information these applicants provided when they joined the service, we can draw the following conclusions.

- Woman 1 would not be compatible with man 1, 3, or 6.
- Woman 2 would not be compatible with man 2 or 4.
- Woman 3 would not be compatible with man 3 or 6.
- Woman 4 would not be compatible with man 4 or 5.

In how many ways can the service successfully match each of the four women with a compatible partner?

12. For $A = \{1, 2, 3, 4, 5\}$ and $B = \{u, v, w, x, y, z\}$, determine the number of one-to-one functions $f: A \to B$ where $f(1) \neq v, w$; $f(2) \neq u, w$; $f(3) \neq x$; and $f(4) \neq v, x, y$.

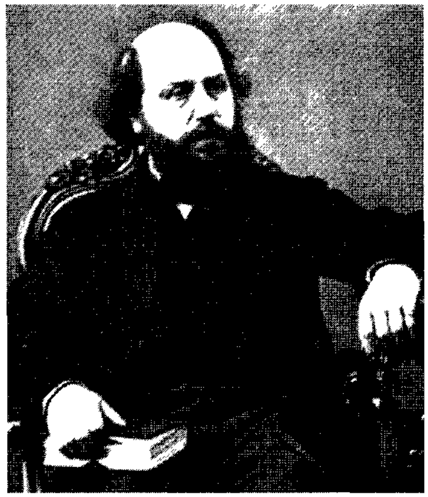
8.6 Summary and Historical Review

In the first and third chapters of this text we were concerned with enumeration problems in which we had to be careful of situations wherein arrangements or selections were overcounted. This situation became even more involved in Chapter 5 when we tried to count the number of onto functions for two finite sets.

With Venn diagrams to lead the way, in this chapter we obtained a pattern called the Principle of Inclusion and Exclusion. Using this principle, we restated each problem in terms of conditions and subsets. Using enumeration formulas on permutations and combinations that were developed earlier, we solved some simpler subproblems and let the principle manage our concern about overcounting. As a result, we were able to solve a variety of problems, some dealing with number theory and one with graph theory. We also proved the formula conjectured earlier in Section 5.3 for the number of onto functions for two finite sets.

This principle has an interesting history, being found in different manuscripts under such names as the "Sieve Method" or the "Principle of Cross Classification." A set-theoretic version of the principle, which concerned itself with set unions and intersections, is found in *Doctrine of Chances* (1718), a text on probability theory by Abraham DeMoivre (1667–1754). Somewhat earlier, in 1708, Pierre Rémond de Montmort (1678–1719) used the idea behind the principle in his solution of the problem generally known as *le problème des rencontres* (matches). (In this old French card game the 52 cards in a first deck are arranged face up in a row — perhaps on a table. Then the 52 cards of a second deck are dealt, with one new card being placed on each of the 52 cards previously arranged on the table top. The score for the game is determined by counting the resulting matches, where both the suit and the face value for each of the two cards must match.)

Credit for the way we developed and dealt with the Principle of Inclusion and Exclusion belongs to James Joseph Sylvester (1814–1897). (This colorful English-born mathematician also made major contributions in the theory of equations; the theory of matrices and determinants; and invariant theory, which he founded with Arthur Cayley (1821–1895). In addition Sylvester founded the *American Journal of Mathematics*, the first American journal established for mathematical research.) The importance of the inclusion-exclusion technique was not generally appreciated, however, until somewhat later, when the publication *Choice and Chance* by W. A. Whitworth [10] made mathematicians more aware of its potential and use.



James Joseph Sylvester (1814–1897)

For more on the application of this principle, examine Chapter 4 of C. L. Liu [4], Chapter 2 of H. J. Ryser [8], or Chapter 8 of A. Tucker [9]. More number-theoretic results related to the principle, including the Möbius inversion formula, can be found in Chapter 2 of M. Hall [1], Chapter X of C. L. Liu [5], and Chapter 16 of G. H. Hardy and E. M. Wright [3]. An extension of this formula is given in the article by G. C. Rota [7].

The article by D. Hanson, K. Seyffarth, and J. H. Weston [2] provides an interesting generalization of the derangement problem discussed in Section 8.3. The ideas behind the rook polynomials and their applications were developed in the late 1930s and during the 1940s and 1950s. Additional material on this topic is found in Chapters 7 and 8 of J. Riordan [6].

REFERENCES

1. Hall, Marshall, Jr. *Combinatorial Theory*. Waltham, Mass.: Blaisdell, 1967.
2. Hanson, Denis, Seyffarth, Karen, and Weston, J. Harley. "Matchings, Derangements, Rencontres." *Mathematics Magazine* 56, no. 4 (September 1983): pp. 224–229.
3. Hardy, Godfrey Harold, and Wright, Edward Maitland. *An Introduction to the Theory of Numbers*, 5th ed. Oxford: Oxford University Press, 1979.
4. Liu, C. L. *Introduction to Combinatorial Mathematics*. New York: McGraw-Hill, 1968.
5. Liu, C. L. *Topics in Combinatorial Mathematics*. Mathematical Association of America, 1972.
6. Riordan, John. *An Introduction to Combinatorial Analysis*. Princeton, N.J.: Princeton University Press, 1980. (Originally published in 1958 by John Wiley & Sons.)
7. Rota, Gian Carlo. "On the Foundations of Combinatorial Theory, I. Theory of Möbius Functions." *Zeitschrift für Wahrscheinlichkeits Theorie 2* (1964): pp. 340–368.
8. Ryser, Herbert J. *Combinatorial Mathematics*. Carus Mathematical Monograph, No. 14. Published by the Mathematical Association of America, distributed by John Wiley & Sons, New York, 1963.
9. Tucker, Alan. *Applied Combinatorics*, 4th ed. New York: Wiley, 2002.
10. Whitworth, William Allen. *Choice and Chance*. Originally published at Cambridge in 1867. Reprint of the 5th ed. (1901), Hafner, New York, 1965.

SUPPLEMENTARY EXERCISES

1. Determine how many $n \in \mathbf{Z}^+$ satisfy $n \leq 500$ and are not divisible by 2, 3, 5, 6, 8, or 10.

2. How many integers n are such that $0 \leq n < 1{,}000{,}000$ and the sum of the digits in n is less than or equal to 37?

3. At next week's church bazaar, Joseph and his cousin Jeffrey must arrange six baseballs, six footballs, six soccer balls, and six volleyballs on the four shelves in the sports booth sponsored by their Boy Scout troop. In how many ways can they do this so that there are at least two, but no more than seven, balls on each shelf? (Here all six balls for any one of the four sports are identical in appearance.)

4. Find the number of positive integers n where $1 \leq n \leq 1000$ and n is *not* a perfect square, cube, or fourth power.

5. In how many ways can we arrange the integers 1, 2, 3, . . . , 8 in a line so that there are no occurrences of the patterns 12, 23, . . . , 78, 81?

6. **a)** If we have k different colors available, in how many ways can we paint the walls of a pentagonal room if adjacent walls are to be painted with different colors?

b) What is the smallest value of k for which such a coloring is possible?

7. Ten students take a physics test in a certain room. When the test is over the students take a break and then return to the room to discuss their answers to the test questions. If there are 14 chairs in this room, in how many ways can the students seat themselves after the break so that no one is in the same chair he, or she, occupied during the test?

8. Using the result of Theorem 8.2, prove that the number of ways we can place s different objects in n distinct containers with m containers each containing exactly r of the objects is

$$\frac{(-1)^m n!\, s!}{m!} \sum_{i=m}^{n} \frac{(-1)^i (n-i)^{s-ir}}{(i-m)!(n-i)!(s-ir)!(r!)^i}.$$

9. If an arrangement of the letters in SURREPTITIOUS is selected at random, what is the probability that it contains (a) (exactly) three pairs of consecutive identical letters? (b) at most three pairs of consecutive identical letters?

10. In how many ways can four w's, four x's, four y's, and four z's be arranged so that there is no consecutive quadruple of the same letter?

11. **a)** Given n distinct objects, in how many ways can we select r of these objects so that each selection includes some particular m of the n objects? (Here $m \leq r \leq n$.)

b) Using the Principle of Inclusion and Exclusion, prove that for $m \leq r \leq n$,

$$\binom{n-m}{n-r} = \sum_{i=0}^{m} (-1)^i \binom{m}{i}\binom{n-i}{r}.$$

12. **a)** Let $\lambda \in \mathbf{Z}^+$. If we have λ different colors available, in how many ways can we color the vertices of the graph shown in Fig. 8.14(a) so that no adjacent vertices share the same color? This result in λ is called the *chromatic polynomial* of the graph, and the smallest value of λ for which the value of this polynomial is positive is called the *chromatic number* of the graph. What is the chromatic number of this graph? (We shall pursue this idea further in Chapter 11.)

b) If there are six colors available, in how many ways can the rooms R_i, $1 \leq i \leq 5$, shown in Fig. 8.14(b) be painted so that rooms with a common doorway, D_j, $1 \leq j \leq 5$, are painted with different colors?

13. Find the number of ways to arrange the letters in LAPTOP so that none of the letters L, A, T, O is in its original position and the letter P is not in the third or sixth position.

14. For $n \in \mathbf{Z}^+$ prove that if $\phi(n) = n - 1$ then n is prime.

15. Let D_{18} denote the set of positive divisors of 18. For $d \in D_{18}$ let $S_d = \{n \mid 0 < n \leq 18 \text{ and } \gcd(n, 18) = d\}$. (a) Show that the collection S_d, $d \in D_{18}$, provides a partition of $\{1, 2, 3, 4, \ldots, 17, 18\}$. (b) Note that $|S_1| = 6 = \phi(18)$ and $|S_2| = 6 = \phi(9)$. For each $d \in D_{18}$, express $|S_d|$ in terms of Euler's phi function.

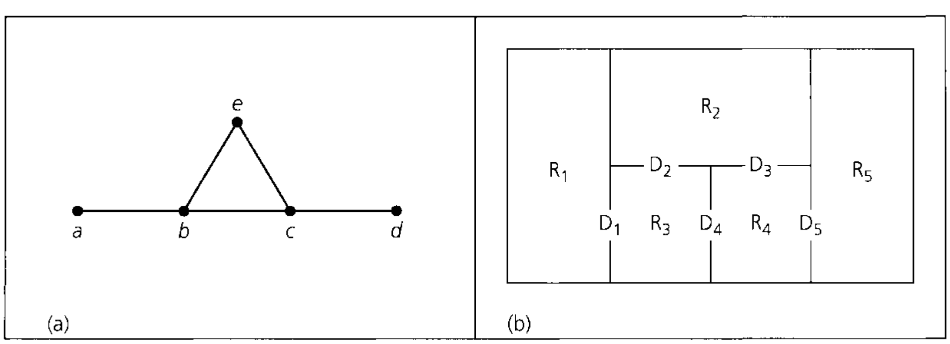


Figure 8.14

16. For $m \in \mathbf{Z}^+$ let $D_m = \{d \in \mathbf{Z}^+ | d \text{ divides } m\}$. For $d \in D_m$ let $S_d = \{n | 0 < n \leq m \text{ and } \gcd(n, m) = d\}$. (a) Show that the collection S_d, $d \in D_m$, provides a partition of $\{1, 2, 3, 4, \ldots, m-1, m\}$. (b) Determine $|S_d|$ for each $d \in D_m$.

17. If $n \in \mathbf{Z}^+$, prove that (a) $\phi(2n) = 2\phi(n)$ when n is even; and (b) $\phi(2n) = \phi(n)$ when n is odd.

18. Let $a, b, c \in \mathbf{Z}^+$ with $c = \gcd(a, b)$. Prove that

$$\phi(ab)\phi(c) = \phi(a)\phi(b)c.$$

19. Caitlyn has 48 different books: 12 each in mathematics, chemistry, physics, and computer science. These books are arranged on four shelves in her office with all books on any one subject on its own shelf. When her office is cleaned, the 48 books are taken down and then replaced on the shelves — once again with all 12 books on any one subject on its own shelf. In how many ways can this be done so that (a) no subject is on its original shelf? (b) one subject is on its original shelf? (c) no subject is on its original shelf and no book is in its original position? [For example, the book originally in the third (from the left) position on the first shelf must not be replaced on the first shelf and must not be in the third (from the left) position on the shelf where it is placed.]

10

Recurrence Relations

In earlier sections of the text we saw some recursive definitions and constructions. In Definitions 5.19, 6.7, 6.12, and 7.9, we obtained concepts at level $n + 1$ (or of size $n + 1$) from comparable concepts at level n (or of size n), after establishing the concept at a first value of n, such as 0 or 1. When we dealt with the Fibonacci and Lucas numbers in Section 4.2, the results at level $n + 1$ turned out to depend on those at levels n and $n - 1$; and for each of these sequences of integers the basis consisted of the first two integers (of the sequence). Now we shall find ourselves in a somewhat similar situation. We shall investigate functions $a(n)$, preferably written as a_n (for $n \geq 0$), where a_n depends on some of the prior terms $a_{n-1}, a_{n-2}, \ldots, a_1, a_0$. This study of what are called either *recurrence relations* or *difference equations* is the discrete counterpart to ideas applied in ordinary differential equations.

Our development will not employ any ideas from differential equations but will start with the notion of a geometric progression. As further ideas are developed, we shall see some of the many applications that make this topic so important.

10.1 The First-Order Linear Recurrence Relation

A *geometric progression* is an infinite sequence of numbers, such as 5, 15, 45, 135, ..., where the division of each term, other than the first, by its immediate predecessor is a constant, called the *common ratio*. For our sequence this common ratio is 3: $15 = 3(5)$, $45 = 3(15)$, and so on. If $a_0, a_1, a_2, \ldots$ is a geometric progression, then $a_1/a_0 = a_2/a_1 = \cdots = a_{n+1}/a_n = \cdots = r$, the common ratio. In our particular geometric progression we have $a_{n+1} = 3a_n$, $n \geq 0$.

The *recurrence relation* $a_{n+1} = 3a_n$, $n \geq 0$, does not define a unique geometric progression. The sequence 7, 21, 63, 189, ... also satisfies the relation. To pinpoint a particular sequence described by $a_{n+1} = 3a_n$, we need to know one of the terms of that sequence. Hence

$$a_{n+1} = 3a_n, \qquad n \geq 0, \qquad a_0 = 5,$$

uniquely defines the sequence 5, 15, 45, ..., whereas

$$a_{n+1} = 3a_n, \qquad n \geq 0, \qquad a_1 = 21,$$

identifies 7, 21, 63, ... as the geometric progression under study.

The equation $a_{n+1} = 3a_n$, $n \geq 0$ is a recurrence relation because the value of a_{n+1} (the present consideration) is dependent on a_n (a prior consideration). Since a_{n+1} depends only on its immediate predecessor, the relation is said to be of *first order*. In particular, this is a *first-order linear homogeneous* recurrence relation *with constant coefficients*. (We'll say more about these ideas later.) The general form of such an equation can be written $a_{n+1} = da_n$, $n \geq 0$, where d is a constant.

Values such as a_0 or a_1, given in addition to the recurrence relations, are called *boundary conditions*. The expression $a_0 = A$, where A is a constant, is also referred to as an *initial condition*. Our examples show the importance of the boundary condition in determining the unique solution.

Let us return now to the recurrence relation

$$a_{n+1} = 3a_n, \qquad n \geq 0, \qquad a_0 = 5.$$

The first four terms of this sequence are

$$\begin{aligned} a_0 &= 5, \\ a_1 &= 3a_0 = 3(5), \\ a_2 &= 3a_1 = 3(3a_0) = 3^2(5), \quad \text{and} \\ a_3 &= 3a_2 = 3(3^2(5)) = 3^3(5). \end{aligned}$$

These results suggest that for each $n \geq 0$, $a_n = 5(3^n)$. This is the *unique solution* of the given recurrence relation. In this solution, the value of a_n is a function of n and there is no longer any dependence on prior terms of the sequence, once we define a_0. To compute a_{10}, for example, we simply calculate $5(3^{10}) = 295{,}245$; there is no need to start at a_0 and build up to a_9 in order to obtain a_{10}.

From this example we are directed to the following. (This result can be established by the Principle of Mathematical Induction.)

The unique solution of the recurrence relation

$$a_{n+1} = da_n, \qquad \text{where } n \geq 0, \quad d \text{ is a constant, } \quad \text{and} \quad a_0 = A,$$

is given by

$$a_n = Ad^n, \qquad n \geq 0.$$

Thus the solution $a_n = Ad^n$, $n \geq 0$, defines a discrete function whose domain is the set **N** of all nonnegative integers.

EXAMPLE 10.1

Solve the recurrence relation $a_n = 7a_{n-1}$, where $n \geq 1$ and $a_2 = 98$.

This is just an alternative form of the relation $a_{n+1} = 7a_n$ for $n \geq 0$ and $a_2 = 98$. Hence the solution has the form $a_n = a_0(7^n)$. Since $a_2 = 98 = a_0(7^2)$, it follows that $a_0 = 2$, and $a_n = 2(7^n)$, $n \geq 0$, is the unique solution.

EXAMPLE 10.2

A bank pays 6% (annual) interest on savings, compounding the interest monthly. If Bonnie deposits \$1000 on the first day of May, how much will this deposit be worth a year later?

The annual interest rate is 6%, so the monthly rate is $6\%/12 = 0.5\% = 0.005$. For $0 \leq n \leq 12$, let p_n denote the value of Bonnie's deposit at the end of n months. Then $p_{n+1} = p_n + 0.005p_n$, where $0.005p_n$ is the interest earned on p_n during month $n+1$, for $0 \leq n \leq 11$, and $p_0 = \$1000$.

The relation $p_{n+1} = (1.005)p_n$, $p_0 = \$1000$, has the solution $p_n = p_0(1.005)^n = \$1000(1.005)^n$. Consequently, at the end of one year, Bonnie's deposit is worth $\$1000(1.005)^{12} = \$1061.68$.

In the next example we find a fifth way to count the number of compositions of a positive integer. The reader may recall that this situation was examined earlier in Examples 1.37, 3.11, 4.12, and 9.12.

EXAMPLE 10.3

Figure 10.1 provides the compositions of 3 and 4. Here we see that compositions (1′)–(4′) of 4 arise from the corresponding compositions of 3 by increasing the last summand (in each corresponding composition of 3) by 1. The other four compositions of 4, namely, (1″)–(4″), are obtained from the compositions of 3 by appending "+1" to each of the corresponding compositions of 3. (The reader may recall seeing such results in Fig. 4.7.)

		(1′)	4
		(2′)	$1+3$
(1)	3	(3′)	$2+2$
(2)	$1+2$	(4′)	$1+1+2$
(3)	$2+1$		
(4)	$1+1+1$	(1″)	$3+1$
		(2″)	$1+2+1$
		(3″)	$2+1+1$
		(4″)	$1+1+1+1$

Figure 10.1

What happens in Fig. 10.1 exemplifies the general situation. So if we let a_n count the number of compositions of n, for $n \in \mathbf{Z}^+$, we find that

$$a_{n+1} = 2a_n, \qquad n \geq 1, \qquad a_1 = 1.$$

However, in order to apply the formula for the unique solution (where $n \geq 0$) to this recurrence relation, we let $b_n = a_{n+1}$. Then we have

$$b_{n+1} = 2b_n, \qquad n \geq 0, \qquad b_0 = 1,$$

so $b_n = b_0(2^n) = 2^n$, and $a_n = b_{n-1} = 2^{n-1}$, $n \geq 1$.

The recurrence relation $a_{n+1} - da_n = 0$ is called *linear* because each subscripted term appears to the first power (as do the variables x and y in the equation of a line in the plane). In a linear relation there are no products such as $a_n a_{n-1}$, which appears in the nonlinear recurrence relation $a_{n+1} - 3a_n a_{n-1} = 0$. However, there are times when a nonlinear recurrence relation can be transformed into a linear one by a suitable algebraic substitution.

EXAMPLE 10.4

Find a_{12} if $a_{n+1}^2 = 5a_n^2$, where $a_n > 0$ for $n \geq 0$, and $a_0 = 2$.

Although this recurrence relation is not linear in a_n, if we let $b_n = a_n^2$, then the new relation $b_{n+1} = 5b_n$ for $n \geq 0$, and $b_0 = 4$, is a linear relation whose solution is $b_n = 4 \cdot 5^n$. Therefore, $a_n = 2(\sqrt{5})^n$ for $n \geq 0$, and $a_{12} = 2(\sqrt{5})^{12} = 31{,}250$.

The general first-order linear recurrence relation with constant coefficients has the form $a_{n+1} + ca_n = f(n)$, $n \geq 0$, where c is a constant and $f(n)$ is a function on the set **N** of nonnegative integers.

When $f(n) = 0$ for all $n \in \mathbf{N}$, the relation is called *homogeneous*; otherwise it is called *nonhomogeneous*. So far we have only dealt with homogeneous relations. Now we shall solve a nonhomogeneous relation. We shall develop specific techniques that work for all linear homogeneous recurrence relations with constant coefficients. However, many different techniques prove useful when we deal with a nonhomogeneous problem, although none allows us to solve everything that can arise.

EXAMPLE 10.5

Perhaps the most popular, though not the most efficient, method of sorting numeric data is a technique called the *bubble sort*. Here the input is a positive integer n and an array $x_1, x_2, x_3, \ldots, x_n$ of real numbers that are to be sorted into ascending order.

The pseudocode procedure in Fig. 10.2 provides an implementation for an algorithm to carry out this sorting process. Here the integer variable i is the counter for the outer **for** loop, whereas the integer variable j is the counter for the inner **for** loop. Finally, the real variable *temp* is used for storage that is needed when an exchange takes place.

```
procedure BubbleSort(n: positive integer; x1,x2,x3,...,xn: real numbers)
begin
  for i := 1 to n - 1 do
    for j := n downto i + 1 do
      if xj < xj-1 then
        begin      {interchange}
          temp := xj-1
          xj-1 := xj
          xj := temp
        end
end
```

Figure 10.2

We compare the last entry, x_n, in the given array with its immediate predecessor, x_{n-1}. If $x_n < x_{n-1}$, we interchange the values stored in x_{n-1} and x_n. In any event we will now have $x_{n-1} \leq x_n$. Then we compare x_{n-1} with its immediate predecessor, x_{n-2}. If $x_{n-1} < x_{n-2}$, we interchange them. We continue the process. After $n - 1$ such comparisons, the smallest number in the list is stored in x_1. We then repeat this process for the $n - 1$ numbers now stored in the (smaller) array $x_2, x_3, \ldots, x_n$. In this way, each time (counted by i) this process is carried out, the smallest number in the remaining sublist "bubbles up" to the front of that sublist.

A small example wherein $n = 5$ and $x_1 = 7$, $x_2 = 9$, $x_3 = 2$, $x_4 = 5$, and $x_5 = 8$ is given in Fig. 10.3 to show how the bubble sort of Fig. 10.2 places a given sequence in ascending order. In this figure each comparison that leads to an interchange is denoted by the symbol ⤸; the symbol } indicates a comparison that results in no interchange.

To determine the time-complexity function $h(n)$ when this algorithm is used on an input (array) of size $n \geq 1$, we count the total number of *comparisons* made in order to sort the n given numbers into ascending order.

If a_n denotes the number of comparisons needed to sort n numbers in this way, then we get the following recurrence relation:

$$a_n = a_{n-1} + (n - 1), \qquad n \geq 2, \qquad a_1 = 0.$$

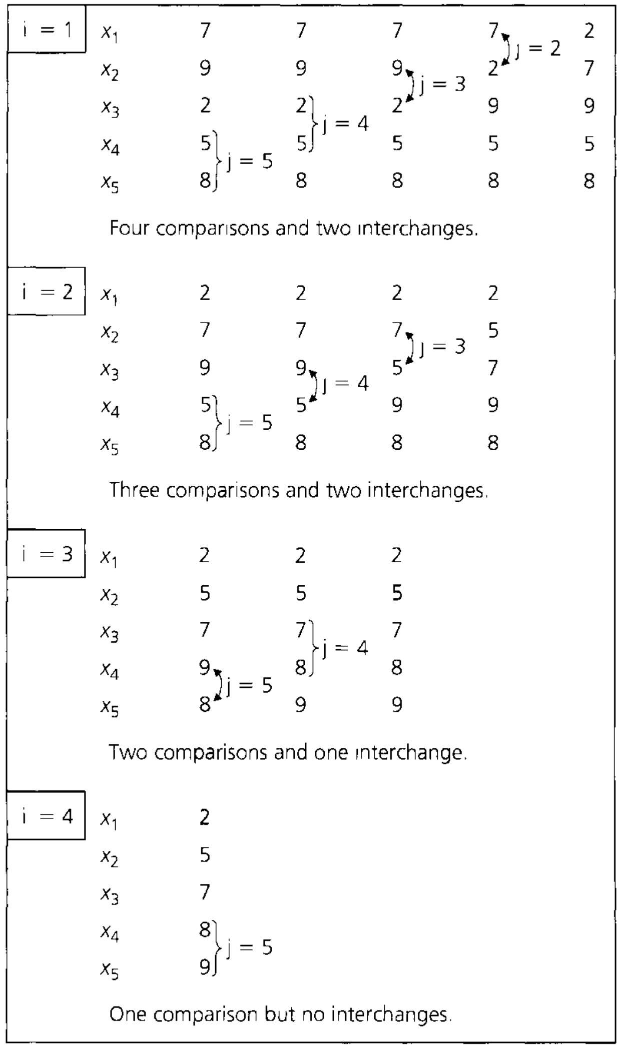


Figure 10.3

This arises as follows. Given a list of n numbers, we make $n - 1$ comparisons to bubble the smallest number up to the start of the list. The remaining sublist of $n - 1$ numbers then requires a_{n-1} comparisons in order to be completely sorted.

This relation is a linear first-order relation with constant coefficients, but the term $n - 1$ makes it nonhomogeneous. Since we have no technique for attacking such a relation, let us list some terms and see whether there is a recognizable pattern.

$$
\begin{aligned}
a_1 &= 0 \\
a_2 &= a_1 + (2 - 1) = 1 \\
a_3 &= a_2 + (3 - 1) = 1 + 2 \\
a_4 &= a_3 + (4 - 1) = 1 + 2 + 3 \\
\cdots & \quad \cdots \qquad \cdots \qquad \cdots
\end{aligned}
$$

In general, $a_n = 1 + 2 + \cdots + (n - 1) = [(n - 1)n]/2 = (n^2 - n)/2$.

As a result, the bubble sort determines the time-complexity function $h: \mathbf{Z}^+ \to \mathbf{R}$ given by $h(n) = a_n = (n^2 - n)/2$. [Here $h(\mathbf{Z}^+) \subset \mathbf{N}$.] Consequently, as a measure of the running time for the algorithm, we write $h \in O(n^2)$. Hence the bubble sort is said to require $O(n^2)$ comparisons.

EXAMPLE 10.6

In part (c) of Example 9.6 we sought the generating function for the sequence 0, 2, 6, 12, 20, 30, 42, ..., and the solution rested upon our ability to recognize that $a_n = n^2 + n$ for each $n \in \mathbf{N}$. If we fail to see this, perhaps we can examine the given sequence and determine whether there is some other pattern that will help us.

Here $a_0 = 0$, $a_1 = 2$, $a_2 = 6$, $a_3 = 12$, $a_4 = 20$, $a_5 = 30$, $a_6 = 42$, and

$$a_1 - a_0 = 2 \qquad a_3 - a_2 = 6 \qquad a_5 - a_4 = 10$$
$$a_2 - a_1 = 4 \qquad a_4 - a_3 = 8 \qquad a_6 - a_5 = 12.$$

These calculations suggest the recurrence relation

$$a_n - a_{n-1} = 2n, \qquad n \geq 1, \qquad a_0 = 0.$$

To solve this relation, we proceed in a slightly different manner from the method we used in Example 10.5. Consider the following n equations:

$$\begin{aligned} a_1 - a_0 &= 2 \\ a_2 - a_1 &= 4 \\ a_3 - a_2 &= 6 \\ \vdots \quad \vdots \quad &\vdots \\ a_n - a_{n-1} &= 2n. \end{aligned}$$

When we add these equations, the sum for the left-hand side will contain a_i and $-a_i$ for all $1 \leq i \leq n - 1$. So we obtain

$$\begin{aligned} a_n - a_0 &= 2 + 4 + 6 + \cdots + 2n = 2(1 + 2 + 3 + \cdots + n) \\ &= 2[n(n+1)/2] = n^2 + n. \end{aligned}$$

Since $a_0 = 0$, it follows that $a_n = n^2 + n$ for all $n \in \mathbf{N}$, as we found earlier in part (c) of Example 9.6.

At this point we shall examine a recurrence relation with a variable coefficient.

EXAMPLE 10.7

Solve the relation $a_n = n \cdot a_{n-1}$, where $n \geq 1$ and $a_0 = 1$.

Writing the first five terms defined by the relation, we have

$$a_0 = 1 \qquad a_2 = 2 \cdot a_1 = 2 \cdot 1 \qquad a_4 = 4 \cdot a_3 = 4 \cdot 3 \cdot 2 \cdot 1$$
$$a_1 = 1 \cdot a_0 = 1 \qquad a_3 = 3 \cdot a_2 = 3 \cdot 2 \cdot 1$$

Therefore, $a_n = n!$ and the solution is the discrete function a_n, which counts the number of permutations of n objects, $n \geq 0$.

While on the subject of permutations, we shall examine a recursive algorithm for generating the permutations of $\{1, 2, 3, \ldots, n-1, n\}$ from those for $\{1, 2, 3, \ldots, n-1\}$.[†] There is only one permutation of $\{1\}$. Examining the permutations of $\{1, 2\}$,

	1	2
2	1	

we see that after writing the permutation 1 twice, we intertwine the number 2 about 1 to get the permutations listed. Writing each of these two permutations three times, we intertwine the number 3 and obtain

	1		2	3
	1	3	2	
3	1		2	
3	2		1	
	2	3	1	
	2		1	3

We see here that the first permutation is 123 and that we obtain each of the next two permutations from its immediate predecessor by interchanging two numbers: 3 and the integer to its left. When 3 reaches the left side of the permutation, we examine the remaining numbers and permute them according to the list of permutations we generated for $\{1, 2\}$. (This makes the procedure recursive.) After that we interchange 3 with the integer on its right until 3 is on the right side of the permutation. We note that if we interchange 1 and 2 in the last permutation, we get 123, the first permutation listed.

Continuing for $S = \{1, 2, 3, 4\}$, we first list each of the six permutations of $\{1, 2, 3\}$ four times. Starting with the permutation 1234, we intertwine the 4 throughout the remaining 23 permutations as indicated in Table 10.1 (on page 454). The only new idea here develops as follows. When progressing from permutation (5) to (6) to (7) to (8), we interchange 4 with the integer to its right. At permutation (8), where 4 has reached the right side, we obtain permutation (9) by keeping the location of 4 fixed and replacing the permutation 132 by 312 from the list of permutations of $\{1, 2, 3\}$. After that we continue as for the first eight permutations until we reach permutation (16), where 4 is again on the right. We then permute 321 to obtain 231 and continue intertwining 4 until all 24 permutations have been generated. Once again, if 1 and 2 are interchanged in the last permutation, we obtain the first permutation in our list.

The chapter references provide more information on recursive procedures for generating permutations and combinations.

We shall close this first section by returning to an earlier idea — the greatest common divisor of two positive integers.

EXAMPLE 10.8

Recursive methods are fundamental in the areas of discrete mathematics and the analysis of algorithms. Such methods arise when we want to solve a given problem by breaking it down, or referring it, to smaller similar problems. In many programming languages this can be implemented by the use of recursive functions and procedures, which are permitted to invoke themselves. This example will provide one such procedure.

[†]The material from here to the end of this section is a digression that uses the idea of recursion. It does not deal with methods for solving recurrence relations and may be omitted with no loss of continuity.

Table 10.1

(1)		1		2		3	4
(2)		1		2	4	3	
(3)		1	4	2		3	
(4)	4	1		2		3	
(5)	4	1		3		2	
(6)		1	4	3		2	
(7)		1		3	4	2	
(8)		1		3		2	4
(9)		3		1		2	4
(10)		3		1	4	2	
(11)		3	4	1		2	
...		...		...		...	
(15)		3		2	4	1	
(16)		3		2		1	4
(17)		2		3		1	4
...		...		...		...	
(22)		2	4	1		3	
(23)		2		1	4	3	
(24)		2		1		3	4

In computing gcd(333, 84) we obtain the following calculations when we use the Euclidean algorithm (presented in Section 4.4).

$$333 = 3(84) + 81 \qquad 0 < 81 < 84 \tag{1}$$

$$84 = 1(81) + 3 \qquad 0 < 3 < 81 \tag{2}$$

$$81 = 27(3) + 0. \tag{3}$$

Since 3 is the last nonzero remainder, the Euclidean algorithm tells us that gcd(333, 84) = 3. However, if we use only the calculations in Eqs. (2) and (3), then we find that gcd(84, 81) = 3. And Eq. (3) alone implies that gcd(81, 3) = 3 because 3 divides 81. Consequently,

$$\gcd(333, 84) = \gcd(84, 81) = \gcd(81, 3) = 3,$$

where the integers involved in the successive calculations get smaller as we go from Eq. (1) to Eq. (2) to Eq. (3).

We also observe that

$$81 = 333 \textbf{ mod } 84 \qquad \text{and} \qquad 3 = 84 \textbf{ mod } 81.$$

Therefore it follows that

$$\gcd(333, 84) = \gcd(84, 333 \textbf{ mod } 84) = \gcd(333 \textbf{ mod } 84, 84 \textbf{ mod } (333 \textbf{ mod } 84)).$$

These results suggest the following recursive method for computing $\gcd(a, b)$, where $a, b \in \mathbf{Z}^+$.

Say we have the input $a, b \in \mathbf{Z}^+$.

Step 1: If $b|a$ (or a **mod** $b = 0$), then $\gcd(a, b) = b$.

Step 2: If $b \nmid a$, then perform the following tasks in the order specified.

i) Set $a = b$.

ii) Set $b = a \textbf{ mod } b$, where the value of a for this assignment is the *old* value of a.

iii) Return to step (1).

These ideas are used in the pseudocode procedure in Fig. 10.4. (The reader may wish to compare this procedure with the one given in Fig. 4.11.)

```
procedure gcd2(a, b: positive integers)
begin
  if a mod b = 0 then
    gcd = b
  else gcd = gcd2(b, a mod b)
end
```

Figure 10.4

EXERCISES 10.1

1. Find a recurrence relation, with initial condition, that uniquely determines each of the following geometric progressions.

a) 2, 10, 50, 250, ...

b) 6, −18, 54, −162, ...

c) 7, 14/5, 28/25, 56/125, ...

2. Find the unique solution for each of the following recurrence relations.

a) $a_{n+1} - 1.5a_n = 0,\ \ n \geq 0$

b) $4a_n - 5a_{n-1} = 0,\ \ n \geq 1$

c) $3a_{n+1} - 4a_n = 0,\ \ n \geq 0,\ \ a_1 = 5$

d) $2a_n - 3a_{n-1} = 0,\ \ n \geq 1,\ \ a_4 = 81$

3. If a_n, $n \geq 0$, is the unique solution of the recurrence relation $a_{n+1} - da_n = 0$, and $a_3 = 153/49$, $a_5 = 1377/2401$, what is d?

4. The number of bacteria in a culture is 1000 (approximately), and this number increases 250% every two hours. Use a recurrence relation to determine the number of bacteria present after one day.

5. If Laura invests $100 at 6% interest compounded quarterly, how many months must she wait for her money to double? (She cannot withdraw the money before the quarter is up.)

6. Paul invested the stock profits he received 15 years ago in an account that paid 8% interest compounded quarterly. If his account now has $7218.27 in it, what was his initial investment?

7. Let $x_1, x_2, \ldots, x_{20}$ be a list of distinct real numbers to be sorted by the bubble-sort technique of Example 10.5. (a) After how many comparisons will the 10 smallest numbers of the original list be arranged in ascending order? (b) How many more comparisons are needed to finish this sorting job?

8. For the implementation of the bubble sort given in Fig. 10.2, the outer **for** loop is executed $n - 1$ times. This occurs regardless of whether any interchanges take place during the execution of the inner **for** loop. Consequently, for $i = k$, where $1 \leq k \leq n - 2$, if the execution of the inner **for** loop results in no interchanges, then the list is in ascending order. So the execution of the outer **for** loop for $k + 1 \leq i \leq n - 1$ is not needed.

a) For the situation described here, how many unnecessary comparisons are made if the execution of the inner **for** loop for $i = k$ $(1 \leq k \leq n - 2)$ results in no interchanges?

b) Write an improved version of the bubble sort shown in Fig. 10.2. (Your result should eliminate the unnecessary comparisons discussed at the start of this exercise.)

c) Using the number of comparisons as a measure of its running time, determine the best-case and the worst-case time complexities for the algorithm implemented in part (b).

9. Say the permutations of {1, 2, 3, 4, 5} are generated by the procedure developed after Example 10.7. (a) What is the last permutation in the list? (b) What two permutations precede 25134? (c) What three permutations follow 25134?

10. For $n > 1$, a permutation $p_1, p_2, p_3, \ldots, p_n$ of the integers $1, 2, 3, \ldots, n$ is called *orderly* if, for each $i = 1, 2, 3, \ldots, n - 1$, there exists a $j > i$ such that $|p_j - p_i| = 1$. [If $n = 2$, the permutations 1, 2 and 2, 1 are both orderly. When $n = 3$ we find that 3, 1, 2 is an orderly permutation, while 2, 3, 1 is not. (Why not?)] (a) List all the orderly permutations for 1, 2, 3. (b) List all the orderly permutations for 1, 2, 3, 4. (c) If p_1, p_2, p_3, p_4, p_5 is an orderly permutation of 1, 2, 3, 4, 5, what value(s) can p_1 be? (d) For $n > 1$, let a_n count the number of orderly permutations for $1, 2, 3, \ldots, n$. Find and solve a recurrence relation for a_n.

10.2 The Second-Order Linear Homogeneous Recurrence Relation with Constant Coefficients

Let $k \in \mathbf{Z}^+$ and $C_0\ (\neq 0), C_1, C_2, \ldots, C_k\ (\neq 0)$ be real numbers. If a_n, for $n \geq 0$, is a discrete function, then

$$C_0 a_n + C_1 a_{n-1} + C_2 a_{n-2} + \cdots + C_k a_{n-k} = f(n), \qquad n \geq k,$$

is a linear recurrence relation (with constant coefficients) of *order* k. When $f(n) = 0$ for all $n \geq 0$, the relation is called *homogeneous*; otherwise, it is called *nonhomogeneous*.

In this section we shall concentrate on the homogeneous relation of order two:

$$C_0 a_n + C_1 a_{n-1} + C_2 a_{n-2} = 0, \qquad n \geq 2.$$

On the basis of our work in Section 10.1, we seek a solution of the form $a_n = cr^n$, where $c \neq 0$ and $r \neq 0$.

Substituting $a_n = cr^n$ into $C_0 a_n + C_1 a_{n-1} + C_2 a_{n-2} = 0$, we obtain

$$C_0 cr^n + C_1 cr^{n-1} + C_2 cr^{n-2} = 0.$$

With $c, r \neq 0$, this becomes $C_0 r^2 + C_1 r + C_2 = 0$, a quadratic equation which is called the *characteristic equation*. The roots r_1, r_2 of this equation determine the following three cases: (a) r_1, r_2 are distinct real numbers; (b) r_1, r_2 form a complex conjugate pair; or (c) r_1, r_2 are real, but $r_1 = r_2$. In all cases, r_1 and r_2 are called the *characteristic roots*.

Case (A): (Distinct Real Roots)

EXAMPLE 10.9

Solve the recurrence relation $a_n + a_{n-1} - 6a_{n-2} = 0$, where $n \geq 2$ and $a_0 = -1$, $a_1 = 8$.

If $a_n = cr^n$ with $c, r \neq 0$, we obtain $cr^n + cr^{n-1} - 6cr^{n-2} = 0$ from which the characteristic equation $r^2 + r - 6 = 0$ follows:

$$0 = r^2 + r - 6 = (r + 3)(r - 2) \Rightarrow r = 2, -3.$$

Since we have two distinct real roots, $a_n = 2^n$ and $a_n = (-3)^n$ are both solutions [as are $b(2^n)$ and $d(-3)^n$, for arbitrary constants b, d]. They are *linearly independent solutions* because one is not a multiple of the other; that is, there is no real constant k such that $(-3)^n = k(2^n)$ for *all* $n \in \mathbf{N}$.† We write $a_n = c_1(2^n) + c_2(-3)^n$ for the *general* solution, where c_1, c_2 are arbitrary constants.

With $a_0 = -1$ and $a_1 = 8$, c_1 and c_2 are determined as follows:

$$-1 = a_0 = c_1(2^0) + c_2(-3)^0 = c_1 + c_2$$
$$8 = a_1 = c_1(2^1) + c_2(-3)^1 = 2c_1 - 3c_2.$$

Solving this system of equations, one finds $c_1 = 1, c_2 = -2$. Therefore, $a_n = 2^n - 2(-3)^n$, $n \geq 0$, is the *unique* solution of the given recurrence relation.

The reader should realize that to determine the unique solution of a second-order linear homogeneous recurrence relation with constant coefficients one needs two initial conditions

†We can also call the solutions $a_n = 2^n$ and $a_n = (-3)^n$ *linearly independent* when the following condition is satisfied: For $k_1, k_2 \in \mathbf{R}$, if $k_1(2^n) + k_2(-3)^n = 0$ for *all* $n \in \mathbf{N}$, then $k_1 = k_2 = 0$.

(values) — that is, the value of a_n for two values of n, very often $n = 0$ and $n = 1$, or $n = 1$ and $n = 2$.

An interesting second-order homogeneous recurrence relation is the *Fibonacci relation.* (This was mentioned earlier in Sections 4.2 and 9.6.)

EXAMPLE 10.10

Solve the recurrence relation $F_{n+2} = F_{n+1} + F_n$, where $n \geq 0$ and $F_0 = 0$, $F_1 = 1$.

As in the previous example, let $F_n = cr^n$, for $c, r \neq 0$, $n \geq 0$. Upon substitution we get $cr^{n+2} = cr^{n+1} + cr^n$. This gives the characteristic equation $r^2 - r - 1 = 0$. The characteristic roots are $r = (1 \pm \sqrt{5})/2$, so the general solution is

$$F_n = c_1\left(\frac{1+\sqrt{5}}{2}\right)^n + c_2\left(\frac{1-\sqrt{5}}{2}\right)^n.$$

To solve for c_1, c_2, we use the given initial values and write $0 = F_0 = c_1 + c_2$, $1 = F_1 = c_1[(1+\sqrt{5})/2] + c_2[(1-\sqrt{5})/2]$. Since $-c_1 = c_2$, we have $2 = c_1(1+\sqrt{5}) - c_1(1-\sqrt{5})$ and $c_1 = 1/\sqrt{5}$. The general solution is given by

$$F_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right], \qquad n \geq 0.$$

When dealing with the Fibonacci numbers one often finds the assignments $\alpha = (1+\sqrt{5})/2$ and $\beta = (1-\sqrt{5})/2$, where α is known as the *golden ratio.* As a result, we find that

$$F_n = \frac{1}{\sqrt{5}}(\alpha^n - \beta^n) = \frac{\alpha^n - \beta^n}{\alpha - \beta}, \qquad n \geq 0.$$

[This representation is referred to as the *Binet form* for F_n, as it was first published in 1843 by Jacques Philippe Marie Binet (1786–1856).]

EXAMPLE 10.11

For $n \geq 0$, let $S = \{1, 2, 3, \ldots, n\}$ (when $n = 0$, $S = \emptyset$), and let a_n denote the number of subsets of S that contain no consecutive integers. Find and solve a recurrence relation for a_n.

For $0 \leq n \leq 4$, we have $a_0 = 1$, $a_1 = 2$, $a_2 = 3$, $a_3 = 5$, and $a_4 = 8$. [For example, $a_3 = 5$ because $S = \{1, 2, 3\}$ has $\emptyset$, $\{1\}$, $\{2\}$, $\{3\}$, and $\{1, 3\}$ as subsets with no consecutive integers (and no other such subsets).] These first five terms are reminiscent of the Fibonacci sequence. But do things change as we continue?

Let $n \geq 2$ and $S = \{1, 2, 3, \ldots, n-2, n-1, n\}$. If $A \subseteq S$ and A is to be counted in a_n, there are two possibilities:

a) $n \in A$: When this happens $(n-1) \notin A$, and $A - \{n\}$ would be counted in a_{n-2}.

b) $n \notin A$: For this case A would be counted in a_{n-1}.

These two cases are exhaustive and mutually disjoint, so we conclude that $a_n = a_{n-1} + a_{n-2}$, where $n \geq 2$ and $a_0 = 1$, $a_1 = 2$, is the recurrence relation for the problem. Now we could solve for a_n, but if we notice that $a_n = F_{n+2}$, $n \geq 0$, then the result of Example 10.10 implies that

$$a_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^{n+2} - \left(\frac{1-\sqrt{5}}{2}\right)^{n+2}\right], \qquad n \geq 0.$$

EXAMPLE 10.12

Suppose we have a $2 \times n$ chessboard, for $n \in \mathbf{Z}^+$. The case for $n = 4$ is shown in part (a) of Fig. 10.5. We wish to cover such a chessboard using 2×1 (vertical) dominoes, which can also be used as 1×2 (horizontal) dominoes. Such dominoes (or tiles) are shown in part (b) of Fig. 10.5.

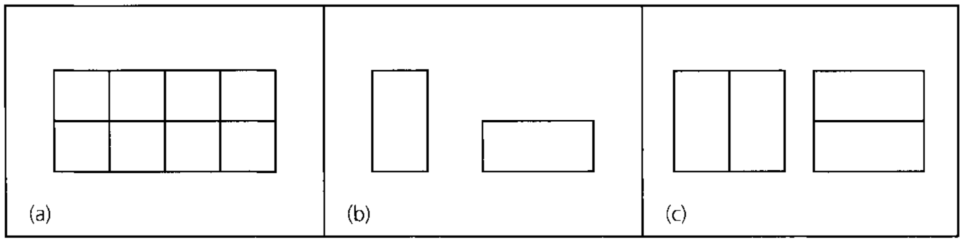


Figure 10.5

For $n \in \mathbf{Z}^+$ we let b_n count the number of ways we can cover (or tile) a $2 \times n$ chessboard using our 2×1 and 1×2 dominoes. Here $b_1 = 1$, for a 2×1 chessboard necessitates one 2×1 (vertical) domino. A 2×2 chessboard can be covered in two ways — using two 2×1 (vertical) dominoes or two 1×2 (horizontal) dominoes, as shown in part (c) of the figure. Hence $b_2 = 2$. For $n \geq 3$, consider the last (nth) column of a $2 \times n$ chessboard. This column can be covered in two ways.

i) By one 2×1 (vertical) domino: Here the remaining $2 \times (n - 1)$ subboard can be covered in b_{n-1} ways.

ii) By the right squares of two 1×2 (horizontal) dominoes placed one above the other: Now the remaining $2 \times (n - 2)$ subboard can be covered in b_{n-2} ways.

Since these two ways have nothing in common and deal with all possibilities, we may write

$$b_n = b_{n-1} + b_{n-2}, \qquad n \geq 3, \qquad b_1 = 1, \qquad b_2 = 2.$$

We find that $b_n = F_{n+1}$, so here is another situation where the Fibonacci numbers arise. The result from Example 10.10 gives us $b_n = (1/\sqrt{5})[((1 + \sqrt{5})/2)^{n+1} - ((1 - \sqrt{5})/2)^{n+1}]$, $n \geq 1$.

EXAMPLE 10.13

At this point we examine an interesting application where the number $\alpha = (1 + \sqrt{5})/2$ plays a major role. This application deals with Gabriel Lamé's work in estimating the number of divisions used in the Euclidean algorithm to find $\gcd(a, b)$, where $a, b \in \mathbf{Z}^+$ with $a \geq b \geq 2$. To find this estimate we need the following property of the Fibonacci numbers, which can be established by the alternative form of the Principle of Mathematical Induction. (A proof is requested in the Section Exercises.)

Property: For $n \geq 3$, $F_n > \alpha^{n-2}$.

Addressing the problem at hand — namely, estimating the number of divisions when the Euclidean algorithm is used to find $\gcd(a, b)$ — we recall the following steps from Theorem 4.7.

Letting $r_0 = a$ and $r_1 = b$, we have

$$\begin{aligned} r_0 &= q_1 r_1 + r_2, & 0 < r_2 < r_1 \\ r_1 &= q_2 r_2 + r_3, & 0 < r_3 < r_2 \\ r_2 &= q_3 r_3 + r_4, & 0 < r_4 < r_3 \\ &\cdots\cdots\cdots\cdots & \\ r_{n-2} &= q_{n-1} r_{n-1} + r_n, & 0 < r_n < r_{n-1} \\ r_{n-1} &= q_n r_n. & \end{aligned}$$

So r_n, the last nonzero remainder, is $\gcd(a, b)$.

From the subscripts on r we see that n divisions have been performed in determining $r_n = \gcd(a, b)$. In addition, $q_i \geq 1$, for all $1 \leq i \leq n - 1$, and $q_n \geq 2$ because $r_n < r_{n-1}$. Examining the n nonzero remainders $r_n, r_{n-1}, r_{n-2}, \ldots, r_2$, and r_1 $(= b)$, we learn that

$$\begin{aligned} & r_n > 0, \text{ so } r_n \geq 1 = F_2. \\ & [(q_n \geq 2) \wedge (r_n \geq 1)] \Rightarrow r_{n-1} = q_n r_n \geq 2 \cdot 1 = 2 = F_3 \\ & r_{n-2} = q_{n-1} r_{n-1} + r_n \geq 1 \cdot r_{n-1} + r_n \geq F_3 + F_2 = F_4 \\ & \cdots\cdots\cdots\cdots \\ & r_2 = q_3 r_3 + r_4 \geq 1 \cdot r_3 + r_4 \geq F_{n-1} + F_{n-2} = F_n \\ & b = r_1 = q_2 r_2 + r_3 \geq 1 \cdot r_2 + r_3 \geq F_n + F_{n-1} = F_{n+1}. \end{aligned}$$

Therefore, if n divisions are performed by the Euclidean algorithm to determine $\gcd(a, b)$, with $a \geq b \geq 2$, then $b \geq F_{n+1}$. So by virtue of the property introduced earlier, we may write $b > \alpha^{(n+1)-2} = \alpha^{n-1} = [(1 + \sqrt{5})/2]^{n-1}$. Consequently, we find now that

$$b > \alpha^{n-1} \Rightarrow \log_{10} b > \log_{10}(\alpha^{n-1}) = (n - 1) \log_{10} \alpha > \frac{n-1}{5},$$

since $\log_{10} \alpha = \log_{10}[(1 + \sqrt{5})/2] \doteq 0.208988 > 0.2 = \frac{1}{5}$.

At this point suppose that $10^{k-1} \leq b < 10^k$, so that the decimal (base 10) representation of b has k digits. Then

$$k = \log_{10} 10^k > \log_{10} b > \frac{n-1}{5}, \quad \text{and} \quad n < 5k + 1.$$

With $n, k \in \mathbf{Z}^+$ we have $n < 5k + 1 \Rightarrow n \leq 5k$, and this last inequality now completes a proof for the following.

Lamé's Theorem: Let $a, b \in \mathbf{Z}^+$ with $a \geq b \geq 2$. Then the number of divisions needed, in the Euclidean algorithm, to determine $\gcd(a, b)$ is at most 5 times the number of decimal digits in b.

Before closing this example, we learn one more fact from Lamé's Theorem. Since $b \geq 2$, it follows that $\log_{10} b \geq \log_{10} 2$, so $5 \log_{10} b \geq 5 \log_{10} 2 = \log_{10} 2^5 = \log_{10} 32 > 1$. From above we know that $n - 1 < 5 \log_{10} b$, so

$$n < 1 + 5 \log_{10} b < 5 \log_{10} b + 5 \log_{10} b = 10 \log_{10} b$$

and $n \in O(\log_{10} b)$. [Hence, the number of divisions needed, in the Euclidean algorithm, to determine $\gcd(a, b)$, for $a, b \in \mathbf{Z}^+$ with $a \geq b \geq 2$, is $O(\log_{10} b)$ — that is, on the order of the number of decimal digits in b.]

Returning to the theme of the section we now examine a recurrence relation in a computer science application.

EXAMPLE 10.14

In many programming languages one may consider those legal arithmetic expressions, *without parentheses*, that are made up of the digits 0, 1, 2, . . . , 9 and the binary operation symbols $+$, $*$, $/$. For example, $3 + 4$ and $2 + 3 * 5$ are legal arithmetic expressions; $8 + * 9$ is not. Here $2 + 3 * 5 = 17$, since there is a hierarchy of operations: Multiplication and division are performed before addition. Operations at the same level are performed in their order of appearance as the expression is scanned from left to right.

For $n \in \mathbf{Z}^+$, let a_n be the number of these (legal) arithmetic expressions that are made up of n symbols. Then $a_1 = 10$, since the arithmetic expressions of one symbol are the 10 digits. Next $a_2 = 100$. This accounts for the expressions 00, 01, . . . , 09, 10, 11, . . . , 99. (There are no unnecessary leading plus signs.) When $n \geq 3$, we consider two cases in order to derive a recurrence relation for a_n:

1) If x is an arithmetic expression of $n - 1$ symbols, the last symbol must be a digit. Adding one more digit to the right of x, we get $10a_{n-1}$ arithmetic expressions of n symbols where the last two symbols are digits.

2) Now let y be an arithmetic expression of $n - 2$ symbols. To obtain an arithmetic expression with n symbols (that is not counted in case 1), we adjoin to the right of y one of the 29 two-symbol expressions $+1, \ldots, +9, +0, *1, \ldots, *9, *0, /1, \ldots, /9$.

From these two cases we have $a_n = 10a_{n-1} + 29a_{n-2}$, where $n \geq 3$ and $a_1 = 10$, $a_2 = 100$. Here the characteristic roots are $5 \pm 3\sqrt{6}$ and the solution is $a_n = (5/(3\sqrt{6})) \cdot [(5 + 3\sqrt{6})^n - (5 - 3\sqrt{6})^n]$ for $n \geq 1$. (Verify this result.)

Another way to complete the solution of this problem is to use the recurrence relation $a_n = 10a_{n-1} + 29a_{n-2}$, with $a_2 = 100$ and $a_1 = 10$, to calculate a value for a_0 — namely, $a_0 = (a_2 - 10a_1)/29 = 0$. The solution for the recurrence relation

$$a_n = 10a_{n-1} + 29a_{n-2}, \qquad n \geq 2, \qquad a_0 = 0, \qquad a_1 = 10$$

is

$$a_n = (5/(3\sqrt{6}))[(5 + 3\sqrt{6})^n - (5 - 3\sqrt{6})^n], \qquad n \geq 0.$$

A second method for counting palindromes arises in our next example.

EXAMPLE 10.15

In Fig. 10.6 we find the palindromes of 3, 4, 5, and 6 — that is, the compositions of 3, 4, 5, and 6 that read the same left to right as right to left. (We saw this concept earlier in Example 9.13.) Consider first the palindromes of 3 and 5. To build the palindromes of 5 from those of 3 we do the following:

i) Add 1 to the first and last summands in a palindrome of 3. This is how we get palindromes (1′) and (2′) for 5 from the respective palindromes (1) and (2) for 3. [*Note:* When we have a one summand palindrome n we get the one summand palindrome $n + 2$. That is how we build palindrome (1′) for 5 from palindrome (1) for 3.]

ii) Append "1+" to the start and "+1" to the end of each palindrome of 3. This technique generates the palindromes (1″) and (2″) for 5 from the respective palindromes (1) and (2) for 3.

(1)	3	(1′)	5	(1)	4	(1′)	6
(2)	1 + 1 + 1	(2′)	2 + 1 + 2	(2)	1 + 2 + 1	(2′)	2 + 2 + 2
		(1″)	1 + 3 + 1	(3)	2 + 2	(3′)	3 + 3
		(2″)	1 + 1 + 1 + 1 + 1	(4)	1 + 1 + 1 + 1	(4′)	2 + 1 + 1 + 2
						(1″)	1 + 4 + 1
						(2″)	1 + 1 + 2 + 1 + 1
						(3″)	1 + 2 + 2 + 1
						(4″)	1 + 1 + 1 + 1 + 1 + 1

Figure 10.6

The situation is similar for building the palindromes of 6 from those of 4.

The preceding observations lead us to the following. For $n \in \mathbf{Z}^+$, let p_n count the number of palindromes of n. Then

$$p_n = 2p_{n-2}, \qquad n \geq 3, \qquad p_1 = 1, \qquad p_2 = 2.$$

Substituting $p_n = cr^n$, for $c, r \neq 0$, $n \geq 1$, into this recurrence relation, the resulting characteristic equation is $r^2 - 2 = 0$. The characteristic roots are $r = \pm\sqrt{2}$, so $p_n = c_1(\sqrt{2})^n + c_2(-\sqrt{2})^n$. From

$$1 = p_1 = c_1(\sqrt{2}) + c_2(-\sqrt{2})$$
$$2 = p_2 = c_1(\sqrt{2})^2 + c_2(-\sqrt{2})^2$$

we find that $c_1 = \left(\frac{1}{2} + \frac{1}{2\sqrt{2}}\right)$, $c_2 = \left(\frac{1}{2} - \frac{1}{2\sqrt{2}}\right)$, so

$$p_n = \left(\frac{1}{2} + \frac{1}{2\sqrt{2}}\right)(\sqrt{2})^n + \left(\frac{1}{2} - \frac{1}{2\sqrt{2}}\right)(-\sqrt{2})^n, \qquad n \geq 1.$$

Unfortunately, this does not look like the result found in Example 9.13. After all, that answer contained no radical terms. However, suppose we consider n even, say $n = 2k$. Then

$$\begin{aligned} p_n &= \left(\frac{1}{2} + \frac{1}{2\sqrt{2}}\right)(\sqrt{2})^{2k} + \left(\frac{1}{2} - \frac{1}{2\sqrt{2}}\right)(-\sqrt{2})^{2k} \\ &= \left(\frac{1}{2} + \frac{1}{2\sqrt{2}}\right)2^k + \left(\frac{1}{2} - \frac{1}{2\sqrt{2}}\right)2^k = 2^k = 2^{n/2} \end{aligned}$$

For n odd, say $n = 2k - 1$, $k \in \mathbf{Z}^+$, we leave it for the reader to show that $p_n = 2^{k-1} = 2^{(n-1)/2}$.

The preceding results can be expressed by $p_n = 2^{\lfloor n/2 \rfloor}$, $n \geq 1$, as we found in Example 9.13.

The recurrence relation for the next example will be set up in two ways. In the first part we shall see how auxiliary variables may be helpful.

EXAMPLE 10.16

Find a recurrence relation for the number of binary sequences of length n that have no consecutive 0's.

a) For $n \geq 1$, let a_n be the number of such sequences of length n. Let $a_n^{(0)}$ count those that end in 0, and $a_n^{(1)}$ those that end in 1. Then $a_n = a_n^{(0)} + a_n^{(1)}$.

We derive a recurrence relation for a_n, $n \geq 1$, by computing $a_1 = 2$ and then considering each sequence x of length $n-1$ (> 0) where x contains no consecutive 0's. If x ends in 1, then we can append a 0 or a 1 to it, giving us $2a_{n-1}^{(1)}$ of the sequences counted by a_n. If the sequence x ends in 0, then only 1 can be appended, resulting in $a_{n-1}^{(0)}$ sequences counted by a_n. Since these two cases exhaust all possibilities and have nothing in common, we have

$$a_n = 2 \cdot a_{n-1}^{(1)} + 1 \cdot a_{n-1}^{(0)}$$

$\downarrow$ **The nth position can be 0 or 1.** $\searrow$ **The nth position can only be 1.**

If we consider any sequence y counted in a_{n-2} we find that the sequence $y1$ is counted in $a_{n-1}^{(1)}$. Likewise, if the sequence $z1$ is counted in $a_{n-1}^{(1)}$, then z is counted in a_{n-2}. Consequently, $a_{n-2} = a_{n-1}^{(1)}$ and

$$a_n = a_{n-1}^{(1)} + \left[a_{n-1}^{(1)} + a_{n-1}^{(0)}\right] = a_{n-1}^{(1)} + a_{n-1} = a_{n-1} + a_{n-2}.$$

Therefore the recurrence relation for this problem is $a_n = a_{n-1} + a_{n-2}$, where $n \geq 3$ and $a_1 = 2$, $a_2 = 3$. (We leave the details of the solution for the reader.)

b) Alternatively, if $n \geq 1$ and a_n counts the number of binary sequences with no consecutive 0's, then $a_1 = 2$ and $a_2 = 3$, and for $n \geq 3$ we consider the binary sequences counted by a_n. There are two possibilities for these sequences:

(Case 1: The nth symbol is 1) Here we find that the preceding $n-1$ symbols form a binary sequence with no consecutive 0's. There are a_{n-1} such sequences.

(Case 2: The nth symbol is 0) Here each such sequence actually ends in 10 and the first $n-2$ symbols provide a binary sequence with no consecutive 0's. In this case there are a_{n-2} such sequences.

Since these two cases cover all the possibilities and have no such sequence in common, we may write

$$a_n = a_{n-1} + a_{n-2}, \qquad n \geq 3, \qquad a_1 = 2, \qquad a_2 = 3,$$

as we found in part (a).

In both part (a) and part (b) we can use the recurrence relation and $a_1 = 2$, $a_2 = 3$ to go back and determine a value for a_0 — namely, $a_0 = a_2 - a_1 = 3 - 2 = 1$. Then we can solve the recurrence relation

$$a_n = a_{n-1} + a_{n-2}, \qquad n \geq 2, \qquad a_0 = 1, \qquad a_1 = 2.$$

Before going any further we want to be sure that the reader understands why a general argument is needed when we develop our recurrence relations. When we are proving a theorem we do *not* draw any general conclusions from a few (or even, perhaps, many) particular instances. The same is true here. The following example should serve to drive this point home.

EXAMPLE 10.17

We start with n identical pennies and let a_n count the number of ways we can arrange these pennies — *contiguous* in each row where each penny above the bottom row touches two pennies in the row below it. (In these arrangements we are *not* concerned with whether any given penny is heads up or heads down.) In Fig. 10.7 we have the possible arrangements for $1 \leq n \leq 6$. From this it follows that

$$a_1 = 1, \quad a_2 = 1, \quad a_3 = 2, \quad a_4 = 3, \quad a_5 = 5, \quad \text{and} \quad a_6 = 8.$$

Consequently, these results might *suggest* that, in general, $a_n = F_n$, the nth Fibonacci number. Unfortunately, we have been led astray, as one finds, for example, that

$$a_7 = 12 \neq 13 = F_7, \quad a_8 = 18 \neq 21 = F_8, \quad \text{and} \quad a_9 = 26 \neq 34 = F_9.$$

(The arrangements in this example were studied by F. C. Auluck in reference [2].)

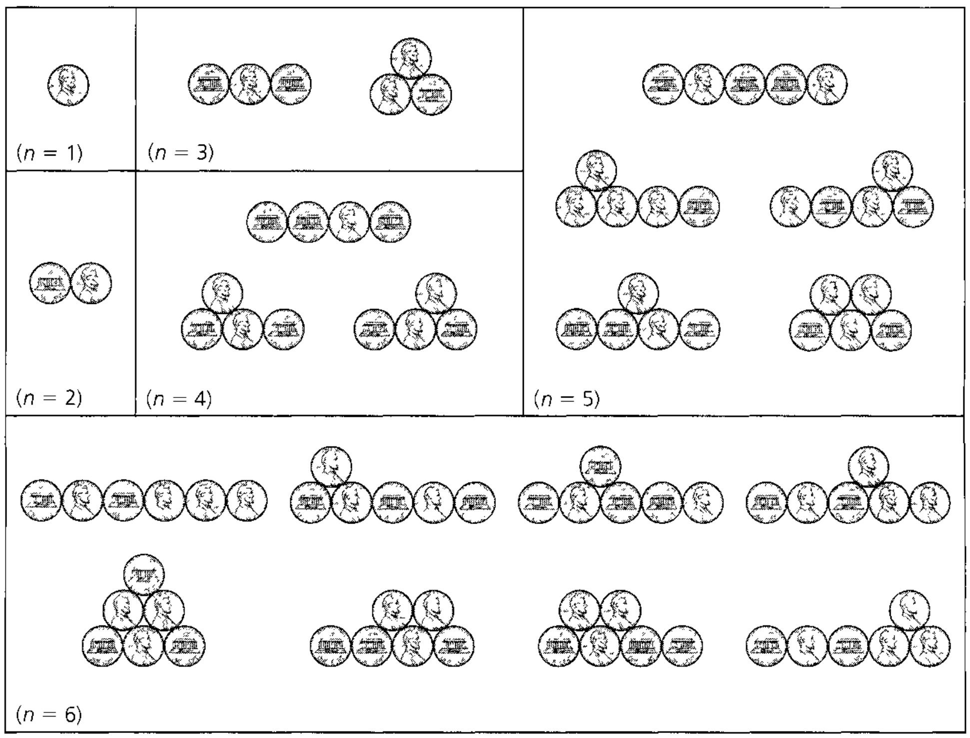


Figure 10.7

The last two examples for case (A) show us how to extend the results for second-order recurrence relations to those of higher order.

EXAMPLE 10.18

Solve the recurrence relation

$$2a_{n+3} = a_{n+2} + 2a_{n+1} - a_n, \quad n \geq 0, \quad a_0 = 0, \quad a_1 = 1, \quad a_2 = 2.$$

Letting $a_n = cr^n$ for $c, r \neq 0$ and $n \geq 0$, we obtain the characteristic equation $2r^3 - r^2 - 2r + 1 = 0 = (2r - 1)(r - 1)(r + 1)$. The characteristic roots are $1/2$, 1, and -1, so the solution is $a_n = c_1(1)^n + c_2(-1)^n + c_3(1/2)^n = c_1 + c_2(-1)^n + c_3(1/2)^n$. [The solutions 1, $(-1)^n$, and $(1/2)^n$ are called linearly independent because it is impossible to express any one of them as a linear combination of the other two.[†]] From $0 = a_0$, $1 = a_1$, and $2 = a_2$, we derive $c_1 = 5/2$, $c_2 = 1/6$, $c_3 = -8/3$. Consequently, $a_n = (5/2) + (1/6)(-1)^n + (-8/3)(1/2)^n$, $n \geq 0$.

EXAMPLE 10.19

For $n \geq 1$ we want to tile a $2 \times n$ chessboard using the two types of tiles shown in part (a) of Fig. 10.8. Letting a_n count the number of such tilings, we find that $a_1 = 1$, since we can tile a 2×1 chessboard (of one column) in only one way — using two 1×1 square tiles. Part (b) of the figure shows us that $a_2 = 5$. Finally, for the 2×3 chessboard there are 11 possible tilings: (i) one that uses six 1×1 square tiles; (ii) eight that use three 1×1 square tiles and one of the larger tiles; and (iii) two that use two of the larger tiles. When $n \geq 4$ we consider the nth column of the $2 \times n$ chessboard. There are three cases to examine:

1) the nth column is covered by two 1×1 square tiles — this case provides a_{n-1} tilings;
2) the $(n - 1)$st and nth columns are tiled with one 1×1 square tile and one larger tile — this case accounts for $4a_{n-2}$ tilings; and
3) the $(n - 2)$nd, $(n - 1)$st, and nth columns are tiled with two of the larger tiles — this results in $2a_{n-3}$ tilings.

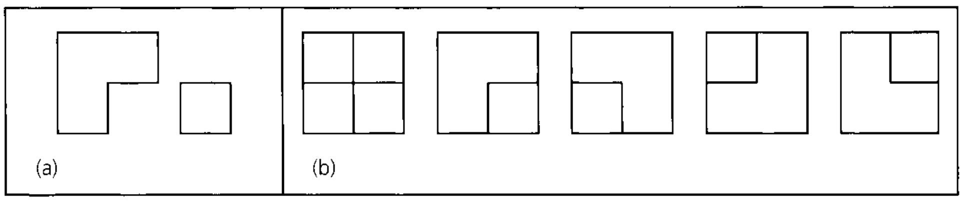


Figure 10.8

These three cases cover all possibilities and no two of the cases have anything in common, so

$$a_n = a_{n-1} + 4a_{n-2} + 2a_{n-3}, \qquad n \geq 4, \qquad a_1 = 1, \qquad a_2 = 5, \qquad a_3 = 11.$$

The characteristic equation $x^3 - x^2 - 4x - 2 = 0$ can be written as $(x + 1)(x^2 - 2x - 2) = 0$, so the characteristic roots are -1, $1 + \sqrt{3}$, and $1 - \sqrt{3}$. Consequently, $a_n = c_1(-1)^n + c_2(1 + \sqrt{3})^n + c_3(1 - \sqrt{3})^n$, $n \geq 1$. From $1 = a_1 = -c_1 + c_2(1 + \sqrt{3}) + c_3(1 - \sqrt{3})$, $5 = a_2 = c_1 + c_2(1 + \sqrt{3})^2 + c_3(1 - \sqrt{3})^2$, and $11 = a_3 = -c_1 + c_2(1 + \sqrt{3})^3 + c_3(1 - \sqrt{3})^3$, we have $c_1 = 1$, $c_2 = 1/\sqrt{3}$, and $c_3 = -1/\sqrt{3}$. So

$$a_n = (-1)^n + (1/\sqrt{3})(1 + \sqrt{3})^n + (-1/\sqrt{3})(1 - \sqrt{3})^n, \qquad n \geq 1.$$

Case (B): (Complex Roots)

Before getting into the case of complex roots, we recall DeMoivre's Theorem:

$$(\cos\theta + i \sin\theta)^n = \cos n\theta + i \sin n\theta, \qquad n \geq 0.$$

[This is part (b) of Exercise 12 of Section 4.1.]

[†]Alternatively, the solutions 1, $(-1)^n$, and $(1/2)^n$ are linearly independent, because if k_1, k_2, k_3 are real numbers, and $k_1(1) + k_2(-1)^n + k_3(1/2)^n = 0$ for *all* $n \in \mathbf{N}$, then $k_1 = k_2 = k_3 = 0$.

If $z = x + iy \in \mathbf{C}$, $z \neq 0$, we can write $z = r(\cos\theta + i\sin\theta)$, where $r = \sqrt{x^2 + y^2}$ and $(y/x) = \tan\theta$, for $x \neq 0$. If $x = 0$, then for $y > 0$,

$$z = yi = yi\sin(\pi/2) = y(\cos(\pi/2) + i\sin(\pi/2)),$$

and for $y < 0$,

$$z = yi = |y|i\sin(3\pi/2) = |y|(\cos(3\pi/2) + i\sin(3\pi/2)).$$

In all cases, $z^n = r^n(\cos n\theta + i\sin n\theta)$, for $n \geq 0$, by DeMoivre's Theorem.

EXAMPLE 10.20

Determine $(1 + \sqrt{3}\,i)^{10}$.

Figure 10.9 shows a geometric way to represent the complex number $1 + \sqrt{3}\,i$ as the point $(1, \sqrt{3})$ in the xy-plane. Here $r = \sqrt{1^2 + (\sqrt{3})^2} = 2$, and $\theta = \pi/3$.

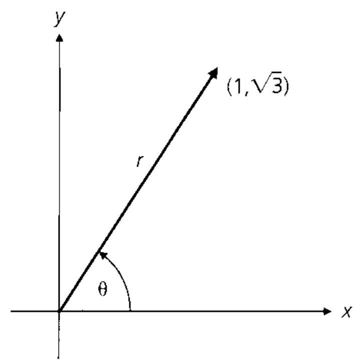


Figure 10.9

So $1 + \sqrt{3}\,i = 2(\cos(\pi/3) + i\sin(\pi/3))$, and

$$\begin{aligned}(1 + \sqrt{3}\,i)^{10} &= 2^{10}(\cos(10\pi/3) + i\sin(10\pi/3)) = 2^{10}(\cos(4\pi/3) + i\sin(4\pi/3))\\ &= 2^{10}((-1/2) - (\sqrt{3}/2)i) = (-2^9)(1 + \sqrt{3}\,i).\end{aligned}$$

We'll use such results in the following examples.

EXAMPLE 10.21

Solve the recurrence relation $a_n = 2(a_{n-1} - a_{n-2})$, where $n \geq 2$ and $a_0 = 1$, $a_1 = 2$.

Letting $a_n = cr^n$, for $c, r \neq 0$, we obtain the characteristic equation $r^2 - 2r + 2 = 0$, whose roots are $1 \pm i$. Consequently, the general solution has the form $c_1(1+i)^n + c_2(1-i)^n$, where c_1 and c_2 presently denote arbitrary *complex* constants. [As in case (A), there are two independent solutions: $(1+i)^n$ and $(1-i)^n$.]

$$1 + i = \sqrt{2}(\cos(\pi/4) + i\sin(\pi/4))$$

and

$$1 - i = \sqrt{2}(\cos(-\pi/4) + i\sin(-\pi/4)) = \sqrt{2}(\cos(\pi/4) - i\sin(\pi/4)).$$

This yields

$$\begin{aligned} a_n &= c_1(1+i)^n + c_2(1-i)^n \\ &= c_1[\sqrt{2}(\cos(\pi/4) + i\sin(\pi/4))]^n + c_2[\sqrt{2}(\cos(-\pi/4) + i\sin(-\pi/4))]^n \\ &= c_1(\sqrt{2})^n(\cos(n\pi/4) + i\sin(n\pi/4)) + c_2(\sqrt{2})^n(\cos(-n\pi/4) + i\sin(-n\pi/4)) \\ &= c_1(\sqrt{2})^n(\cos(n\pi/4) + i\sin(n\pi/4)) + c_2(\sqrt{2})^n(\cos(n\pi/4) - i\sin(n\pi/4)) \\ &= (\sqrt{2})^n[k_1\cos(n\pi/4) + k_2\sin(n\pi/4)], \end{aligned}$$

where $k_1 = c_1 + c_2$ and $k_2 = (c_1 - c_2)i$.

$$1 = a_0 = [k_1\cos 0 + k_2\sin 0] = k_1$$
$$2 = a_1 = \sqrt{2}[1\cdot\cos(\pi/4) + k_2\sin(\pi/4)], \text{ or } 2 = 1 + k_2, \text{ and } k_2 = 1.$$

The solution for the given initial conditions is then given by

$$a_n = (\sqrt{2})^n[\cos(n\pi/4) + \sin(n\pi/4)], \qquad n \geq 0.$$

[*Note*: This solution contains no complex numbers. A small point may bother the reader here. How did we start with c_1, c_2 complex and end up with $k_1 = c_1 + c_2$ and $k_2 = (c_1 - c_2)i$ real? This happens if c_1, c_2 are complex conjugates.]

Let us now examine an application from linear algebra.

EXAMPLE 10.22

For $b \in \mathbf{R}^+$, consider the $n \times n$ determinant† D_n given by

$$\begin{vmatrix} b & b & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ b & b & b & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & b & b & b & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & b & b & b & \cdots & 0 & 0 & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & \cdots & b & b & b & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & b & b & b & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & b & b & b \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & b & b \end{vmatrix}$$

Find the value of D_n as a function of n.

Let a_n, $n \geq 1$, denote the value of the $n \times n$ determinant D_n. Then

$$a_1 = |b| = b \quad \text{and} \quad a_2 = \begin{vmatrix} b & b \\ b & b \end{vmatrix} = 0 \quad \left(\text{and} \quad a_3 = \begin{vmatrix} b & b & 0 \\ b & b & b \\ 0 & b & b \end{vmatrix} = -b^3\right).$$

†The expansion of determinants is discussed in Appendix 2.

Expanding D_n by its first row, we have $D_n =$

$$b\underbrace{\begin{vmatrix} b & b & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ b & b & b & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & b & b & b & \cdots & 0 & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdots & b & b & b & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & b & b & b \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & b & b \end{vmatrix}}_{\text{(This is } D_{n-1}.)} - b\begin{vmatrix} b & b & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & b & b & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & b & b & b & \cdots & 0 & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdots & b & b & b & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & b & b & b \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & b & b \end{vmatrix}$$

When we expand the second determinant by its first column, we find that $D_n = bD_{n-1} - (b)(b)D_{n-2} = bD_{n-1} - b^2D_{n-2}$. This translates into the relation $a_n = ba_{n-1} - b^2a_{n-2}$, for $n \geq 3$, $a_1 = b$, $a_2 = 0$.

If we let $a_n = cr^n$ for c, $r \neq 0$ and $n \geq 1$, the characteristic equation produces the roots $b[(1/2) \pm i\sqrt{3}/2]$.

Hence

$$\begin{aligned} a_n &= c_1[b((1/2) + i\sqrt{3}/2)]^n + c_2[b((1/2) - i\sqrt{3}/2)]^n \\ &= b^n[c_1(\cos(\pi/3) + i\sin(\pi/3))^n + c_2(\cos(\pi/3) - i\sin(\pi/3))^n] \\ &= b^n[k_1\cos(n\pi/3) + k_2\sin(n\pi/3)]. \end{aligned}$$

$b = a_1 = b[k_1\cos(\pi/3) + k_2\sin(\pi/3)]$, so $1 = k_1(1/2) + k_2(\sqrt{3}/2)$, or $k_1 + \sqrt{3}\,k_2 = 2$.

$0 = a_2 = b^2[k_1\cos(2\pi/3) + k_2\sin(2\pi/3)]$, so $0 = (k_1)(-1/2) + k_2(\sqrt{3}/2)$, or $k_1 = \sqrt{3}\,k_2$.

Hence $k_1 = 1$, $k_2 = 1/\sqrt{3}$ and the value of D_n is

$$b^n[\cos(n\pi/3) + (1/\sqrt{3})\sin(n\pi/3)].$$

Case (C): (Repeated Real Roots)

EXAMPLE 10.23

Solve the recurrence relation $a_{n+2} = 4a_{n+1} - 4a_n$, where $n \geq 0$ and $a_0 = 1$, $a_1 = 3$.

As in the other two cases, we let $a_n = cr^n$, where c, $r \neq 0$ and $n \geq 0$. Then the characteristic equation is $r^2 - 4r + 4 = 0$ and the characteristic roots are both $r = 2$. (So $r = 2$ is called "a root of multiplicity 2.") Unfortunately, we now lack two independent solutions: 2^n and 2^n are definitely multiples of each other. We need one more independent solution. Let us try $g(n)2^n$ where $g(n)$ is not a constant. Substituting this into the given relation yields

$$g(n+2)2^{n+2} = 4g(n+1)2^{n+1} - 4g(n)2^n$$

or

$$g(n+2) = 2g(n+1) - g(n). \tag{1}$$

One finds that $g(n) = n$ satisfies Eq. (1).† So $n2^n$ is a second independent solution. (It is independent because it is impossible to have $n2^n = k2^n$ for *all* $n \geq 0$ if k is a constant.)

†Actually, the general solution is $g(n) = an + b$, for arbitrary constants a, b, with $a \neq 0$. Here we chose $a = 1$ and $b = 0$ to make $g(n)$ as simple as possible.

The general solution is of the form $a_n = c_1(2^n) + c_2n(2^n)$. With $a_0 = 1, a_1 = 3$ we find $a_n = 2^n + (1/2)n(2^n) = 2^n + n(2^{n-1}), n \geq 0$.

In general, if $C_0a_n + C_1a_{n-1} + C_2a_{n-2} + \cdots + C_ka_{n-k} = 0$, with C_0 ($\neq 0$), $C_1, C_2, \ldots, C_k$ ($\neq 0$) real constants, and r a characteristic root of multiplicity m, where $2 \leq m \leq k$, then the part of the general solution that involves the root r has the form

$$A_0r^n + A_1nr^n + A_2n^2r^n + \cdots + A_{m-1}n^{m-1}r^n = (A_0 + A_1n + A_2n^2 + \cdots + A_{m-1}n^{m-1})r^n,$$

where $A_0, A_1, A_2, \ldots, A_{m-1}$ are arbitrary constants.

Our last example involves a little probability.

EXAMPLE 10.24

If a first case of measles is recorded in a certain school system, let p_n denote the probability that at least one case is reported during the nth week after the first recorded case. School records provide evidence that $p_n = p_{n-1} - (0.25)p_{n-2}$, where $n \geq 2$. Since $p_0 = 0$ and $p_1 = 1$, if the first case (of a new outbreak) is recorded on Monday, March 3, 2003, when did the probability for the occurrence of a new case decrease to less than 0.01 for the first time?

With $p_n = cr^n$ for $c, r \neq 0$, the characteristic equation for the recurrence relation is $r^2 - r + (1/4) = 0 = (r - (1/2))^2$. The general solution has the form $p_n = (c_1 + c_2n)(1/2)^n$, $n \geq 0$. For $p_0 = 0$, $p_1 = 1$, we get $c_1 = 0, c_2 = 2$, so $p_n = n2^{-n+1}, n \geq 0$.

The first integer n for which $p_n < 0.01$ is 12. Hence, it was not until the week of May 19, 2003, that the probability of another new case occurring was less than 0.01.

EXERCISES 10.2

1. Solve the following recurrence relations. (No final answer should involve complex numbers.)

a) $a_n = 5a_{n-1} + 6a_{n-2},\ n \geq 2,\ a_0 = 1,\ a_1 = 3$

b) $2a_{n+2} - 11a_{n+1} + 5a_n = 0,\ n \geq 0,\ a_0 = 2,\ a_1 = -8$

c) $a_{n+2} + a_n = 0,\ n \geq 0,\ a_0 = 0,\ a_1 = 3$

d) $a_n - 6a_{n-1} + 9a_{n-2} = 0,\ n \geq 2,\ a_0 = 5,\ a_1 = 12$

e) $a_n + 2a_{n-1} + 2a_{n-2} = 0,\ n \geq 2,\ a_0 = 1,\ a_1 = 3$

2. a) Verify the final solutions in Examples 10.14 and 10.23.

b) Solve the recurrence relation in Example 10.16.

3. If $a_0 = 0$, $a_1 = 1$, $a_2 = 4$, and $a_3 = 37$ satisfy the recurrence relation $a_{n+2} + ba_{n+1} + ca_n = 0$, where $n \geq 0$ and b, c are constants, determine b, c and solve for a_n.

4. Find and solve a recurrence relation for the number of ways to park motorcycles and compact cars in a row of n spaces if each cycle requires one space and each compact needs two. (All cycles are identical in appearance, as are the cars, and we want to use up all the n spaces.)

5. Answer the question posed in Exercise 4 if (a) the motorcycles come in two distinct models; (b) the compact cars come in three different colors; and (c) the motorcycles come in two distinct models and the compact cars come in three different colors.

6. Answer the questions posed in Exercise 5 if empty spaces are allowed.

7. In Exercise 12 of Section 4.2 we learned that $F_0 + F_1 + F_2 + \cdots + F_n = \sum_{i=0}^{n} F_i = F_{n+2} - 1$. This is one of many such properties of the Fibonacci numbers that were discovered by the French mathematician François Lucas (1842–1891). Although we established the result by the Principle of Mathematical Induction, we see that it is easy to develop this formula by adding the system of $n + 1$ equations

$$\begin{aligned} F_0 &= F_2 - F_1 \\ F_1 &= F_3 - F_2 \\ \cdots &\quad \cdots \quad \cdots \\ F_{n-1} &= F_{n+1} - F_n \\ F_n &= F_{n+2} - F_{n+1}. \end{aligned}$$

Develop formulas for each of the following sums, and then check the general result by the Principle of Mathematical Induction.

a) $F_1 + F_3 + F_5 + \cdots + F_{2n-1}$, where $n \in \mathbf{Z}^+$

b) $F_0 + F_2 + F_4 + \cdots + F_{2n}$, where $n \in \mathbf{Z}^+$

8. a) Prove that

$$\lim_{n\to\infty} \frac{F_{n+1}}{F_n} = \frac{1+\sqrt{5}}{2}.$$

(This limit has come to be known as the *golden ratio* and is often designated by α, as we mentioned in Example 10.10.)

b) Consider a regular pentagon $ABCDE$ inscribed in a circle, as shown in Fig. 10.10.

i) Use the law of sines and the double angle formula for the sine to show that $AC/AX = 2\cos 36°$.

ii) As $\cos 18° = \sin 72°$ $= 4 \sin 18° \cos 18°(1 - 2\sin^2 18°)$ (Why?), show that $\sin 18°$ is a root of the polynomial equation $8x^3 - 4x + 1 = 0$, and deduce that $\sin 18° = (\sqrt{5} - 1)/4$.

c) Verify that $AC/AX = (1 + \sqrt{5})/2$.

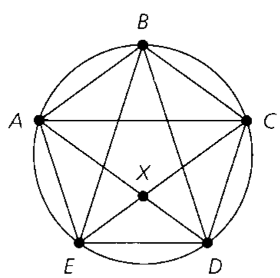


Figure 10.10

9. For $n \geq 0$, let a_n count the number of ways a sequence of 1's and 2's will sum to n. For example, $a_3 = 3$ because (1) 1, 1, 1; (2) 1, 2; and (3) 2, 1 sum to 3. Find and solve a recurrence relation for a_n.

10. For $\Sigma = \{0, 1\}$, let $A \subseteq \Sigma^*$, where $A = \{00, 1\}$. For $n \geq 1$, let a_n count the number of strings in A^* of length n. Find and solve a recurrence relation for a_n. (The reader may wish to refer to Exercise 25 for Section 6.1.)

11. a) For $n \geq 1$, let a_n count the number of binary strings of length n, where there are no consecutive 1's. Find and solve a recurrence relation for a_n.

b) For $n \geq 1$, let b_n count the number of binary strings of length n, where there are no consecutive 1's and the first and last bit of the string are not both 1. Find and solve a recurrence relation for b_n.

12. Suppose that poker chips come in four colors — red, white, green, and blue. Find and solve a recurrence relation for the number of ways to stack n of these poker chips so that there are no consecutive blue chips.

13. An alphabet Σ consists of the four numeric characters 1, 2, 3, 4, and the seven alphabetic characters a, b, c, d, e, f, g. Find and solve a recurrence relation for the number of words of length n (in Σ^*), where there are no consecutive (identical or distinct) alphabetic characters.

14. An alphabet Σ consists of seven numeric characters and k alphabetic characters. For $n \geq 0$, a_n counts the number of strings (in Σ^*) of length n that contain no consecutive (identical or distinct) alphabetic characters. If $a_{n+2} = 7a_{n+1} + 63a_n$, $n \geq 0$, what is the value of k?

15. Solve the recurrence relation $a_{n+2} = a_{n+1}a_n$, $n \geq 0$, $a_0 = 1$, $a_1 = 2$.

16. For $n \geq 1$, let a_n be the number of ways to write n as an ordered sum of positive integers, where each summand is at least 2. (For example, $a_5 = 3$ because here we may represent 5 by 5, by 2 + 3, and by 3 + 2.) Find and solve a recurrence relation for a_n.

17. a) For a fixed nonnegative integer n, how many compositions of $n + 3$ have no 1 as a summand?

b) For the compositions in part (a), how many start with (i) 2; (ii) 3; (iii) k, where $2 \leq k \leq n + 1$?

c) How many of the compositions in part (a) start with $n + 2$ or $n + 3$?

d) How are the results in parts (a)–(c) related to the formula derived at the start of Exercise 7?

18. Determine the points of intersection of the parabola $y = x^2 - 1$ and the line $y = x$.

19. Find the points of intersection of the hyperbola $y = 1 + \frac{1}{x}$ and the line $y = x$.

20. a) For $\alpha = (1 + \sqrt{5})/2$, show that $\alpha^2 = \alpha + 1$.

b) If $n \in \mathbf{Z}^+$, prove that $\alpha^n = \alpha F_n + F_{n-1}$.

21. Let F_n denote the nth Fibonacci number, for $n \geq 0$, and let $\alpha = (1 + \sqrt{5})/2$. For $n \geq 3$, prove that (a) $F_n > \alpha^{n-2}$ and (b) $F_n < \alpha^{n-1}$.

22. a) For $n \in \mathbf{Z}^+$, let a_n count the number of palindromes of $2n$. Then $a_{n+1} = 2a_n$, $n \geq 1$, $a_1 = 2$. Solve this first-order recurrence relation for a_n.

b) For $n \in \mathbf{Z}^+$, let b_n count the number of palindromes of $2n - 1$. Set up and solve a first-order recurrence relation for b_n.

(You may want to compare your solutions here with those given in Examples 9.13 and 10.15.)

23. Consider ternary strings — that is, strings where 0, 1, 2 are the only symbols used. For $n \geq 1$, let a_n count the number of ternary strings of length n where there are no consecutive 1's and no consecutive 2's. Find and solve a recurrence relation for a_n.

24. For $n \geq 1$, let a_n count the number of ways to tile a $2 \times n$ chessboard using horizontal (1×2) dominoes [which can also be used as vertical (2×1) dominoes] and square (2×2) tiles. Find and solve a recurrence relation for a_n.

25. In how many ways can one tile a 2×10 chessboard using dominoes and square tiles (as in Exercise 24) if the dominoes come in four colors and the square tiles come in five colors?

26. Let $\Sigma = \{0, 1\}$ and $A = \{0, 01, 11\} \subseteq \Sigma^*$. For $n \geq 1$, let a_n count the number of strings in A^* of length n. Find and solve a recurrence relation for a_n.

27. Let $\Sigma = \{0, 1\}$ and $A = \{0, 01, 011, 111\} \subseteq \Sigma^*$. For $n \geq 1$, let a_n count the number of strings in A^* of length n. Find and solve a recurrence relation for a_n.

28. Let $\Sigma = \{0, 1\}$ and $A = \{0, 01, 011, 0111, 1111\} \subseteq \Sigma^*$. For $n \geq 1$, let a_n count the number of strings in A^* of length n. Find and solve a recurrence relation for a_n.

29. A particle moves horizontally to the right. For $n \in \mathbf{Z}^+$, the distance the particle travels in the $(n + 1)$st second is equal to twice the distance it travels during the nth second. If x_n, $n \geq 0$, denotes the position of the particle at the start of the $(n + 1)$st second, find and solve a recurrence relation for x_n, where $x_0 = 1$ and $x_1 = 5$.

30. For $n \geq 1$, let D_n be the following $n \times n$ determinant.

$$\begin{vmatrix} 2 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 \\ . & . & . & . & . & \cdots & . & . & . & . \\ 0 & 0 & 0 & 0 & 0 & \cdots & 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 2 \end{vmatrix}$$

Find and solve a recurrence relation for the value of D_n.

31. Solve the recurrence relation $a_{n+2}^2 - 5a_{n+1}^2 + 4a_n^2 = 0$, where $n \geq 0$ and $a_0 = 4$, $a_1 = 13$.

32. Determine the constants b and c if $a_n = c_1 + c_2(7^n)$, $n \geq 0$, is the general solution of the relation $a_{n+2} + ba_{n+1} + ca_n = 0$, $n \geq 0$.

33. Prove that any two consecutive Fibonacci numbers are relatively prime.

34. Write a computer program (or develop an algorithm) to determine whether a given nonnegative integer is a Fibonacci number.

10.3 The Nonhomogeneous Recurrence Relation

We now turn to the recurrence relations

$$a_n + C_1 a_{n-1} = f(n), \qquad n \geq 1, \tag{1}$$

$$a_n + C_1 a_{n-1} + C_2 a_{n-2} = f(n), \qquad n \geq 2, \tag{2}$$

where C_1 and C_2 are constants, $C_1 \neq 0$ in Eq. (1), $C_2 \neq 0$, and $f(n)$ is not identically 0. Although there is no general method for solving all nonhomogeneous relations, for certain functions $f(n)$ we shall find a successful technique.

We start with the special case for Eq. (1), when $C_1 = -1$. For the nonhomogeneous relation $a_n - a_{n-1} = f(n)$, we have

$$\begin{aligned} a_1 &= a_0 + f(1) \\ a_2 &= a_1 + f(2) = a_0 + f(1) + f(2) \\ a_3 &= a_2 + f(3) = a_0 + f(1) + f(2) + f(3) \\ &\vdots \\ a_n &= a_{n-1} + f(n) = a_0 + f(1) + \cdots + f(n) = a_0 + \sum_{i=1}^{n} f(i). \end{aligned}$$

We can solve this type of relation in terms of n, if we can find a suitable summation formula for $\sum_{i=1}^{n} f(i)$.

EXAMPLE 10.25

Solve the recurrence relation $a_n - a_{n-1} = 3n^2$, where $n \geq 1$ and $a_0 = 7$.

Here $f(n) = 3n^2$, so the unique solution is

$$a_n = a_0 + \sum_{i=1}^{n} f(i) = 7 + 3\sum_{i=1}^{n} i^2 = 7 + \frac{1}{2}(n)(n+1)(2n+1).$$

When a formula for the summation is not known, the following procedure will handle Eq. (1) for certain functions $f(n)$, regardless of the value of C_1 ($\neq 0$). It also works for the second-order nonhomogeneous relation in Eq. (2) — again, for certain functions $f(n)$. Known as the *method of undetermined coefficients*, it relies on the associated homogeneous relation obtained when $f(n)$ is replaced by 0.

For either of Eq. (1) or Eq. (2), we let $a_n^{(h)}$ denote the general solution of the associated homogeneous relation, and we let $a_n^{(p)}$ be a solution of the given nonhomogeneous relation. The term $a_n^{(p)}$ is called a *particular* solution. Then $a_n = a_n^{(h)} + a_n^{(p)}$ is the general solution of the given relation. To determine $a_n^{(p)}$ we use the form of $f(n)$ to suggest a form for $a_n^{(p)}$.

EXAMPLE 10.26

Solve the recurrence relation $a_n - 3a_{n-1} = 5(7^n)$, where $n \geq 1$ and $a_0 = 2$.

The solution of the associated homogeneous relation is $a_n^{(h)} = c(3^n)$. Since $f(n) = 5(7^n)$, we seek a particular solution $a_n^{(p)}$ of the form $A(7^n)$. As $a_n^{(p)}$ is to be a solution of the given nonhomogeneous relation, we place $a_n^{(p)} = A(7^n)$ into the given relation and find that $A(7^n) - 3A(7^{n-1}) = 5(7^n)$, $n \geq 1$. Dividing by 7^{n-1}, we find that $7A - 3A = 5(7)$, so $A = 35/4$, and $a_n^{(p)} = (35/4)7^n = (5/4)7^{n+1}$, $n \geq 0$. The general solution is $a_n = c(3^n) + (5/4)7^{n+1}$. With $2 = a_0 = c + (5/4)(7)$, it follows that $c = -27/4$ and $a_n = (5/4)(7^{n+1}) - (1/4)(3^{n+3})$, $n \geq 0$.

EXAMPLE 10.27

Solve the recurrence relation $a_n - 3a_{n-1} = 5(3^n)$, where $n \geq 1$ and $a_0 = 2$.

As in Example 10.26, $a_n^{(h)} = c(3^n)$, but here $a_n^{(h)}$ and $f(n)$ are not linearly independent. As a result we consider a particular solution $a_n^{(p)}$ of the form $Bn(3^n)$. (What happens if we substitute $a_n^{(p)} = B(3^n)$ into the given relation?)

Substituting $a_n^{(p)} = Bn3^n$ into the given relation yields

$$Bn(3^n) - 3B(n-1)(3^{n-1}) = 5(3^n), \quad \text{or} \quad Bn - B(n-1) = 5, \quad \text{so} \quad B = 5.$$

Hence $a_n = a_n^{(h)} + a_n^{(p)} = (c + 5n)3^n$, $n \geq 0$. With $a_0 = 2$, the unique solution is $a_n = (2 + 5n)(3^n)$, $n \geq 0$.

From the two preceding examples we generalize as follows.

Consider the nonhomogeneous first-order relation

$$a_n + C_1 a_{n-1} = kr^n,$$

where k is a constant and $n \in \mathbf{Z}^+$. If r^n is *not* a solution of the associated homogeneous relation

$$a_n + C_1 a_{n-1} = 0,$$

then $a_n^{(p)} = Ar^n$, where A is a constant. When r^n is a solution of the associated homogeneous relation, then $a_n^{(p)} = Bnr^n$, for B a constant.

Now consider the case of the nonhomogeneous second-order relation

$$a_n + C_1 a_{n-1} + C_2 a_{n-2} = kr^n,$$

where k is a constant. Here we find that

a) $a_n^{(p)} = Ar^n$, for A a constant, if r^n is not a solution of the associated homogeneous relation;

b) $a_n^{(p)} = Bnr^n$, where B is a constant, if $a_n^{(h)} = c_1 r^n + c_2 r_1^n$, where $r_1 \neq r$; and

c) $a_n^{(p)} = Cn^2 r^n$, for C a constant, when $a_n^{(h)} = (c_1 + c_2 n) r^n$.

EXAMPLE 10.28

The Towers of Hanoi. Consider n circular disks (having different diameters) with holes in their centers. These disks can be stacked on any of the pegs shown in Fig. 10.11. In the figure, $n = 5$ and the disks are stacked on peg 1 with no disk resting upon a smaller one. The objective is to transfer the disks one at a time so that we end up with the original stack on peg 3. Each of pegs 1, 2, and 3 may be used as a temporary location for any disk(s), but at no time are we allowed to have a larger disk on top of a smaller one on any peg. What is the minimum number of moves needed to do this for n disks?

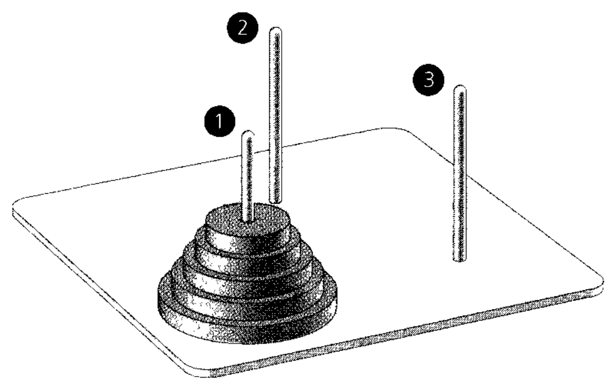


Figure 10.11

For $n \geq 0$, let a_n count the *minimum* number of moves it takes to transfer n disks from peg 1 to peg 3 in the manner described. Then, for $n + 1$ disks we can do the following:

a) Transfer the top n disks from peg 1 to peg 2 according to the directions that are given. This takes at least a_n moves.

b) Transfer the largest disk from peg 1 to peg 3. This takes one move.

c) Finally, transfer the n disks on peg 2 onto the largest disk, now on peg 3 — once again following the specified directions. This also requires at least a_n moves.

Consequently, at this point we know that a_{n+1} is no more than $2a_n + 1$ — that is, $a_{n+1} \leq 2a_n + 1$. But could there be a method where we actually have $a_{n+1} < 2a_n + 1$? Alas, no! For at some point the largest disk (the one at the bottom of the original stack — on peg 1) must be moved to peg 3. This move requires that peg 3 has no disks on it. So this largest disk may only be moved to peg 3 after the n smaller disks have moved to peg 2 [where they are stacked in increasing size from the smallest (on the top) to the largest (on the bottom)]. Getting these n smaller disks moved, accordingly, requires at least a_n moves. The largest disk must be moved at least once to get it to peg 3. Then, to get the n smaller disks on top of the largest disk (all on peg 3), according to the requirements, requires at least a_n more steps. So $a_{n+1} \geq a_n + 1 + a_n = 2a_n + 1$.

With $2a_n + 1 \leq a_{n+1} \leq 2a_n + 1$, we now obtain the relation $a_{n+1} = 2a_n + 1$, where $n \geq 0$ and $a_0 = 0$.

For $a_{n+1} - 2a_n = 1$, we know that $a_n^{(h)} = c(2^n)$. Since $f(n) = 1 = (1)^n$ is not a solution of $a_{n+1} - 2a_n = 0$, we set $a_n^{(p)} = A(1)^n = A$ and find from the given relation that $A = 2A + 1$, so $A = -1$ and $a_n = c(2^n) - 1$. From $a_0 = 0 = c - 1$ it then follows that $c = 1$, so $a_n = 2^n - 1, n \geq 0$.

The next example arises from the mathematics of finance.

EXAMPLE 10.29

Pauline takes out a loan of S dollars that is to be paid back in T periods of time. If r is the interest rate per period for the loan, what (constant) payment P must she make at the end of each period?

We let a_n denote the amount still owed on the loan at the end of the nth period (following the nth payment). Then at the end of the $(n + 1)$st period, the amount Pauline still owes on her loan is a_n (the amount she owed at the end of the nth period) $+ ra_n$ (the interest that accrued during the $(n + 1)$st period) $- P$ (the payment she made at the end of the $(n + 1)$st period). This gives us the recurrence relation

$$a_{n+1} = a_n + ra_n - P, \qquad 0 \leq n \leq T - 1, \qquad a_0 = S, \qquad a_T = 0.$$

For this relation $a_n^{(h)} = c(1 + r)^n$, while $a_n^{(p)} = A$ since no constant is a solution of the associated homogeneous relation. With $a_n^{(p)} = A$ we find $A - (1 + r)A = -P$, so $A = P/r$. From $a_0 = S$, we obtain $a_n = (S - (P/r))(1 + r)^n + (P/r), 0 \leq n \leq T$.

Since $0 = a_T = (S - (P/r))(1 + r)^T + (P/r)$, it follows that

$$(P/r) = ((P/r) - S)(1 + r)^T \quad \text{and} \quad P = (Sr)[1 - (1 + r)^{-T}]^{-1}.$$

We now consider a problem in the analysis of algorithms.

EXAMPLE 10.30

For $n \geq 1$, let S be a set containing 2^n real numbers.

The following procedure is used to determine the maximum and minimum elements of S. We wish to determine the number of comparisons made between pairs of elements in S during the execution of this procedure.

If a_n denotes the number of needed comparisons, then $a_1 = 1$. When $n = 2$, $|S| = 2^2 = 4$, so $S = \{x_1, x_2, y_1, y_2\} = S_1 \cup S_2$ where $S_1 = \{x_1, x_2\}$, $S_2 = \{y_1, y_2\}$. Since $a_1 = 1$, it takes one comparison to determine the maximum and minimum elements in each of S_1, S_2. Comparing the minimum elements of S_1 and S_2 and then their maximum elements, we learn the maximum and minimum elements in S and find that $a_2 = 4 = 2a_1 + 2$. In general, if $|S| = 2^{n+1}$, we write $S = S_1 \cup S_2$ where $|S_1| = |S_2| = 2^n$. To determine the maximum and minimum elements in each of S_1 and S_2 requires a_n comparisons. Comparing the maximum (minimum) elements of S_1 and S_2 requires one more comparison; consequently, $a_{n+1} = 2a_n + 2, n \geq 1$.

Here $a_n^{(h)} = c(2^n)$ and $a_n^{(p)} = A$, a constant. Substituting $a_n^{(p)}$ into the relation, we find that $A = 2A + 2$, or $A = -2$. So $a_n = c2^n - 2$, and with $a_1 = 1 = 2c - 2$, we obtain $c = 3/2$. Therefore $a_n = (3/2)(2^n) - 2$.

A note of caution! The existence of this procedure, which requires $(3/2)(2^n) - 2$ comparisons, does *not* exclude the possibility that we could achieve the same results via another remarkably clever method that requires fewer comparisons.

An example on counting certain strings of length 10, for the quaternary alphabet $\Sigma = \{0, 1, 2, 3\}$, provides a slight twist to what we've been doing so far.

EXAMPLE 10.31

For the alphabet $\Sigma = \{0, 1, 2, 3\}$, there are $4^{10} = 1{,}048{,}576$ strings of length 10 (in Σ^{10}, or Σ^*). Now we want to know how many of these more than 1 million strings contain an even number of 1's.

Instead of being so specific about the length of the strings, we will start by letting a_n count those strings among the 4^n strings in Σ^n where there are an even number of 1's. To determine how the strings counted by a_n, for $n \geq 2$, are related to those counted by a_{n-1}, consider the nth symbol of one of these strings of length n (where there is an even number of 1's). Two cases arise:

1) The nth symbol is 0, 2, or 3: Here the preceding $n - 1$ symbols provide one of the strings counted by a_{n-1}. So this case provides $3a_{n-1}$ of the strings counted by a_n.
2) The nth symbol is 1: In this case, there must be an odd number of 1's among the first $n - 1$ symbols. There are 4^{n-1} strings of length $n - 1$ and we want to avoid those that have an even number of 1's — there are $4^{n-1} - a_{n-1}$ such strings. Consequently, this second case gives us $4^{n-1} - a_{n-1}$ of the strings counted by a_n.

These two cases are exhaustive and mutually disjoint, so we may write

$$a_n = 3a_{n-1} + (4^{n-1} - a_{n-1}) = 2a_{n-1} + 4^{n-1}, \qquad n \geq 2.$$

Here $a_1 = 3$ (for the strings 0, 2, and 3). We find that $a_n^{(h)} = c(2^n)$ and $a_n^{(p)} = A(4^{n-1})$. Upon substituting $a_n^{(p)}$ into the above relation we have $A(4^{n-1}) = 2A(4^{n-2}) + 4^{n-1}$, so $4A = 2A + 4$ and $A = 2$. Hence, $a_n = c(2^n) + 2(4^{n-1})$, $n \geq 2$. From $3 = a_1 = 2c + 2$ it follows that $c = 1/2$, so $a_n = 2^{n-1} + 2(4^{n-1})$, $n \geq 1$.

When $n = 10$, we learn that of the $4^{10} = 1{,}048{,}576$ strings in Σ^{10}, there are $2^9 + 2(4^9) = 524{,}800$ that contain an even number of 1's.

Before continuing we realize that the answer here for a_n can be checked by using the exponential generating function $f(x) = \sum_{n=0}^{\infty} a_n \frac{x^n}{n!}$ (where $a_0 = 1$). From the techniques developed in Section 9.4 we have

$$\begin{aligned}
f(x) &= \left(1 + x + \frac{x^2}{2!} + \cdots\right)\left(1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots\right)\left(1 + x + \frac{x^2}{2!} + \cdots\right)\left(1 + x + \frac{x^2}{2!} + \cdots\right) \\
&= e^x \cdot \left(\frac{e^x + e^{-x}}{2}\right) \cdot e^x \cdot e^x \\
&= \left(\frac{1}{2}\right) e^{4x} + \left(\frac{1}{2}\right) e^{2x} \\
&= \left(\frac{1}{2}\right) \sum_{n=0}^{\infty} \frac{(4x)^n}{n!} + \left(\frac{1}{2}\right) \sum_{n=0}^{\infty} \frac{(2x)^n}{n!}.
\end{aligned}$$

Here $a_n =$ the coefficient of $\frac{x^n}{n!}$ in $f(x) = \left(\frac{1}{2}\right) 4^n + \left(\frac{1}{2}\right) 2^n = 2^{n-1} + 2(4^{n-1})$, as above.

EXAMPLE 10.32

In 1904, the Swedish mathematician Helge von Koch (1870–1924) created the intriguing curve now known as the Koch "snowflake" curve. The construction of this curve starts with an equilateral triangle, as shown in part (a) of Fig. 10.12, where the triangle has side 1, perimeter 3, and area $\sqrt{3}/4$. (Recall that an equilateral triangle of side s has perimeter $3s$ and area $s^2\sqrt{3}/4$.) The triangle is then transformed into the Star of David in Fig. 10.12(b) by removing the middle one-third of each side (of the original equilateral triangle) and attaching a new equilateral triangle whose side has length $1/3$. So as we go from part (a) to part (b) in the figure, each side of length 1 is transformed into 4 sides of length $1/3$, and we get a 12-sided polygon of area $(\sqrt{3}/4) + (3)(\sqrt{3}/4)(1/3)^2 = \sqrt{3}/3$. Continuing the process, we transform the figure of part (b) into that of part (c) by removing the middle one-third of each of the 12 sides in the Star of David and attaching an equilateral triangle of side $1/9$ $(= (1/3)^2)$. Now we have [in Fig. 10.12(c)] a $4^2(3)$-sided polygon whose area is

$$(\sqrt{3}/3) + (4)(3)(\sqrt{3}/4)[(1/3)^2]^2 = 10\sqrt{3}/27.$$

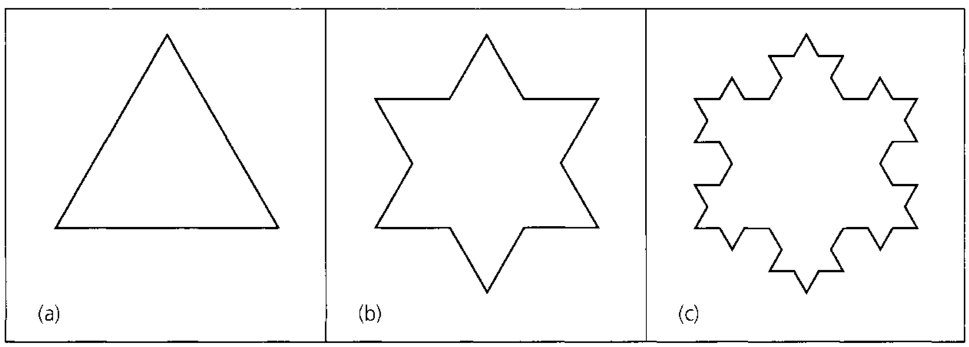


Figure 10.12

For $n \geq 0$, let a_n denote the area of the polygon P_n obtained from the original equilateral triangle after we apply n transformations of the type described above [the first from P_0 in Fig. 10.12(a) to P_1 in Fig. 10.12(b) and the second from P_1 in Fig. 10.12(b) to P_2 in Fig. 10.12(c)]. As we go from P_n (with $4^n(3)$ sides) to P_{n+1} (with $4^{n+1}(3)$ sides), we find that

$$a_{n+1} = a_n + (4^n(3))(\sqrt{3}/4)(1/3^{n+1})^2 = a_n + (1/(4\sqrt{3}))(4/9)^n$$

because in transforming P_n into P_{n+1} we remove the middle one-third of each of the $4^n(3)$ sides of P_n and attach an equilateral triangle of side $(1/3^{n+1})$.

The homogeneous part of the solution for this first-order nonhomogeneous recurrence relation is $a_n^{(h)} = A(1)^n = A$. Since $(4/9)^n$ is not a solution of the associated homogeneous relation, the particular solution is given by $a_n^{(p)} = B(4/9)^n$, where B is a constant. Substituting this into the recurrence relation $a_{n+1} = a_n + (1/(4\sqrt{3}))(4/9)^n$, we find that $B = (-9/5)(1/(4\sqrt{3}))$. Consequently,

$$a_n = A + (-9/5)(1/(4\sqrt{3}))(4/9)^n = A - (1/(5\sqrt{3}))(4/9)^{n-1}, \qquad n \geq 0.$$

Since $\sqrt{3}/4 = a_0 = A - (1/(5\sqrt{3}))(4/9)^{-1}$, it follows that $A = 6/(5\sqrt{3})$ and

$$a_n = (6/(5\sqrt{3})) - (1/(5\sqrt{3}))(4/9)^{n-1} = (1/(5\sqrt{3}))[6 - (4/9)^{n-1}], \qquad n \geq 0.$$

[As n grows larger, we find that $(4/9)^{n-1}$ tends to 0 and a_n approaches $6/(5\sqrt{3})$. We can also obtain this value by continuing the calculations we had before we introduced our recurrence relation, thus noting that this limiting area is also given by

$$\begin{aligned}(\sqrt{3}/4) &+ (\sqrt{3}/4)(3)(1/3)^2 + (\sqrt{3}/4)(4)(3)(1/3^2)^2 + (\sqrt{3}/4)(4^2)(3)(1/3^3)^2 + \cdots \\ &= (\sqrt{3}/4) + (\sqrt{3}/4)(3)\sum_{n=0}^{\infty} 4^n(1/3^{n+1})^2 = (\sqrt{3}/4) + (1/(4\sqrt{3}))\sum_{n=0}^{\infty}(4/9)^n \\ &= (\sqrt{3}/4) + (1/(4\sqrt{3}))[1/(1-(4/9))] = (\sqrt{3}/4) + (1/(4\sqrt{3}))(9/5) = 6/(5\sqrt{3})\end{aligned}$$

by using the result for the sum of a geometric series from part (b) of Example 9.5.]

EXAMPLE 10.33

For $n \geq 1$, let $X_n = \{1, 2, 3, \ldots, n\}$; $\mathcal{P}(X_n)$ denotes the power set of X_n. We want to determine a_n, the number of edges in the Hasse diagram for the partial order $(\mathcal{P}(X_n), \subseteq)$. Here $a_1 = 1$ and $a_2 = 4$, and from Fig. 10.13 it follows that

$$a_3 = 2a_2 + 2^2.$$

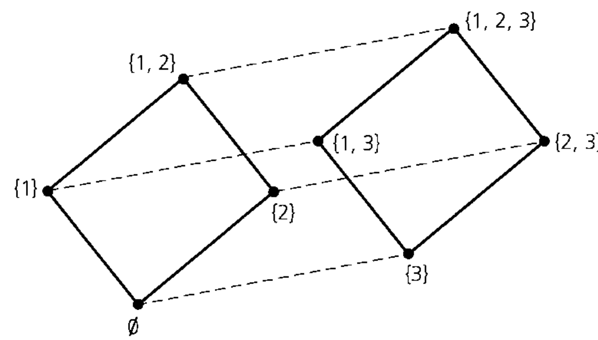


Figure 10.13

This is because the Hasse diagram for $(\mathcal{P}(X_3), \subseteq)$ contains the a_2 edges in the Hasse diagram for $(\mathcal{P}(X_2), \subseteq)$ as well as the a_2 edges in the Hasse diagram for the partial order $(\{\{3\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}, \subseteq)$. [Note the identical structure shared by the partial orders $(\mathcal{P}(\{1, 2\}), \subseteq)$ and $(\{\{3\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}, \subseteq)$.] In addition, there are 2^2 other (dashed) edges — one for each subset of $\{1, 2\}$. Now for $n \geq 1$, consider the Hasse diagrams for the partial orders $(\mathcal{P}(X_n), \subseteq)$ and $(\{T \cup \{n+1\} | T \in \mathcal{P}(X_n)\}, \subseteq)$. For each $S \in \mathcal{P}(X_n)$, draw an edge from S in $(\mathcal{P}(X_n), \subseteq)$ to $S \cup \{n+1\}$ in $(\{T \cup \{n+1\} | T \in \mathcal{P}(X_n)\}, \subseteq)$. The result is the Hasse diagram for $(\mathcal{P}(X_{n+1}), \subseteq)$. From the construction we see that

$$a_{n+1} = 2a_n + 2^n, \qquad n \geq 1, \qquad a_1 = 1.$$

The solution to this recurrence relation, with the given condition $a_1 = 1$, is $a_n = n2^{n-1}$, $n \geq 1$.

Each of our next two examples deals with a second-order relation.

EXAMPLE 10.34

Solve the recurrence relation

$$a_{n+2} - 4a_{n+1} + 3a_n = -200, \qquad n \geq 0, \qquad a_0 = 3000, \qquad a_1 = 3300.$$

Here $a_n^{(h)} = c_1(3^n) + c_2(1^n) = c_1(3^n) + c_2$. Since $f(n) = -200 = -200(1^n)$ is a solution of the associated homogeneous relation, here $a_n^{(p)} = An$ for some constant A. This leads us to

$$A(n+2) - 4A(n+1) + 3An = -200, \quad \text{so} \quad -2A = -200, \quad A = 100.$$

Hence $a_n = c_1(3^n) + c_2 + 100n$. With $a_0 = 3000$ and $a_1 = 3300$, we have $a_n = 100(3^n) + 2900 + 100n$, $n \geq 0$.

Before proceeding any further, a point needs to be made about the role of technology in solving recurrence relations. When a computer algebra system is available, we are spared much of the drudgery of computation. Consequently, all our effort can be directed to analyzing the situation at hand and setting up the recurrence relation with its initial condition(s). Once this is done our job is just about finished. A line or two of code will often do the trick! For example, the Maple code in Fig. 10.14 shows how one can readily solve the recurrence relations of Examples 10.33 and 10.34.

```
> rsolve({a(n+1)=2*a(n)+2^n,a(1)=1},a(n));
```

$$-\frac{2^n}{2} + \left(\frac{n}{2} + \frac{1}{2}\right)2^n$$

```
> simplify(%);
```

$$2^{(n-1)}\,n$$

```
> rsolve({a(n+2)=4*a(n+1)+3*a(n)=-200,a(0)=3000,a(1)=3300},a(n));
```

$$100\,3^n + 2900 + 100\,n$$

Figure 10.14

EXAMPLE 10.35

In part (a) of Fig. 10.15 we have an *iterative* algorithm (written as a pseudocode procedure) for computing the nth Fibonacci number, for $n \geq 0$. Here the input is a nonnegative integer n and the output is the Fibonacci number F_n. The variables *i*, *fib*, *last*, *next_to_last*, and *temp* are integer variables. In this algorithm we calculate F_n (in this case for $n \geq 0$) by first assigning or computing all of the previous values $F_0, F_1, F_2, \ldots, F_{n-1}$. Here the number of additions needed to determine F_n is 0 for $n = 0, 1$ and $n - 1$ (within the **for** loop) for $n \geq 2$.

Part (b) of Fig. 10.15 provides a pseudocode procedure to implement a *recursive* algorithm for calculating F_n for $n \in \mathbf{N}$. Here the variable *fib* is likewise an integer variable. For this procedure we wish to determine a_n, the number of additions performed in computing F_n, $n \geq 0$. We find that $a_0 = 0$, $a_1 = 0$, and from the shaded line in the procedure — namely,

```
fib := FibNum2(n - 1) + FibNum2(n - 2)                (*)
```

we obtain the nonhomogeneous recurrence relation

$$a_n = a_{n-1} + a_{n-2} + 1, \qquad n \geq 2,$$

where the summand of 1 is due to the addition in Eq. (*).

```
procedure FibNum1(n: nonnegative integer)
begin
  if n = 0 then
    fib := 0
  else if n = 1 then
    fib := 1
  else
    begin
      last := 1
      next_to_last := 0
      for i := 2 to n do
        begin
          temp := last
          last := last + next_to_last
          next_to_last := temp
        end
      fib := last
    end
end                                        (a)
```

```
procedure FibNum2(n: nonnegative integer)
begin
  if n = 0 then
    fib := 0
  else if n = 1 then
    fib := 1
  else
    fib := FibNum2(n - 1) + FibNum2(n - 2)
end                                        (b)
```

Figure 10.15

Here we find that $a_n^{(h)} = c_1\left(\frac{1+\sqrt{5}}{2}\right)^n + c_2\left(\frac{1-\sqrt{5}}{2}\right)^n$ and that $a_n^{(p)} = A$, a constant. Upon substituting $a_n^{(p)}$ into the nonhomogeneous recurrence relation we find that

$$A = A + A + 1,$$

so $A = -1$ and $a_n = c_1\left(\frac{1+\sqrt{5}}{2}\right)^n + c_2\left(\frac{1-\sqrt{5}}{2}\right)^n - 1$.

Since $a_0 = 0$ and $a_1 = 0$ it follows that

$$c_1 + c_2 = 1 \quad \text{and} \quad c_1\left(\frac{1+\sqrt{5}}{2}\right) + c_2\left(\frac{1-\sqrt{5}}{2}\right) = 1.$$

From these equations we learn that $c_1 = (1+\sqrt{5})/(2\sqrt{5})$, $c_2 = (\sqrt{5}-1)/(2\sqrt{5})$. Therefore,

$$\begin{aligned} a_n &= \left(\frac{1+\sqrt{5}}{2\sqrt{5}}\right)\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2\sqrt{5}}\right)\left(\frac{1-\sqrt{5}}{2}\right)^n - 1 \\ &= \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^{n+1} - 1. \end{aligned}$$

As n gets larger $[(1-\sqrt{5})/2]^{n+1}$ approaches 0 since $|(1-\sqrt{5})/2|<1$, and $a_n \doteq (1/\sqrt{5})[(1+\sqrt{5})/2]^{n+1} = ((1+\sqrt{5})/(2\sqrt{5}))((1+\sqrt{5})/2)^n$.

Consequently, we can see that, as the value of n increases, the first procedure requires far less computation than the second one does.

We now summarize and extend the solution techniques already discussed in Examples 10.26 through 10.35.

Given a linear nonhomogeneous recurrence relation (with constant coefficients) of the form $C_0a_n + C_1a_{n-1} + C_2a_{n-2} + \cdots + C_ka_{n-k} = f(n)$, where $C_0 \neq 0$ and $C_k \neq 0$, let $a_n^{(h)}$ denote the homogeneous part of the solution a_n.

1) If $f(n)$ is a constant multiple of one of the forms in the first column of Table 10.2 and is not a solution of the associated homogeneous relation, then $a_n^{(p)}$ has the form shown in the second column of Table 10.2. (Here A, B, A_0, A_1, $A_2, \ldots, A_{t-1}$, A_t are constants determined by substituting $a_n^{(p)}$ into the given relation; t, r, and θ are also constants.)

Table 10.2

	$a_n^{(p)}$
c, a constant	A, a constant
n	$A_1n + A_0$
n^2	$A_2n^2 + A_1n + A_0$
n^t, $t \in \mathbf{Z}^+$	$A_tn^t + A_{t-1}n^{t-1} + \cdots + A_1n + A_0$
r^n, $r \in \mathbf{R}$	Ar^n
$\sin \theta n$	$A \sin \theta n + B \cos \theta n$
$\cos \theta n$	$A \sin \theta n + B \cos \theta n$
n^tr^n	$r^n(A_tn^t + A_{t-1}n^{t-1} + \cdots + A_1n + A_0)$
$r^n \sin \theta n$	$Ar^n \sin \theta n + Br^n \cos \theta n$
$r^n \cos \theta n$	$Ar^n \sin \theta n + Br^n \cos \theta n$

2) When $f(n)$ comprises a sum of constant multiples of terms such as those in the first column of the table for item (1), and none of these terms is a solution of the associated homogeneous relation, then $a_n^{(p)}$ is made up of the sum of the corresponding terms in the column headed by $a_n^{(p)}$. For example, if $f(n) = n^2 + 3\sin 2n$ and no summand of $f(n)$ is a solution of the associated homogeneous relation, then $a_n^{(p)} = (A_2n^2 + A_1n + A_0) + (A \sin 2n + B \cos 2n)$.

3) Things get trickier if a summand $f_1(n)$ of $f(n)$ is a solution of the associated homogeneous relation. This happens, for example, when $f(n)$ contains summands such as cr^n or $(c_1 + c_2n)r^n$ and r is a characteristic root. If $f_1(n)$ causes this problem, we multiply the trial solution $(a_n^{(p)})_1$ corresponding to $f_1(n)$ by the smallest power of n, say n^s, for which no summand of $n^s f_1(n)$ is a solution of the associated homogeneous relation. Then $n^s\ (a_n^{(p)})_1$ is the corresponding part of $a_n^{(p)}$.

In order to check some of our preceding remarks on particular solutions for nonhomogeneous recurrence relations, the next application provides us with a situation that can be solved in more than one way.

EXAMPLE 10.36

For $n \geq 2$, suppose that there are n people at a party and that each of these people shakes hands (exactly one time) with all of the other people there (and no one shakes hands with himself or herself). If a_n counts the total number of handshakes, then

$$a_{n+1} = a_n + n, \qquad n \geq 2, \qquad a_2 = 1, \tag{3}$$

because when the $(n+1)$st person arrives, he or she will shake hands with the n other people who have already arrived.

According to the results in Table 10.2, we might think that the trial (particular) solution for Eq. (3) is $A_1 n + A_0$, for constants A_0 and A_1. But here the associated homogeneous relation is $a_{n+1} = a_n$, or $a_{n+1} - a_n = 0$, for which $a_n^{(h)} = c(1^n) = c$, where c denotes an arbitrary constant. Therefore, the summand A_0 (in $A_1 n + A_0$) is a solution of the associated homogeneous relation. Consequently, the third remark (given with Table 10.2) tells us that we must multiply $A_1 n + A_0$ by the smallest power of n for which we no longer have any constant summand. This is accomplished by multiplying $A_1 n + A_0$ by n^1, and so we find here that

$$a_n^{(p)} = A_1 n^2 + A_0 n.$$

When we substitute this result into Eq. (3) we have

$$A_1(n+1)^2 + A_0(n+1) = A_1 n^2 + A_0 n + n,$$

or $A_1 n^2 + (2A_1 + A_0)n + (A_1 + A_0) = A_1 n^2 + (A_0 + 1)n$.

By comparing the coefficients on like powers of n we find that

(n^2): $A_1 = A_1$;

(n): $2A_1 + A_0 = A_0 + 1$; and

(n^0): $A_1 + A_0 = 0$.

Consequently, $A_1 = 1/2$ and $A_0 = -1/2$, so $a_n^{(p)} = (1/2)n^2 + (-1/2)n$ and $a_n = a_n^{(h)} + a_n^{(p)} = c + (1/2)(n)(n-1)$. Since $a_2 = 1$, it follows from $1 = a_2 = c + (1/2)(2)(1)$ that $c = 0$, and $a_n = (1/2)(n)(n-1)$, for $n \geq 2$.

We can also obtain this result by considering the n people in the room and realizing that each possible handshake corresponds with a selection of size 2 from this set of size n — and there are $\binom{n}{2} = (n!)/(2!(n-2)!) = (1/2)(n)(n-1)$ such selections. [Or we can consider the n people as vertices of an undirected graph (with no loops) where an edge corresponds with a handshake. Our answer is then the number of edges in the complete graph K_n, and there are $\binom{n}{2} = (1/2)(n)(n-1)$ such edges.]

Our last example further demonstrates how we may use the results in Table 10.2.

EXAMPLE 10.37

a) Consider the nonhomogeneous recurrence relation

$$a_{n+2} - 10a_{n+1} + 21a_n = f(n), \qquad n \geq 0.$$

Here the homogeneous part of the solution is

$$a_n^{(h)} = c_1(3^n) + c_2(7^n),$$

for arbitrary constants c_1, c_2.

In Table 10.3 we list the form for the particular solution for certain choices of $f(n)$. Here the values of the 11 constants A_i, for $0 \leq i \leq 10$, are determined by substituting $a_n^{(p)}$ into the given nonhomogeneous recurrence relation.

Table 10.3

$f(n)$	$a_n^{(p)}$
5	A_0
$3n^2 - 2$	$A_3n^2 + A_2n + A_1$
$7(11^n)$	$A_4(11^n)$
$31(r^n), r \neq 3, 7$	$A_5(r^n)$
$6(3^n)$	A_6n3^n
$2(3^n) - 8(9^n)$	$A_7n3^n + A_8(9^n)$
$4(3^n) + 3(7^n)$	$A_9n3^n + A_{10}n7^n$

b) The homogeneous component of the solution for

$$a_n + 4a_{n-1} + 4a_{n-2} = f(n), \qquad n \geq 2,$$

is

$$a_n^{(h)} = c_1(-2)^n + c_2n(-2)^n,$$

where c_1, c_2 denote arbitrary constants. Consequently,

1) if $f(n) = 5(-2)^n$, then $a_n^{(p)} = An^2(-2)^n$;
2) if $f(n) = 7n(-2)^n$, then $a_n^{(p)} = n^2(-2)^n(A_1n + A_0)$; and
3) if $f(n) = -11n^2(-2)^n$, then $a_n^{(p)} = n^2(-2)^n(B_2n^2 + B_1n + B_0)$.
(Here, the constants A, A_0, A_1, B_0, B_1, and B_2 are determined by substituting $a_n^{(p)}$ into the given nonhomogeneous recurrence relation.)

EXERCISES 10.3

1. Solve each of the following recurrence relations.

a) $a_{n+1} - a_n = 2n + 3, \quad n \geq 0, \quad a_0 = 1$
b) $a_{n+1} - a_n = 3n^2 - n, \quad n \geq 0, \quad a_0 = 3$
c) $a_{n+1} - 2a_n = 5, \quad n \geq 0, \quad a_0 = 1$
d) $a_{n+1} - 2a_n = 2^n, \quad n \geq 0, \quad a_0 = 1$

2. Use a recurrence relation to derive the formula for $\sum_{i=0}^{n} i^2$.

3. a) Let n lines be drawn in the plane such that each line intersects every other line but no three lines are ever coincident. For $n \geq 0$, let a_n count the number of regions into which the plane is separated by the n lines. Find and solve a recurrence relation for a_n.

b) For the situation in part (a), let b_n count the number of infinite regions that result. Find and solve a recurrence relation for b_n.

4. On the first day of a new year, Joseph deposits \$1000 in an account that pays 6% interest compounded monthly. At the beginning of each month he adds \$200 to his account. If he continues to do this for the next four years (so that he makes 47 additional deposits of \$200), how much will his account be worth exactly four years after he opened it?

5. Solve the following recurrence relations.

a) $a_{n+2} + 3a_{n+1} + 2a_n = 3^n, \quad n \geq 0, \quad a_0 = 0, \quad a_1 = 1$
b) $a_{n+2} + 4a_{n+1} + 4a_n = 7, \quad n \geq 0, \quad a_0 = 1, \quad a_1 = 2$

6. Solve the recurrence relation $a_{n+2} - 6a_{n+1} + 9a_n = 3(2^n) + 7(3^n)$, where $n \geq 0$ and $a_0 = 1$, $a_1 = 4$.

7. Find the general solution for the recurrence relation $a_{n+3} - 3a_{n+2} + 3a_{n+1} - a_n = 3 + 5n$, $n \geq 0$.

8. Determine the number of n-digit quaternary (0, 1, 2, 3) sequences in which there is never a 3 anywhere to the right of a 0.

9. Meredith borrows \$2500, at 12% compounded monthly, to buy a computer. If the loan is to be paid back over two years, what is his monthly payment?

10. The general solution of the recurrence relation $a_{n+2} + b_1a_{n+1} + b_2a_n = b_3n + b_4$, $n \geq 0$, with b_i constant for $1 \leq i \leq 4$, is $c_12^n + c_23^n + n - 7$. Find b_i for each $1 \leq i \leq 4$.

11. Solve the following recurrence relations.

a) $a_{n+2}^2 - 5a_{n+1}^2 + 6a_n^2 = 7n, \quad n \geq 0, \quad a_0 = a_1 = 1$
b) $a_n^2 - 2a_{n-1} = 0, \quad n \geq 1, \quad a_0 = 2$ (Let $b_n = \log_2 a_n$, $n \geq 0$.)

12. Let $\Sigma = \{0, 1, 2, 3\}$. For $n \geq 1$, let a_n count the number of strings in Σ^n containing an odd number of 1's. Find and solve a recurrence relation for a_n.

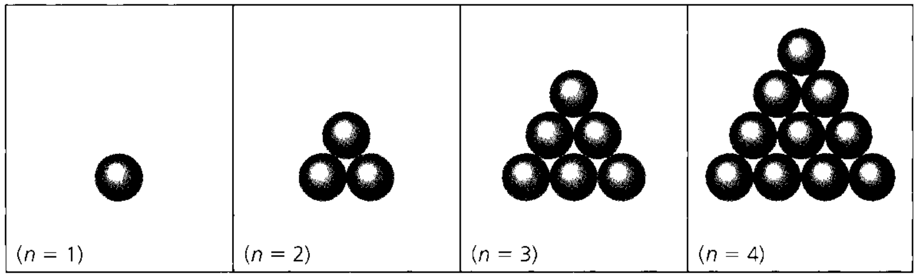


Figure 10.16

13. **a)** For the binary string 001110, there are three runs: 00, 111, and 0. Meanwhile, the string 000111 has only two runs: 000 and 111; while the string 010101 determines the six runs: 0, 1, 0, 1, 0, 1. For $n = 1$, we consider two binary strings, namely, 0 and 1 — these two strings (of length 1) determine a total of two runs. There are four binary strings of length $n = 2$ and these strings determine 1 (for 00) + 2 (for 01) + 2 (for 10) + 1 (for 11) = 6 runs. Find and solve a recurrence relation for t_n, the total number of runs determined by the 2^n binary strings of length n, where $n \geq 1$.

b) Answer the question posed in part (a) for quaternary strings of length n. (Here the alphabet comprises 0, 1, 2, 3.)

c) Generalize the results of parts (a) and (b).

14. **a)** For $n \geq 1$, the nth *triangular number* t_n is defined by $t_n = 1 + 2 + \cdots + n = n(n+1)/2$. Find and solve a recurrence relation for s_n, $n \geq 1$, where $s_n = t_1 + t_2 + \cdots + t_n$, the sum of the first n triangular numbers. [The reader may wish to compare the result obtained here with the formula given in Example 4.5 or with the result requested in part (b) of Exercise 8 of Section 9.5.]

b) In an organic laboratory, Kelsey synthesizes a crystalline structure that is made up of 10,000,000 triangular layers of atoms. The first layer of the structure has one atom, the second layer has three atoms, and, in general, the nth layer has $1 + 2 + \cdots + n = t_n$ atoms. (Consider each layer, other than the last, as if it were placed upon the spaces that result among the neighboring atoms of the succeeding layer. See Fig. 10.16.) (i) How many atoms are there in one of these crystalline structures? (ii) How many atoms are packed (strictly) between the 10,000th and 100,000th layer?

15. Write a computer program (or develop an algorithm) to solve the problem of the Towers of Hanoi. For $n \in \mathbf{Z}^+$, the program should provide the necessary steps for transferring the n disks from peg 1 to peg 3 under the restrictions specified in Example 10.28.

10.4 The Method of Generating Functions

With all the different cases we had to consider for the nonhomogeneous linear recurrence relation, we now get some assistance from the generating function. This technique will find both the homogeneous and the particular solutions for a_n, and it will incorporate the given initial conditions as well. Furthermore, we'll be able to do even more with this method.

We demonstrate the method in the following examples.

EXAMPLE 10.38

Solve the relation $a_n - 3a_{n-1} = n$, $n \geq 1$, $a_0 = 1$.

This relation represents an infinite set of equations:

$$\begin{array}{ll} (n = 1) & a_1 - 3a_0 = 1 \\ (n = 2) & a_2 - 3a_1 = 2 \\ \vdots & \vdots \quad \vdots \end{array}$$

Multiplying the first of these equations by x, the second by x^2, and so on, we obtain

$$\begin{array}{lll} (n=1) & a_1x^1 - 3a_0x^1 = 1x^1 \\ (n=2) & a_2x^2 - 3a_1x^2 = 2x^2 \\ \vdots & \vdots \quad \vdots \end{array}$$

Adding this second set of equations, we find that

$$\sum_{n=1}^{\infty} a_nx^n - 3\sum_{n=1}^{\infty} a_{n-1}x^n = \sum_{n=1}^{\infty} nx^n. \tag{1}$$

We want to solve for a_n in terms of n. To accomplish this, let $f(x) = \sum_{n=0}^{\infty} a_nx^n$ be the (ordinary) generating function for the sequence $a_0, a_1, a_2 \ldots$. Then Eq. (1) can be rewritten as

$$(f(x) - a_0) - 3x\sum_{n=1}^{\infty} a_{n-1}x^{n-1} = \sum_{n=1}^{\infty} nx^n \left(= \sum_{n=0}^{\infty} nx^n\right). \tag{2}$$

Since $\sum_{n=1}^{\infty} a_{n-1}x^{n-1} = \sum_{n=0}^{\infty} a_nx^n = f(x)$ and $a_0 = 1$, the left-hand side of Eq. (2) becomes $(f(x) - 1) - 3xf(x)$.

Before we can proceed, we need the generating function for the sequence $0, 1, 2, 3, \ldots$. Recall from part (c) of Example 9.5 that

$$\frac{x}{(1-x)^2} = x + 2x^2 + 3x^3 + \cdots, \quad \text{so}$$

$$(f(x) - 1) - 3xf(x) = \frac{x}{(1-x)^2}, \quad \text{and} \quad f(x) = \frac{1}{(1-3x)} + \frac{x}{(1-x)^2(1-3x)}.$$

Using a partial fraction decomposition, we find that

$$\frac{x}{(1-x)^2(1-3x)} = \frac{A}{(1-x)} + \frac{B}{(1-x)^2} + \frac{C}{(1-3x)},$$

or

$$x = A(1-x)(1-3x) + B(1-3x) + C(1-x)^2.$$

From the following assignments for x, we get

$$\begin{array}{lll} (x=1): & 1 = B(-2), & B = -\frac{1}{2}. \\ \left(x = \frac{1}{3}\right): & \frac{1}{3} = C\left(\frac{2}{3}\right)^2, & C = \frac{3}{4}. \\ (x=0): & 0 = A + B + C, & A = -(B+C) = -\frac{1}{4}. \end{array}$$

Therefore,

$$\begin{aligned} f(x) &= \frac{1}{1-3x} + \frac{(-1/4)}{(1-x)} + \frac{(-1/2)}{(1-x)^2} + \frac{(3/4)}{(1-3x)} \\ &= \frac{(7/4)}{(1-3x)} + \frac{(-1/4)}{(1-x)} + \frac{(-1/2)}{(1-x)^2}. \end{aligned}$$

We find a_n by determining the coefficient of x^n in each of the three summands.

a) $(7/4)/(1-3x) = (7/4)[1/(1-3x)]$
$= (7/4)[1 + (3x) + (3x)^2 + (3x)^3 + \cdots]$, and the coefficient of x^n is $(7/4)3^n$.

b) $(-1/4)/(1-x) = (-1/4)[1 + x + x^2 + \cdots]$, and the coefficient of x^n here is $(-1/4)$.

c) $(-1/2)/(1-x)^2 = (-1/2)(1-x)^{-2}$
$= (-1/2)\left[\binom{-2}{0} + \binom{-2}{1}(-x) + \binom{-2}{2}(-x)^2 + \binom{-2}{3}(-x)^3 + \cdots\right]$
and the coefficient of x^n is given by $(-1/2)\binom{-2}{n}(-1)^n = (-1/2)(-1)^n\binom{2+n-1}{n}\cdot(-1)^n = (-1/2)(n+1)$.

Therefore $a_n = (7/4)3^n - (1/2)n - (3/4)$, $n \geq 0$. (Note that there is no special concern here with $a_n^{(p)}$. Also, the same answer is obtained by using the techniques of Section 10.3.)

In our next example we extend what we learned in Example 10.38 to a second-order relation. This time we present the solution within a list of instructions one can follow in order to apply the generating-function method.

EXAMPLE 10.39

Consider the recurrence relation

$$a_{n+2} - 5a_{n+1} + 6a_n = 2, \qquad n \geq 0, \qquad a_0 = 3, \qquad a_1 = 7.$$

1) We first multiply this given relation by x^{n+2} because $n+2$ is the largest subscript that appears. This gives us

$$a_{n+2}x^{n+2} - 5a_{n+1}x^{n+2} + 6a_n x^{n+2} = 2x^{n+2}.$$

2) Then we sum all of the equations represented by the result in step (1) and obtain

$$\sum_{n=0}^{\infty} a_{n+2}x^{n+2} - 5\sum_{n=0}^{\infty} a_{n+1}x^{n+2} + 6\sum_{n=0}^{\infty} a_n x^{n+2} = 2\sum_{n=0}^{\infty} x^{n+2}.$$

3) In order to have each of the subscripts on a match the corresponding exponent on x, we rewrite the equation in step (2) as

$$\sum_{n=0}^{\infty} a_{n+2}x^{n+2} - 5x\sum_{n=0}^{\infty} a_{n+1}x^{n+1} + 6x^2\sum_{n=0}^{\infty} a_n x^{n} = 2x^2\sum_{n=0}^{\infty} x^{n}.$$

Here we also rewrite the power series on the right-hand side of the equation in a form that will permit us to use what we learned in Section 2 of Chapter 9.

4) Let $f(x) = \sum_{n=0}^{\infty} a_n x^n$ be the generating function for the solution. The equation in step (3) now takes the form

$$(f(x) - a_0 - a_1x) - 5x(f(x) - a_0) + 6x^2 f(x) = \frac{2x^2}{1-x},$$

or

$$(f(x) - 3 - 7x) - 5x(f(x) - 3) + 6x^2 f(x) = \frac{2x^2}{1-x}.$$

5) Solving for $f(x)$ we have

$$(1-5x+6x^2)f(x) = 3-8x+\frac{2x^2}{1-x} = \frac{3-11x+10x^2}{1-x},$$

from which it follows that

$$f(x) = \frac{3-11x+10x^2}{(1-5x+6x^2)(1-x)} = \frac{(3-5x)(1-2x)}{(1-3x)(1-2x)(1-x)} = \frac{3-5x}{(1-3x)(1-x)}.$$

A partial-fraction decomposition (by hand, or via a computer algebra system) gives us

$$f(x) = \frac{2}{1-3x}+\frac{1}{1-x} = 2\sum_{n=0}^{\infty}(3x)^n + \sum_{n=0}^{\infty} x^n.$$

Consequently, $a_n = 2(3^n)+1,\ n \geq 0$.

We consider a third example, which has a familiar result.

EXAMPLE 10.40

Let $n \in \mathbf{N}$. For $r \geq 0$, let $a(n, r) =$ the number of ways we can select, with repetitions allowed, r objects from a set of n distinct objects.

For $n \geq 1$, let $\{b_1, b_2, \ldots, b_n\}$ be the set of these objects, and consider object b_1. Exactly one of two things can happen.

a) The object b_1 is never selected. Hence the r objects are selected from $\{b_2, \ldots, b_n\}$. This we can do in $a(n-1, r)$ ways.

b) The object b_1 is selected at least once. Then we must select $r-1$ objects from $\{b_1, b_2, \ldots, b_n\}$, so we can continue to select b_1 in addition to the one selection of it we've already made. There are $a(n, r-1)$ ways to accomplish this.

Then $a(n, r) = a(n-1, r) + a(n, r-1)$ because these two cases cover all possibilities and are mutually disjoint.

Let $f_n = \sum_{r=0}^{\infty} a(n, r)x^r$ be the generating function for the sequence $a(n, 0), a(n, 1), a(n, 2), \ldots$. [Here f_n is an abbreviation for $f_n(x)$.] From $a(n, r) = a(n-1, r) + a(n, r-1)$, where $n \geq 1$ and $r \geq 1$, it follows that

$$a(n, r)x^r = a(n-1, r)x^r + a(n, r-1)x^r \quad \text{and}$$

$$\sum_{r=1}^{\infty} a(n, r)x^r = \sum_{r=1}^{\infty} a(n-1, r)x^r + \sum_{r=1}^{\infty} a(n, r-1)x^r.$$

Realizing that $a(n, 0) = 1$ for $n \geq 0$ and $a(0, r) = 0$ for $r > 0$, we write

$$f_n - a(n, 0) = f_{n-1} - a(n-1, 0) + x\sum_{r=1}^{\infty} a(n, r-1)x^{r-1},$$

so $f_n - 1 = f_{n-1} - 1 + xf_n$. Therefore, $f_n - xf_n = f_{n-1}$, or $f_n = f_{n-1}/(1-x)$.

If $n = 5$, for example, then

$$f_5 = \frac{f_4}{(1-x)} = \frac{1}{(1-x)}\cdot\frac{f_3}{(1-x)} = \frac{f_3}{(1-x)^2} = \frac{f_2}{(1-x)^3} = \frac{f_1}{(1-x)^4}$$
$$= \frac{f_0}{(1-x)^5} = \frac{1}{(1-x)^5},$$

since $f_0 = a(0, 0) + a(0, 1)x + a(0, 2)x^2 + \cdots = 1 + 0 + 0 + \cdots$.

In general, $f_n = 1/(1-x)^n = (1-x)^{-n}$, so $a(n, r)$ is the coefficient of x^r in $(1-x)^{-n}$, which is $\binom{-n}{r}(-1)^r = \binom{n+r-1}{r}$.

[Here we dealt with a recurrence relation for $a(n, r)$, a discrete function of the two (integer) variables $n, r \geq 0$.]

Our last example shows how generating functions may be used to solve a system of recurrence relations.

EXAMPLE 10.41

This example provides an approximate model for the propagation of high- and low-energy neutrons as they strike the nuclei of fissionable material (such as uranium) and are absorbed. Here we deal with a fast reactor where there is no moderator (such as water). (In reality, all the neutrons have fairly high energy and there are not just two energy levels. There is a continuous spectrum of energy levels, and these neutrons at the upper end of the spectrum are called the high-energy neutrons. The higher-energy neutrons tend to produce more new neutrons than the lower-energy ones.)

Consider the reactor at time 0 and suppose one high-energy neutron is injected into the system. During each time interval thereafter (about 1 microsecond, or 10^{-6} second) the following events occur.

a) When a high-energy neutron interacts with a nucleus (of fissionable material), upon absorption this results (one microsecond later) in two new high-energy neutrons and one low-energy one.

b) For interactions involving a low-energy neutron, only one neutron of each energy level is produced.

Assuming that all free neutrons interact with nuclei one microsecond after their creation, find functions of n such that

$$a_n = \text{the number of high-energy neutrons},$$

$$b_n = \text{the number of low-energy neutrons},$$

in the reactor after n microseconds, $n \geq 0$.

Here we have $a_0 = 1$, $b_0 = 0$ and the *system* of recurrence relations

$$a_{n+1} = 2a_n + b_n \tag{3}$$

$$b_{n+1} = a_n + b_n. \tag{4}$$

Let $f(x) = \sum_{n=0}^{\infty} a_n x^n$, $g(x) = \sum_{n=0}^{\infty} b_n x^n$ be the generating functions for the sequences $\{a_n | n \geq 0\}$, $\{b_n | n \geq 0\}$, respectively. From Eqs. (3) and (4), when $n \geq 0$

$$a_{n+1}x^{n+1} = 2a_n x^{n+1} + b_n x^{n+1} \tag{3$'$}$$

$$b_{n+1}x^{n+1} = a_n x^{n+1} + b_n x^{n+1}. \tag{4$'$}$$

Summing Eq. (3)′ over all $n \geq 0$, we have

$$\sum_{n=0}^{\infty} a_{n+1}x^{n+1} = 2x \sum_{n=0}^{\infty} a_n x^n + x \sum_{n=0}^{\infty} b_n x^n. \tag{3$''$}$$

In similar fashion, Eq. (4)′ yields

$$\sum_{n=0}^{\infty} b_{n+1}x^{n+1} = x \sum_{n=0}^{\infty} a_n x^n + x \sum_{n=0}^{\infty} b_n x^n. \tag{4$''$}$$

Introducing the generating functions at this point, we get

$$f(x) - a_0 = 2xf(x) + xg(x) \tag{3)''}$$

$$g(x) - b_0 = xf(x) + xg(x), \tag{4)''}$$

a system of equations relating the generating functions. Solving this system, we find that

$$f(x) = \frac{1-x}{x^2-3x+1} = \left(\frac{5+\sqrt{5}}{10}\right)\left(\frac{1}{\gamma - x}\right) + \left(\frac{5-\sqrt{5}}{10}\right)\left(\frac{1}{\delta - x}\right) \quad \text{and}$$

$$g(x) = \frac{x}{x^2-3x+1} = \left(\frac{-5-3\sqrt{5}}{10}\right)\left(\frac{1}{\gamma - x}\right) + \left(\frac{-5+3\sqrt{5}}{10}\right)\left(\frac{1}{\delta - x}\right),$$

where

$$\gamma = \frac{3+\sqrt{5}}{2}, \qquad \delta = \frac{3-\sqrt{5}}{2}.$$

Consequently,

$$a_n = \left(\frac{5+\sqrt{5}}{10}\right)\left(\frac{3-\sqrt{5}}{2}\right)^{n+1} + \left(\frac{5-\sqrt{5}}{10}\right)\left(\frac{3+\sqrt{5}}{2}\right)^{n+1} \quad \text{and}$$

$$b_n = \left(\frac{-5-3\sqrt{5}}{10}\right)\left(\frac{3-\sqrt{5}}{2}\right)^{n+1} + \left(\frac{-5+3\sqrt{5}}{10}\right)\left(\frac{3+\sqrt{5}}{2}\right)^{n+1}, \qquad n \geq 0.$$

EXERCISES 10.4

1. Solve the following recurrence relations by the method of generating functions.

a) $a_{n+1} - a_n = 3^n, \ n \geq 0, \ a_0 = 1$

b) $a_{n+1} - a_n = n^2, \ n \geq 0, \ a_0 = 1$

c) $a_{n+2} - 3a_{n+1} + 2a_n = 0, \ n \geq 0, \ a_0 = 1, \ a_1 = 6$

d) $a_{n+2} - 2a_{n+1} + a_n = 2^n, \ n \geq 0, \ a_0 = 1, \ a_1 = 2$

2. For n distinct objects, let $a(n, r)$ denote the number of ways we can select, without repetition, r of the n objects when $0 \leq r \leq n$. Here $a(n, r) = 0$ when $r > n$. Use the recurrence relation $a(n, r) = a(n-1, r-1) + a(n-1, r)$, where $n \geq 1$ and $r \geq 1$, to show that $f(x) = (1+x)^n$ generates $a(n, r)$, $r \geq 0$.

3. Solve the following systems of recurrence relations.

a) $a_{n+1} = -2a_n - 4b_n$
$b_{n+1} = 4a_n + 6b_n$
$n \geq 0, \ a_0 = 1, \ b_0 = 0$

b) $a_{n+1} = 2a_n - b_n + 2$
$b_{n+1} = -a_n + 2b_n - 1$
$n \geq 0, \ a_0 = 0, \ b_0 = 1$

10.5 A Special Kind of Nonlinear Recurrence Relation (Optional)

Thus far our study of recurrence relations has dealt with linear relations with constant coefficients. The study of nonlinear recurrence relations and of relations with variable coefficients is not a topic we shall pursue except for one special nonlinear relation that lends itself to the method of generating functions.

We shall develop the method in a counting problem on data structures. Before doing so, however, we first observe that if $f(x) = \sum_{i=0}^{\infty} a_i x^i$ is the generating function for $a_0, a_1, a_2, \ldots$, then $[f(x)]^2$ generates $a_0a_0, a_0a_1 + a_1a_0, a_0a_2 + a_1a_1 + a_2a_0, \ldots,$

$a_0a_n + a_1a_{n-1} + a_2a_{n-2} + \cdots + a_{n-1}a_1 + a_na_0, \ldots,$ the convolution of the sequence $a_0, a_1, a_2, \ldots,$ with itself.

EXAMPLE 10.42

In Sections 3.4 and 5.1, we encountered the idea of a tree diagram. In general, a *tree* is an undirected graph that is connected and has no loops or cycles. Here we examine rooted binary trees.

In Fig. 10.17 we see two such trees, where the circled vertex denotes the *root*. These trees are called *binary* because from each vertex there are at most two edges (called *branches*) descending (since a rooted tree is a directed graph) from that vertex.

In particular, these rooted binary trees are *ordered* in the sense that a left branch descending from a vertex is considered different from a right branch descending from that vertex. For the case of three vertices, the five possible ordered rooted binary trees are shown in Fig. 10.18. (If no attention were paid to order, then the last four rooted trees would be the same structure.)

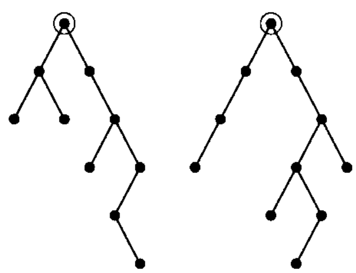


Figure 10.17

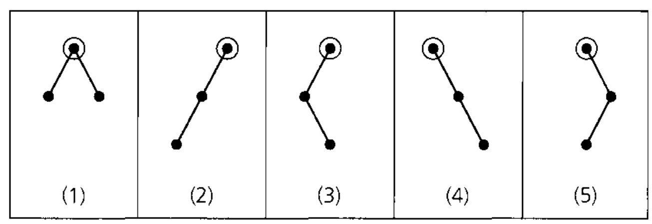


Figure 10.18

Our objective is to count, for $n \geq 0$, the number b_n of rooted ordered binary trees on n vertices. Assuming that we know the values of b_i for $0 \leq i \leq n$, in order to obtain b_{n+1} we select one vertex as the root and note, as in Fig. 10.19, that the substructures descending on the left and right of the root are smaller (rooted ordered binary) trees whose total number of vertices is n. These smaller trees are called *subtrees* of the given tree. Among these possible subtrees is the empty subtree, of which there is only 1 $(= b_0)$.

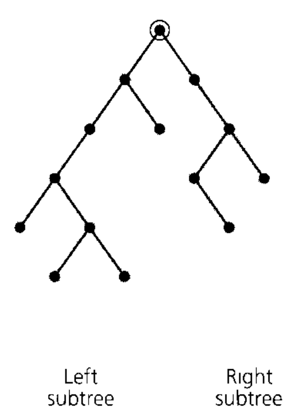


Figure 10.19

Now consider how the n vertices in these two subtrees can be divided up.

(1) 0 vertices on the left, n vertices on the right. This results in $b_0 b_n$ overall substructures to be counted in b_{n+1}.

(2) 1 vertex on the left, $n-1$ vertices on the right, yielding $b_1 b_{n-1}$ rooted ordered binary trees on $n+1$ vertices.

. . . . . .

$(i+1)$ i vertices on the left, $n-i$ on the right, for a count of $b_i b_{n-i}$ toward b_{n+1}.

. . . . . .

$(n+1)$ n vertices on the left and none on the right, contributing $b_n b_0$ of the trees.

Hence, for all $n \geq 0$,

$$b_{n+1} = b_0 b_n + b_1 b_{n-1} + b_2 b_{n-2} + \cdots + b_{n-1} b_1 + b_n b_0,$$

and

$$\sum_{n=0}^{\infty} b_{n+1} x^{n+1} = \sum_{n=0}^{\infty} (b_0 b_n + b_1 b_{n-1} + \cdots + b_{n-1} b_1 + b_n b_0) x^{n+1}. \qquad (1)$$

Now let $f(x) = \sum_{n=0}^{\infty} b_n x^n$ be the generating function for $b_0, b_1, b_2, \ldots$. We rewrite Eq. (1) as

$$(f(x) - b_0) = x \sum_{n=0}^{\infty} (b_0 b_n + b_1 b_{n-1} + \cdots + b_n b_0) x^n = x[f(x)]^2.$$

This brings us to the quadratic [in $f(x)$]

$$x[f(x)]^2 - f(x) + 1 = 0, \quad \text{so} \quad f(x) = [1 \pm \sqrt{1-4x}]/(2x).$$

But $\sqrt{1-4x} = (1-4x)^{1/2} = \binom{1/2}{0} + \binom{1/2}{1}(-4x) + \binom{1/2}{2}(-4x)^2 + \cdots$, where the coefficient of x^n, $n \geq 1$, is

$$\begin{aligned}
\binom{1/2}{n}(-4)^n &= \frac{(1/2)((1/2)-1)((1/2)-2)\cdots((1/2)-n+1)}{n!}(-4)^n \\
&= (-1)^{n-1}\frac{(1/2)(1/2)(3/2)\cdots((2n-3)/2)}{n!}(-4)^n \\
&= \frac{(-1)2^n(1)(3)\cdots(2n-3)}{n!} \\
&= \frac{(-1)2^n(n!)(1)(3)\cdots(2n-3)(2n-1)}{(n!)(n!)(2n-1)} \\
&= \frac{(-1)(2)(4)\cdots(2n)(1)(3)\cdots(2n-1)}{(2n-1)(n!)(n!)} = \frac{(-1)}{(2n-1)}\binom{2n}{n}.
\end{aligned}$$

In $f(x)$ we select the negative radical; otherwise, we would have negative values for the b_n's. Then

$$f(x) = \frac{1}{2x}\left[1 - \left[1 - \sum_{n=1}^{\infty} \frac{1}{(2n-1)}\binom{2n}{n} x^n\right]\right],$$

and b_n, the coefficient of x^n in $f(x)$, is half the coefficient of x^{n+1} in

$$\sum_{n=1}^{\infty} \frac{1}{(2n-1)}\binom{2n}{n} x^n.$$

So

$$b_n = \frac{1}{2}\left[\frac{1}{2(n+1)-1}\right]\binom{2(n+1)}{n+1} = \frac{(2n)!}{(n+1)!(n!)} = \frac{1}{(n+1)}\binom{2n}{n}.$$

The numbers b_n are called the *Catalan numbers* — the same sequence of numbers we encountered in Section 1.5. As we mentioned earlier (following Example 1.42), these numbers are named after the Belgian mathematician Eugène Charles Catalan (1814–1894), who used them in determining the number of ways to parenthesize the expression $x_1x_2x_3 \cdots x_n$. The first nine Catalan numbers are $b_0 = 1, b_1 = 1, b_2 = 2, b_3 = 5, b_4 = 14, b_5 = 42, b_6 = 132, b_7 = 429$, and $b_8 = 1430$.

We continue now with a second application of the Catalan numbers. This is based on an example given by Shimon Even. (See page 86 of reference [6].)

EXAMPLE 10.43

An important data structure that arises in computer science is the *stack*. This structure allows the storage of data items according to the following restrictions.

1) All insertions take place at *one* end of the structure. This is called the *top* of the stack, and the insertion process is referred to as the *push* procedure.
2) All deletions from the (nonempty) stack also take place from the top. We call the deletion process the *pop* procedure.

Since the *last* item inserted *in* this structure is the *first* item that can then be popped *out* of it, the stack is often referred to as a "last-in-first-out" (LIFO) structure.

Intuitive models for this data structure include a pile of poker chips on a table, a stack of trays in a cafeteria, and the discard pile used in playing certain card games. In all three of these cases, we can only (1) insert a new entry at the top of the pile or stack or (2) take (delete) the entry at the top of the (nonempty) pile or stack.

Here we shall use this data structure, with its push and pop procedures, to help us permute the (ordered) list $1, 2, 3, \ldots, n$, for $n \in \mathbf{Z}^+$. The diagram in Fig. 10.20 shows how each integer of the input $1, 2, 3, \ldots, n$ must be pushed onto the top of the stack in the order given. However, we may pop an entry from the top of the (nonempty) stack at any time. But once an entry is popped from the stack, it may not be returned to either the top of the stack or the input left to be pushed onto the stack. The process continues until no entry is left in the stack. Thus the ordered sequence of elements popped from the stack determines a permutation of $1, 2, 3, \ldots, n$.

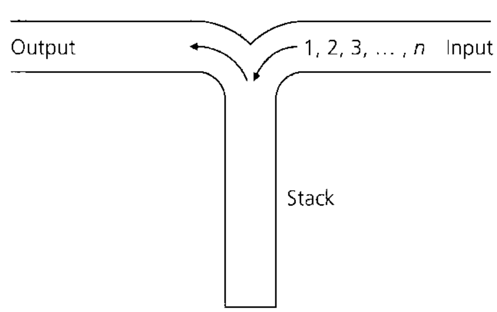


Figure 10.20

If $n = 1$, our input list consists of only the integer 1. We insert 1 at the top of the (empty) stack and then pop it out. This results in the permutation 1.

For $n = 2$, there are two permutations possible for 1, 2, and we can get both of them using the stack.

1) To get 1, 2 we place 1 at the top of the (empty) stack and then pop it. Then 2 is placed at the top of the (empty) stack and it is popped.
2) The permutation 2, 1 is obtained when 1 is placed at the top of the (empty) stack and 2 is then pushed onto the top of this (nonempty) stack. Upon popping first 2 from the top of the stack, and then 1, we obtain 2, 1.

Turning to the case where $n = 3$, we find that we can obtain only five of the $3! = 6$ possible permutations of 1, 2, 3 in this situation. For example, the permutation 2, 3, 1 results when we take the following steps.

- Place 1 at the top of the (empty) stack.
- Push 2 onto the top of the stack (on top of 1).
- Pop 2 from the stack.
- Push 3 onto the top of the stack (on top of 1).
- Pop 3 from the stack.
- Pop 1 from the stack, leaving it empty.

The reason we fail to obtain all six permutations of 1, 2, 3 is that we cannot generate the permutation 3, 1, 2 using the stack. For in order to have 3 in the first position of the permutation, we must build the stack by first pushing 1 onto the (empty) stack, then pushing 2 onto the top of the stack (on top of 1), and finally pushing 3 onto the stack (on top of 2). After 3 is popped from the top of the stack, we get 3 as the first number in our permutation. But with 2 now at the top of the stack, we cannot pop 1 until after 2 has been popped, so the permutation 3, 1, 2 cannot be generated.

When $n = 4$, there are 14 permutations of the (ordered) list 1, 2, 3, 4 that can be generated by the stack method. We list them in the four columns of Table 10.4 according to the location of 1 in the permutation.

Table 10.4

1, 2, 3, 4	2, 1, 3, 4	2, 3, 1, 4	2, 3, 4, 1
1, 2, 4, 3	2, 1, 4, 3	3, 2, 1, 4	2, 4, 3, 1
1, 3, 2, 4			3, 2, 4, 1
1, 3, 4, 2			3, 4, 2, 1
1, 4, 3, 2			4, 3, 2, 1

1) There are five permutations with 1 in the first position, because after 1 is pushed onto and popped from the stack, there are five ways to permute 2, 3, 4 using the stack.
2) When 1 is in the second position, 2 must be in the first position. This is because we pushed 1 onto the (empty) stack, then pushed 2 on top of it and then popped 2 and then 1. There are two permutations in column 2, because 3, 4 can be permuted in two ways on the stack.

3) For column 3 we have 1 in position three. We note that the only numbers that can precede it are 2 and 3, which can be permuted on the stack (with 1 on the bottom) in two ways. Then 1 is popped, and we push 4 onto the (empty) stack and then pop it.

4) In the last column we obtain five permutations: After we push 1 onto the top of the (empty) stack, there are five ways to permute 2, 3, 4 using the stack (with 1 on the bottom). Then 1 is popped from the stack to complete the permutation.

On the basis of these observations, for $1 \leq i \leq 4$, let a_i count the number of ways to permute the integers $1, 2, 3, \ldots, i$ (or any list of i consecutive integers) using the stack. Also, we define $a_0 = 1$ since there is only one way to permute nothing, using the stack. Then

$$a_4 = a_0a_3 + a_1a_2 + a_2a_1 + a_3a_0,$$

where

a) Each summand a_ja_k satisfies $j + k = 3$.

b) The subscript j tells us that there are j integers to the left of 1 in the permutation — in particular, for $j \geq 1$, these are the integers from 2 to $j + 1$, inclusive.

c) The subscript k indicates that there are k integers to the right of 1 in the permutation — for $k \geq 1$, these are the integers from $4 - (k - 1)$ to 4.

This permutation problem can now be generalized to any $n \in \mathbf{N}$, so that

$$a_{n+1} = a_0a_n + a_1a_{n-1} + a_2a_{n-2} + \cdots + a_{n-1}a_1 + a_na_0,$$

with $a_0 = 1$. From the result in Example 10.42 we know that

$$a_n = \frac{1}{(n+1)}\binom{2n}{n}.$$

Now let us make one final observation about the permutations in Table 10.4. Consider, for example, the permutation 3, 2, 4, 1. How did this permutation come about? First 1 is pushed onto the empty stack. This is then followed by pushing 2 on top of 1 and then pushing 3 on top of 2. Now 3 is popped from the top of the stack, leaving 2 and 1; then 2 is popped from the top of the stack, leaving just 1. At this point 4 is pushed on top of 1 and then popped, leaving 1 on the stack. Finally, 1 is popped from the (top of the) stack, leaving the stack empty. So the permutation 3, 2, 4, 1 comes about from the following sequence of four pushes and four pops:

push, push, push, pop, pop, push, pop, pop.

Now replace each "push" with a "1" and each "pop" with a "0". The result is the sequence

1 1 1 0 0 1 0 0.

Similarly, the permutation 1, 3, 4, 2 is determined by the sequence

push, pop, push, push, pop, push, pop, pop

and this corresponds with the sequence

1 0 1 1 0 1 0 0.

In fact, each permutation in Table 10.4 gives rise to a sequence of four 1's and four 0's. But there are $8!/(4!\,4!) = 70$ ways to list four 1's and four 0's. Do these 14 sequences have some special property? Yes! As we go from left to right in each of these sequences, the number of 1's (pushes) is never exceeded by the number of 0's (pops) [just like in part (b) of Example 1.43 — another situation counted by the Catalan numbers].

Our last example for this section is comparable to Example 10.17. Once again we see that we must guard against trying to obtain a general result without a general argument — no matter what a few special cases might suggest.

EXAMPLE 10.44

Here we start with n *distinct objects* and, for $n \geq 1$, we distribute them among at most n *identical containers*, but we do not allow more than three objects in any container, and we are not concerned about how the objects are arranged within any one container. We let a_n count the number of these distributions, and from Fig. 10.21 we see that

$$a_0 = 1, \qquad a_1 = 1, \qquad a_2 = 2, \qquad a_3 = 5, \quad \text{and} \quad a_4 = 14.$$

It appears that we might have the first five terms in the sequence of Catalan numbers. Unfortunately, the pattern breaks down and we find, for example, that

$$a_5 = 46 \neq 42 \text{ (the sixth Catalan number)} \quad \text{and}$$
$$a_6 = 166 \neq 132 \text{ (the seventh Catalan number).}$$

(The distributions in this example were studied by F. L. Miksa, L. Moser, and M. Wyman in reference [22].)

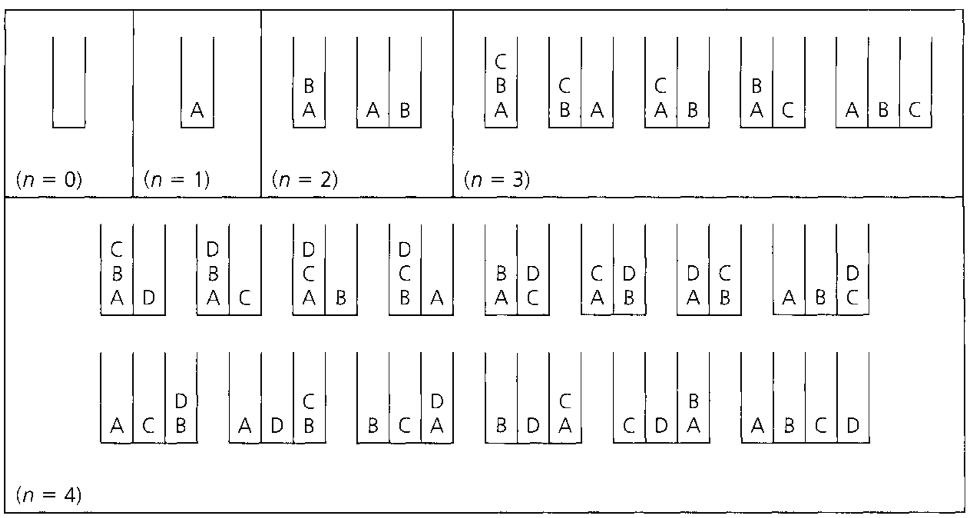


Figure 10.21

Other examples that involve the Catalan numbers can be found in the chapter references.

EXERCISES 10.5

1. For the rooted ordered binary trees of Example 10.42, calculate b_4 and draw all of these four-vertex structures.

2. Verify that for all $n \geq 0$,

$$\frac{1}{2}\left(\frac{1}{2n+1}\right)\binom{2n+2}{n+1} = \left(\frac{1}{n+1}\right)\binom{2n}{n}.$$

3. Show that for all $n \geq 2$,

$$\binom{2n-1}{n} - \binom{2n-1}{n-2} = \frac{1}{(n+1)}\binom{2n}{n}.$$

4. Which of the following permutations of 1, 2, 3, 4, 5, 6, 7, 8 can be obtained using the stack (of Example 10.43)?

a) 4, 2, 3, 1, 5, 6, 7, 8 **b)** 5, 4, 3, 6, 2, 1, 8, 7
c) 4, 5, 3, 2, 1, 8, 6, 7 **d)** 3, 4, 2, 1, 7, 6, 8, 5

5. Suppose that the integers 1, 2, 3, 4, 5, 6, 7, 8 are permuted using the stack (of Example 10.43). (a) How many permutations are possible? (b) How many permutations have 1 in position 4 and 5 in position 8? (c) How many permutations have 1 in position 6? (d) How many permutations start with 321?

6. This exercise deals with a problem that was first proposed by Leonard Euler. The problem examines a given convex polygon of n (≥ 3) sides — that is, a polygon of n sides that satisfies the property: For all points P_1, P_2 within the interior of the polygon, the line segment joining P_1 and P_2 also lies within the interior of the polygon. Given a convex polygon of n sides, Euler wanted to count the number of ways the interior of the polygon could be triangulated (subdivided into triangles) by drawing diagonals that do not intersect.

For a convex polygon of $n \geq 3$ sides, let t_n count the number of ways the interior of the polygon can be triangulated by drawing nonintersecting diagonals.

a) Define $t_2 = 1$ and verify that

$$t_{n+1} = t_2 t_n + t_3 t_{n-1} + \cdots + t_{n-1} t_3 + t_n t_2.$$

b) Express t_n as a function of n.

7. In Fig. 10.22 we have two of the five ways in which we can triangulate the interior of a convex pentagon with no intersecting diagonals. Here we have labeled four of the sides — with the letters a, b, c, d — as well as the five vertices. In part (i) we use the labels on sides a and b to give us the label ab on the diagonal connecting vertices 2 and 4. This is because this diagonal (labeled ab), together with the sides a and b, provides us with one of the interior triangles for this triangulation of the convex pentagon. Then the diagonal ab and the side c give rise to the label $(ab)c$ on the diagonal determined by vertices 2 and 5 — and the sides labeled ab, c and $(ab)c$ provide a second interior triangle for this triangulation. Continuing in this way, we label the base edge connecting vertices 1 and 2 with $((ab)c)d$ — one of the five ways we can introduce parentheses in order to obtain the three products (of two numbers at a time) needed to compute $abcd$. The triangulation in part (ii) of the figure corresponds with the parenthesized product $(ab)(cd)$.

a) Determine the parenthesized product involving a, b, c, d for the other three triangulations of the convex pentagon.

b) Find the parenthesized product for each of the triangulated convex hexagons in parts (iii) and (iv) of Fig. 10.22

[From part (a) we learn that there are five ways to parenthesize the expression $abcd$ (and five ways to triangulate a convex pentagon). Part (b) shows us two of the 14 ways one can introduce parentheses for the expression $abcde$ (and triangulate a convex hexagon). In general, there are $\frac{1}{n+1}\binom{2n}{n}$ ways to parenthesize the expression $x_1 x_2 x_3 \cdots x_{n-1} x_n x_{n+1}$. It was in solving this problem that Eugène Charles Catalan discovered the sequence that now bears his name.]

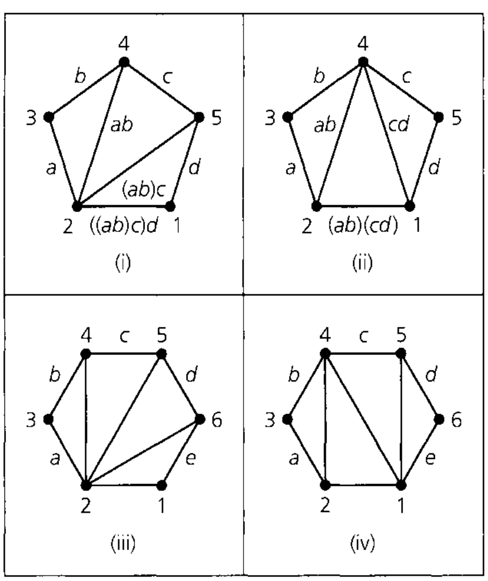


Figure 10.22

8. For $n \geq 0$,

$$b_n = \left(\frac{1}{n+1}\right)\binom{2n}{n}$$

is the nth Catalan number.

a) Show that for all $n \geq 0$,

$$b_{n+1} = \frac{2(2n+1)}{(n+2)} b_n.$$

b) Use the result of part (a) to write a computer program (or develop an algorithm) that calculates the first 15 Catalan numbers.

9. For $n \geq 0$, evenly distribute $2n$ points on the circumference of a circle, and label these points cyclically with the integers $1, 2, 3, \ldots, 2n$. Let a_n be the number of ways in which these $2n$ points can be paired off as n chords where no two chords intersect. (The case for $n = 3$ is shown in Fig. 10.23.) Find and solve a recurrence relation for a_n, $n \geq 0$.

10. For $n \in \mathbf{N}$, consider all paths from $(0, 0)$ to $(2n, 0)$ using the moves N: $(x, y) \to (x+1, y+1)$ and S: $(x, y) \to (x+1, y-1)$, where any such path can never fall below the x-axis. The five paths (generally called *mountain ranges*) for $n = 3$ are shown in Fig. 10.24. How many mountain ranges are there for each $n \in \mathbf{N}$? (Verify your claim!)

11. For $n \in \mathbf{Z}^+$, let $f: \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$, where f is monotone increasing [that is, $1 \leq i < j \leq n \Rightarrow f(i) \leq f(j)$] and $f(i) \geq i$ for all $1 \leq i \leq n$. (a) Determine the five monotone increasing functions $f: \{1, 2, 3\} \to \{1, 2, 3\}$, where $f(i) \geq i$ for all $1 \leq i \leq 3$. (b) Use the graphs of the functions from part (a) to set up a one-to-one correspondence with the paths from $(0, 0)$ to $(3, 3)$ using the moves R: $(x, y) \to (x+1, y)$, U: $(x, y) \to (x, y+1)$, where each such path never falls below the line $y = x$. (The reader may

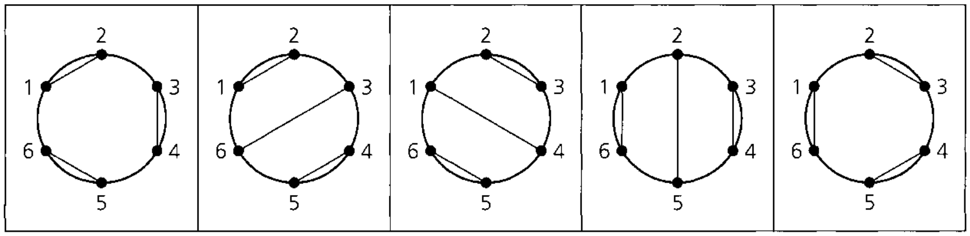


Figure 10.23

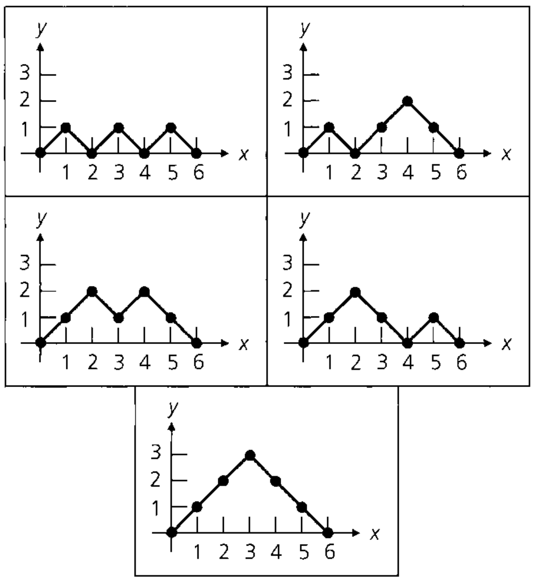


Figure 10.24

wish to check Exercise 3 for Section 1.5.) (c) If the paths in part (b) are rotated clockwise through 45°, what results do we find? (d) How many monotone increasing functions f have domain and codomain equal to $\{1, 2, 3, \ldots, n\}$, for $n \in \mathbf{Z}^+$, and satisfy $f(i) \geq i$ for all $1 \leq i \leq n$?

12. For $n \in \mathbf{Z}^+$, let $g: \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$, where $g(i) \leq i$ for all $1 \leq i \leq n$. (a) Determine the five functions $g: \{1, 2, 3\} \to \{1, 2, 3\}$ where $g(i) \leq i$ for all $1 \leq i \leq 3$. (b) Set up a one-to-one correspondence between the functions in part (a) here and those in part (a) of the previous exercise. [You want a one-to-one correspondence that will generalize when you examine the functions $f, g: \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$, $n \in \mathbf{Z}^+$, where $f(i) \geq i$ and $g(i) \leq i$ for all $1 \leq i \leq n$.] (c) How many functions g have domain and codomain equal to $\{1, 2, 3, \ldots, n\}$, for $n \in \mathbf{Z}^+$, and satisfy $g(i) \leq i$ for all $1 \leq i \leq n$?

13. For $n \in \mathbf{N}$, consider the arrangements of pennies built on a contiguous row of n pennies. Each penny that is not in the bottom row (of n pennies) rests upon the two pennies below it, and there is no concern about whether heads or tails appears. The situation for $n = 3$ is shown in Fig. 10.25. How many such arrangements are there for a contiguous row of n pennies, $n \in \mathbf{N}$?

14. For $n \in \mathbf{N}$, let s_n count the number of ways one can travel from (0, 0) to (n, n) using the moves R: $(x, y) \to (x + 1, y)$, U: $(x, y) \to (x, y + 1)$, D: $(x, y) \to (x + 1, y + 1)$, where the path can never rise above the line $y = x$. (a) Determine s_2. (b) How is s_2 related to the Catalan numbers b_0, b_1, b_2? (c) How is s_3 related to b_0, b_1, b_2, b_3? What is s_3? (d) For $n \in \mathbf{N}$, how is s_n related to $b_0, b_1, b_2, \ldots, b_n$? (The numbers $s_0, s_1, s_2, \ldots$ are known as the *Schröder numbers*.)

15. A one-to-one function $f: \{1, 2, 3, \ldots, n\} \to \{1, 2, 3, \ldots, n\}$ is often called a *permutation*. Such a permutation is termed a *rise/fall* permutation when $f(1) < f(2)$, $f(2) > f(3)$, $f(3) < f(4)$, ... For example, if $n = 4$ the five permutations 1324 (where $f(1) = 1$, $f(2) = 3$, $f(3) = 2$, $f(4) = 4$), 1423, 2314, 2413, and 3412 are the rise/fall permutations (for 1, 2, 3, 4). This we denote by writing $E_4 = 5$, where, in general, E_n counts the number of rise/fall permutations for $1, 2, 3, \ldots, n$. The numbers $E_0, E_1, E_2, E_3, \ldots$ are called the *Euler numbers* (not to be confused with the Eulerian numbers in Example 4.21). We define $E_0 = 1$ and find that $E_1 = 1$, $E_2 = 1$.

a) Find the rise/fall permutations for 1, 2, 3. What is E_3?

b) Find the rise/fall permutations for 1, 2, 3, 4, 5. What is E_5?

c) Explain why in each rise/fall permutation of $1, 2, 3, \ldots n$, we find n at position $2i$ for some $1 \leq i \leq \lfloor n/2 \rfloor$, if $n > 1$.

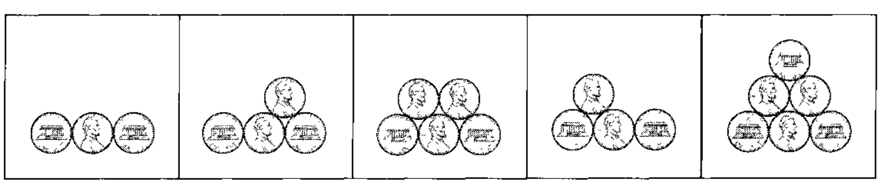


Figure 10.25

d) For $n \geq 2$, show that

$$E_n = \sum_{i=1}^{\lfloor n/2 \rfloor} \binom{n-1}{2i-1} E_{2i-1} E_{n-2i}, \quad E_0 = E_1 = 1.$$

e) Where do we find 1 in a rise/fall permutation of $1, 2, 3, \ldots, n$?

f) For $n \geq 1$, show that

$$E_n = \sum_{i=0}^{\lfloor (n-1)/2 \rfloor} \binom{n-1}{2i} E_{2i} E_{n-2i-1}, \quad E_0 = 1.$$

g) Prove that for $n \geq 2$,

$$E_n = \left(\frac{1}{2}\right) \sum_{i=0}^{n-1} \binom{n-1}{i} E_i E_{n-i-1}, \quad E_0 = E_1 = 1.$$

h) Use the result in part (g) to find E_6 and E_7.

i) Find the Maclaurin series expansion for $f(x) = \sec x + \tan x$. Conjecture (no proof required) the sequence for which this is the exponential generating function.

10.6 Divide-and-Conquer Algorithms (Optional)†

One of the most important and widely applicable types of efficient algorithms is based on a *divide-and-conquer* approach. Here the strategy, in general, is to solve a given problem of size n ($n \in \mathbf{Z}^+$) by

1) Solving the problem for a small value of n directly (this provides an initial condition for the resulting recurrence relation).
2) Breaking the general problem of size n into a smaller problems of the same type and (approximately) the same size — either $\lceil n/b \rceil$ or $\lfloor n/b \rfloor$,‡ where $a, b \in \mathbf{Z}^+$ with $1 \leq a < n$ and $1 < b < n$.

Then we solve the a smaller problems and use their solutions to construct a solution for the original problem of size n. We shall be especially interested in cases where n is a power of b, and $b = 2$.

We shall study those divide-and-conquer algorithms where

1) The time to solve the initial problem of size $n = 1$ is a constant $c \geq 0$, and
2) The time to break the given problem of size n into a smaller (similar) problems, together with the time to combine the solutions of these smaller problems to get a solution for the given problem, is $h(n)$, a function of n.

Our concern here will actually be with the time-complexity function $f(n)$ for these algorithms. Consequently, we shall use the notation $f(n)$ here, instead of the subscripted notation a_n that we used in the earlier sections of this chapter.

The conditions that have now been stated lead to the following recurrence relation.

$$f(1) = c,$$
$$f(n) = af(n/b) + h(n), \qquad \text{for } n = b^k, \qquad k \geq 1.$$

We note that the domain of f is $\{1, b, b^2, b^3, \ldots\} = \{b^i | i \in \mathbf{N}\} \subset \mathbf{Z}^+$.

†The material in this section may be skipped with no loss of continuity. It will be used in Section 12.3 to determine the time-complexity function for the merge sort algorithm. However, the result there will also be obtained for a special case of the merge sort by another method that does not use the material developed in this section.

‡For each $x \in \mathbf{R}$, recall that $\lceil x \rceil$ denotes the *ceiling* of x and $\lfloor x \rfloor$ the *floor* of x, or *greatest integer* in x, where

a) $\lfloor x \rfloor = \lceil x \rceil = x$, for $x \in \mathbf{Z}$.
b) $\lfloor x \rfloor$ = the integer directly to the left of x, for $x \in \mathbf{R} - \mathbf{Z}$.
c) $\lceil x \rceil$ = the integer directly to the right of x, for $x \in \mathbf{R} - \mathbf{Z}$.

In our first result, the solution of this recurrence relation is derived for the case where $h(n)$ is the constant c.

THEOREM 10.1 Let $a, b, c \in \mathbf{Z}^+$ with $b \geq 2$, and let $f: \mathbf{Z}^+ \to \mathbf{R}$. If

$$f(1) = c, \quad \text{and}$$
$$f(n) = af(n/b) + c, \qquad \text{for } n = b^k, \qquad k \geq 1,$$

then for all $n = 1, b, b^2, b^3, \ldots,$

1) $f(n) = c(\log_b n + 1)$, when $a = 1$, and

2) $f(n) = \dfrac{c(an^{\log_b a} - 1)}{a - 1}$, when $a \geq 2$.

Proof: For $k \geq 1$ and $n = b^k$, we write the following system of k equations. [Starting with the second equation, we obtain each of these equations from its immediate predecessor by (i) replacing each occurrence of n in the prior equation by n/b and (ii) multiplying the resulting equation in (i) by a.]

$$\begin{aligned} f(n) &= af(n/b) + c \\ af(n/b) &= a^2 f(n/b^2) + ac \\ a^2 f(n/b^2) &= a^3 f(n/b^3) + a^2 c \\ \vdots \quad & \quad \vdots \qquad\qquad \vdots \\ a^{k-2} f(n/b^{k-2}) &= a^{k-1} f(n/b^{k-1}) + a^{k-2} c \\ a^{k-1} f(n/b^{k-1}) &= a^k f(n/b^k) + a^{k-1} c \end{aligned}$$

We see that each of the terms $af(n/b), a^2 f(n/b^2), \ldots, a^{k-1} f(n/b^{k-1})$ occurs one time as a summand on both the left-hand and right-hand sides of these equations. Therefore, upon adding both sides of the k equations and canceling these common summands, we obtain

$$f(n) = a^k f(n/b^k) + [c + ac + a^2 c + \cdots + a^{k-1} c].$$

Since $n = b^k$ and $f(1) = c$, we have

$$\begin{aligned} f(n) &= a^k f(1) + c[1 + a + a^2 + \cdots + a^{k-1}] \\ &= c[1 + a + a^2 + \cdots + a^{k-1} + a^k]. \end{aligned}$$

1) If $a = 1$, then $f(n) = c(k + 1)$. But $n = b^k \iff \log_b n = k$, so $f(n) = c(\log_b n + 1)$, for $n \in \{b^i | i \in \mathbf{N}\}$.

2) When $a \geq 2$, then $f(n) = \dfrac{c(1 - a^{k+1})}{1 - a} = \dfrac{c(a^{k+1} - 1)}{a - 1}$, from identity 4 of Table 9.2. Now $n = b^k \iff \log_b n = k$, so

$$a^k = a^{\log_b n} = (b^{\log_b a})^{\log_b n} = (b^{\log_b n})^{\log_b a} = n^{\log_b a},$$

and

$$f(n) = \frac{c(an^{\log_b a} - 1)}{(a - 1)}, \qquad \text{for } n \in \{b^i | i \in \mathbf{N}\}.$$

EXAMPLE 10.45

a) Let $f: \mathbf{Z}^+ \to \mathbf{R}$, where

$$f(1) = 3, \quad \text{and}$$
$$f(n) = f(n/2) + 3, \qquad \text{for } n = 2^k, \qquad k \in \mathbf{Z}^+.$$

So by part (1) of Theorem 10.1, with $c = 3$, $b = 2$, and $a = 1$, it follows that $f(n) = 3(\log_2 n + 1)$ for $n \in \{1, 2, 4, 8, 16, \ldots\}$.

b) Suppose that $g: \mathbf{Z}^+ \to \mathbf{R}$ with

$$g(1) = 7, \quad \text{and}$$
$$g(n) = 4g(n/3) + 7, \qquad \text{for } n = 3^k, \qquad k \in \mathbf{Z}^+.$$

Then with $c = 7$, $b = 3$, and $a = 4$, part (2) of Theorem 10.1 implies that $g(n) = (7/3)(4n^{\log_3 4} - 1)$, when $n \in \{1, 3, 9, 27, 81, \ldots\}$.

c) Finally, consider $h: \mathbf{Z}^+ \to \mathbf{R}$, where

$$h(1) = 5, \quad \text{and}$$
$$h(n) = 7h(n/7) + 5, \qquad \text{for } n = 7^k, \qquad k \in \mathbf{Z}^+.$$

Once again we use part (2) of Theorem 10.1, this time with $a = b = 7$ and $c = 5$. Here we learn that $h(n) = (5/6)(7n^{\log_7 7} - 1) = (5/6)(7n - 1)$ for $n \in \{1, 7, 49, 343, \ldots\}$.

Considering Theorem 10.1, we must unfortunately realize that although we know about f for $n \in \{1, b, b^2, \ldots\}$, we cannot say anything about the value of f for the integers in $\mathbf{Z}^+ - \{1, b, b^2, \ldots\}$. So at this time we are unable to deal with the concept of f as a time-complexity function. To overcome this, we now generalize Definition 5.23, wherein the idea of function dominance was first introduced.

Definition 10.1

Let $f, g: \mathbf{Z}^+ \to \mathbf{R}$ with S an infinite subset of $\mathbf{Z}^+$. We say that *g dominates f on S* (or *f is dominated by g on S*) if there exist constants $m \in \mathbf{R}^+$ and $k \in \mathbf{Z}^+$ such that $|f(n)| \leq m|g(n)|$ for all $n \in S$, where $n \geq k$.

Under these conditions we also say that $f \in O(g)$ *on S*.

EXAMPLE 10.46

Let $f: \mathbf{Z}^+ \to \mathbf{R}$ be defined so that

$$f(n) = n, \qquad \text{for } n \in \{1, 3, 5, 7, \ldots\} = S_1,$$
$$f(n) = n^2, \qquad \text{for } n \in \{2, 4, 6, 8, \ldots\} = S_2.$$

Then $f \in O(n)$ on S_1 and $f \in O(n^2)$ on S_2. However, we *cannot* conclude that $f \in O(n)$.

EXAMPLE 10.47

From Example 10.45, it now follows from Definition 10.1 that

a) $f \in O(\log_2 n)$ on $\{2^k | k \in \mathbf{N}\}$ **b)** $g \in O(n^{\log_3 4})$ on $\{3^k | k \in \mathbf{N}\}$

c) $h \in O(n)$ on $\{7^k | k \in \mathbf{N}\}$.

Using Definition 10.1, we now consider the following corollaries for Theorem 10.1. The first is a generalization of the first two results given in Example 10.47.

COROLLARY 10.1 Let $a, b, c \in \mathbf{Z}^+$ with $b \geq 2$, and let $f: \mathbf{Z}^+ \to \mathbf{R}$. If

$$f(1) = c, \quad \text{and}$$
$$f(n) = af(n/b) + c, \qquad \text{for } n = b^k, \qquad k \geq 1,$$

then

1) $f \in O(\log_b n)$ on $\{b^k | k \in \mathbf{N}\}$, when $a = 1$, and
2) $f \in O(n^{\log_b a})$ on $\{b^k | k \in \mathbf{N}\}$, when $a \geq 2$.

Proof: This proof is left as an exercise for the reader.

This second corollary changes the equal signs of Theorem 10.1 to inequalities. As a result, the codomain of f must be restricted from $\mathbf{R}$ to $\mathbf{R}^+ \cup \{0\}$.

COROLLARY 10.2 For $a, b, c \in \mathbf{Z}^+$ with $b \geq 2$, let $f: \mathbf{Z}^+ \to \mathbf{R}^+ \cup \{0\}$. If

$$f(1) \leq c, \quad \text{and}$$
$$f(n) \leq af(n/b) + c, \qquad \text{for } n = b^k, \qquad k \geq 1,$$

then for all $n = 1, b, b^2, b^3, \ldots,$

1) $f \in O(\log_b n)$, when $a = 1$, and
2) $f \in O(n^{\log_b a})$, when $a \geq 2$.

Proof: Consider the function $g: \mathbf{Z}^+ \to \mathbf{R}^+ \cup \{0\}$, where

$$g(1) = c, \quad \text{and}$$
$$g(n) = ag(n/b) + c, \qquad \text{for } n \in \{1, b, b^2, \ldots\}.$$

By Corollary 10.1,

$$g \in O(\log_b n) \quad \text{on} \quad \{b^k | k \in \mathbf{N}\}, \quad \text{when } a = 1, \quad \text{and}$$
$$g \in O(n^{\log_b a}) \quad \text{on} \quad \{b^k | k \in \mathbf{N}\}, \quad \text{when } a \geq 2.$$

We claim that $f(n) \leq g(n)$ for all $n \in \{1, b, b^2, \ldots\}$. To prove our claim, we induct on k where $n = b^k$. If $k = 0$, then $n = b^0 = 1$ and $f(1) \leq c = g(1)$ — so the result is true for this first case. Assuming the result is true for some $t \in \mathbf{N}$, we have $f(n) = f(b^t) \leq g(b^t) = g(n)$, for $n = b^t$. Then for $k = t + 1$ and $n = b^k = b^{t+1}$, we find that

$$f(n) = f(b^{t+1}) \leq af(b^{t+1}/b) + c = af(b^t) + c \leq ag(b^t) + c = g(b^{t+1}) = g(n).$$

Therefore, it follows by the Principle of Mathematical Induction that $f(n) \leq g(n)$ for all $n \in \{1, b, b^2, \ldots\}$. Consequently, $f \in O(g)$ on $\{b^k | k \in \mathbf{N}\}$, and the corollary follows because of our earlier statement about g.

Up to this point, our study of divide-and-conquer algorithms has been predominantly theoretical. It is high time we gave an example in which these ideas can be applied. The following result will confirm one of our earlier examples.

EXAMPLE 10.48

For $n = 1, 2, 4, 8, 16, \ldots$, let $f(n)$ count the number of comparisons needed to find the maximum and minimum elements in a set $S \subset \mathbf{R}$, where $|S| = n$ and the procedure in Example 10.30 is used.

If $n = 1$, then the maximum and minimum elements are the same element. Therefore, no comparisons are necessary and $f(1) = 0$.

If $n > 1$, then $n = 2^k$ for some $k \in \mathbf{Z}^+$, and we partition S as $S_1 \cup S_2$ where $|S_1| = |S_2| = n/2 = 2^{k-1}$. It takes $f(n/2)$ comparisons to find the maximum M_i and the minimum m_i for each set S_i, $i = 1, 2$. For $n \geq 4$, knowing m_1, M_1, m_2, and M_2, we then compare m_1 with m_2 and M_1 and M_2 to determine the minimum and maximum elements in S. Therefore,

$$f(n) = 2f(n/2) + 1, \qquad \text{when } n = 2, \quad \text{and}$$
$$f(n) = 2f(n/2) + 2, \qquad \text{when } n = 4, 8, 16, \ldots.$$

Unfortunately, these results do not provide the hypotheses of Theorem 10.1. However, if we change our equations into the inequalities

$$f(1) \leq 2$$
$$f(n) \leq 2f(n/2) + 2, \qquad \text{for } n = 2^k, \qquad k \geq 1,$$

then by Corollary 10.2 the time-complexity function $f(n)$, measured by the number of comparisons made in this recursive procedure, satisfies $f \in O(n^{\log_2 2}) = O(n)$, for all $n = 1, 2, 4, 8, \ldots$.

We can examine the relationship between this example and Example 10.30 even further. From that earlier result, we know that if $|S| = n = 2^k, k \geq 1$, then the number of comparisons $f(n)$ we need (in the given procedure) to find the maximum and minimum elements in S is $(3/2)(2^k) - 2$. (*Note*: Our statement here replaces the variable n of Example 10.30 by the variable k.)

Since $n = 2^k$, we find that we can now write

$$f(1) = 0$$
$$f(n) = f(2^k) = (3/2)(2^k) - 2 = (3/2)n - 2, \qquad \text{for } n = 2, 4, 8, 16, \ldots.$$

Hence $f \in O(n)$ for $n \in \{2^k | k \in \mathbf{N}\}$, just as we obtained above using Corollary 10.2.

All of our results have required that $n = b^k$, for some $k \in \mathbf{N}$, so it is only natural to ask whether we can do anything in the case where n is allowed to be an arbitrary positive integer. To find out, we introduce the following idea.

Definition 10.2

A function $f: \mathbf{Z}^+ \to \mathbf{R}^+ \cup \{0\}$ is called *monotone increasing* if for all $m, n \in \mathbf{Z}^+, m < n \Rightarrow f(m) \leq f(n)$.

This permits us to consider results for all $n \in \mathbf{Z}^+$ — under certain circumstances.

THEOREM 10.2

Let $f: \mathbf{Z}^+ \to \mathbf{R}^+ \cup \{0\}$ be monotone increasing, and let $g: \mathbf{Z}^+ \to \mathbf{R}$. For $b \in \mathbf{Z}^+, b \geq 2$, suppose that $f \in O(g)$ for all $n \in S = \{b^k | k \in \mathbf{N}\}$. Under these conditions,

a) If $g \in O(\log n)$, then $f \in O(\log n)$.

b) If $g \in O(n \log n)$, then $f \in O(n \log n)$.

c) If $g \in O(n^r)$, then $f \in O(n^r)$, for $r \in \mathbf{R}^+ \cup \{0\}$.

Proof: We shall prove part (a) and leave parts (b) and (c) for the Section Exercises. Before starting, we should note that the base for the logarithms in parts (a) and (b) is any positive real number greater than 1.

Since $f \in O(g)$ on S, and $g \in O(\log n)$, we at least have $f \in O(\log n)$ on S. Therefore, by Definition 10.1, there exist constants $m \in \mathbf{R}^+$ and $s \in \mathbf{Z}^+$ such that $f(n) = |f(n)| \leq m|\log n| = m \log n$ for all $n \in S$, $n \geq s$. We need to find a constant $M \in \mathbf{R}^+$ such that $f(n) \leq M \log n$ for *all* $n \geq s$, not just those $n \in S$.

First let us agree to choose s large enough so that $\log s \geq 1$. Now let $n \in \mathbf{Z}^+$, where $n \geq s$ but $n \notin S$. Then there exists $k \in \mathbf{Z}^+$ such that $s \leq b^k < n < b^{k+1}$. Since f is monotone increasing and positive,

$$\begin{aligned} f(n) \leq f(b^{k+1}) &\leq m \log(b^{k+1}) = m[\log(b^k) + \log b] \\ &= m \log(b^k) + m \log b \\ &< m \log(b^k) + m \log b \log(b^k) \\ &= m(1 + \log b) \log(b^k) \\ &< m(1 + \log b) \log n. \end{aligned}$$

So with $M = m(1 + \log b)$ we find that for all $n \in \mathbf{Z}^+ - S$, if $n \geq s$ then $f(n) < M \log n$. Hence $f(n) \leq M \log n$ for *all* $n \in \mathbf{Z}^+$, where $n \geq s$, and $f \in O(\log n)$.

We shall now use the result of Theorem 10.2 in determining the time-complexity function $f(n)$ for a searching algorithm known as *binary search.*

In Example 5.70 we analyzed an algorithm wherein an array $a_1, a_2, a_3, \ldots, a_n$ of integers was searched for the presence of a particular integer called *key*. At that time the array entries were not given in any particular order, so we simply compared the value of *key* with those of the array elements $a_1, a_2, a_3, \ldots, a_n$. This would not be very efficient, however, if we knew that $a_1 < a_2 < a_3 < \cdots < a_n$. (After all, one does not search a telephone book for the telephone number of a particular person by starting at page 1 and examining every name in succession. The alphabetical ordering of the last names is used to speed up the searching process.) Let us look at a particular example.

EXAMPLE 10.49

Consider the array $a_1, a_2, a_3, \ldots, a_7$ of integers, where $a_1 = 2$, $a_2 = 4$, $a_3 = 5$, $a_4 = 7$, $a_5 = 10$, $a_6 = 17$, and $a_7 = 20$, and let *key* $= 9$. We search this array as follows:

1) Compare *key* with the entry at the center of the array; here it is $a_4 = 7$. Since *key* $> a_4$, we now concentrate on the remaining elements in the subarray a_5, a_6, a_7.
2) Now compare *key* with the center element a_6. Since *key* $= 9 < 17 = a_6$, we now turn to the subarray (of a_5, a_6, a_7) that consists of those elements smaller than a_6. Here this is only the element a_5.
3) Comparing *key* with a_5, we find that *key* $\neq a_5$, so *key* is not present in the given array $a_1, a_2, a_3, \ldots, a_7$.

From the results of Example 10.49, we make the following observations for a general (ordered) array of integers (or real numbers). Let $a_1, a_2, a_3, \ldots, a_n$ denote the given array, and let *key* denote the integer (or real number) for which we are searching. Unlike our array in Example 5.70, here

$$a_1 < a_2 < a_3 < \cdots < a_n.$$

1) First we compare the value of *key* with the array entry at or near the center. This entry is $a_{(n+1)/2}$ for n odd or $a_{n/2}$ for n even.

   Whether n is even or odd, the array element subscripted by $c = \lfloor (n+1)/2 \rfloor$ is the center, or near center, element. Note that at this point 1 is the value of the smallest subscript for the array subscripts, whereas n is the value of the largest subscript.

2) If *key* is a_c, we are finished. If not, then
   a) If *key* exceeds a_c, we search (with this dividing process) the subarray $a_{c+1}, a_{c+2}, \ldots, a_n$.
   b) If *key* is smaller than a_c, then the dividing process is applied in searching the subarray $a_1, a_2, \ldots, a_{c-1}$.

The preceding observations have been used in developing the pseudocode procedure in Fig. 10.26. Here the input is an ordered array $a_1, a_2, a_3, \ldots, a_n$ of integers, or real numbers, in ascending order, the positive integer n (for the number of entries in the given array), and the value of the integer variable *key*. If the array elements are integers (real numbers), then *key* should be an integer (real number). The variables s and l are integer variables used for storing the smallest and largest subscripts for the subscripts of the array or subarray being searched. The integer variable c stores the index for the array (subarray) element at, or near, the center of the array (subarray). In general, $c = \lfloor (s+l)/2 \rfloor$. The integer variable *location* stores the subscript of the array entry where *key* is located; the value of *location* is 0 when *key* is not present in the given array.

```
procedure BinarySearch(n: positive integer; key,a1,a2,a3,...,an: integers)
begin
s := 1   {s is the smallest subscript of the subarray being searched}
l := n   {l is the largest subscript of the subarray being searched}
location := 0
while s ≤ l do
  begin
    c := ⌊(s + l)/2⌋
    if key = ac then
      begin
        location := c
        s := l + 1
      end
    else if key < ac then
      l := c - 1
    else s := c + 1
  end
end
```

Figure 10.26

We want to measure the (worst-case) time complexity for the algorithm implemented in Fig. 10.26. Here $f(n)$ will count the maximum number of comparisons (between *key* and a_c) needed to determine whether the given number *key* appears in the ordered array $a_1, a_2, a_3, \ldots, a_n$.

- For $n = 1$, *key* is compared to a_1 and $f(1) = 1$.
- When $n = 2$, in the worst case *key* is compared to a_1 and then to a_2, so $f(2) = 2$.
- In the case of $n = 3$, $f(3) = 2$ (in the worst case).
- When $n = 4$, the worst case occurs when *key* is first compared to a_2 and then a binary search of a_3, a_4 follows. Searching a_3, a_4 requires (in the worst case) $f(2)$ comparisons. So $f(4) = 1 + f(2) = 3$.

At this point we see that $f(1) \le f(2) \le f(3) \le f(4)$, and we conjecture that f is a monotone increasing function. To verify this, we shall use the Principle of Mathematical Induction in its alternative form. Here we assume that for all $i, j \in \{1, 2, 3, \ldots, n\}$, $i < j \Rightarrow f(i) \le f(j)$. Now consider the integer $n + 1$. We have two cases to examine.

1) $n + 1$ is odd: Here we write $n = 2k$ and $n + 1 = 2k + 1$, for some $k \in \mathbf{Z}^+$. In the worst case, $f(n + 1) = f(2k + 1) = 1 + f(k)$, where 1 counts the comparison of *key* with a_{k+1}, and $f(k)$ counts the (maximum) number of comparisons needed in a binary search of the subarray $a_1, a_2, \ldots, a_k$ or the subarray $a_{k+2}, a_{k+3}, \ldots, a_{2k+1}$.
Now $f(n) = f(2k) = 1 + \max\{f(k - 1), f(k)\}$. Since $k - 1, k < n$, by the induction hypothesis we have $f(k - 1) \le f(k)$, so $f(n) = 1 + f(k) = f(n + 1)$.

2) $n + 1$ is even: At this time we have $n + 1 = 2r$, for some $r \in \mathbf{Z}^+$, and in the worst case, $f(n + 1) = 1 + \max\{f(r - 1), f(r)\} = 1 + f(r)$, by the induction hypothesis. Therefore,

$$f(n) = f(2r - 1) = 1 + f(r - 1) \le 1 + f(r) = f(n + 1).$$

Consequently, the function f is monotone increasing.

Now it is time to determine the worst-case time complexity for the binary search algorithm, using the function $f(n)$. Since

$$f(1) = 1, \quad \text{and}$$
$$f(n) = f(n/2) + 1, \qquad \text{for } n = 2^k, \qquad k \ge 1,$$

it follows from Theorem 10.1 (with $a = 1$, $b = 2$, and $c = 1$) that

$$f(n) = \log_2 n + 1, \quad \text{and} \quad f \in O(\log_2 n) \qquad \text{for } n \in \{1, 2, 4, 8, \ldots\}.$$

But with f monotone increasing, from Theorem 10.2 it now follows that $f \in O(\log_2 n)$ (for all $n \in \mathbf{Z}^+$). Consequently, binary search is an $O(\log_2 n)$ algorithm, whereas the searching algorithm of Example 5.70 is $O(n)$. Therefore, as the value of n increases, binary search is the more efficient algorithm — but then it requires the additional condition that the array be ordered.

This section has introduced some of the basic ideas in the study of divide-and-conquer algorithms. It also extends the material first introduced on computational complexity and the analysis of algorithms in Sections 5.7 and 5.8.

The Section Exercises include some extensions of the results developed in this section. The reader who wants to pursue this topic further should find the chapter references both helpful and interesting.

EXERCISES 10.6

1. In each of the following, $f: \mathbf{Z}^+ \to \mathbf{R}$. Solve for $f(n)$ relative to the given set S, and determine the appropriate "big-Oh" form for f on S.

a) $f(1) = 5$
$f(n) = 4f(n/3) + 5, \quad n = 3, 9, 27, \ldots$
$S = \{3^i | i \in \mathbf{N}\}$

b) $f(1) = 7$
$f(n) = f(n/5) + 7, \quad n = 5, 25, 125, \ldots$
$S = \{5^i | i \in \mathbf{N}\}$

2. Let $a, b, c \in \mathbf{Z}^+$ with $b \geq 2$, and let $d \in \mathbf{N}$. Prove that the solution for the recurrence relation

$$f(1) = d$$
$$f(n) = af(n/b) + c, \qquad n = b^k, \qquad k \geq 1$$

satisfies

a) $f(n) = d + c \log_b n$, for $n = b^k$, $k \in \mathbf{N}$, when $a = 1$.

b) $f(n) = dn^{\log_b a} + (c/(a-1))[n^{\log_b a} - 1]$, for $n = b^k$, $k \in \mathbf{N}$, when $a \geq 2$.

3. Determine the appropriate "big-Oh" forms for f on $\{b^k | k \in \mathbf{N}\}$ in parts (a) and (b) of Exercise 2.

4. In each of the following, $f: \mathbf{Z}^+ \to \mathbf{R}$. Solve for $f(n)$ relative to the given set S, and determine the appropriate "big-Oh" form for f on S.

a) $f(1) = 0$
$f(n) = 2f(n/5) + 3, \quad n = 5, 25, 125, \ldots$
$S = \{5^i | i \in \mathbf{N}\}$

b) $f(1) = 1$
$f(n) = f(n/2) + 2, \quad n = 2, 4, 8, \ldots$
$S = \{2^i | i \in \mathbf{N}\}$

5. Consider a tennis tournament for n players, where $n = 2^k$, $k \in \mathbf{Z}^+$. In the first round $n/2$ matches are played, and the $n/2$ winners advance to round 2, where $n/4$ matches are played. This halving process continues until a winner is determined.

a) For $n = 2^k$, $k \in \mathbf{Z}^+$, let $f(n)$ count the total number of matches played in the tournament. Find and solve a recurrence relation for $f(n)$ of the form

$$f(1) = d$$
$$f(n) = af(n/2) + c, \qquad n = 2, 4, 8, \ldots,$$

where a, c, and d are constants.

b) Show that your answer in part (a) also solves the recurrence relation

$$f(1) = d$$
$$f(n) = f(n/2) + (n/2), \qquad n = 2, 4, 8, \ldots.$$

6. Complete the proofs for Corollary 10.1 and parts (b) and (c) of Theorem 10.2.

7. What is the best-case time-complexity function for binary search?

8. a) Modify the procedure in Example 10.48 as follows: For any $S \subset \mathbf{R}$, where $|S| = n$, partition S as $S_1 \cup S_2$, where $|S_1| = |S_2|$, for n even, and $|S_1| = 1 + |S_2|$, for n odd. Show that if $f(n)$ counts the number of comparisons needed (in this procedure) to find the maximum and minimum elements of S, then f is a monotone increasing function.

b) What is the appropriate "big-Oh" form for the function f of part (a)?

9. In Corollary 10.2 we were concerned with finding the appropriate "big-Oh" form for a function $f: \mathbf{Z}^+ \to \mathbf{R}^+ \cup \{0\}$ where

$$f(1) \leq c, \qquad \text{for } c \in \mathbf{Z}^+$$
$$f(n) \leq af(n/b) + c,$$
$$\text{for } a, b \in \mathbf{Z}^+ \text{ with } b \geq 2, \text{ and } n = b^k, k \in \mathbf{Z}^+.$$

Here the constant c in the second inequality is interpreted as the amount of time needed to break down the given problem of size n into a smaller (similar) problems of size n/b and to combine the a solutions of these smaller problems in order to get a solution for the original problem of size n. Now we shall examine a situation wherein this amount of time is no longer constant but depends on n.

a) Let $a, b, c \in \mathbf{Z}^+$, with $b \geq 2$. Let $f: \mathbf{Z}^+ \to \mathbf{R}^+ \cup \{0\}$ be a monotone increasing function, where

$$f(1) \leq c$$
$$f(n) \leq af(n/b) + cn, \qquad \text{for } n = b^k, \qquad k \in \mathbf{Z}^+.$$

Use an argument similar to the one given (for equalities) in Theorem 10.1 to show that for all $n = 1, b, b^2, b^3, \ldots,$

$$f(n) \leq cn \sum_{i=0}^{k} (a/b)^i.$$

b) Use the result of part (a) to show that $f \in O(n \log n)$, where $a = b$. (The base for the log function here is any real number greater than 1.)

c) When $a \neq b$, show that part (a) implies that

$$f(n) \leq \left(\frac{c}{a-b}\right)(a^{k+1} - b^{k+1}).$$

d) From part (c), prove that (i) $f \in O(n)$, when $a < b$; and (ii) $f \in O(n^{\log_b a})$, when $a > b$. [*Note*: The "big-Oh" form for f here and in part (b) is for f on $\mathbf{Z}^+$, not just $\{b^k | k \in \mathbf{N}\}$.]

10. In this exercise we briefly introduce the *Master Theorem*. (For more on this result, including a proof, we refer the reader to pp. 73–84 of reference [5] by T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein.)

Consider the recurrence relation

$$f(1) = 1,$$
$$f(n) = af(n/b) + h(n),$$

where $n \in \mathbf{Z}^+$, $n > 1$, $a \in \mathbf{Z}^+$, $a < n$, and $b \in \mathbf{R}^+$, $1 < b < n$. The function h accounts for the time (or cost) of dividing the given problem of size n into a smaller (similar) problems of size approximately n/b and then combining the results from