# 2
# Combinatorics

*[Combinatorics] has emerged as a new subject standing at the crossroads between pure and applied mathematics, the center of bustling activity, a simmering pot of new problems and exciting speculations.*

— Gian-Carlo Rota, [243, p. vii]

The formal study of combinatorics dates at least to Gottfried Wilhelm Leibniz's *Dissertatio de Arte Combinatoria* in the seventeenth century. The last half-century, however, has seen a huge growth in the subject, fueled by problems and applications from many fields of study. Applications of combinatorics arise, for example, in chemistry, in studying arrangements of atoms in molecules and crystals; biology, in questions about the structure of genes and proteins; physics, in problems in statistical mechanics; communications, in the design of codes for encryption, compression, and correction of errors; and especially computer science, for instance in problems of scheduling and allocating resources, and in analyzing the efficiency of algorithms.

Combinatorics is, in essence, the study of arrangements: pairings and groupings, rankings and orderings, selections and allocations. There are three principal branches in the subject. *Enumerative combinatorics* is the science of counting. Problems in this subject deal with determining the number of possible arrangements of a set of objects under some particular constraints. *Existential combinatorics* studies problems concerning the existence of arrangements that possess some specified property. *Constructive combinatorics* is the design and study of algorithms for creating arrangements with special properties.

J.M. Harris et al., *Combinatorics and Graph Theory*, DOI: 10.1007/978-0-387-79711-3_2,

Combinatorics is closely related to the theory of graphs. Many problems in graph theory concern arrangements of objects and so may be considered as combinatorial problems. For example, the theory of matchings and Ramsey theory, both studied in the previous chapter, have the flavor of existential combinatorics, and we continue their study later in this chapter. Also, combinatorial techniques are often employed to address problems in graph theory. For example, in Section 2.5 we determine another method for finding the chromatic polynomial of a graph.

We focus on topics in enumerative combinatorics through most of this chapter, but turn to some questions in existential combinatorics in Sections 2.4 and 2.10, and to some problems in constructive combinatorics in Sections 2.9 and 2.10. Throughout this chapter we study arrangements of finite sets. Chapter 3 deals with arrangements and combinatorial problems involving infinite sets. Our study in this chapter includes the investigation of the following questions.

- Should a straight beat a flush in the game of poker? What about a full house?
- Suppose a lazy professor collects a quiz from each student in a class, then shuffles the papers and redistributes them randomly to the class for grading. How likely is it that no one receives his or her own quiz to grade?
- How many ways are there to make change for a dollar?
- How many different necklaces with twenty beads can be made using rhodonite, rose quartz, and lapis lazuli beads, if a necklace can be worn in any orientation?
- How many seating arrangements are possible for $n$ guests attending a wedding reception in a banquet room with $k$ round tables?
- Suppose 100 medical students rank 100 positions for residencies at hospitals in order of preference, and the hospitals rank the students in order of preference. Is there a way to assign the students to the hospitals in such a way that no student and hospital prefer each other to their assignment? Is there an efficient algorithm for finding such a matching?
- Is it possible to find a collection of $n \geq 3$ points in the plane, not all on the same line, so that every line that passes through two of the points in fact passes through a third? Or, if we require instead that no three points lie on the same line, can we arrange a large number of points so that no subset of them forms the vertices of a convex octagon?

## 2.1 Some Essential Problems

*The mere formulation of a problem is far more essential than its solution…*

— Albert Einstein

We begin our study of combinatorics with two essential observations that underlie many counting strategies and techniques. The first is a simple observation about counting when presented with a number of alternative choices.

**Sum Rule.** *Suppose $S_1$, $S_2$, …, $S_m$ are mutually disjoint finite sets, and $|S_i| = n_i$ for $1 \le i \le m$. Then the number of ways to select one object from any of the sets $S_1$, $S_2$, …, $S_m$ is the sum $n_1 + n_2 + \cdots + n_m$.*

We often use the sum rule implicitly when solving a combinatorial problem when we break the set of possible outcomes into several disjoint cases, each of which can be analyzed separately. For example, suppose a coy college athlete tells us that his two-digit jersey number is divisible by 3, its first digit is odd, and its second digit is less than its first. How many numbers satisfy these criteria? A natural approach is to break the problem into five cases based on the first digit. Analyzing each of 1, 3, 5, 7, and 9 in turn, we find the possibilities are $\{\}$, $\{30\}$, $\{51, 54\}$, $\{72, 75\}$, or $\{90, 93, 96\}$, so there are eight possible jersey numbers in all.

The second essential observation concerns counting problems where selections are made in sequence.

**Product Rule.** *Suppose $S_1$, $S_2$, …, $S_m$ are finite sets, and $|S_i| = n_i$ for $1 \le i \le m$. Then the number of ways to select one element from $S_1$, followed by one element from $S_2$, and so on, ending with one element from $S_m$, is the product $n_1 n_2 \cdots n_m$, provided that the selections are independent, that is, the elements chosen from $S_1$, …, $S_{i-1}$ have no bearing on the selection from $S_i$, for each $i$.*

For example, consider the number of $m$-letter acronyms that can be formed using the full alphabet. To construct such an acronym, we make $m$ choices in sequence, one for each position, and each choice has no effect on any subsequent selection. Thus, by the product rule, the number of such acronyms is $26^m$.

We can apply a similar strategy to count the number of valid phone numbers in the U.S. and Canada. Under the North American Numbering Plan, a phone number has ten digits, consisting of an area code, then an exchange, then a station code. The three-digit area code cannot begin with 0 or 1, and its second digit can be any number except 9. The three-digit exchange cannot begin with 0 or 1, and the station code can be any four-digit number. Using the product rule, we find that the number of valid phone numbers under this plan is $(8 \cdot 9 \cdot 10) \cdot (8 \cdot 10^2) \cdot 10^4 =$ 5 760 000 000.

One might object that certain three-digit numbers are service codes reserved for special use in many areas, like 411 for information and 911 for emergencies. Let's compute the number of valid phone numbers for which neither the area code nor the exchange end with the digits 11. The amended number of area codes is then $8(9 \cdot 10 - 1) = 712$, and for exchanges we obtain $8 \cdot 99 = 792$. Thus, the number of valid phone numbers is $712 \cdot 792 \cdot 10^4 = 5\,639\,040\,000$.

We can use the product rule to solve three basic combinatorial problems.

**Problem 1.** How many ways are there to order a collection of $n$ different objects?

For example, how many ways are there to arrange the cards in a standard deck of 52 playing cards by shuffling? How many different batting orders are possible among the nine players on a baseball team? How many ways are there to arrange ten books on a shelf?

To order a collection of $n$ objects, we need to pick one object to be first, then another one to be second, and another one third, and so on. There are $n$ different choices for the first object, then $n-1$ remaining choices for the second, and $n-2$ for the third, and so forth, until just one choice remains for the last object. The total number of ways to order the $n$ objects is therefore the product of the integers between 1 and $n$. This number, called $n$ *factorial*, is written $n!$. An ordering, or rearrangement, of $n$ objects is often called a *permutation* of the objects. Thus, the number of permutations of $n$ items is $n!$.

Our second problem generalizes the first one.

**Problem 2.** How many ways are there to make an ordered list of $k$ objects from a collection of $n$ different objects?

For example, how many ways can a poll rank the top 20 teams in a college sport if there are 100 teams in the division? How many ways can a band arrange a play list of twelve songs if they know only 25 different songs?

Applying the same reasoning used in the first problem, we find that the answer to Problem 2 is the product $n(n-1)(n-2)\cdots(n-k+1)$, or $n!/(n-k)!$. This number is sometimes denoted by $P(n,k)$, but products like this occur frequently in combinatorics, and a more descriptive notation is often used to designate them.

We define the *falling factorial power* $x^{\underline{k}}$ as a product of $k$ terms beginning with $x$, with each successive term one less than its predecessor:

$$x^{\underline{k}} = x(x-1)(x-2)\cdots(x-k+1) = \prod_{i=0}^{k-1}(x-i). \tag{2.1}$$

The expression $x^{\underline{k}}$ is pronounced "$x$ to the $k$ falling." Similarly, we define the *rising factorial power* $x^{\overline{k}}$ ("$x$ to the $k$ rising") by

$$x^{\overline{k}} = x(x+1)(x+2)\cdots(x+k-1) = \prod_{i=0}^{k-1}(x+i). \tag{2.2}$$

Thus, we see that $P(n,k) = n^{\underline{k}} = (n-k+1)^{\overline{k}}$, and $n! = n^{\underline{n}} = 1^{\overline{n}}$. Also, the expressions $n^{\underline{0}}$, $n^{\overline{0}}$, and $0!$ all represent products having no terms at all. Multiplying any expression by such an empty product should not disturb the value of the expression, so the value of each of these degenerate products is taken to be 1.

Our third problem concerns unordered selections.

**Problem 3.** How many ways are there to select $k$ objects from a collection of $n$ objects, if the order of selection is irrelevant?

For example, how many different hands are possible in the game of poker? A poker hand consists of five cards drawn from a standard deck of 52 different cards. The order of the cards in a hand is unimportant, since players can rearrange their cards freely.

The solution to Problem 3 is usually denoted by $\binom{n}{k}$, or sometimes $C(n,k)$. The expression $\binom{n}{k}$ is pronounced "$n$ choose $k$."

We can find a formula for $\binom{n}{k}$ by using our solutions to Problems 1 and 2. Since there are $k!$ different ways to order a collection of $k$ objects, it follows that the product $\binom{n}{k}k!$ is the number of possible ordered lists of $k$ objects selected from the same collection of $n$ objects. Therefore,

$$\binom{n}{k} = \frac{n^{\underline{k}}}{k!} = \frac{n!}{k!(n-k)!}. \tag{2.3}$$

The numbers $\binom{n}{k}$ are called *binomial coefficients*, for reasons discussed in the next section. The binomial coefficients are ubiquitous in combinatorics, and we close this section with a few applications of these numbers.

1. The number of different hands in poker is $\binom{52}{5} = 52^{\underline{5}}/5! = 2\,598\,960$. The number of different thirteen-card hands in the game of bridge is $\binom{52}{13} = 635\,013\,559\,600$.

2. To play the Texas lottery game Lotto Texas, a gambler selects six different numbers between 1 and 54. The order of selection is unimportant. The number of possible lottery tickets is therefore $\binom{54}{6} = 25\,827\,165$.

3. Suppose we need to travel $m$ blocks east and $n$ blocks south in a regular grid of city streets. How many paths are there to our destination if we travel only east and south?

   We can represent a path to our destination as a sequence $b_1, b_2, \ldots, b_{n+m}$, where $b_i$ represents the direction we are traveling during the $i$th block of our route. Exactly $m$ of the terms in this sequence must be "east," and there are precisely $\binom{m+n}{m}$ ways to select $m$ positions in the sequence to have this value. The remaining $n$ positions in the sequence must all be "south," so the number of possible paths is $\binom{m+n}{m} = \frac{(m+n)!}{m!n!}$.

4. A standard deck of playing cards consists of four suits (spades, hearts, clubs, and diamonds), each with thirteen cards. Each of the cards in a suit has a different face value: a number between 2 and 10, or a jack, queen, king, or ace. How many poker hands have exactly three cards with the same face value?

   We can answer this question by considering how to construct such a hand through a sequence of simple steps. First, select one of the thirteen different face values. Second, choose three of the four cards in the deck having this value. Third, pick two cards from the 48 cards having a different face value. By the product rule, the number of possibilities is

$$\binom{13}{1}\binom{4}{3}\binom{48}{2} = 58\,656.$$

Poker aficionados will recognize that this strategy counts the number of ways to deal either of two different hands in the game: the "three of a kind" and the stronger "full house." A full house consists of a matched triple together with a matched pair, for example, three jacks and two aces; a three of a kind has only a matched triple. The number of ways to deal a full house is

$$\binom{13}{1}\binom{4}{3}\binom{12}{1}\binom{4}{2} = 3744,$$

since choosing a matched pair involves first selecting one of twelve different remaining face values, then picking two of the four cards having this value. The number of three of a kind hands is therefore $58\,656 - 3744 = 54\,912$.

We can also compute this number directly by modifying our first strategy. To avoid the possibility of selecting a matched pair in the last step, we can replace the term $\binom{48}{2} = 48 \cdot 47/2$ by $48 \cdot 44/2$, since the face value of the last card should not match any other card selected. Indeed, we calculate $13 \cdot 4 \cdot 48 \cdot 44/2 = 54\,912$. Notice that dividing by 2 is required in the last step, since the last two cards may be selected in any order.

## Exercises

1. In the C++ programming language, a variable name must start with a letter or the underscore character (_), and succeeding characters must be letters, digits, or the underscore character. Uppercase and lowercase letters are considered to be different characters.

    (a) How many variable names with exactly five characters can be formed in C++?

    (b) How many are there with at most five characters?

    (c) How many are there with at most five characters, if they must read exactly the same forwards and backwards? For example, `kayak` and `T55T` are admissible, but `Kayak` is not.

2. Assume that a vowel is one of the five letters `A`, `E`, `I`, `O`, or `U`.

    (a) How many eleven-letter sequences from the alphabet contain exactly three vowels?

    (b) How many of these have at least one repeated letter?

3. There are 30 teams in the National Basketball Association: 15 in the Western Conference, and 15 in the Eastern Conference.

   (a) Suppose each of the teams in the league has one pick in the first round of the NBA draft. How many ways are there to arrange the order of the teams selecting in the draft?

   (b) Suppose that each of the first three positions in the draft must be awarded to one of the fourteen teams that did not advance to the playoffs that year. How many ways are there to assign the first three positions in the draft?

   (c) How many ways are there for eight teams from each conference to advance to the playoffs, if order is unimportant?

   (d) Suppose that every team has three centers, four guards, and five forwards. How many ways are there to select an all-star team with the same composition from the Western Conference?

4. According to the *Laws of the Game* of the International Football Association, a full football (soccer) team consists of eleven players, one of whom is the goalkeeper. The other ten players fall into one of three outfield positions: defender, midfielder, and striker. There is no restriction on the number of players at each of these positions, as long as the total number of outfield players is ten.

   (a) How many different configurations are there for a full football team? For example, one team may field four strikers, three midfielders, and three defenders, in addition to the goalkeeper. Another may play five strikers, no midfielders, and five defenders, plus the goalkeeper.

   (b) Repeat the previous problem if there must be at least two players at each outfield position.

   (c) How many ways can a coach assign eleven different players to one of the four positions, if there must be exactly one goalkeeper, but there is no restriction on the number of players at each outfield position?

5. A political science quiz has two parts. In the first, you must present your opinion of the four most influential secretaries-general in the history of the United Nations in a ranked list. In the second, you must name ten members of the United Nations security council in any order, including at least two permanent members of the council. If there have been eight secretaries-general in U.N. history, and there are fifteen members of the U.N. security council, including the five permanent members, how many ways can you answer the quiz, assuming you answer both parts completely?

6. A midterm exam in phenomenology has two parts. The first part consists of ten multiple choice questions. Each question has four choices, labeled (a), (b), (c), and (d), and one may pick any combination of responses on each

of these questions. For example, one could choose just (a) alone on one question, or both (b) and (c), or all four possibilities, or none of them. In the second part, one may choose either to answer eight true/false questions, or to select the proper definition of each of seven terms from a list of ten possible definitions. Every question must be answered on whichever part is chosen, but one is not allowed to complete both portions. How many ways are there to complete the exam?

7. A ballot lists ten candidates for city council, eight candidates for the school board, and five bond issues. The ballot instructs voters to choose up to four people running for city council, rank up to three candidates for the school board, and approve or reject each bond issue. How many different ballots may be cast, if partially completed (or empty) ballots are allowed?

8. Compute the number of ways to deal each of the following five-card hands in poker.

   (a) Straight: the values of the cards form a sequence of consecutive integers. A jack has value 11, a queen 12, and a king 13. An ace may have a value of 1 or 14, so A 2 3 4 5 and 10 J Q K A are both straights, but K A 2 3 4 is not. Furthermore, the cards in a straight cannot all be of the same suit (a flush).

   (b) Flush: All five cards have the same suit (but not in addition a straight).

   (c) Straight flush: both a straight and a flush. Make sure that your counts for straights and flushes do not include the straight flushes.

   (d) Four of a kind.

   (e) Two distinct matching pairs (but not a full house).

   (f) Exactly one matching pair (but no three of a kind).

   (g) At least one card from each suit.

   (h) At least one card from each suit, but no two values matching.

   (i) Three cards of one suit, and the other two of another suit, like three hearts and two spades.

9. In the lottery game Texas Two Step, a player selects four different numbers between 1 and 35 in step 1, then selects an additional "bonus ball" number in the same range in step 2. The latter number is not considered to be part of the set selected in step 1, and in fact it may match one of the numbers selected there.

   (a) A resident of College Station always selects a bonus ball number that is different from any of the numbers he picks in step 1. How many of the possible Texas Two Step tickets have this property?

(b) In Rhode Island's lottery game Wild Money, a gambler picks a set of five numbers between $1$ and $35$. Is the number of possible tickets in this game the same as the number of tickets in Texas Two Step where the bonus ball number is different from the other numbers? Determine the ratio of the number of possible tickets in Wild Money to the number in the restricted Texas Two Step.

10. (a) A superstitious resident of Amarillo always picks three even numbers and three odd numbers when playing Lotto Texas. What fraction of all possible lottery tickets have this property?

    (b) Suppose in a more general lottery game one selects six numbers between $1$ and $2n$. What fraction of all lottery tickets have the property that half the numbers are odd and half are even?

    (c) What is the limiting value of this probability as $n$ grows large?

11. Suppose a positive integer $N$ factors as $N = p_1^{n_1} p_2^{n_2} \cdots p_m^{n_m}$, where $p_1$, $p_2$, …, $p_m$ are distinct prime numbers and $n_1$, $n_2$, …, $n_m$ are all positive integers. How many different positive integers are divisors of $N$?

12. Assume that a positive integer cannot have $0$ as its leading digit.

    (a) How many five-digit positive integers have no repeated digits at all?

    (b) How many have no consecutive repeated digits?

    (c) How many have at least one run of consecutive repeated digits (for example, $23324$, $45551$, or $151155$, but not $12121$)?

13. How many positive integers are there whose representation in base $8$ has exactly eight octal digits, at most one of which is odd? An octal digit is a number between $0$ and $7$, inclusive. Assume that the octal representation of a positive integer cannot start with a zero.

14. Let $\Delta$ be the *difference operator*: $\Delta(f(x)) = f(x+1) - f(x)$. Show that

$$\Delta(x^{\underline{n}}) = nx^{\underline{n-1}},$$

and use this to prove that

$$\sum_{k=0}^{m-1} k^{\underline{n}} = \frac{m^{\underline{n+1}}}{n+1}.$$

## 2.2 Binomial Coefficients

*About binomial theorem I'm teeming with a lot o' news,*
*With many cheerful facts about the square of the hypotenuse.*
— Gilbert and Sullivan, *The Pirates of Penzance*

The binomial coefficients possess a number of interesting arithmetic properties. In this section we study some of the most important identities associated with these numbers. Because binomial coefficients occur so frequently in this subject, knowing these essential identities will be helpful in our later studies.

The first identity generalizes our formula (2.3).

**Expansion.** *If $n$ is a nonnegative integer and $k$ is an integer, then*

$$\binom{n}{k} = \begin{cases} \dfrac{n!}{k!(n-k)!} & \text{if } 0 \le k \le n, \\ 0 & \text{otherwise.} \end{cases} \tag{2.4}$$

Designating the value of $\binom{n}{k}$ to be 0 when $k < 0$ or $k > n$ is sensible, for there are no ways to select fewer than zero or more than $n$ items from a collection of $n$ objects.

Notice that every subset of $k$ objects selected from a set of $n$ objects leaves a complementary collection of $n - k$ objects that are not selected. Counting the number of subsets with $k$ objects is therefore the same as counting the number of subsets with $n-k$ objects. This observation leads us to our second identity, which is easy to verify using the expansion formula.

**Symmetry.** *If $n$ is a nonnegative integer and $k$ is an integer, then*

$$\binom{n}{k} = \binom{n}{n-k}. \tag{2.5}$$

Before presenting the next identity, let us consider again the problem of counting poker hands. Suppose the ace of spades is the most desirable card in the deck (it certainly is in American Western movies), and we would like to know the number of five-card hands that include this card. The answer is the number of ways to select four cards from the other $51$ cards in the deck, namely, $\binom{51}{4}$. We can also count the number of hands that do not include the ace of spades. This is the number of ways to pick five cards from the other 51, that is, $\binom{51}{5}$. But every poker hand either includes the ace of spades or does not, so

$$\binom{52}{5} = \binom{51}{5} + \binom{51}{4}.$$

More generally, suppose we distinguish one particular object in a collection of $n$ objects. The number of unordered collections of $k$ of the objects that include the distinguished object is $\binom{n-1}{k-1}$; the number of collections that do not include this special object is $\binom{n-1}{k}$. We therefore obtain the following identity.

**Addition.** *If $n$ is a positive integer and $k$ is any integer, then*

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}. \tag{2.6}$$

We can prove this identity more formally using the expansion identity. It is easy to check that the identity holds for $k \le 0$ or $k \ge n$. If $0 < k < n$, we have

$$\begin{aligned}\binom{n-1}{k} + \binom{n-1}{k-1} &= \frac{(n-1)!}{k!(n-1-k)!} + \frac{(n-1)!}{(k-1)!(n-k)!} \\ &= \frac{((n-k)+k)(n-1)!}{k!(n-k)!} \\ &= \frac{n!}{k!(n-k)!} \\ &= \binom{n}{k}.\end{aligned}$$

We can use this identity to create a table of binomial coefficients. Let $n \ge 0$ index the rows of the table, and let $k \ge 0$ index the columns. Begin by entering 1 in the first position of each row, since $\binom{n}{0} = 1$ for $n \ge 0$; then use (2.6) to compute the entries in successive rows of the table. The resulting pattern of numbers is called *Pascal's triangle*, after Blaise Pascal, who studied many of its properties in his *Traité du Triangle Arithmétique*, written in 1654. (See [85] for more information on its history.) The first few rows of Pascal's triangle are shown in Figure 2.1.

| $\binom{n}{k}$ | $k=0$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | $2^n$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n=0$ | 1 | | | | | | | | | | | 1 |
| 1 | 1 | 1 | | | | | | | | | | 2 |
| 2 | 1 | 2 | 1 | | | | | | | | | 4 |
| 3 | 1 | 3 | 3 | 1 | | | | | | | | 8 |
| 4 | 1 | 4 | 6 | 4 | 1 | | | | | | | 16 |
| 5 | 1 | 5 | 10 | 10 | 5 | 1 | | | | | | 32 |
| 6 | 1 | 6 | 15 | 20 | 15 | 6 | 1 | | | | | 64 |
| 7 | 1 | 7 | 21 | 35 | 35 | 21 | 7 | 1 | | | | 128 |
| 8 | 1 | 8 | 28 | 56 | 70 | 56 | 28 | 8 | 1 | | | 256 |
| 9 | 1 | 9 | 36 | 84 | 126 | 126 | 84 | 36 | 9 | 1 | | 512 |
| 10 | 1 | 10 | 45 | 120 | 210 | 252 | 210 | 120 | 45 | 10 | 1 | 1024 |

TABLE 2.1. Pascal's triangle for binomial coefficients, $\binom{n}{k}$.

The next identity explains the origin of the name for these numbers: They appear as coefficients when expanding powers of the binomial expression $x + y$.

**The Binomial Theorem.** *If $n$ is a nonnegative integer, then*

$$(x+y)^n = \sum_k \binom{n}{k} x^k y^{n-k}. \tag{2.7}$$

The notation $\sum_k$ means that the sum extends over all integers $k$. Thus, the right side of (2.7) is formally an infinite sum, but all terms with $k < 0$ or $k > n$ are zero by the expansion identity, so there are only $n + 1$ nonzero terms in this sum.

*Proof.* We prove this identity by induction on $n$. For $n = 0$, both sides evaluate to $1$. Suppose then that the identity holds for a fixed nonnegative integer $n$. We need to verify that it holds for $n + 1$. Using our inductive hypothesis, then distributing the remaining factor of $(x + y)$, we obtain

$$\begin{aligned}(x+y)^{n+1} &= (x+y)\sum_k \binom{n}{k} x^k y^{n-k} \\ &= \sum_k \binom{n}{k} x^{k+1} y^{n-k} + \sum_k \binom{n}{k} x^k y^{n+1-k}.\end{aligned}$$

Now we reindex the first sum, replacing each occurrence of $k$ by $k - 1$. Since the original sum extends over all values of $k$, the reindexed sum does, too. Thus

$$\begin{aligned}(x+y)^{n+1} &= \sum_k \binom{n}{k-1} x^k y^{n+1-k} + \sum_k \binom{n}{k} x^k y^{n+1-k} \\ &= \sum_k \left(\binom{n}{k-1} + \binom{n}{k}\right) x^k y^{n+1-k} \\ &= \sum_k \binom{n+1}{k} x^k y^{n+1-k},\end{aligned}$$

by the addition identity. This completes the induction, and we conclude that the identity holds for all $n \geq 0$. □

We note two important consequences of the binomial theorem. First, setting $x = y = 1$ in (2.7), we obtain

$$\sum_k \binom{n}{k} = 2^n. \tag{2.8}$$

Thus, summing across the $n$th row in Pascal's triangle yields $2^n$, and there are therefore exactly $2^n$ different subsets of a set of $n$ elements. These row sums are included in Table 2.1.

Second, setting $x = -1$ and $y = 1$ in (2.7), we find that the alternating sum across any row of Pascal's triangle is zero, except of course for the top row:

$$\sum_k (-1)^k \binom{n}{k} = \begin{cases} 0 & \text{if } n \geq 1, \\ 1 & \text{if } n = 0. \end{cases} \tag{2.9}$$

This is obvious from the symmetry relation when $n$ is odd, but less clear when $n$ is even.

These two consequences of the binomial theorem concern sums over the lower index of binomial coefficients. The next identity tells us the value of a sum over the upper index.

**Summing on the Upper Index.** *If $m$ and $n$ are nonnegative integers, then*

$$\sum_{k=0}^{n} \binom{k}{m} = \binom{n+1}{m+1}. \tag{2.10}$$

*Proof.* We use induction on $n$ to verify this identity. For $n = 0$, each side equals $1$ if $m = 0$, and each side is $0$ if $m > 0$. Suppose then that the identity holds for some fixed nonnegative integer $n$. We must show that it holds for the case $n + 1$. Let $m$ be a nonnegative integer. We obtain

$$\begin{aligned}\sum_{k=0}^{n+1} \binom{k}{m} &= \binom{n+1}{m} + \sum_{k=0}^{n} \binom{k}{m} \\ &= \binom{n+1}{m} + \binom{n+1}{m+1} \\ &= \binom{n+2}{m+1}.\end{aligned}$$

By induction, the identity holds for all $n \geq 0$. □

To illustrate one last identity, we study the Lotto Texas game in more detail. Recall that a player selects six different numbers between $1$ and $54$ to enter the lottery. The largest prize is awarded to anyone matching all six numbers picked in a random drawing by lottery officials, but smaller prizes are given to players matching at least three of these numbers. To determine fair amounts for these smaller prizes, the state lottery commission needs to know the number of possible tickets that match exactly $k$ of the winning numbers, for every $k$.

Clearly, there is just one way to match all six winning numbers. There are $\binom{6}{5} = 6$ ways to pick five of the six winning numbers, and $48$ ways to select one losing number, so there are $6 \cdot 48 = 288$ tickets that match five numbers. Selecting four of the winning numbers and two of the losing numbers makes $\binom{6}{4}\binom{48}{2} = 16\,920$ possible tickets, and in general we see that the number of tickets that match exactly $k$ of the winning numbers is $\binom{6}{k}\binom{48}{6-k}$. By summing over $k$, we count every possible ticket exactly once, so

$$\binom{54}{6} = \sum_{k} \binom{6}{k}\binom{48}{6-k}.$$

More generally, if a lottery game requires selecting $m$ numbers from a set of $m+n$ numbers, we obtain the identity

$$\binom{m+n}{m} = \sum_{k} \binom{m}{k}\binom{n}{m-k}.$$

That is, the number of possible tickets equals the sum over $k$ of the number of ways to match exactly $k$ of the $m$ winning numbers and $m-k$ of the $n$ losing numbers. More generally still, suppose a lottery game requires a player to select $\ell$ numbers on a ticket, and each drawing selects $m$ winning numbers. Using the same reasoning, we find that

$$\binom{m+n}{\ell} = \sum_k \binom{m}{k}\binom{n}{\ell-k}.$$

Now replace $\ell$ by $\ell+p$ and reindex the sum, replacing $k$ by $k+p$, to obtain the following identity.

**Vandermonde's Convolution.** *If $m$ and $n$ are nonnegative integers and $\ell$ and $p$ are integers, then*

$$\binom{m+n}{\ell+p} = \sum_k \binom{m}{p+k}\binom{n}{\ell-k}. \tag{2.11}$$

Notice that the lower indices in the binomial coefficients on the right side sum to a constant.

**Exercises**

1. Use a combinatorial argument to prove that there are exactly $2^n$ different subsets of a set of $n$ elements. (Do not use the binomial theorem.)

2. Prove the absorption/extraction identity: If $n$ is a positive integer and $k$ is a nonzero integer, then
$$\binom{n}{k} = \frac{n}{k}\binom{n-1}{k-1}. \tag{2.12}$$

3. Use algebraic methods to prove the cancellation identity: If $n$ and $k$ are nonnegative integers and $m$ is an integer with $m \le n$, then
$$\binom{n}{k}\binom{k}{m} = \binom{n}{m}\binom{n-m}{k-m}. \tag{2.13}$$
This identity is very useful when the left side appears in a sum over $k$, since the right side has only a single occurrence of $k$.

4. Suppose that a museum curator with a collection of $n$ paintings by Jackson Pollack needs to select $k$ of them for display, and needs to pick $m$ of these to put in a particularly prominent part of the display. Show how to count the number of possible combinations in two ways so that the cancellation identity appears.

5. Prove the parallel summation identity: If $m$ and $n$ are nonnegative integers, then
$$\sum_{k=0}^{n} \binom{m+k}{k} = \binom{m+n+1}{n}. \tag{2.14}$$

6. Prove the hexagon identity: If $n$ is a positive integer and $k$ is an integer, then
$$\binom{n-1}{k-1}\binom{n}{k+1}\binom{n+1}{k} = \binom{n-1}{k}\binom{n}{k-1}\binom{n+1}{k+1}. \tag{2.15}$$
Why is it called the hexagon identity?

7. Compute the value of the following sums. Your answer should be an expression involving one or two binomial coefficients.

   (a) $\displaystyle\sum_{k} \binom{80}{k}\binom{k+1}{31}$.

   (b) $\displaystyle\sum_{k\geq 0} \frac{1}{k+1}\binom{99}{k}\binom{200}{120-k}$.

   (c) $\displaystyle\sum_{k=100}^{201} \sum_{j=100}^{k} \binom{201}{k+1}\binom{j}{100}$.

   (d) $\displaystyle\sum_{k} \binom{n}{k}^2$, for a nonnegative integer $n$.

   (e) $\displaystyle\sum_{k\leq m} (-1)^k \binom{n}{k}$, for an integer $m$ and a nonnegative integer $n$.

8. Prove the binomial theorem for falling factorial powers,
$$(x+y)^{\underline{n}} = \sum_{k} \binom{n}{k} x^{\underline{k}} y^{\underline{n-k}},$$
and for rising factorial powers,
$$(x+y)^{\overline{n}} = \sum_{k} \binom{n}{k} x^{\overline{k}} y^{\overline{n-k}}.$$

9. Let $n$ be a nonnegative integer. Suppose $f(x)$ and $g(x)$ are functions defined for all real numbers $x$, and that both functions are $n$ times differentiable. Let $f^{(k)}(x)$ denote the $k$th derivative of $f(x)$, so $f^{(0)}(x) = f(x)$, $f^{(1)}(x) = f'(x)$, and $f^{(2)}(x) = f''(x)$. Let $h(x) = f(x)g(x)$. Show that
$$h^{(n)}(x) = \sum_{k} \binom{n}{k} f^{(k)}(x) g^{(n-k)}(x).$$

10. In the Virginia lottery game Win For Life, an entry consists of a selection of six different numbers between 1 and 42, and each drawing selects seven different numbers in this range. How many different entries can match at least three of the drawn numbers?

11. The state of Florida administers several lottery games. In Florida Lotto, a player picks a set of six numbers between 1 and 53. In Fantasy 5, a gambler chooses a set of five numbers between 1 and 36. In which game is a player more likely to match at least two numbers against the ones drawn?

## 2.3 Multinomial Coefficients

*Alba, alma, ball, balm, bama, blam, lama, lamb, ma'am, mall, malm, mama, ...*

— Words constructed from letters in "Alma, Alabama"

Suppose we want to know the number of ways to place $n$ different objects into two boxes, one marked $A$ and the other marked $B$, in such a way that box $A$ receives a specified number $a$ of the objects, and box $B$ gets the remaining $b$ objects, so $a + b = n$. Assume that the order of placement of the objects in each box is immaterial, and denote the total number of such arrangements by $\binom{n}{a,b}$. We can compute this number easily by using our knowledge of binomial coefficients. Since each valid distribution corresponds to a different subset of $a$ objects for box $A$, we see that $\binom{n}{a,b}$ is simply the binomial coefficient $\binom{n}{a}$ (or $\binom{n}{b}$). Thus, $\binom{n}{a,b} = \frac{n!}{a!b!}$.

Now imagine we have three boxes, labeled $A$, $B$, and $C$, and suppose we want to know the number of ways to place a prescribed number $a$ of the objects in box $A$, a given number $b$ in box $B$, and the remaining $c = n - a - b$ in box $C$. Again, assume the order of placement of objects in each box is irrelevant, and denote this number by $\binom{n}{a,b,c}$. Since each arrangement can be described by first selecting $a$ elements from the set of $n$ for box $A$, and then picking $b$ objects from the remaining $n - a$ for box $B$, we see by the product rule that

$$\begin{aligned} \binom{n}{a,b,c} &= \binom{n}{a}\binom{n-a}{b} \\ &= \frac{n!}{a!(n-a)!} \cdot \frac{(n-a)!}{b!(n-a-b)!} \\ &= \frac{n!}{a!\,b!\,c!}. \end{aligned} \tag{2.16}$$

The number $\binom{n}{a,b,c}$ is called a *trinomial coefficient*.

We can generalize this problem for an arbitrary number of boxes. Suppose we have $n$ objects, together with $m$ boxes labeled 1, 2, ..., $m$, and suppose $k_1$, $k_2$, ..., $k_m$ are nonnegative integers satisfying $k_1 + k_2 + \cdots + k_m = n$. We define the *multinomial coefficient* $\binom{n}{k_1,k_2,\ldots,k_m}$ to be the number of ways to place $k_1$ of the objects in box 1, $k_2$ in box 2, and so on, without regard to the order of the objects in each box. Then an argument similar to our analysis for trinomial coefficients shows that

$$\begin{aligned}\binom{n}{k_1,\ldots,k_m} &= \binom{n}{k_1}\binom{n-k_1}{k_2}\binom{n-k_1-k_2}{k_3}\cdots \\ &\qquad\cdot\binom{n-k_1-\cdots-k_{m-2}}{k_{m-1}} \\ &= \frac{n!}{k_1!\,k_2!\cdots k_m!}.\end{aligned} \tag{2.17}$$

Multinomial coefficients often arise in a natural way in combinatorial problems. While we can always reduce questions about multinomial coefficients to problems about binomial coefficients or factorials by using (2.17), it is often useful to handle them directly. We derive some important formulas for multinomial coefficients in this section. These generalize some of the statements about binomial coefficients from Section 2.2. We begin with a more general formula for expanding multinomial coefficients in terms of factorials.

**Expansion.** *If $n$ is a nonnegative integer, and $k_1$, …, $k_m$ are integers satisfying $k_1+\cdots+k_m=n$, then*

$$\binom{n}{k_1,\ldots,k_m} = \begin{cases}\dfrac{n!}{k_1!\cdots k_m!} & \textit{if each } k_i \ge 0,\\ 0 & \textit{otherwise.}\end{cases} \tag{2.18}$$

Taking $\binom{n}{k_1,\ldots,k_m}=0$ when at least one of the $k_i$ is negative is certainly sensible, since it is impossible to place a negative number of objects in a box.

Second, it is clear that rearranging the numbers $k_1$, …, $k_m$ does not affect the value of the multinomial coefficient $\binom{n}{k_1,\ldots,k_m}$, since this just corresponds to relabeling the boxes. We can state this in the following way.

**Symmetry.** *Suppose $\pi(1)$, …, $\pi(m)$ is a permutation of $\{1,\ldots,m\}$. Then*

$$\binom{n}{k_1,\ldots,k_m} = \binom{n}{k_{\pi(1)},\ldots,k_{\pi(m)}}. \tag{2.19}$$

Third, we can observe a simple addition law. Let $\alpha$ be one of the objects from the set of $n$. It must be placed in one of the boxes. If we place $\alpha$ in box 1, then there are $\binom{n-1}{k_1-1,k_2,\ldots,k_m}$ ways to arrange the remaining $n-1$ objects to create a valid arrangement. If we set $\alpha$ in box 2, then there are $\binom{n-1}{k_1,k_2-1,k_3,\ldots,k_m}$ to complete the assignment of objects to boxes. Continuing in this way, we obtain the following identity.

**Addition.** *If $n$ is a positive integer and $k_1 + \cdots + k_m = n$, then*

$$\binom{n}{k_1, \ldots, k_m} = \binom{n-1}{k_1 - 1, k_2, \ldots, k_m} + \binom{n-1}{k_1, k_2 - 1, k_3, \ldots, k_m} + \cdots + \binom{n-1}{k_1, k_2, \ldots, k_{m-1}, k_m - 1}. \tag{2.20}$$

In the last section, the addition identity for $m = 2$ produced Pascal's triangle for the binomial coefficients. We can use a similar strategy to generate a geometric arrangement of the trinomial coefficients when $m = 3$, which we might call *Pascal's pyramid*. The top level of the pyramid corresponds to $n = 0$, just as in Pascal's triangle, and here we place a single 1, for $\binom{0}{0,0,0}$. The next level holds the numbers for $n = 1$, and we place the three 1s in a triangular formation, just below the $n = 0$ datum at the apex, for the numbers $\binom{1}{1,0,0}$, $\binom{1}{0,1,0}$, and $\binom{1}{0,0,1}$. In general, we use the addition formula (2.20) to compute the numbers in level $n$ from those in level $n - 1$, and we place the value of $\binom{n}{a,b,c}$ in level $n$ just below the triangular arrangement of numbers $\binom{n-1}{a-1,b,c}$, $\binom{n-1}{a,b-1,c}$, and $\binom{n-1}{a,b,c-1}$ in level $n-1$. Figure 2.1 shows the first few levels of the pyramid of trinomial coefficients. Here, the position of each number in level $n$ is shown relative to the positions of the numbers in level $n - 1$, each of which is marked with a triangle ($\triangle$).

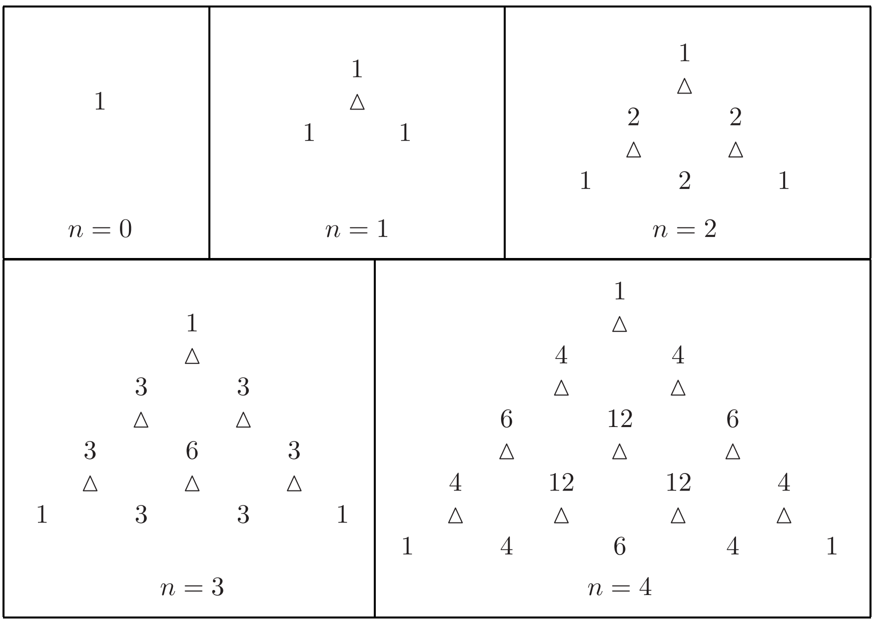



FIGURE 2.1. The first five levels of Pascal's pyramid.

We can use the addition identity to obtain an important generalization of the binomial theorem for multinomial coefficients.

**The Multinomial Theorem.** *If $n$ is a nonnegative integer, then*

$$(x_1 + \cdots + x_m)^n = \sum_{k_1+\cdots+k_m=n} \binom{n}{k_1, \ldots, k_m} x_1^{k_1} \cdots x_m^{k_m}. \tag{2.21}$$

Here, the notation $\sum_{k_1+\cdots+k_m=n}$ means that the sum extends over all integer $m$-tuples $(k_1, \ldots, k_m)$ whose sum is $n$. Of course, there are infinitely many such $m$-tuples, but only finitely many produce a nonzero term by the Expansion identity, so this is in effect a finite sum. We prove (2.21) for the case $m = 3$; the general case is left as an exercise.

*Proof.* The formula

$$(x + y + z)^n = \sum_{a+b+c=n} \binom{n}{a, b, c} x^a y^b z^c \tag{2.22}$$

certainly holds for $n = 0$, so suppose that it is valid for $n$. We compute

$$\begin{aligned}
(x+y + z)^{n+1} &= (x + y + z) \sum_{a+b+c=n} \binom{n}{a, b, c} x^a y^b z^c \\
&= \sum_{a+b+c=n} \binom{n}{a, b, c} x^{a+1} y^b z^c + \sum_{a+b+c=n} \binom{n}{a, b, c} x^a y^{b+1} z^c \\
&\quad + \sum_{a+b+c=n} \binom{n}{a, b, c} x^a y^b z^{c+1} \\
&= \sum_{a+b+c=n+1} \left( \binom{n}{a-1, b, c} + \binom{n}{a, b-1, c} + \binom{n}{a, b, c-1} \right) x^a y^b z^c \\
&= \sum_{a+b+c=n+1} \binom{n+1}{a, b, c} x^a y^b z^c,
\end{aligned}$$

so (2.22) holds for all $n \geq 0$. ☐

Some additional formulas for multinomial coefficients are developed in the exercises. Some of these may be obtained by selecting particular values for $x_1, \ldots,$ $x_m$ in the multinomial theorem.

We close this section by describing a common way that multinomial coefficients appear in combinatorial problems. Suppose we need to count the number of ways to order a collection of $n$ objects. If all the objects are different, then the answer is simply $n!$, but what if our collection includes some duplicate objects? Such a collection is called a *multiset*. Certainly we expect fewer different arrangements when there are some duplicate objects. For example, there are just six ways to line up four poker chips, two of which are red and the other two blue: $rrbb$, $rbrb$, $rbbr$, $brrb$, $brbr$, and $bbrr$.

Suppose we have a multiset of size $n$ that includes exactly $k_1$ identical copies of one object, $k_2$ instances of another, and so on, ending with $k_m$ duplicates of the last object, so $k_1 + \cdots + k_m = n$. In any ordering of these $n$ objects, we may rearrange the $k_i$ copies of object $i$ in any way without disturbing the arrangement. Since we can do this for any of the $m$ objects independently, it follows that each distinct ordering of the items occurs $k_1!k_2!\cdots k_m!$ times among the $n!$ ways that one could arrange the objects if they had been distinguishable. Therefore, the number of distinct arrangements of our multiset is

$$\frac{n!}{k_1!\cdots k_m!} = \binom{n}{k_1, \ldots, k_m}.$$

We could also obtain this formula by using our first combinatorial model for the multinomial coefficients. Suppose we have $n$ ping-pong balls, numbered $1$ through $n$, and $m$ boxes, each labeled with a different object from our multiset. The number of ways to distribute the balls among the boxes, with $k_1$ in box 1, $k_2$ in box 2, and so on, is the multinomial coefficient $\binom{n}{k_1,\ldots,k_m}$. But each arrangement corresponds to an ordering of the elements of our multiset: The numbers in box $i$ indicate the positions of object $i$ in the listing.

We have thus answered the analogue of Problem 1 from Section 2.1 for multisets. We can also study a generalization of Problem 2: How many ways are there to make an ordered list of $r$ objects from a multiset of $n$ objects, if the multiset comprises $k_i$ copies of object $i$ for $1 \le i \le m$? Our approach to this problem depends on the $k_i$ and $r$, so we'll study an example. Suppose a contemplative resident of Alma, Alabama, wants to know the number of ways to rearrange the letters of her home town and state, ignoring differences in case. There are eleven letters in all: six `A`s, one `B`, two `L`s, and two `M`s, so she computes the total number to be $\binom{11}{6,1,2,2} = \frac{11!}{6!2!2!} = 13\,860$.

Suppose she also wants to know the number of four-letter sequences of letters that can be formed from the same string, `ALMAALABAMA`, like the ones in the list that open this section, only they do not have to be English words. This is the multiset version of Problem 2 with $n = 11$, $r = 4$, $m = 4$, $k_1 = 6$, $k_2 = 1$, and $k_3 = k_4 = 2$. We can solve this by constructing each sequence in two steps: first, select four elements from the multiset; second, count the number of ways to order that subcollection. We can group the possible sub-multisets according to their pattern of repeated elements. For example, consider the subcollections that have two copies of one object, and two copies of another. Denote this pattern by $wwxx$. There are $\binom{3}{2} = 3$ ways to select values for $w$ and $x$, since we must pick two of the three letters `A`, `L`, and `M`. Each of these subcollections can be ordered in any of $\binom{4}{2,2} = 6$ ways, so the pattern $wwxx$ produces $3 \cdot 6 = 18$ possible four-letter sequences in all. There are five possible patterns for a four-element multiset, which we can denote $wwww$, $wwwx$, $wwxx$, $wwxy$, and $wxyz$. The analysis of each one is summarized in the following table.

| Pattern | Sub-multisets | Orderings per sub-multiset | Total |
|---|---|---|---|
| $wwww$ | 1 | 1 | 1 |
| $wwwx$ | $\binom{3}{1}$ | $\binom{4}{3,1}$ | 12 |
| $wwxx$ | $\binom{3}{2}$ | $\binom{4}{2,2}$ | 18 |
| $wwxy$ | $\binom{3}{1}\binom{3}{2}$ | $\binom{4}{2,1,1}$ | 108 |
| $wxyz$ | 1 | $4!$ | 24 |

Summing the values in the rightmost column, we find that there are exactly 163 ways to form a four-letter sequence from the letters in Alma, Alabama.

## Exercises

1. Prove the addition identity for multinomial coefficients (2.20) by using the expansion identity (2.18).

2. For nonnegative integers $a$, $b$, and $c$, let $P(a,b,c)$ denote the number of paths in three-dimensional space that begin at the origin, end at $(a,b,c)$, and consist entirely of steps of unit length, each of which is parallel to a coordinate axis. Prove that $P(a,b,c) = \binom{a+b+c}{a,b,c}$.

3. Prove the multinomial theorem (2.21) for an arbitrary positive integer $m$.

4. Prove the following identities for sums of multinomial coefficients, if $m$ and $n$ are positive integers.

   (a) $$\sum_{k_1+\cdots+k_m} \binom{n}{k_1,\ldots,k_m} = m^n.$$

   (b) $$\sum_{k_1+\cdots+k_m} \binom{n}{k_1,\ldots,k_m} (-1)^{k_2+k_4+\cdots+k_{2\ell}} = \begin{cases} 0 & \text{if } m = 2\ell, \\ 1 & \text{if } m = 2\ell+1. \end{cases}$$

5. Prove that if $n$ is a nonnegative integer and $k$ is an integer, then
$$\sum_j \binom{n}{j,k,n-j-k} = 2^{n-k}\binom{n}{k}.$$

6. Prove the multinomial theorem for falling factorial powers,
$$(x_1+\cdots+x_m)^{\underline{n}} = \sum_{k_1+\cdots+k_m=n} \binom{n}{k_1,\ldots,k_m} x_1^{\underline{k_1}} \cdots x_m^{\underline{k_m}},$$
and for rising factorial powers,
$$(x_1+\cdots+x_m)^{\overline{n}} = \sum_{k_1+\cdots+k_m=n} \binom{n}{k_1,\ldots,k_m} x_1^{\overline{k_1}} \cdots x_m^{\overline{k_m}}.$$
You may find it helpful to consider the trinomial case first.

7. Use a combinatorial argument to establish the following analogue of Vandermonde's convolution for trinomial coefficients. If $m$ and $n$ are nonnegative integers, and $a+b+c=m+n$, then

$$\sum_{\alpha+\beta+\gamma=m} \binom{m}{\alpha,\beta,\gamma}\binom{n}{a-\alpha,b-\beta,c-\gamma} = \binom{m+n}{a,b,c}.$$

8. State an analogue of Vandermonde's convolution for multinomial coefficients, and use a combinatorial argument to establish it.

9. Compute the number of $r$-letter sequences that can be formed by using the letters in each location below, for each given value of $r$. Ignore differences in case.

   (a) Bug Tussle, TX: $r=3$, $r=4$, $r=11$.

   (b) Cooleemee, NC: $r=4$, $r=10$, $r=11$.

   (c) Oconomowoc, WI: $r=4$, $r=11$, $r=12$.

   (d) Unalaska, Alaska: $r=3$, $r=4$, $r=14$.

   (e) Walla Walla, WA: $r=4$, $r=5$, $r=12$.

10. Certainly there are more four-letter sequences that can be formed by using the letters in Bobo, Mississippi, than can be formed by using the letters in Soso, Mississippi. Is the difference more or less than the distance between these two cities in miles, which is 267?

11. A band of combinatorial ichthyologists asserts that the number of five-letter sequences that can be formed using the letters of the Hawaiian long-nosed butterfly fish, the lauwiliwilinukunuku'oi'oi, is more than twice as large as the number of five-letter sequences that can be created using the name of the state fish of Hawaii, the painted triggerfish humuhumunukunukuapua'a. Prove or disprove their claim by computing the exact number in each case.

## 2.4 The Pigeonhole Principle

*I am just here for anyone that's for the pigeons.*

— Mike Tyson, Phoenix City Council meeting, June 1, 2005,
reported in *The Arizona Republic*

We now turn to a simple, but powerful, idea in combinatorial reasoning known as the *pigeonhole principle*. We can state it in the following way.

**Theorem 2.1** (Pigeonhole Principle). *Let $n$ be a positive integer. If more than $n$ objects are distributed among $n$ containers, then some container must contain more than one object.*

The proof is simple—if each container held at most one object, then there would be at most $n$ objects in all.

This mathematical idea is also called the *box principle* (especially in number theory texts), which is sensible enough, since we can imagine the containers as boxes. In German, it is the *drawer principle*, logically enough, after Dirichlet's original term, the *Schubfachprinzip*. It may seem odd to think of our containers as pigeon roosts, but the name probably originally referred to the "pigeonholes" one sees in those old desks with lots of square nooks for squirreling away papers. (One imagines however that the origins of the term may be the subject of some, well, squabbling....) So while the traditional name may be somewhat antiquated, at least the avian nomenclature saves us from talking about Dirichlet's drawers.

The pigeonhole principle is very useful in establishing the existence of a particular configuration or combination in many mathematical contexts. We begin with a few simple examples.

1. Suppose $400$ freshmen enroll in introductory calculus one term. Then two must have the same birthday. Here, the pigeonholes are calendar days, so $n = 366$.

2. In honor of champion pugilist (and pigeon enthusiast) Mike Tyson, suppose that $n$ boxers schedule a round-robin tournament, so each fighter meets every other in a bout, and afterwards no contestant is undefeated. Then each boxer has between $1$ and $n - 1$ wins, so two boxers must have the same record in the tournament.

3. It is estimated that the average full head of hair has $100\,000$ to $150\,000$ strands of hair. Let's assume that the most hirsute among us has less than $250\,000$ strands of hair on their head. The city of Phoenix has over $1.5$ million residents, so it follows that there must be at least two residents with exactly the same number of hairs on their head. Moreover, since only a fraction of the population is bald, the statement surely remains true if we exclude those with no hair at all. (Sorry, Iron Mike.)

In this last problem, we can in fact conclude considerably more. The population of Phoenix is more than six times the maximum number of hairs per head, and a moment's thought reveals that there must in fact exist at least *six* people in Phoenix with identical hair counts. We can thus state a more powerful pigeonhole principle.

**Theorem 2.2** (Generalized Pigeonhole Principle)**.** *Let $m$ and $n$ be positive integers. If more than $mn$ objects are distributed among $n$ containers, then at least one container must contain at least $m + 1$ objects.*

The proof is again easy—if each container held at most $m$ objects then the total number of objects would be at most $mn$. An alternative formulation of this statement appears in the exercises. Next, we establish the following arithmetic variation on the pigeonhole principle.

**Theorem 2.3.** *Suppose $a_1$, $a_2$, ..., $a_n$ is a sequence of real numbers with mean $\mu$, so $\mu = (a_1 + \cdots + a_n)/n$. Then there exist integers $i$ and $j$, with $1 \le i, j \le n$, such that $a_i \le \mu$ and $a_j \ge \mu$.*

The proof is again straightforward—if every element of the sequence were strictly greater than $\mu$, then we would have $a_1 + \cdots + a_n > n\mu$, a contradiction. Thus, an integer $i$ must exist with $a_i \le \mu$. A similar argument establishes the existence of $j$.

While the pigeonhole principle and the variations we describe here are all quite simple to state and verify, this idea plays a central role in establishing many decidedly nontrivial statements in mathematics. We conclude this section with two examples.

### Monotonic Subsequences

We say a sequence $a_1, \ldots, a_n$ is *increasing* if $a_1 \le a_2 \le \cdots \le a_n$, and *strictly increasing* if $a_1 < a_2 < \cdots < a_n$. We define *decreasing* and *strictly decreasing* in the same way. Consider first an arrangement of the integers between 1 and 10, for example,

$$3, 5, 8, 10, 6, 1, 9, 2, 7, 4. \tag{2.23}$$

Scan the list for an increasing subsequence of maximal length. Above, we find $(3, 5, 8, 10)$, $(3, 5, 8, 9)$, $(3, 5, 6, 7)$, and $(3, 5, 6, 9)$ all qualify with length 4. Next, scan the list for a decreasing subsequence of maximal length. Here, the best we can do is length 3, achieved by $(8, 6, 1)$, $(8, 6, 2)$, $(8, 6, 4)$, $(10, 6, 2)$, $(10, 6, 4)$, $(10, 7, 4)$, and $(9, 7, 4)$. Is it possible to find an arrangement of the integers from 1 to 10 that simultaneously avoids both an increasing subsequence of length 4 and a decreasing subsequence of length 4? The following theorem asserts that this is not possible. Its statement dates to an early and influential paper of Erdős and Szekeres [94], the same one cited in Section 1.8 for its contribution to the development of Ramsey theory.

**Theorem 2.4.** *Suppose $m$ and $n$ are positive integers. A sequence of more than $mn$ real numbers must contain either an increasing subsequence of length at least $m + 1$, or a strictly decreasing subsequence of length at least $n + 1$.*

*Proof.* Suppose that $r_1, r_2, \ldots, r_{mn+1}$ is a sequence of real numbers which contains neither an increasing subsequence of length $m + 1$, nor a strictly decreasing subsequence of length $n + 1$. For each integer $i$ with $1 \le i \le mn + 1$, let $a_i$ denote the length of the longest increasing subsequence in this sequence of numbers whose first term is $r_i$, and let $d_i$ denote the length of the longest strictly decreasing subsequence beginning with this term. For example, for the sequence (2.23) we see that $a_2 = 3$ (for $5, 8, 10$ or $5, 8, 9$), and $d_2 = 2$ (for $5, 1$ or $5, 2$). By our hypothesis, we know that $1 \le a_i \le m$ and $1 \le d_i \le n$ for each $i$, and thus there are only $mn$ different possible values for the ordered pair $(a_i, d_i)$. However, there are $mn + 1$ such ordered pairs, so by the pigeonhole principle there exist two integers $j$ and $k$ with $j < k$ such that $a_j = a_k$ and $d_j = d_k$. Denote this pair by $(\alpha, \delta)$, so $\alpha = a_j = a_k$ and $\delta = d_j = d_k$. Now let $r_k, r_{i_2}, \ldots, r_{i_\alpha}$ denote a maximal increasing subsequence beginning with $r_k$ and let $r_k, r_{i'_2}, \ldots, r_{i'_\delta}$ denote a maximal strictly decreasing subsequence beginning with this term. If $r_j \le r_k$, then $r_j, r_k, r_{i_2}, \ldots, r_{i_\alpha}$ is an increasing subsequence of length $\alpha + 1$ beginning with $r_j$. On the other hand, if $r_j > r_k$, then $r_j, r_k, r_{i'_2}, \ldots, r_{i'_\delta}$ is a strictly decreasing subsequence of length $\delta + 1$ beginning with $r_j$. In either case, we reach a contradiction. □

Of course, we can replace "increasing" with "strictly increasing" and simultaneously "strictly decreasing" with "decreasing" in this statement.

### Approximating Irrationals by Rationals

Let $\alpha$ be an irrational number. Since every real interval $[a, b]$ with $a < b$ contains infinitely many rational numbers, certainly there exist rational numbers arbitrarily close to $\alpha$. Suppose however we restrict the rationals we may select to the set of fractions with bounded denominator. How closely can we approximate $\alpha$ now? More specifically, given an irrational number $\alpha$ and a positive integer $Q$, does there exist a rational number $p/q$ with $1 \le q \le Q$ and $\left|\alpha - \frac{p}{q}\right|$ especially small? How small can we guarantee?

At first glance, if we select a random denominator $q$ in the range $[1, Q]$, then certainly $\alpha$ lies in some interval $(\frac{k}{q}, \frac{k+1}{q})$, for some integer $k$, so its distance to the nearest multiple of $1/q$ is at most $1/2q$. We might therefore expect that on average we would observe a distance of about $1/4q$, for randomly selected $q$. In view of Theorem 2.3, we might then expect that approximations with distance at most $1/4q$ must exist. In fact, however, we can establish a much stronger result by using the pigeonhole principle. The following important theorem is due to Dirichlet and his *Schubfachprinzip*.

We first require some notation. For a real number $x$, let $\lfloor x \rfloor$ denote the *floor* of $x$, or *integer part* of $x$. It is defined to be the largest integer $m$ satisfying $m \le x$. Similarly, the *ceiling* of $x$, denoted by $\lceil x \rceil$, is the smallest integer $m$ satisfying $x \le m$. Last, the *fractional part* of $x$, denoted by $\{x\}$, is defined by $\{x\} = x - \lfloor x \rfloor$. For example, for $x = \pi$ we have $\lfloor \pi \rfloor = 3$, $\lceil \pi \rceil = 4$, and $\{\pi\} = 0.14159\ldots$; for $x = 1$ we obtain $\lfloor 1 \rfloor = \lceil 1 \rceil = 1$ and $\{1\} = 0$.

**Theorem 2.5** (Dirichlet's Approximation Theorem)**.** *Suppose $\alpha$ is an irrational real number, and $Q$ is a positive integer. Then there exists a rational number $p/q$ with $1 \le q \le Q$ satisfying*

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q(Q+1)}.$$

*Proof.* Divide the real interval $[0, 1]$ into $Q + 1$ subintervals of equal length:

$$\left[0, \frac{1}{Q+1}\right), \left[\frac{1}{Q+1}, \frac{2}{Q+1}\right), \ldots, \left[\frac{Q-1}{Q+1}, \frac{Q}{Q+1}\right), \left[\frac{Q}{Q+1}, 1\right].$$

Since each of the $Q+2$ real numbers

$$0, \{\alpha\}, \{2\alpha\}, \ldots, \{Q\alpha\}, 1 \tag{2.24}$$

lies in $[0,1]$, by the pigeonhole principle at least two of them must lie in the same subinterval. Each of the numbers in (2.24) can be written in a unique way as $r\alpha - s$ with $r$ and $s$ integers and $0 \le r \le Q$, so it follows that there exist integers $r_1$, $r_2$, $s_1$, and $s_2$, with $0 \le r_1, r_2 \le Q$, such that

$$|(r_2\alpha - s_2) - (r_1\alpha - s_1)| < \frac{1}{Q+1}.$$

Since only 0 and 1 in our list have the same $r$-value, we can assume that $r_1 \ne r_2$, so suppose $r_1 < r_2$. Let $q = r_2 - r_1$, so $1 \le q \le Q$, and let $p = s_2 - s_1$. Then $p$ and $q$ satisfy

$$|q\alpha - p| < \frac{1}{Q+1},$$

and the conclusion follows upon dividing through by $q$. □

Since $q \le Q$, we immediately obtain that the rational number $p/q$ guaranteed by the theorem satisfies

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^2+q}. \tag{2.25}$$

Exercise 11 asks you to show that there exist infinitely many rational numbers $p/q$ that satisfy this inequality for a fixed irrational number $\alpha$.

**Exercises**

1. Show that at any party with at least two people, there must exist at least two people in the group who know the same number of other guests at the party. Assume that each pair of people at the party are either mutual friends or mutual strangers.

2. Prove the following version of the pigeonhole principle. Let $m$ and $n$ be positive integers. If $m$ objects are distributed in some way among $n$ containers, then at least one container must hold at least $1 + \lfloor (m-1)/n \rfloor$ objects.

3. Prove the following more general version of the pigeonhole principle. Suppose that $m_1$, $m_2$, …, $m_n$ are all positive integers, let $M = m_1 + m_2 + \cdots + m_n - n + 1$, and suppose each of $n$ containers is labeled with an integer between 1 and $n$. Prove that if $M$ objects are distributed in some way among the $n$ containers, then there exists an integer $i$ between 1 and $n$ such that the container labeled with $i$ contains at least $m_i$ objects.

4. The top four pitchers on a college baseball team combine for 297 strikeouts over the course of a season. If each pitcher had at least 40 strikeouts over the

course of the season, and the fourth-best pitcher had less than $50$ strikeouts, how many strikeouts could the best pitcher have made over the season? Your answer should be a range of possible numbers.

5. Find the smallest value of $m$ so that the following statement is valid: Any collection of $m$ distinct positive integers must contain at least two numbers whose sum or difference is a multiple of 10. Prove that your value is best possible.

6. Suppose $A = (a_1, a_2, \ldots, a_n)$ is a sequence of positive real numbers. Let $H(A)$ denote the *harmonic mean* of $A$, defined by

$$H(A) = n\left(\sum_{i=1}^{n} \frac{1}{a_i}\right)^{-1}.$$

Show there exist integers $i$ and $j$, with $1 \le i, j \le n$, satisfying

$$a_i \le H(A) \le a_j.$$

7. Suppose the integers from $1$ to $n$ are arranged in some order around a circle, and let $k$ be an integer with $1 \le k \le n$. Show that there must exist a sequence of $k$ adjacent numbers in the arrangement whose sum is at least $\lceil k(n+1)/2 \rceil$.

8. Suppose the integers from $1$ to $n$ are arranged in some order around a circle, and let $k$ be an integer with $1 \le k \le n$. Show that there must exist a sequence of $k$ adjacent numbers in the arrangement whose product is at least $\left\lceil (n!)^{k/n} \right\rceil$.

9. Let $n$ be a positive integer. Exhibit an arrangement of the integers between $1$ and $n^2$ which has no increasing or decreasing subsequence of length $n+1$.

10. Let $m$ and $n$ be positive integers. Exhibit an arrangement of the integers between 1 and $mn$ which has no increasing subsequence of length $m+1$, and no decreasing subsequence of length $n+1$.

11. Let $\alpha$ be an irrational number. Prove that there exist infinitely many rational numbers $p/q$ satisfying (2.25).

12. Let $n$ be a positive integer, and let $b \ge 2$ be an integer.

    (a) Show that there exists a nonzero multiple $N$ of $n$ whose base-$b$ representation consists entirely of 0s and 1s. (No partial credit will be awarded for the case $b = 2$!) Hint: Consider the sequence of numbers $\sum_{i=0}^{k} b^i$ for a number of values of $k$.

    (b) Show that there exists a multiple $N$ of $n$ whose base-$b$ representation consists entirely of 1s if and only if no prime number $p$ which divides $b$ is a factor of $n$.

(c) Suppose that the greatest common divisor of $b$ and $n$ is 1, and let $d_1, \ldots, d_m$ be a sequence of integers with $0 \le d_i < b$ for each $i$ and $d_1 \neq 0$. Show that there exists a multiple $N$ of $n$ whose base-$b$ representation is obtained by juxtaposing some integral number of copies of the base-$b$ digit sequence $d_1 d_2 \cdots d_m$.

13. Let $a_1, a_2, \ldots, a_n$ be a sequence of integers. Show that there exist integers $j$ and $k$ with $1 \le j \le k \le n$ such that the sum $\sum_{i=j}^{k} a_i$ is a multiple of $n$.

14. (Bloch and Pólya [28].) For a positive integer $d$, let $\mathcal{N}_d$ denote the set of polynomials with degree at most $d-1$ whose coefficients are all 0 or 1. For example, $\mathcal{N}_3 = \{0, 1, x, x^2, 1+x, 1+x^2, x+x^2, 1+x+x^2\}$.

(a) Let $f^{(k)}(x)$ denote the $k$th derivative of $f(x)$. Show that if $f \in \mathcal{N}_d$ then $f^{(k-1)}(1) \le d^{\underline{k}}/k$.

(b) Let $m$ be a positive integer. Determine an upper bound on the number of different possible $m$-tuples $(f(1), f'(1), \ldots, f^{(m-1)}(1))$ achieved by polynomials $f(x) \in \mathcal{N}_d$.

(c) Prove that if $d > 1$ and

$$\frac{d}{\log_2 d} > \binom{m+1}{2},$$

then there exists a polynomial $h(x)$ of degree at most $d-1$ whose coefficients are all 0, 1, or $-1$, and which is divisible by $(x-1)^m$.

## 2.5 The Principle of Inclusion and Exclusion

*What we here have to do is to conceive, and invent a notation for, all the possible combinations which any number of class terms can yield; and then find some mode of symbolic expression which shall indicate which of these various compartments are empty or occupied...*

— John Venn, [275, p. 23]

Suppose there are 50 beads in a drawer: 25 are glass, 30 are red, 20 are spherical, 18 are red glass, 12 are glass spheres, 15 are red spheres, and 8 are red glass spheres. How many beads are neither red, nor glass, nor spheres?

We can answer this question by organizing all of this information using a Venn diagram with three overlapping sets: $G$ for glass beads, $R$ for red beads, and $S$ for spherical beads. See Figure 2.2. We are given that there are eight red glass spheres, so start by labeling the common intersection of the sets $G$, $R$, and $S$ in the diagram with 8. Then the region just above this one must have ten elements, since there are 18 red glass beads, and exactly eight of these are spherical. Continuing in this way, we determine the size of each of the sets represented in the diagram, and we conclude that there are exactly twelve beads in the drawer that are neither red, nor glass, nor spheres.

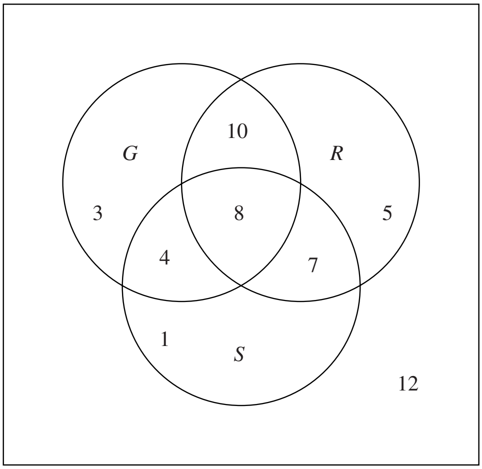



FIGURE 2.2. A solution using a Venn diagram.

Alternatively, we can answer this question by determining the size of the set $G \cup R \cup S$ (does this make us counting GURUS?). Summing the number of elements in the sets $G$, $R$, and $S$ produces a number that is too large, since this sum counts the beads that are in more than one of these sets at least twice. We can try to compensate by subtracting the number of elements in the sets $G \cap R$, $G \cap S$, and $R \cap S$ from the sum. This produces a total that is too small, since the beads that have all three attributes are counted three times in the first step, then subtracted three times in the second step. Thus, we must add the number of elements in $G \cap R \cap S$ to the sum, and we find that

$$|G \cup R \cup S| = |G| + |R| + |S| - |G \cap R| - |G \cap S| - |R \cap S| + |G \cap R \cap S| \,.$$

Letting $N_0$ denote the number of beads with none of the three attributes, we then compute

$$\begin{aligned} N_0 &= 50 - |G \cup R \cup S| \\ &= 50 - |G| - |R| - |S| + |G \cap R| + |G \cap S| + |R \cap S| - |G \cap R \cap S| \\ &= 50 - 25 - 30 - 20 + 18 + 12 + 15 - 8 \\ &= 12. \end{aligned}$$

This suggests a general technique for solving some similar combinatorial problems. Suppose we have a collection of $N$ distinct objects, and each object may satisfy one or more properties that we label $a_1, a_2, \ldots, a_r$. Let $N(a_i)$ denote the number of objects having property $a_i$, let $N(a_i a_j)$ signify the number having both property $a_i$ and property $a_j$, and in general let $N(a_{i_1}a_{i_2}\ldots a_{i_m})$ represent the number satisfying the $m$ properties $a_{i_1}, \ldots, a_{i_m}$. Let $N_0$ denote the number of objects having none of the properties. We prove the following theorem.

**Theorem 2.6** (Principle of Inclusion and Exclusion)**.** *Using the notation above,*

$$\begin{aligned} N_0 = N - \sum_i N(a_i) + \sum_{i<j} N(a_i a_j) - \sum_{i<j<k} N(a_i a_j a_k) + \cdots \\ + (-1)^m \sum_{i_1<\cdots<i_m} N(a_{i_1}\ldots a_{i_m}) + \cdots + (-1)^r N(a_1 a_2 \ldots a_r). \end{aligned} \tag{2.26}$$

*Proof.* Suppose an object satisfies none of the properties. Then the expression on the right side counts it precisely once, in the $N$ term. On the other hand, suppose an object satisfies precisely $m$ of the properties, with $m$ a positive number. Then it is counted once in the $N$ term, $m$ times in the $\sum N(a_i)$ term, $\binom{m}{2}$ times in the second sum, and in general $\binom{m}{k}$ times in the $k$th sum. Therefore, the total contribution on the right side from this object is

$$\sum_k (-1)^k \binom{m}{k} = 0$$

by (2.9). This completes the proof. □

We consider four applications of this counting principle.

### The Euler $\varphi$ Function

Two integers are said to be *relatively prime* if their greatest common divisor is 1. If $n$ is a positive integer, let $\varphi(n)$ be the number of positive integers $m \leq n$ that are relatively prime to $n$. This function, called the Euler $\varphi$ function or the Euler totient function, is important in number theory. We can derive a formula for this function by using the principle of inclusion and exclusion.

We must name a set and list a number of properties such that the number of elements in the set satisfying none of the properties is $\varphi(n)$. Suppose $n$ is divisible by precisely $r$ different primes, which we label $p_1$ through $p_r$. Select $\{1, 2, \ldots, n\}$ as the set, and let $a_i$ be the property "is divisible by $p_i$." Then $N_0 = \varphi(n)$, and it is easy to compute the terms on the right side of the equation in Theorem 2.6: $N = n$, $N(a_i) = n/p_i$, $N(a_i a_j) = n/(p_i p_j)$, and so on. Therefore,

$$\begin{aligned} \varphi(n) &= n - \sum_i \frac{n}{p_i} + \sum_{i<j} \frac{n}{p_i p_j} - \sum_{i<j<k} \frac{n}{p_i p_j p_k} + \cdots + (-1)^r \frac{n}{p_1 p_2 \cdots p_r} \\ &= n \prod_{i=1}^{r} \left(1 - \frac{1}{p_i}\right). \end{aligned}$$

Exercise 7 asks you to verify the last step.

For example, the primes dividing 24 are 2 and 3, so $\varphi(24) = 24(1-\frac{1}{2})(1-\frac{1}{3}) = 8$. The eight numbers between 1 and 24 that are relatively prime to 24 are 1, 5, 7, 11, 13, 17, 19, and 23.

### Counting Prime Numbers

Suppose $m$ is a composite positive integer, so $m$ can be written as a product of two integers that are both greater than 1: $m = ab$ with $1 < a \le b$. Then $a^2 \le m$, so $a \le \sqrt{m}$, and so $m$ must be divisible by a prime number $p$ with $p \le \sqrt{m}$.

We can use this observation, together with Theorem 2.6, to count the prime numbers between 1 and a given positive integer $n$. We start with the set of integers $\{1, 2, \ldots, n\}$, and use the theorem to count the number of elements that remain when multiples of prime numbers $p \le \sqrt{n}$ are excluded from the set. Since every composite number $m \le n$ has a prime factor $p \le \sqrt{m}$, excluding all of these numbers removes all the composite numbers from the set.

For example, for $n = 120$, the largest prime less than or equal to $\sqrt{n}$ is the fourth prime number, 7, so we require just four properties in Theorem 2.6 to exclude all the composite numbers in the set $\{1, 2, \ldots, 120\}$. The four properties are $a_1 =$ "is even," $a_2 =$ "is divisible by 3," $a_3 =$ "is divisible by 5," and $a_4 =$ "is divisible by 7." We compute $N(a_1) = 120/2 = 60$, $N(a_2) = 120/3 = 40$, $N(a_3) = 120/5 = 24$, and $N(a_4) = \lfloor 120/7 \rfloor = 24$. (The quantity $\lfloor x \rfloor$ was defined on page 153.)

Continuing our calculation, we compute $N(a_1 a_2) = \lfloor 120/6 \rfloor = 20$, then $N(a_1 a_3) = \lfloor 120/10 \rfloor = 12$, etc., and find that $N_0 = 120 - (60 + 40 + 24 + 17) + (20 + 12 + 8 + 8 + 5 + 3) - (4 + 2 + 1 + 1) + 0 = 27$. But this is not the number of prime numbers between 1 and 120, for our method excludes the primes 2, 3, 5, and 7, and includes the nonprime 1. Accounting for these exceptions, we find that the number of primes between 1 and 120 is $N_0 + 4 - 1 = 30$.

### Chromatic Polynomials

Let $G$ be a graph. Recall that its chromatic polynomial $c_G(x)$ measures the number of ways to color the vertices of $G$ using at most $x$ colors in such a way that no two vertices connected by an edge have the same color. We can use Theorem 2.6 to compute chromatic polynomials.

Suppose $G$ has $n$ vertices, and consider the set of colorings of the vertices of $G$ using at most $x$ colors, so the number of colorings in this set is $N = x^n$. To find $c_G(x)$, we must exclude all of the inadmissible colorings from this set. For each edge $e_i$ in the graph, select property $a_i$ to be "edge $e_i$ connects two vertices that have the same color." In this way, the colorings in the set that satisfy none of the properties are precisely the admissible colorings, so $N_0 = c_G(x)$.

For example, we compute the chromatic polynomial for the complete graph $K_3$ using this strategy. This graph has three edges, so we take $r = 3$ in the theorem. We compute $N(a_1) = N(a_2) = N(a_3) = x^2$, since every coloring satisfying one of the properties has two vertices with the same color, and the third vertex may be any color. Also, $N(a_1 a_2) = N(a_2 a_3) = N(a_1 a_3) = N(a_1 a_2 a_3) = x$, as every coloring satisfying more than one property must have all vertices colored identically. Thus, $c_{K_3}(x) = N_0 = x^3 - 3x^2 + 3x - x = x(x-1)(x-2) = x^{\underline{3}}$.

### Derangements

Suppose a lazy professor gives a quiz to a class of $n$ students, then collects the papers, shuffles them, and redistributes them randomly to the class for grading. The professor would prefer that no student receives his or her own paper to grade. What is the probability that this occurs? Is this probability substantially different for different class sizes? What do you think the limiting probability is as $n \to \infty$? Notice that as $n$ grows larger, there are more ways for at least one person to receive his or her own quiz back, but perhaps this increase is swamped by the growth of the total number of permutations possible.

Suppose we have $n$ objects in an initial configuration. A permutation of these objects in which the position of each object differs from its initial position is called a *derangement* of the objects. Since $n!$ denotes the number of permutations of $n$ objects, following [133] we denote the number of derangements of $n$ objects by $n¡$ (and since $n!$ is often pronounced "$n$ bang," perhaps $n¡$ should be pronounced "$n$ gnab").

We compute $n¡$ for some small values of $n$. For $n = 0$, there is just one permutation, and it vacuously satisfies the derangement condition, so $0¡ = 1$. There is only one permutation of a single object, and it is not a derangement, so $1¡ = 0$. Only one of the two permutations of two objects is a derangement, so $2¡ = 1$, and exactly two of the six permutations of three objects satisfies the condition: If our original arrangement is $[1, 2, 3]$, then the derangements are $[2, 3, 1]$ and $[3, 1, 2]$. Thus $3¡ = 2$. We find that $4¡ = 9$: The derangements of $[1, 2, 3, 4]$ are $[2, 1, 4, 3]$, $[2, 3, 4, 1]$, $[2, 4, 1, 3]$, $[3, 1, 4, 2]$, $[3, 4, 1, 2]$, $[3, 4, 2, 1]$, $[4, 1, 2, 3]$, $[4, 3, 1, 2]$, and $[4, 3, 2, 1]$. Thus, the probability that a random permutation of a fixed number $n$ of objects is a derangement is respectively 1, 0, $\frac{1}{2}$, $\frac{1}{3}$, and $\frac{3}{8}$ for $n = 0$ through 4.

We can use Theorem 2.6 to determine a formula for $n¡$. We select the original set to be the collection of all permutations of $n$ objects, and for $1 \le i \le n$ let $a_i$ denote the property that element $i$ remains in its original position in a permutation. Then $N_0$ is the number of permutations where no elements remain in their original position, so $N_0 = n¡$.

To compute $N(a_i)$, we see that element $i$ is fixed, but the other $n - 1$ elements may be arranged arbitrarily, so $N(a_i) = (n-1)!$. Similarly, $N(a_i a_j) = (n-2)!$ for $i < j$, $N(a_i a_j a_k) = (n-3)!$ for $i < j < k$, and so on. Therefore,

$$\begin{aligned} n¡ &= n! - \sum_{i} (n-1)! + \sum_{i<j} (n-2)! - \cdots \\ &\quad + (-1)^m \sum_{i_1 < \cdots < i_m} (n-m)! + \cdots + (-1)^n. \end{aligned}$$

Since the number of different $m$-tuples $(i_1, i_2, \ldots, i_m)$ with $1 \le i_1 < i_2 < \cdots < i_m \le n$ is $\binom{n}{m}$, we obtain

$$\begin{aligned} n¡ &= n! - \binom{n}{1}(n-1)! + \binom{n}{2}(n-2)! - \cdots \\ &\quad + (-1)^m \binom{n}{m}(n-m)! + \cdots + (-1)^n \\ &= \sum_m (-1)^m \binom{n}{m}(n-m)! \\ &= n! \sum_{m=0}^{n} \frac{(-1)^m}{m!}. \end{aligned}$$

Thus, the probability that a permutation of $n$ objects is a derangement is

$$\frac{n¡}{n!} = \sum_{m=0}^{n} \frac{(-1)^m}{m!},$$

and in the limit,

$$\lim_{n\to\infty} \frac{n¡}{n!} = \sum_{m\ge 0} \frac{(-1)^m}{m!} = e^{-1}. \tag{2.27}$$

Our lazy professor obtains a desired configuration about $36.8\%$ of the time, for sizable classes.

**Exercises**

1. A noted vexillologist tells you that 30 of the $50$ U.S. state flags have blue as a background color, twelve have stripes, 26 exhibit a plant or animal, nine have both blue in the background and stripes, 23 have both blue in the background and feature a plant or animal, and three have both stripes and a plant or animal. One of the flags in this last category (California) does not have any blue in the background. How many state flags have no blue in the background, no stripes, and no plant or animal featured?

2. Suppose $50$ socks lie in a drawer. Each one is either white or black, ankle-high or knee-high, and either has a hole or doesn't. 22 socks are white, four of these have a hole, and one of these four is knee-high. Ten white socks are knee-high, ten black socks are knee-high, and five knee-high socks have a hole. Exactly three ankle-high socks have a hole.

   (a) Use Theorem 2.6 to determine the number of black, ankle-high socks with no holes.

   (b) Draw a Venn diagram that shows the number of socks with each combination of characteristics.

3. The buffet line at a local steakhouse has 35 dishes. Sixteen dishes contain meat, fourteen dishes are fried, and of the dishes with meat, eight contain vegetables and seven are fried. Of the fried dishes, five contain a vegetable. Just two dishes are fried and contain both meat and a vegetable, and ten dishes (principally in the dessert section) contain neither meat nor a vegetable and are not fried. Use Theorem 2.6 to determine how many dishes contain vegetables.

4. A sneaky registrar reports the following information about a group of 400 students. There are 180 taking a math class, 200 taking an English class, 160 taking a biology class, and 250 in a foreign language class. 80 are enrolled in both math and English, 90 in math and biology, 120 in math and a foreign language, 70 in English and biology, 140 in English and a foreign language, and 60 in biology and a foreign language. Also, there are 25 in math, English, and a foreign language, 30 in math, English, and biology, 40 in math, biology, and a foreign language, and fifteen in English, biology, and a foreign language. Finally, the sum of the number of students with a course in all four subjects, plus the number of students with a course in none of the four subjects, is 100. Use Theorem 2.6 to determine the number of students that are enrolled in all four subjects simultaneously: math, biology, English, and a foreign language.

5. On a busy evening a number of guests visit a gourmet restaurant, and everyone orders something. 140 guests order a beverage, 190 order an entree, 100 order an appetizer, 90 order a dessert, 65 order a beverage and an appetizer, 125 order a beverage and an entree, 60 order a beverage and a dessert, 85 order an entree and an appetizer, 75 order an entree and a dessert, 60 order an appetizer and a dessert, 40 order a beverage, appetizer, and dessert, 55 order a beverage, entree, and dessert, 45 order an appetizer, entree, and dessert, 35 order a beverage, entree, and appetizer, and ten order all four types of items. Use Theorem 2.6 to determine the number of guests who visited the restaurant that evening.

6. Use Theorem 2.6 to determine the number of five-card hands drawn from a standard deck that contain at least one card from each of the four suits.

7. Let $\alpha_1, \alpha_2, \ldots, \alpha_r$ be real numbers. Show that

$$\prod_{i=1}^{r}(1-\alpha_i) = 1 - \sum_{i}\alpha_i + \sum_{i<j}\alpha_i\alpha_j - \sum_{i<j<k}\alpha_i\alpha_j\alpha_k + \cdots + (-1)^r\alpha_1\alpha_2\cdots\alpha_r.$$

8. (a) Show that $\varphi(mn) = \varphi(m)\varphi(n)$ if $m$ and $n$ are relatively prime.

   (b) Show that $\varphi(mn) \neq \varphi(m)\varphi(n)$ if $m$ and $n$ are not relatively prime. Is one quantity always larger than the other in this case?

(c) Determine all integers $n$ satisfying $\varphi(n) = 12$, 13, or 14.

9. Use Theorem 2.6 to count the number of prime numbers less than 168.

10. Use Theorem 2.6 to determine the chromatic polynomial for each of the following graphs.

    (a) The yield sign (add a single edge to the bipartite graph $K_{1,3}$).

    (b) The bipartite graph $K_{2,3}$.

11. Is the probability that a permutation of $n$ objects is a derangement substantially different for $n = 12$ and $n = 120$? Quantify your answer.

12. (Deranged twins.) Suppose $n + 2$ people are seated behind a long table facing an audience to staff a panel discussion. Two of the people are identical twins, wearing identical clothing. At intermission, the panelists decide to rearrange themselves so that it will be apparent to the audience that everyone has moved to a different seat when the panel reconvenes. Each twin can therefore take neither her own former place, nor her twin's. Let $T_n$ denote the number of different ways to derange the panel in this way.

    (a) Compute $T_0$, $T_1$, $T_2$, and $T_3$.

    (b) Compute $T_4$.

    (c) Determine a formula for $T_n$, and check that your formula produces $T_{10} = 72\,755\,370$.

    (d) Compute the value of $\lim\limits_{n\to\infty} \dfrac{T_n}{(n+2)!}$.

13. Suppose our lazy professor collects a quiz and a homework assignment from a class of $n$ students one day, then distributes both the quizzes and the homework assignments back to the class in a random fashion for grading. Each student receives one quiz and one homework assignment to grade.

    (a) What is the probability that every student receives someone else's quiz to grade, and someone else's homework to grade?

    (b) What is the probability that no student receives both their own quiz and their own homework assignment to grade? In this case, some students may receive their own quiz, and others may receive their own homework assignment.

    (c) Compute the limiting probability as $n \to \infty$ in each case.

14. Let $N_m$ denote the number of objects from a collection of $N$ objects that possess exactly $m$ of the properties $a_1, a_2, \ldots, a_r$. Generalize the principle of inclusion and exclusion by showing that

$$N_m = \sum_{k=m}^{r} (-1)^{k-m} \binom{k}{m} s_k, \tag{2.28}$$

where

$$s_k = \sum_{i_1 < \cdots < i_k} N(a_{i_1} \ldots a_{i_k}). \tag{2.29}$$

## 2.6 Generating Functions

*And own no other function: each your doing,*
*So singular in each particular,*
*Crowns what you are doing in the present deed,*
*That all your acts are queens.*
— William Shakespeare, *The Winter's Tale*, Act IV, Scene IV

Given a sequence $\{a_k\}$ with $k \geq 0$, its *generating function* $G(x)$ is defined by

$$G(x) = \sum_{k \geq 0} a_k x^k. \tag{2.30}$$

Thus, $G(x)$ is a polynomial if $\{a_k\}$ is a finite sequence, and a power series if $\{a_k\}$ is infinite. For example, if $a_k = (-1)^k/k!$, then $G(x)$ is the Maclaurin series for $e^{-x}$, and if $a_k = \binom{n}{k}$ for a fixed nonnegative integer $n$, then $G(x) = (1+x)^n$, by the binomial theorem.

To illustrate how generating functions can be used to solve combinatorial problems, let us consider again the problem of determining the number of $k$-element subsets of an $n$-element set. Fix $n$, and let $a_k$ denote this number. Of course, we showed in Section 2.1 that $a_k = n^{\underline{k}}/k!$; here we derive this formula again using generating functions.

Suppose we wish to enumerate all subsets of the $n$-element set. To construct one subset, we must pick which elements to include in the subset and which to exclude. Let us denote the choice to omit an element by $x^0$, and the choice to include it by $x^1$. Using "+" to represent "or," the choice to include or exclude one element then is denoted by $x^0 + x^1$. We must make $n$ such choices to construct a subset, so using multiplication to denote "and," the expression $(x^0 + x^1)^n$ models the choices required to make a subset.

Since "and" distributes over "or" just as multiplication distributes over addition, we may expand this expression using standard rules of arithmetic to obtain representations for all $2^n$ subsets. For example, when $n = 3$, we obtain

$$\begin{aligned}\left(x^0 + x^1\right)^3 &= x^0x^0x^0 + x^0x^0x^1 + x^0x^1x^0 + x^0x^1x^1 \\ &\quad + x^1x^0x^0 + x^1x^0x^1 + x^1x^1x^0 + x^1x^1x^1.\end{aligned}$$

The first term represents the empty subset, the second signifies the subset containing just the third item in the original set, etc. Writing 1 for $x^0$ and $x$ for $x^1$ and treating the expression as a polynomial, we find that $(1+x)^3 = 1+3x+3x^2+x^3$, and the coefficient of $x^k$ is the number of subsets of a three-element set having exactly $k$ items.

In general, we find that the generating function for the sequence $\{a_k\}$ is $(1+x)^n$, so $a_k = \binom{n}{k} = n^{\underline{k}}/k!$, by the binomial theorem. Since our proof of the binomial theorem relies only on basic facts of arithmetic, this argument provides an independent derivation for the number of $k$-element subsets of a set with $n$ elements.

This example illustrates the general strategy for using generating functions to solve combinatorial problems. First, express the problem in terms of determining one or more values of an unknown sequence $\{a_k\}$. Second, determine a generating function for this sequence, writing the monomial $x^k$ to represent selecting an object $k$ times, then using addition to represent alternative choices and multiplication to represent sequential choices. Third, use analytic methods to expand the generating function and determine the values of the encoded sequence.

For example, suppose a drawer contains twelve beads: three red, four blue, and five green, and suppose we wish to determine the number of ways to select six beads from a drawer, if beads of the same color are indistinguishable, and the order of selection is irrelevant. Let $a_k$ denote the number of ways to select $k$ beads from the drawer. Then the generating function for this sequence is

$$\begin{aligned} G(x) &= (1+x+x^2+x^3)(1+x+x^2+x^3+x^4) \\ &\qquad \cdot (1+x+x^2+x^3+x^4+x^5) \\ &= 1+3x+6x^2+10x^3+14x^4+17x^5+18x^6 \\ &\qquad +17x^7+14x^8+10x^9+6x^{10}+3x^{11}+x^{12}. \end{aligned}$$

For example, we see from this that there are exactly $a_6 = 18$ ways to select six beads from the drawer. Indeed, we can check this by constructing all such selections:

$$\begin{array}{llllll} rrrggg, & rrrggb, & rrrgbb, & rrrbbb, & rrgggg, & rrgggb, \\ rrggbb, & rrgbbb, & rrbbbb, & rggggb, & rgggbb, & rggbbb, \\ rgbbbb, & rbbbbb, & ggggbb, & gggbbb, & ggbbbb, & gbbbbb. \end{array} \tag{2.31}$$

In the following sections, we explore the power of this method by studying several combinatorial problems.

**Exercises**

1. In this problem, we verify that the arithmetic operations performed in generating functions model the logical selections made in combinatorial problems. Write $a^k$ to denote selecting $k$ copies of object $a$.

   (a) Clearly, there are exactly four different subsets of the set $\{a, b\}$. We can model the construction of the different possible subsets of this two-element set by considering two choices: Pick $a$ or not, and then pick $b$ or not. Thus, we can denote all the possible choices by writing: ($a^0$ or $a^1$) and ($b^0$ or $b^1$). Expand this expression using the logical rule "($P$ or $Q$) and $R \equiv$ ($P$ and $R$) or ($Q$ and $R$)". Continue expanding until you obtain an expression of the form "$C_1$ or $C_2$ or $C_3$ or $C_4$," where each $C_i$ is a logical expression involving only and's.

(b) Rewrite exactly the same logical computation, but now use $x^0$ in place of $a^0$ or $b^0$, $x^1$ in place of $a^1$ or $b^1$, $+$ instead of "or", and $*$ instead of "and". Then simplify the expression by combining exponents in the usual way.

(c) Repeat this procedure for the three-element set $\{a, b, c\}$.

(d) Repeat this procedure for sub-multisets of the three-element multiset $\{a, a, b\}$.

2. Suppose a drawer contains three red beads, four blue beads, and five green beads. Use a generating function to determine the number of ways to select six beads if one must select at least one red bead, an odd number of blue beads, and an even number of green beads. Then check your answer using the combinations shown in (2.31). Assume that beads of the same color are indistinguishable, and that the order of selection is irrelevant.

3. Suppose a drawer contains ten red beads, eight blue beads, and eleven green beads. Determine a generating function that encodes the answer to each of the following problems.

   (a) The number of ways to select $k$ beads from the drawer.

   (b) The number of ways to select $k$ beads if one must obtain an even number of red beads, an odd number of blue beads, and a prime number of green beads.

   (c) The number of ways to select $k$ beads if one must obtain exactly two red beads, at least five blue beads, and at most four green beads.

### 2.6.1 Double Decks

*I don't like the games you play, professor.*

— Roger Thornhill, in *North by Northwest*

How many five-card poker hands can be dealt from a double deck? Assume that the two decks are identical. More generally, how many ways are there to select $m$ items from $n$ different items, where each item can be selected at most twice? Let us denote this number by $t_{n,m}$, and let $G_n(x)$ be the generating function for the sequence $\{t_{n,m}\}$ for $m \geq 0$ and $n$ fixed.

We find that $G_n(x) = \left(1 + x + x^2\right)^n$, since each object may be selected zero times, one time, or two times. To find $t_{n,m}$, we must determine a formula for the coefficient of $x^m$ in $G_n(x)$. This is simply a matter of applying the binomial theorem twice:

$$
\begin{aligned}
G_n(x) &= \bigl(1 + (x + x^2)\bigr)^n \\
&= \sum_k \binom{n}{k} \left(x + x^2\right)^k \\
&= \sum_k \binom{n}{k} x^k \sum_j \binom{k}{j} x^j \\
&= \sum_k \sum_j \binom{n}{k} \binom{k}{j} x^{j+k} \\
&= \sum_m \left( \sum_j \binom{n}{m-j} \binom{m-j}{j} \right) x^m,
\end{aligned}
$$

where we obtained the last line by substituting $m$ for $j + k$. Therefore,

$$
t_{n,m} = \sum_{j=0}^{\lfloor m/2 \rfloor} \binom{n}{m-j} \binom{m-j}{j}. \tag{2.32}
$$

The number of five-card poker hands that can be dealt from a double deck is then

$$
\binom{52}{5}\binom{5}{0} + \binom{52}{4}\binom{4}{1} + \binom{52}{3}\binom{3}{2} = 3\,748\,160.
$$

There is a simple combinatorial explanation for this expression. A five-card hand dealt from a double deck may have zero, one, or two cards repeated. There are $\binom{52}{5}$ hands with no cards repeated, $\binom{52}{4}\binom{4}{1}$ hands with exactly one card repeated, and $\binom{52}{3}\binom{3}{2}$ hands with exactly two cards repeated. A similar explanation applies for the general formula (2.32).

**Exercises**

1. Derive (2.32) by using the multinomial theorem to expand $G_n(x)$.

2. Use a combinatorial argument to count the number of different five-card hands that can be dealt from a triple deck, then the number of five-card hands that can be dealt from a quadruple deck.

3. Use a combinatorial argument to count the number of different six-card hands that can be dealt from $r$ combined decks, for each positive integer $r$.

4. Use a generating function to determine the number of ways to select a hand of $m$ cards from a triple deck, if there are $n$ distinct cards in a single deck. Verify that your expression produces the correct answers when $n = 52$ and $m = 5$ or $m = 6$.

### 2.6.2 Counting with Repetition

*Then, shalt thou count to three, no more, no less. Three shalt be the number thou shalt count, and the number of the counting shall be three. Four shalt thou not count, nor either count thou two, excepting that thou then proceed to three. Five is right out.*
— *Monty Python and the Holy Grail*

Suppose there is an inexhaustible supply of each of $n$ different objects. How many ways are there to select $m$ objects from the $n$ different objects, if you are allowed to select each object as many times as you like?

Let $a_{n,m}$ denote this number. Evidently, for fixed $n$, the generating function for $\{a_{n,m}\}_{m\geq 0}$ is

$$G_n(x) = \left(1 + x + x^2 + \cdots\right)^n = \left(\frac{1}{1-x}\right)^n,$$

since the sum is just a geometric series in $x$. This raises questions on convergence, for this formula is valid only for $|x| < 1$. We largely ignore these analytic issues, since we treat generating functions as formal series.

Thus, to find a formula for $a_{n,m}$, we must find the coefficient of $x^m$ in $G_n(x)$.

Let us consider a more general problem. Let $f(x) = (1 + x)^\alpha$, where $\alpha$ is a real number. Then $f'(0) = \alpha$, $f''(0) = \alpha(\alpha - 1)$, and in general, $f^{(k)}(0) = \alpha^{\underline{k}}$. Therefore, the Maclaurin series for $f(x)$ is

$$(1 + x)^\alpha = \sum_{k\geq 0} \frac{\alpha^{\underline{k}}}{k!} x^k.$$

Define the *generalized binomial coefficient* by

$$\binom{\alpha}{k} = \begin{cases} \alpha^{\underline{k}}/k! & \text{if } k \geq 0, \\ 0 & \text{if } k < 0. \end{cases} \tag{2.33}$$

Note that $\binom{\alpha}{k}$ equals the ordinary binomial coefficient whenever $\alpha$ is a nonnegative integer. We have the following theorem.

**Theorem 2.7** (Generalized Binomial Theorem)**.** *If* $|x| < 1$ *or* $\alpha$ *is a nonnegative integer, then*

$$(1 + x)^\alpha = \sum_k \binom{\alpha}{k} x^k. \tag{2.34}$$

The proof of convergence may be found in many analysis texts, where it is often proved as a consequence of Bernstein's theorem on convergence of Taylor series (see for instance [11]). We do not supply the proof here.

Before solving our problem concerning selection with unlimited repetition, we note a useful identity for generalized binomial coefficients.

**Negating the Upper Index.** *If $\alpha$ is a real number and $k$ is an integer, then*

$$\binom{\alpha}{k} = (-1)^k \binom{k-\alpha-1}{k}. \tag{2.35}$$

*Proof.* For $k < 0$, the identity is clear. For $k \geq 0$, we have

$$\binom{\alpha}{k} = \frac{1}{k!}\prod_{i=0}^{k-1}(\alpha - i).$$

Reindex this product, replacing each $i$ by $k-1-i$, to obtain

$$\begin{aligned}\binom{\alpha}{k} &= \frac{1}{k!}\prod_{i=0}^{k-1}(\alpha-(k-i-1))\\ &= \frac{(-1)^k}{k!}\prod_{i=0}^{k-1}(k-1-i-\alpha)\\ &= (-1)^k\binom{k-\alpha-1}{k}.\end{aligned}$$ □

We may now solve our problem of determining $a_{n,m}$. We compute

$$\begin{aligned}G_n(x) &= (1-x)^{-n}\\ &= \sum_m \binom{-n}{m}(-x)^m\\ &= \sum_m \binom{n+m-1}{m}x^m,\end{aligned}$$

and therefore the number of ways to select $m$ objects from a collection of $n$ different objects, with repetition allowed, is

$$a_{n,m} = \binom{n+m-1}{m}. \tag{2.36}$$

For example, the number of five-card poker hands that can be dealt from a stack of five or more decks is $\binom{56}{5} = 3\,819\,816$.

Finally, suppose we lay all $52$ cards of a standard deck face up on a table. How many ways can we place five identical poker chips on the cards if we allow more than one chip to be placed on each card? To solve this, notice that each possible placement of chips corresponds to a hand of five cards, where repeated cards are allowed: If $k$ chips lie on a particular card, place that card into the hand $k$ times. Further, every such five-card hand can be represented by a judicious placement of chips. Therefore, the answer is the same as that of the previous example, $\binom{56}{5}$.

In general, the number of ways to place $m$ identical objects into $n$ distinguishable bins is the same as the number of ways to select $m$ objects from a set of $n$ objects with repetition allowed: The answer to both problems is $\binom{n+m-1}{m}$.

**Exercises**

1. Prove the addition identity for generalized binomial coefficients: If $\alpha$ is a real number and $k$ is an integer, then
$$\binom{\alpha}{k} = \binom{\alpha-1}{k} + \binom{\alpha-1}{k-1}.$$

2. Prove the absorption/extraction identity for generalized binomial coefficients: If $\alpha$ is a real number and $k$ is a nonzero integer, then
$$\binom{\alpha}{k} = \frac{\alpha}{k}\binom{\alpha-1}{k-1}.$$

3. Prove the cancellation identity for generalized binomial coefficients: If $\alpha$ is a real number and $k$ and $m$ are integers, then
$$\binom{\alpha}{k}\binom{k}{m} = \binom{\alpha}{m}\binom{\alpha-m}{k-m}.$$

4. Prove the parallel summation identity for generalized binomial coefficients: If $\alpha$ is a real number and $n$ is an integer, then
$$\sum_{k=0}^{n}\binom{\alpha+k}{k} = \binom{\alpha+n+1}{n}.$$

5. Suppose that an unlimited number of jelly beans is available in each of five different colors: red, green, yellow, white, and black.
   (a) How many ways are there to select twenty jelly beans?
   (b) How many ways are there to select twenty jelly beans if we must select at least two jelly beans of each color?

6. A catering company brings fifty identical hamburgers to a party with twenty guests.
   (a) How many ways can the hamburgers be divided among the guests, if none is left over?
   (b) How many ways can the hamburgers be divided among the guests, if every guest receives at least one hamburger, and none is left over?
   (c) Repeat these problems if there may be burgers left over.

7. A zodiac sign is one of twelve constellations that the sun travels through (from the vantage point of the earth) over the course of a year. Each person has a zodiac sign based on the position of sun on their birth date. The *astrological configuration* of a party with $n$ guests is a list of twelve numbers that records the number of guests with each sign, so the first number records the number of people with the sign Capricorn, the second, Aquarius, ..., the last, Sagittarius.

(a) How many different astrological configurations are possible for $n = 100$?

(b) How many astrological configurations are possible for $n = 100$, if each component is at least $5$?

8. Two lottery systems are proposed for a new state lottery. In the first system, players select six *different* numbers from $\{1, 2, \ldots, 50\}$. In the second system, players select six numbers from $\{1, 2, \ldots, 45\}$, *and* may select any number as many times as they want. (In the second system, each ball selected in the lottery drawing is replaced before another ball is selected.) Which system has more possible tickets?

9. Suppose $100$ identical tickets for rides are distributed among $40$ children at a carnival.

   (a) How many ways can the tickets be distributed, if each child receives at least two tickets, and all the tickets are distributed?

   (b) How many ways can the tickets be distributed, if each child receives at least one ticket, and some tickets may be left over?

   (c) Suppose one child has twelve tickets, and each ticket may be used on any of six different rides. How many ways can the child spend her tickets, if she can choose any ride any number of times, and the order of choice is unimportant?

### 2.6.3 Changing Money

*Jesus went into the temple, and began to cast out them that sold and bought in the temple, and overthrew the tables of the moneychangers …*

— *Mark 11:15*

We now turn to a problem popularized by the analyst and combinatorialist George Pólya [225]: How many ways are there to change a dollar? That is, how many combinations of pennies, nickels, dimes, quarters, half-dollars, and dollar coins total \$1? Our discussion of this problem follows the treatment of Graham, Knuth, and Patashnik [133].

Let us define $a_k$ to be the number of ways to make $k$ cents in change, and let $A(x)$ be a generating function for $a_k$: $A(x) = \sum_k a_k x^k$. Before analyzing this problem, pause a moment and make a guess. Do you think $a_{50}$ is more than $50$ or less than $50$? Is $a_{100}$ more than $100$ or less than $100$? How fast do you think $a_k$ grows as a function of $k$? Is it a polynomial in $k$? Exponential in $k$? Perhaps something between these?

To create a pile of change, we must make six choices, selecting a number of pennies, then nickels, then dimes, quarters, half-dollars, and dollars. We can model our choice of pennies by the sum

$$1 + x + x^2 + x^3 + \cdots = \frac{1}{1-x}.$$

One might be tempted to use the same expression to model the different choices for each of the other coins, since we can pick any number of nickels, and any number of dimes, etc., but this would be incorrect. This would yield a generating function for the number of ways to select $k$ coins from a set of six different coins, not the number of ways to form $k$ cents. Instead, when choosing nickels, we select either zero cents, or five cents, or ten cents, and so on, so this selection is modeled as

$$1 + x^5 + x^{10} + x^{15} + \cdots = \frac{1}{1-x^5}.$$

Therefore, the number of ways to make $k$ cents using either pennies or nickels is given by the generating function

$$\frac{1}{(1-x)(1-x^5)}.$$

Continuing in this way, we find that

$$A(x) = \frac{1}{(1-x)(1-x^5)(1-x^{10})(1-x^{25})(1-x^{50})(1-x^{100})}, \tag{2.37}$$

so we merely need to find the coefficient of $a_k$ in the Maclaurin series for $A(x)$! This sounds rather daunting, so let us determine a few values of $a_k$ by hand first.

Let $\{p_k\}$ denote the number of ways to make $k$ cents using only pennies, so $p_k = 1$ for all $k$. Let $P(x)$ be the generating function for $\{p_k\}$, so $P(x) = 1/(1-x)$. Let $n_k$ be the number of ways to make $k$ cents using either pennies or nickels, so its generating function is

$$N(x) = \frac{P(x)}{1-x^5}.$$

Thus $N(x) = P(x) + x^5 N(x)$, and by equating coefficients we find that

$$n_k = \begin{cases} p_k & \text{if } 0 \le k \le 4, \\ p_k + n_{k-5} & \text{if } k \ge 5. \end{cases}$$

In the same way, let $d_k$ denote the number of ways to make $k$ cents using pennies, nickels, or dimes, and let $D(x)$ be its generating function. We then have $D(x) = N(x) + x^{10} D(x)$, and so

$$d_k = \begin{cases} n_k & \text{if } 0 \le k \le 9, \\ n_k + d_{k-10} & \text{if } k \ge 10. \end{cases}$$

There is a simple combinatorial interpretation for this equation. If $k < 10$, then we can choose only nickels and pennies to form $k$ cents, so $d_k = n_k$ in this case. If $k \geq 10$, we may form $k$ cents using only nickels and pennies, or we can choose one dime, then form the remaining $k-10$ cents using dimes, nickels, and pennies. Thus $d_k = n_k + d_{k-10}$ in this case.

Similarly, using $q_k$ for allowing quarters, $h_k$ for half dollars, and finally $a_k$ for dollar coins, we have

$$q_k = \begin{cases} d_k & \text{if } 0 \leq k \leq 24, \\ d_k + q_{k-25} & \text{if } k \geq 25; \end{cases}$$

$$h_k = \begin{cases} q_k & \text{if } 0 \leq k \leq 49, \\ q_k + h_{k-50} & \text{if } k \geq 50; \end{cases}$$

$$a_k = \begin{cases} h_k & \text{if } 0 \leq k \leq 99, \\ h_k + a_{k-100} & \text{if } k \geq 100. \end{cases}$$

We may use these formulas to construct Table 2.2 below showing the number of ways to make $k$ cents with the different coin sets.

| $k$ | 0 | 5 | 10 | 15 | 20 | 25 | 30 | 35 | 40 | 45 | 50 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $p_k$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| $n_k$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| $d_k$ | 1 | 2 | 4 | 6 | 9 | 12 | 16 | 20 | 25 | 30 | 36 |
| $q_k$ | 1 | | | | | 13 | | | | | 49 |
| $h_k$ | 1 | | | | | | | | | | 50 |
| $a_k$ | 1 | | | | | | | | | | |

| $k$ | 55 | 60 | 65 | 70 | 75 | 80 | 85 | 90 | 95 | 100 |
|---|---|---|---|---|---|---|---|---|---|---|
| $p_k$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| $n_k$ | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 |
| $d_k$ | 42 | 49 | 56 | 64 | 72 | 81 | | 100 | | 121 |
| $q_k$ | | | | | 121 | | | | | 242 |
| $h_k$ | | | | | | | | | | 292 |
| $a_k$ | | | | | | | | | | 293 |

TABLE 2.2. Computing the number of ways to make $k$ cents in change.

We find that there are precisely $50$ ways to make $50$ cents in change, and 293 ways to make one dollar in change.

This is a fairly efficient method to determine $a_k$, since apparently we can calculate this number using at most $5k$ arithmetic operations. But we can do much better! We can compute $a_k$ using at most a constant number of arithmetic operations, regardless of the value of $k$. To show this, let us first simplify $A(x)$ by exploiting the fact that all but one of the exponents in (2.37) is a multiple of 5. Let

$$B(x) = \frac{1}{(1-x)^2(1-x^2)(1-x^5)(1-x^{10})(1-x^{20})},$$

so that

$$A(x) = (1+x+x^2+x^3+x^4)B(x^5).$$

Writing $b_k$ for the coefficient of $x^k$ in the Maclaurin series for $B(x)$ and equating coefficients, we find that

$$b_k = a_{5k} = a_{5k+1} = a_{5k+2} = a_{5k+3} = a_{5k+4}.$$

But this makes sense (or perhaps cents?), since the last few cents can be represented only using pennies. Now

$$B(x) = \frac{C(x)}{(1-x^{20})^6},$$

where

$$\begin{aligned}
C(x) = {} & \left(1+x+\cdots+x^{19}\right)^2\left(1+x^2+\cdots+x^{18}\right)\left(1+x^5+x^{10}+x^{15}\right) \\
& \cdot \left(1+x^{10}\right) \\
= {} & x^{81} + 2x^{80} + 4x^{79} + 6x^{78} + 9x^{77} + 13x^{76} + 18x^{75} + 24x^{74} + 31x^{73} \\
& + 39x^{72} + 50x^{71} + 62x^{70} + 77x^{69} + 93x^{68} + 112x^{67} + 134x^{66} \\
& + 159x^{65} + 187x^{64} + 218x^{63} + 252x^{62} + 287x^{61} + 325x^{60} + 364x^{59} \\
& + 406x^{58} + 449x^{57} + 493x^{56} + 538x^{55} + 584x^{54} + 631x^{53} + 679x^{52} \\
& + 722x^{51} + 766x^{50} + 805x^{49} + 845x^{48} + 880x^{47} + 910x^{46} + 935x^{45} \\
& + 955x^{44} + 970x^{43} + 980x^{42} + 985x^{41} + 985x^{40} + 980x^{39} + 970x^{38} \\
& + 955x^{37} + 935x^{36} + 910x^{35} + 880x^{34} + 845x^{33} + 805x^{32} + 766x^{31} \\
& + 722x^{30} + 679x^{29} + 631x^{28} + 584x^{27} + 538x^{26} + 493x^{25} + 449x^{24} \\
& + 406x^{23} + 364x^{22} + 325x^{21} + 287x^{20} + 252x^{19} + 218x^{18} + 187x^{17} \\
& + 159x^{16} + 134x^{15} + 112x^{14} + 93x^{13} + 77x^{12} + 62x^{11} + 50x^{10} \\
& + 39x^9 + 31x^8 + 24x^7 + 18x^6 + 13x^5 + 9x^4 + 6x^3 + 4x^2 + 2x + 1.
\end{aligned}$$

We know from the previous section that

$$\frac{1}{(1-z)^n} = \sum_k \binom{n+k-1}{n-1} z^k,$$

so

$$B(x) = C(x)\sum_k \binom{k+5}{5} x^{20k}. \tag{2.38}$$

Therefore, writing $C(x) = \sum_k c_k x^k$, we have $a_{100} = b_{20} = c_0\binom{6}{5} + c_{20}\binom{5}{5} = 6 + 287 = 293$, and

$$\begin{aligned} a_{1000} &= b_{200} \\ &= c_0\binom{15}{5} + c_{20}\binom{14}{5} + c_{40}\binom{13}{5} + c_{60}\binom{12}{5} + c_{80}\binom{11}{5} \\ &= 2\,103\,596. \end{aligned}$$

Our expression for computing $a_k$ is a sum having at most five terms, so this method allows us to compute $a_k$ using only a constant number of operations.

Finally, consider the crazy system of coinage where there is a coin minted worth $n$ cents for every $n \geq 1$. Let $p_n$ denote the number of ways to make $n$ cents in change in this system. For example, $p_4 = 5$, since we can make four cents by using four pennies, or one two-cent piece and two pennies, or one three-cent piece and one penny, or two two-cent pieces, or one four-cent piece. By representing these five possibilities as the sums $1 + 1 + 1 + 1$, $2 + 1 + 1$, $3 + 1$, $2 + 2$, and $4$, we see that $p_n$ is the number of ways to write $n$ as a sum of one of more positive integers, disregarding the order of the summands. Such a representation is called a *partition* of $n$. Evidently the generating function $P(x)$ for the sequence of partitions is given by the infinite product

$$P(x) = \prod_{k \geq 1} \frac{1}{1 - x^k}. \tag{2.39}$$

We explore this generating function and the sequence $\{p_n\}$ in Section 2.8.1.

**Exercises**

1. Use (2.38) to compute $a_{2009}$, the number of ways to make \$20.09 in change.

2. How many ways are there to select 100 coins from an inexhaustible supply of pennies, nickels, dimes, quarters, half-dollars, and dollar coins?

3. Show that the number of ways to make $10m$ cents in change using only pennies, nickels, and dimes is $(m + 1)^2$.

4. Show that $a_k$ can be computed using equation (2.38) using at most 60 arithmetic operations. Optimize your method to show that $a_k$ can be computed using at most 31 arithmetic operations.

5. Prove that $a_k$ grows like $k^5$ by showing that there exist positive constants $c$ and $C$ such that $ck^5 < a_k < Ck^5$ for sufficiently large $k$.

6. The following coins were in circulation in the United States in 1875: the Indian-head penny, a bronze two-cent piece (last minted in 1873), a silver three-cent piece (also last minted in 1873), a nickel three-cent piece, the

shield nickel (worth five cents), the seated liberty half-dime, dime, twenty-cent piece (produced for only four years beginning in 1875), quarter, half-dollar, and silver dollar, and the Indian-head gold dollar. (We ignore the trade dollar, minted for circulation between 1873 and 1878, as it was issued for overseas trade. This coin holds the distinction of being the only U.S. coin to be demonetized.)

(a) How many ways were there to make twenty cents in change in 1875? How about twenty-five cents? Compute these values using the tabular method of this section.

(b) Write down a generating function in the form of a rational function for the number of ways to make $k$ cents in change in 1875, then use a computer algebra system to find the number of ways to make one dollar in change in 1875.

7. (Inspired in part by [133, ex. 7.21].) A ransom note demands:

(i) \$10000 in unmarked fifty- and hundred-dollar bills, and

(ii) the number of ways to award the cash.

You realize that both old-fashioned and redesigned anticounterfeit bills are available in both denominations.

(a) Answer the second demand of the ransom note. For extra credit, answer the first demand ⌣̈.

(b) Find a closed form for the number of ways to make $50m$ dollars using the two kinds of fifty- and hundred-dollar bills, for a nonnegative integer $m$.

8. In 2010, there are six different kinds of nickels in general circulation in the U.S., and six different kinds of pennies. Four of the varieties of nickels were issued in 2004 and 2005 and commemorate the bicentennial of the Lewis and Clark expedition—their respective designs on the reverse show a handshake, a boat, a bison, and an ocean view; the other two show president Jefferson's home, Monticello. Four of the pennies were issued in 2009 to commemorate the bicentennial of Lincoln's birth, with each design evoking a different period of the life of the U.S. president.

(a) Determine a generating function in closed form for the number of ways $a_k$ to make $k$ cents in change using only pennies and nickels available in 2010, counting each design as a different coin.

(b) Determine a finite sequence $c_0, c_1, \ldots, c_n$ so that

$$a_k = \sum_{j=\lceil (k-n)/5 \rceil}^{\lfloor k/5 \rfloor} c_{k-5j} \binom{j+11}{11}.$$

(c) Use the formula to determine $a_5$, and verify that your answer is correct.

(d) Use the formula to determine $a_{23}$, $a_{24}$, and $a_{25}$.

9. In 2010, there are fifty commemorative quarters in general circulation in the U.S., one for each state, and sixteen different presidential dollar coins, showing Washington through Lincoln on the obverse. Prove that the number of ways to make $25k$ cents in change using just these 66 different coins is

$$\sum_{a+b+2c=k} (-1)^{b+c} \binom{65+a}{a} \binom{15+b}{b} \binom{15+c}{c}.$$

Then use this formula to determine the number of ways to change one dollar using just these coins.

10. A hungry math major visits the school's cafeteria and wants to know the number of ways $s_k$ to take $k$ servings of food, including at least one main course, an even number (possibly zero) of side vegetables, an odd number of rolls, and at least two desserts. The cafeteria's food can be distinguished only in the coarsest way: Every dish is either a main course, a side vegetable, a roll, or a dessert. There is an unlimited supply of each kind of dish available.

(a) Determine a closed form for the generating function $\sum_k s_k x^k$.

(b) Show that

$$s_k = \binom{\lfloor \frac{k+1}{2} \rfloor}{3} + \binom{\lceil \frac{k+1}{2} \rceil}{3}.$$

The quantities $\lfloor x \rfloor$ and $\lceil x \rceil$ are defined on page 153.

### 2.6.4 Fibonacci Numbers

*Attention! Attention! Ladies and gentlemen, attention! There is a herd of killer rabbits headed this way and we desperately need your help!*

— *Night of the Lepus*

Hey, shouldn't that be a *colony* of killer rabbits?

Leonardo of Pisa, better known as Fibonacci, proposed the following harey problem in 1202. Assume that the rabbit population grows according to the following rules.

1. Every pair of adult rabbits produces a pair of baby rabbits, one of each gender, every month.

2. Baby rabbits become adult rabbits at age one month and produce their first offspring at age two months.

3. Rabbits are immortal.

Starting with a single pair of baby rabbits at the start of the first month, how many pairs of rabbits are there after $k$ months?

Let $F_k$ denote this number. In the first month, there is one pair of baby rabbits, so $F_1 = 1$. Likewise, $F_2 = 1$, as there is one pair of adult rabbits in the second month. In the third month, we have one baby pair and one adult pair, so $F_3 = 2$, and in the fourth month, the babies become adults and the adults produce another pair of offspring, so there is one pair of babies and two pairs of adults: $F_4 = 3$. Continuing in this way, we record the population in the following table.

| $k$ | Baby pairs | Adult pairs | $F_k$ |
|---|---|---|---|
| 0 | 0 | 0 | 0 |
| 1 | 1 | 0 | 1 |
| 2 | 0 | 1 | 1 |
| 3 | 1 | 1 | 2 |
| 4 | 1 | 2 | 3 |
| 5 | 2 | 3 | 5 |
| 6 | 3 | 5 | 8 |
| 7 | 5 | 8 | 13 |

Notice that the number of pairs of adults in month $k$ equals the total number of pairs of rabbits in month $k-1$. This is $F_{k-1}$. Also, the number of pairs of baby rabbits in month $k$ equals the number of adult pairs in month $k-1$, which is the total number of pairs in month $k-2$. This is $F_{k-2}$. Therefore,

$$F_k = F_{k-1} + F_{k-2}, \quad k \geq 2. \tag{2.40}$$

This recurrence, together with the initial conditions $F_0 = 0$ and $F_1 = 1$, determines the Fibonacci sequence $\{F_k\} = \{0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, \ldots\}$. This sequence appears frequently in combinatorial problems.

In this section we determine a closed form for $F_k$ by analyzing its generating function. We will adapt this technique to solve other recurrences later in this chapter.

Let $G(x)$ be the generating function for $\{F_k\}$. Then

$$\begin{aligned}
G(x) &= \sum_{k\geq 0} F_k x^k \\
&= F_0 + F_1 x + \sum_{k\geq 2} F_k x^k \\
&= x + \sum_{k\geq 2} (F_{k-1} + F_{k-2}) x^k \\
&= x + x \sum_{k\geq 2} F_{k-1} x^{k-1} + x^2 \sum_{k\geq 2} F_{k-2} x^{k-2} \\
&= x + x \sum_{k\geq 1} F_k x^k + x^2 \sum_{k\geq 0} F_k x^k,
\end{aligned}$$

and so

$$G(x) = x + xG(x) + x^2G(x).$$

Therefore,

$$G(x) = \frac{-x}{x^2 + x - 1},$$

and thus $F_k$ is the coefficient of $x^k$ in the Maclaurin series for this rational function. How can we determine this series without all the messy differentiation?

The trick is using partial fractions to write $G(x)$ as a sum of simpler rational functions. Write $x^2 + x - 1 = (x + \varphi)(x + \hat{\varphi})$, where $\varphi$ is the golden ratio, $\varphi = (1 + \sqrt{5})/2$, and $\hat{\varphi} = (1 - \sqrt{5})/2$. Write

$$\frac{-x}{x^2 + x - 1} = \frac{A}{x + \varphi} + \frac{B}{x + \hat{\varphi}}$$

and solve to find that $A = -\varphi/\sqrt{5}$ and $B = \hat{\varphi}/\sqrt{5}$. Thus

$$\begin{aligned} G(x) &= \frac{1}{\sqrt{5}} \left( \frac{\hat{\varphi}}{x + \hat{\varphi}} - \frac{\varphi}{x + \varphi} \right) \\ &= \frac{1}{\sqrt{5}} \left( \frac{1}{1 + x/\hat{\varphi}} - \frac{1}{1 + x/\varphi} \right) \\ &= \frac{1}{\sqrt{5}} \left( \frac{1}{1 - \varphi x} - \frac{1}{1 - \hat{\varphi} x} \right), \end{aligned}$$

since $\varphi\hat{\varphi} = -1$. Now the two terms on the right are closed forms for simple geometric series, so

$$G(x) = \frac{1}{\sqrt{5}} \sum_{k \geq 0} (\varphi^k - \hat{\varphi}^k) x^k,$$

and therefore

$$F_k = \frac{\varphi^k - \hat{\varphi}^k}{\sqrt{5}}. \tag{2.41}$$

Notice that $|\hat{\varphi}| < 1$, so $F_k \sim \varphi^k/\sqrt{5}$: a large number of rabbits indeed.

## Exercises

1. In each of the following problems, first compute the value the expression for a few small values of $n$. Then use your data to conjecture a general formula. Last, prove that your formula is correct.

   (a) $\displaystyle\sum_{k=0}^{n} F_k$.

   (b) $\displaystyle\sum_{k=0}^{n} F_{2k}$.

(c) $\sum_{k=1}^{n} F_{2k-1}$, if $n \geq 1$.

(d) $F_{n+1}F_{n-1} - F_n^2$, if $n \geq 1$.

2. Solve the recurrence $a_k = 2a_{k-1} + 3a_{k-2}$, if $a_0 = 0$ and $a_1 = 8$.

3. Suppose $a_0 = 0$, $a_1 = 5$, and $a_k = a_{k-1} + 6a_{k-2}$ for $k \geq 2$. Compute a closed form for the generating function of the sequence $\{a_k\}$. Then use this to determine a formula for $a_k$.

4. Solve the recurrence $a_k = 2a_{k-1} + 2a_{k-2}$, if $a_0 = 0$ and $a_1 = 1$.

5. Prove the following identities involving Fibonacci numbers.

   (a) $F_{m+n} = F_m F_{n+1} + F_{m-1}F_n$, if $m \geq 1$ and $n \geq 0$.

   (b) $F_n^2 + F_{n+1}^2 = F_{2n+1}$.

   (c) $F_{n+1}^2 - F_{n-1}^2 = F_{2n}$, if $n \geq 1$.

   (d) $\sum_{k=0}^{n} F_k^2 = F_n F_{n+1}$.

   (e) $\sum_{k=0}^{n} (-1)^{n-k} F_k = F_{n-1} - (-1)^n$, if $n \geq 1$.

   (f) $\sum_{k=0}^{n} (-1)^{n-k} k F_k = (n+1)F_{n-1} - F_{n-2} - 2(-1)^n$, if $n \geq 2$.

6. Prove that if $m$ and $n$ are nonnegative integers, then $F_m$ divides $F_{mn}$.

7. The *Lucas numbers* are defined by $L_0 = 2$, $L_1 = 1$, and $L_k = L_{k-1} + L_{k-2}$ for $k \geq 2$. Find a formula for $L_k$ in terms of $\varphi$ and $\hat{\varphi}$.

8. Prove the following identities involving Lucas and Fibonacci numbers.

   (a) $L_n = F_{n+1} + F_{n-1}$, if $n \geq 1$.

   (b) $\sum_{k=0}^{n} L_k^2 = L_n L_{n+1} + 2$.

   (c) $\sum_{k=0}^{n} (-1)^k L_{n-k} = L_{n-1} + 3(-1)^n$, if $n \geq 1$.

   (d) $F_{2n} = F_n L_n$.

   (e) $L_{2n} = L_n^2 - 2(-1)^n$.

9. The *Perrin sequence* is defined by $a_0 = 3$, $a_1 = 0$, $a_2 = 2$, and $a_k = a_{k-2} + a_{k-3}$ for $k \geq 3$. The *Padovan sequence* is defined by $b_0 = 0$, $b_1 = 1$, $b_2 = 1$, and $b_k = b_{k-2} + b_{k-3}$ for $k \geq 3$.

(a) Find generating functions in the form of rational functions for the Perrin sequence and the Padovan sequence.

(b) Prove that $a_k = r^k + \alpha^k + \overline{\alpha}^k$, where $r$, $\alpha$, and $\overline{\alpha}$ are the three complex roots of $x^3 - x - 1$. Conclude that $a_k \sim r^k$.

The Perrin sequence has an interesting property: If $p$ is a prime number, then $p$ divides the $p$th term in the Perrin sequence, $p \mid a_p$. This was first noted by Lucas in 1878 [192–194] (perhaps Lucas would have been interested in Exercise 6 of Section 2.6.3). Thus we obtain a test for composite numbers: If $n$ does not divide $a_n$, then $n$ is not prime. Unfortunately, the converse is false: There are infinitely many composite $n$ with the property that $n \mid a_n$. This was proved by Grantham [137].

10. In the children's game of hopscotch, a player hops across an array of squares drawn on the ground, landing on only one foot whenever there is just one square at a position, and landing on both feet when there are two. If every position has either one or two squares, how many different hopscotch games have exactly $n$ squares? Figure 2.3 shows the five different hopscotch games having four squares.

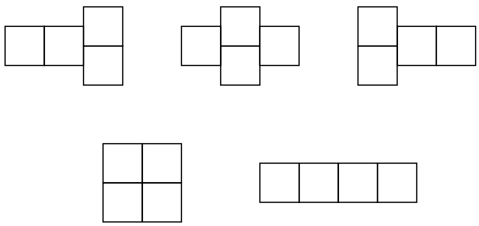


FIGURE 2.3. Hopscotch games with four squares.

11. Use a combinatorial argument and Exercise 10 to prove that

$$F_n = \sum_k \binom{n-k-1}{k}.$$

### 2.6.5 Recurrence Relations

*O me! O life!... of the questions of these recurring;*
— Walt Whitman, *Leaves of Grass*

In the "Tower of Hanoi" puzzle, one begins with a pyramid of $k$ disks stacked around a center pole, with the disks arranged from largest diameter on the bottom to smallest diameter on top. There are also two empty poles that can accept disks. The object of the puzzle is to move the entire stack of disks to one of the other poles, subject to three constraints:

1. Only one disk may be moved at a time.

2. Disks can be placed only on one of the three poles.

3. A larger disk cannot be placed on a smaller one.

How many moves are required to move the entire stack of $k$ disks onto another pole? Let $a_k$ denote this number. Clearly, $a_1 = 1$. To move $k$ disks, we must first move the $k-1$ top disks to one of the other poles, then move the bottom disk to the third pole, then move the stack of $k-1$ disks to that pole, so $a_k = 2a_{k-1} + 1$ for $k \geq 1$. Thus, $a_2 = 3$, $a_3 = 7$, $a_4 = 15$, and it appears that $a_k = 2^k - 1$.

We can certainly verify this formula by induction, but we wish to show how recurrences of this form can be solved by using generating functions. Consider the more general recurrence

$$a_k = ba_{k-1} + c, \quad k \geq 1,$$

where $b$ and $c$ are constants. This is a linear recurrence relation, since $a_k$ is a linear function of the preceding values of the sequence. (The Fibonacci recurrence is also a linear recurrence relation.) If $c$ is zero, we call the recurrence *homogeneous*; otherwise, it is *inhomogeneous*.

Let $G(x)$ be the generating function for $\{a_k\}$. Then

$$\begin{aligned} G(x) &= \sum_{k\geq 0} a_k x^k \\ &= a_0 + \sum_{k\geq 1} \left(ba_{k-1}x^k + cx^k\right) \\ &= a_0 + bx\sum_{k\geq 0} a_k x^k + cx\sum_{k\geq 0} x^k \\ &= a_0 + bxG(x) + \frac{cx}{1-x}, \end{aligned}$$

and so

$$G(x) = \frac{cx}{(1-bx)(1-x)} + \frac{a_0}{1-bx}.$$

Assuming $b \neq 1$, we compute

$$\frac{cx}{(1-bx)(1-x)} = \frac{c}{b-1}\left(\frac{1}{1-bx} - \frac{1}{1-x}\right),$$

so

$$\begin{aligned} G(x) &= \left(a_0 + \frac{c}{b-1}\right)\left(\frac{1}{1-bx}\right) - \frac{c}{b-1}\left(\frac{1}{1-x}\right) \\ &= \left(a_0 + \frac{c}{b-1}\right)\sum_{k\geq 0} b^k x^k - \frac{c}{b-1}\sum_{k\geq 0} x^k, \end{aligned}$$

and therefore

$$a_k = \left(a_0 + \frac{c}{b-1}\right) b^k - \frac{c}{b-1}. \tag{2.42}$$

For example, to find the number of moves needed to solve the Tower of Hanoi puzzle, we set $a_0 = 0$, $b = 2$, and $c = 1$ to obtain $a_k = 2^k - 1$. Also, if we set $b = -\frac{1}{2}$ and $c = 2$, we find that $a_k = (-1)^k(a_0 - \frac{4}{3})/2^k + \frac{4}{3}$, so $a_k$ approaches $\frac{4}{3}$ as $k$ grows large, independent of the initial value $a_0$.

We conclude with a short list of useful generating functions. Since

$$\frac{1}{1-x} = \sum_{k\geq 0} x^k, \tag{2.43}$$

we differentiate both sides to find that

$$\frac{1}{(1-x)^2} = \sum_{k\geq 1} kx^{k-1},$$

and so

$$\frac{x}{(1-x)^2} = \sum_{k\geq 0} kx^k. \tag{2.44}$$

Thus we obtain a closed form for the generating function of the identity sequence $\{k\}$. We take up the problem of determining a generating function for $\{k^n\}$, for any fixed positive integer $n$, in Section 2.8.5.

Finally, we integrate both sides of (2.43) to obtain the generating function for $\{1/k\}$:

$$-\ln(1-x) = \sum_{k\geq 1} \frac{x^k}{k}. \tag{2.45}$$

**Exercises**

1. Find a recurrence relation for the maximal number of regions of the plane separated by $k$ straight lines, then solve it.

2. Solve for $a_k$ in terms of $a_0$ and the other parameters in each of the following recurrence relations.

   (a) $a_k = a_{k-1} + c$.

   (b) $a_k = ba_{k-1} + cb^k$.

   (c) $a_k = ba_{k-1} + cr^k$, assuming $b \neq r$.

   (d) $a_k = ba_{k-1} + cr^k + d$, assuming $b \notin \{1, r\}$.

   (e) $a_k = ba_{k-1} + ck$, assuming $b \neq 1$.

   (f) $a_k = ba_{k-1} + ck + d$, assuming $b \neq 1$.

3. Find a closed form for the generating function of the sequence $\{k^2\}_{k\geq 0}$.

4. Let $v_n$ denote the number of ways that $3n$ different people can split up into $n$ three-person teams for a volleyball tournament, and let $v_0 = 1$. Assume that team members are unordered, so the team $\{a, b, c\}$ is the same as the team $\{c, a, b\}$, and assume that the teams are unordered, so putting $\{a, b, c\}$ on the first team and $\{d, e, f\}$ on the second is the same as putting $\{d, e, f\}$ on the first team and $\{a, b, c\}$ on the second. Determine a recurrence relation for $v_n$, then use it to compute $v_4$.

5. Let $d_k$ denote the minimal degree of a polynomial with $\{0, 1\}$ coefficients that is divisible by $(x + 1)^k$. For example, certainly $d_1 = 1$, since $f_1(x) = x + 1$ has the required properties, and $d_2 \leq 4$, since $f_2(x) = (x + 1)(x^3 + 1) = x^4 + x^3 + x + 1$ is permissible (in fact, $d_2 = 4$).

   (a) Determine an upper bound on $d_3$ by multiplying $f_2(x)$ by a suitable binomial of the form $x^r + 1$, choosing $r$ as small as possible. Then iterate this process to obtain upper bounds for $d_4$ and $d_5$.

   (b) Observe that one can obtain an upper bound on $d_k$ in general by constructing a polynomial of the form

   $$f_k(x) = \prod_{i=1}^{k} (x^{r_i} + 1)$$

   for a judiciously selected sequence $\{r_i\}$. Describe how to calculate $\{r_i\}$, and compute the values of this sequence for $i \leq 7$.

   (c) Determine a linear, homogeneous recurrence relation for the sequence $\{r_i\}$.

   (d) Compute a closed formula for $r_i$.

   (e) Determine an upper bound for $d_k$.

6. A *binary sequence* is a sequence in which each term is 0 or 1. Determine a recurrence relation for the number of binary sequences of length $n$ that do not contain two adjacent 1s, then find a simple expression for this number.

7. Let $t_n$ denote the number of binary sequences of length $n$ that do not contain three adjacent 1s.

   (a) Determine a recurrence relation for $t_n$, and enough initial values to generate the sequence.

   (b) Determine a closed form for the generating function

   $$T(x) = \sum_{n \geq 0} t_n x^n.$$

   (c) Define $t_n^*$ by $t_0^* = t_1^* = 0$, $t_2^* = 1$, and $t_n^* = t_{n-3}$ for $n \geq 3$. Determine a closed form for $T^*(x) = \sum_{n \geq 0} t_n^* x^n$. The numbers $\{t_n^*\}$ are known as the *tribonacci numbers*.

8. For a fixed positive integer $m$, let $s_{m,n}$ denote the number of binary sequences of length $n$ that do not contain $m$ adjacent 1s.

   (a) Determine a recurrence relation in $n$ for $s_{m,n}$, and enough initial values to generate the sequence.

   (b) Show that the generating function $S_m(x)$ for $\{s_{n,m}\}_{n\geq 0}$ is

   $$S_m(x) = \frac{1 - x^m}{x^{m+1} - 2x + 1}.$$

   Hint: First define a sequence $s^*_{m,n}$ from $s_{m,n}$ in the same manner as Exercise 7c. Then find its generating function $S^*_m(x)$, and use this to determine $S(x)$. The numbers $\{s^*_{m,n}\}_{n\geq 0}$ are known as the *generalized Fibonacci numbers of order* $m$, or the $m$*-generalized Fibonacci numbers*.

### 2.6.6 Catalan Numbers

*zero, un, dos, tres, quatre, cinc, sis, set, vuit, nou, deu, onze, dotze, tretze, catorze, quinze, setze, disset, divuit, dinou, vint.*

How many ways are there to compute a product of $k + 1$ matrices? Matrix multiplication is associative but not commutative, so this is the number of ways to place $k - 1$ pairs of parentheses in the product $x_0 x_1 \ldots x_k$ in such a way that the order of multiplications is completely specified. Let $C_k$ denote this number.

Let us first compute a few values of $C_k$. There is only one way to compute the product of one or two matrices. There are two ways to group a product of three matrices, $(x_0 x_1)x_2$ and $x_0(x_1 x_2)$, and there are five ways for a product of four matrices: $((x_0 x_1)x_2)x_3$, $(x_0(x_1 x_2))x_3$, $(x_0 x_1)(x_2 x_3)$, $x_0((x_1 x_2)x_3)$, and $x_0(x_1(x_2 x_3))$. A bit more work gives us 14 ways to compute a product of five matrices: There are five ways if one pair of parentheses is $x_0(x_1 x_2 x_3 x_4)$, another five for $(x_0 x_1 x_2 x_3)x_4$, two for $(x_0 x_1)(x_2 x_3 x_4)$, and two more for $(x_0 x_1 x_2)(x_3 x_4)$. We record these numbers in the following table.

| $k$ | $C_k$ |
|---|---|
| 0 | 1 |
| 1 | 1 |
| 2 | 2 |
| 3 | 5 |
| 4 | 14 |

Can we determine a recurrence relation for $C_k$?

Suppose we group the terms so that the last multiplication occurs between $x_i$ and $x_{i+1}$:

$$(x_0 x_1 \ldots x_i)(x_{i+1} \ldots x_k).$$

Then there are $C_i$ ways to group the terms in the first part of the product, and $C_{k-1-i}$ ways for the second part, so there are $C_i C_{k-1-i}$ ways to group the remaining terms in this case. Summing over $i$, we obtain the following formula for the total number of ways to group the $k+1$ terms:

$$C_k = \sum_{i=0}^{k-1} C_i C_{k-1-i}, \quad k \geq 1. \tag{2.46}$$

We compute

$$\begin{aligned}
C_1 &= C_0 C_0 = 1, \\
C_2 &= C_0 C_1 + C_1 C_0 = 2, \\
C_3 &= C_0 C_2 + C_1 C_1 + C_2 C_0 = 5, \\
C_4 &= C_0 C_3 + C_1 C_2 + C_2 C_1 + C_3 C_0 = 14, \\
C_5 &= C_0 C_4 + C_1 C_3 + C_2 C_2 + C_3 C_1 + C_4 C_0 = 42.
\end{aligned}$$

We would like to solve this recurrence to find a formula for $C_k$, so let us define the generating function for this sequence,

$$G(x) = \sum_{k \geq 0} C_k x^k.$$

Unlike other recurrences we have studied, this one is not linear, and has a variable number of terms. To solve it, we require one fact concerning products of generating functions.

If $A(x) = \sum_{k \geq 0} a_k x^k$ and $B(x) = \sum_{k \geq 0} b_k x^k$, then

$$A(x)B(x) = \sum_{k \geq 0} \left( \sum_{i=0}^{k} a_i b_{k-i} \right) x^k.$$

Let $c_k = \sum_{i=0}^{k} a_i b_{k-i}$. The sequence $\{c_k\}$ is called the *convolution* of the sequences $\{a_k\}$ and $\{b_k\}$. Thus, the generating function of the convolution of two sequences is the product of the generating functions of the sequences.

Using this fact, we find that

$$\begin{aligned}
G(x) &= \sum_{k \geq 0} C_k x^k \\
&= C_0 + \sum_{k \geq 1} \left( \sum_{i=0}^{k-1} C_i C_{k-1-i} \right) x^k \\
&= 1 + x \sum_{k \geq 0} \left( \sum_{i=0}^{k} C_i C_{k-i} \right) x^k \\
&= 1 + xG(x)^2,
\end{aligned}$$

since $\{\sum_{i=0}^{k} C_i C_{k-i}\}$ is the convolution of $\{C_k\}$ with itself. Thus,

$$xG(x)^2 - G(x) + 1 = 0,$$

and so

$$G(x) = \frac{1 \pm \sqrt{1-4x}}{2x}.$$

Only one of these functions can be the generating function for $\{C_k\}$, and it must satisfy

$$\lim_{x \to 0} G(x) = C_0 = 1.$$

It is easy to check that the correct function is

$$G(x) = \frac{1 - \sqrt{1-4x}}{2x}.$$

We now expand $G(x)$ as a Maclaurin series to find a formula for $C_k$. Using the generalized binomial theorem and the identity for negating the upper index, we find that

$$\begin{aligned}
(1-4x)^{1/2} &= \sum_{k \geq 0} \binom{1/2}{k} (-4x)^k \\
&= \sum_{k \geq 0} \binom{k-3/2}{k} 4^k x^k \\
&= 1 + \sum_{k \geq 1} \binom{k-3/2}{k} 4^k x^k \\
&= 1 + 4x \sum_{k \geq 0} \binom{k-1/2}{k+1} 4^k x^k.
\end{aligned}$$

Therefore,

$$G(x) = -2 \sum_{k \geq 0} \binom{k-1/2}{k+1} 4^k x^k,$$

and so

$$C_k = -2^{2k+1} \binom{k-1/2}{k+1}.$$

We can find a much simpler form for $C_k$. Expanding the generalized binomial coefficient and multiplying each term in the product by 2, we compute that

$$\begin{aligned}
C_k &= -\frac{2^{2k+1}}{(k+1)!} \prod_{i=0}^{k} \left(k - \frac{1}{2} - i\right) \\
&= -\frac{2^k}{(k+1)!} \prod_{i=0}^{k} (2k-1-2i).
\end{aligned}$$

The product consists of all the odd numbers between $-1$ and $2k-1$, so

$$\begin{aligned} C_k &= \frac{2^k}{(k+1)!}\prod_{i=1}^{k}(2i-1) \\ &= \frac{2^k}{(k+1)!}\prod_{i=1}^{k}\frac{(2i-1)(2i)}{2i} \\ &= \frac{1}{k!(k+1)!}\prod_{i=1}^{k}(2i-1)(2i). \end{aligned}$$

The remaining product is simply $(2k)!$, so

$$C_k = \frac{(2k)!}{k!(k+1)!} = \frac{1}{k+1}\binom{2k}{k}. \tag{2.47}$$

$C_k$ is called the $k$th *Catalan number*.

Incidentally, since $C_k$ is an integer, we have shown that $k+1$ always divides the binomial coefficient $\binom{2k}{k}$. Can you find an independent arithmetic proof of this fact?

Sloane and Plouffe [258] remark that the Catalan numbers are perhaps the second most frequently occurring numbers in combinatorics, after the binomial coefficients. Indeed, Stanley [262, ex. 6.19] lists 66 different combinatorial interpretations of these numbers! We close with another problem whose solution involves the Catalan numbers.

A *rooted tree* is a tree with a distinguished vertex called the *root*. The vertices in a rooted tree form a hierarchy, with the root at the highest level, and the level of every other vertex determined by its distance from the root. Some familiar terms are often used to describe relationships between vertices in a rooted tree: If $v$ and $w$ are adjacent vertices and $v$ lies closer to the root than $w$, then $v$ is the *parent* of $w$, and $w$ is a *child* of $v$. Likewise, one may define siblings, grandparents, cousins, and other family relationships in a rooted tree.

We say that a rooted tree is *strictly binary* if every parent vertex has exactly two children. How many strictly binary trees are there with $k$ parent vertices? Do not take symmetry into account: If two trees are mirror images of one another, count both configurations. Figure 2.4 shows that there are five trees with three parent vertices.

It is easy to see that the number of strictly binary trees with $k$ parent vertices is $C_k$. By Exercise 2, every such tree has $k+1$ leaves. Label these vertices with $x_0$ through $x_k$ from left to right in the tree. Then the tree determines an order of multiplication for the $x_i$. For example, the five trees in Figure 2.4 correspond to the multiplications $((x_0x_1)x_2)x_3$, $(x_0(x_1x_2))x_3$, $(x_0x_1)(x_2x_3)$, $x_0((x_1x_2)x_3)$, and $x_0(x_1(x_2x_3))$, respectively. Binary trees like these are often used in computer science to designate the order of evaluation of arithmetic expressions.

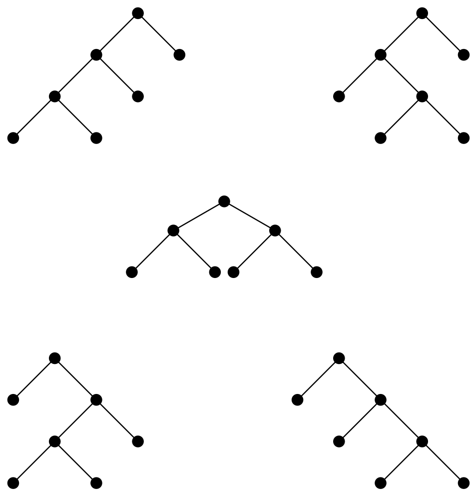


FIGURE 2.4. Strictly binary trees with three parent vertices.

**Exercises**

1. Show that every vertex in a rooted tree has at most one parent.

2. Show that a strictly binary tree having exactly $k$ parent vertices has exactly $k+1$ leaves.

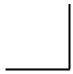
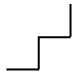
3. A *diagonal* of a convex polygon is a line segment connecting two non-adjacent vertices of the polygon. Let $p_n$ denote the number of ways to decompose a convex polygon having $n$ vertices into triangles by drawing $n-3$ diagonals that do not cross inside the polygon. Assume that the vertices of the polygon are labeled, so that triangulations with different orientations are counted separately.

   (a) Determine $p_3$, $p_4$, $p_5$, and $p_6$ by showing all the possible triangulations.

   (b) Let $v$ be a fixed vertex of a polygon with $n = 7$ sides. Count all the triangulations of the heptagon by considering two cases: (i) $v$ is not an endpoint of any of the four diagonals added in a triangulation, and (ii) $v$ is an endpoint of at least one of the diagonals. Use this to determine the value of $p_7$ without drawing every possible triangulation.

   (c) Determine a formula for $p_n$.

4. A *staircase* of size $n$ is a path in the plane from the origin to the point $(n, n)$ consisting of exactly $n$ horizontal and $n$ vertical steps, each of length 1, with the added condition that the path never rises above the line $y = x$. Let $s_n$ denote the number of staircases of size $n$. For example, $s_1 = 1$ since the only staircase is . Also, $s_2 = 2$ since the only possible staircases are

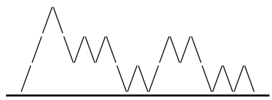
 and  .

(a) Determine $s_3$ and $s_4$ by drawing all the staircases of these sizes.

(b) For $1 \leq k \leq n$, let $s_{n,k}$ denote the number of staircases of size $n$ for which the first intersection of the path with the $y = x$ line (after the origin) is the point $(k, k)$. Determine a formula for $s_{n,k}$ in terms of the numbers $s_m$.

(c) Determine a recurrence relation for $s_n$, then a formula.

5. Let $r_n$ denote the number of mountain ridgelines one can draw using $n$ ascending steps and $n$ descending steps. A ridgeline must start and end on the horizon, and may never dip below the horizon. For example, $r_1 = 1$ since the only ridgeline is $/\backslash$, and the following illustrates a valid ridgeline with $n = 11$.



Two ridgelines that are mirror images of one another count as different arrangements.

(a) Determine $r_2$, $r_3$, and $r_4$ by drawing all of the possible ridgelines.

(b) Use a combinatorial argument to determine a recurrence relation for $r_n$, then find a formula for $r_n$.

6. Suppose $2k$ people are seated around a table. How many ways are there for the $k$ pairs of people to shake hands simultaneously across the table in such a way that no arms cross?

7. Show that the coefficient of $x^k$ in the Maclaurin series expansion of $(1 - (1 - 3x)^{1/3})/x$ is

$$\frac{1}{(k+1)!} \prod_{i=1}^{k} (3i - 1).$$

8. Use an arithmetic argument to show that $(2k)!$ is divisible by $k!(k + 1)!$. Hint: First compute the number of times a prime number $p$ divides $m!$.

## 2.7 Pólya's Theory of Counting

*Who are you who are so wise in the ways of science?*

— Sir Bedivere, in *Monty Python and the Holy Grail*

How many ways can King Arthur and his knights sit at the round table? How many different necklaces with $n$ beads can be formed using $m$ different kinds of beads?

Both these questions ask for a number of combinations in the presence of symmetry. Since there is no distinguished position at a round table, seating Arthur first, then Gawain, Percival, Bedivere, Tristram, and Galahad clockwise around the table yields the same configuration as seating Tristram first, then Galahad, Arthur, Gawain, Percival, and Bedivere in clockwise order. Similarly, we should consider two necklaces to be identical if we can transform one into the other by rotating the necklace or by turning it over.

Before answering these questions, let us first rephrase them in the language of group theory.

### 2.7.1 Permutation Groups

*I haven't fought just one person in a long time. I've been specializing in groups.*

— Fezzik, in *The Princess Bride*

A *group* consists of a set $G$ together with a binary operator $\circ$ defined on this set. The set and the operator must satisfy four properties.

- Closure. For every $a$ and $b$ in $G$, $a \circ b$ is in $G$.
- Associativity. For every $a$, $b$, and $c$ in $G$, $a \circ (b \circ c) = (a \circ b) \circ c$.
- Identity. There exists an element $e$ in $G$ that satisfies $e \circ a = a \circ e = a$ for every $a$ in $G$. The element $e$ is called the *identity* of $G$.
- Inverses. For every element $a$ in $G$, there exists an element $b$ in $G$ such that $a \circ b = b \circ a = e$. The element $b$ is called the *inverse* of $a$.

In addition, if $a \circ b = b \circ a$ for every $a$ and $b$ in $G$, we say that $G$ is an *abelian*, or *commutative*, group.

For example, the set of integers forms a group under addition. The identity element is $0$, since $0 + i = i + 0 = i$ for every integer $i$, and the inverse of the integer $i$ is the integer $-i$. Similarly, the set of nonzero rational numbers forms a group under multiplication (with identity element 1), as does the set of nonzero real numbers.

We can also construct groups of permutations. A permutation of $n$ objects may be described by a function $\pi$ defined on the set $\{1, 2, \ldots, n\}$ by ordering the objects in some fashion, then taking $\pi(i) = j$ if the $i$th object in the ordering occupies the $j$th position in the permutation. For example, the permutation $[c, d, a, e, b]$ of the list $[a, b, c, d, e]$ is represented by the function $\pi$ defined on the set $\{1, 2, 3, 4, 5\}$, with $\pi(1) = 3$, $\pi(2) = 5$, $\pi(3) = 1$, $\pi(4) = 2$, and $\pi(5) = 4$. Notice that a function $\pi : \{1, \ldots, n\} \to \{1, \ldots, n\}$ arising from a permutation has the property that $\pi(i) \neq \pi(j)$ whenever $i \neq j$. Such a function is called an *injective*, or *one-to-one*, function. The map $\pi$ also has the property that for every $m$ with $1 \leq m \leq n$, there exists a number $i$ such that $\pi(i) = m$. A function like this that maps to every element in its range is called *surjective*, or *onto*, and a function that is both injective and surjective is said to be a *bijection*. Thus, every permutation of $n$ objects corresponds to a bijection $\pi$ on the set $\{1, 2, \ldots n\}$, and every such bijection corresponds to a permutation.

Let $S_n$ denote the set of all bijections on the set $\{1, 2, \ldots, n\}$. Exercise 4 asks you to verify that this set forms a group under the operation of composition of functions. For example, the identity element of the group is the identity map $\pi_0$, defined by $\pi_0(k) = k$ for each $k$, since $\pi \circ \pi_0 = \pi_0 \circ \pi = \pi$ for every $\pi$ in $S_n$. This group is called the *symmetric group* on $n$ elements.

The size of the group $S_n$ is the number of permutations of $n$ objects, so $|S_n| = n!$. Because of our correspondence, we normally refer to an element of $S_n$ as a permutation, rather than a bijection.

To specify a particular permutation $\pi$ in $S_n$, we need to name the value of $\pi(k)$ for each $k$. This is often written in two rows as follows:

$$\begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ \pi(1) & \pi(2) & \pi(3) & \ldots & \pi(n) \end{pmatrix}.$$

For example,

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 1 & 2 & 4 \end{pmatrix}$$

denotes the permutation described earlier.

We can describe the permutations in a more succinct manner by using *cycle notation*. For example, in the permutation above, $\pi$ sends 1 to 3 and 3 to 1, and sends 2 to 5, 5 to 4, and 4 to 2. So we can think of $\pi$ as a combination of two *cycles*, $1 \to 3 \to 1$ and $2 \to 5 \to 4 \to 2$, and denote it by

$$(13)(254).$$

Of course, we could also denote this same permutation by the cycles $(542)(31)$, so to make our notation unique, we make two demands. First, the cycle containing 1 must appear first, followed by the cycle containing the smallest number not appearing in the first cycle, and so on. Second, we require the first number listed in each cycle to be the smallest number appearing in that cycle. To simplify the notation, cycles of length 1 are usually omitted, so $(1253)(4)$ is written more simply as $(1253)$. The identity permutation is denoted by $(1)$.

The composition of two permutations is computed from right to left. For example, suppose $\pi_1 = (13)(254)$ and $\pi_2 = (15423)$. We determine the composition $\pi_1 \circ \pi_2$ by applying $\pi_2$ first, then $\pi_1$. Since $\pi_2$ sends 1 to 5, and $\pi_1$ sends 5 to 4, the composition $\pi_1 \circ \pi_2$ then sends 1 to 4. In the same way, we see that $\pi_1 \circ \pi_2$ sends 4 to 5, 5 to 2, 2 to 1, and 3 to 3. Thus, $\pi_1 \circ \pi_2 = (1452)$. In cycle notation, we denote the composition of two permutations by juxtaposing their cycles, so

$$\pi_1 \circ \pi_2 = (13)(254)\,(15423) = (1452).$$

Notice that the cycles for $\pi_1$ appear first, so products of cycles are always computed from right to left. Also, we calculate that $\pi_2 \circ \pi_1 = (15423)\,(13)(254) = (2435)$, so in general $S_n$ is not an abelian group.

A subset $H$ of a group $G$ is called a *subgroup* of $G$ if $H$ is itself a group under the same binary operation. The group $S_n$ contains many subgroups; for example, $\{(1), (12)\}$ is a subgroup of $S_n$ for every $n \geq 2$. We investigate three particularly important subgroups of $S_n$.

### The Cyclic Group

If $\pi$ is a permutation in $S_n$ and $m$ is a nonnegative integer, let $\pi^m$ denote the permutation obtained by composing $\pi$ with itself $m$ times, so $\pi^0 = (1)$, and $\pi^3 = \pi \circ \pi \circ \pi$. Let

$$\langle \pi \rangle = \{\pi^m \ : \ m \geq 0\}, \tag{2.48}$$

so that $\langle \pi \rangle$ is a subset of $S_n$. In fact (Exercise 5), $\langle \pi \rangle$ is a subgroup of $S_n$, and we call this group the cyclic subgroup generated by $\pi$ in $S_n$.

The *cyclic group* $C_n$ is the subgroup of the symmetric group $S_n$ generated by the permutation $(123 \cdots n)$, so

$$C_n = \langle (123 \cdots n) \rangle. \tag{2.49}$$

Clearly, $C_n$ contains $n$ elements, since $n$ applications of the generating permutation are required to return to the identity permutation. For example, $(1234)^2 = (13)(24)$, $(1234)^3 = (1432)$, and $(1234)^4 = (1)$, so

$$C_4 = \{(1),\ (1234),\ (13)(24),\ (1432)\}. \tag{2.50}$$

The group $C_n$ may be realized as the group of rotational symmetries of a regular polygon having $n$ sides. For example, each of the permutations of (2.50) corresponds to a permutation of the vertices of Figure 2.5 obtained by rotating the square by $0$, $90$, $180$, or $270$ degrees.

### The Dihedral Group

The *dihedral group* $D_n$ is the group of symmetries of a regular polygon with $n$ sides, including reflections as well as rotations. Since $C_n$ consists of just the rotational symmetries of such a figure, evidently $C_n$ is a subgroup of $D_n$.

Referring to Figure 2.5, we see that $D_4$ consists of the four rotations of $C_4$, plus the four reflections $(12)(34)$, $(14)(23)$, $(13)$, and $(24)$. The first two permutations represent reflections about the vertical and horizontal axes of symmetry of the square; the last two represent flips about the diagonal axes of symmetry. In general, if $n$ is even, we obtain $n/2$ reflections through axes of symmetry that pass through opposite vertices, and $n/2$ reflections through axes that pass through midpoints of opposite edges. Combining these with the $n$ rotations of $C_n$, we find that $|D_n| = 2n$ in this case.

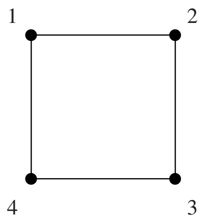



FIGURE 2.5. A square with labeled vertices.

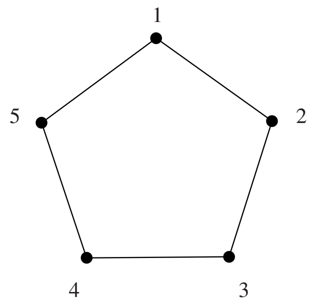



FIGURE 2.6. A regular pentagon with labeled vertices.

Using Figure 2.6, we find that $D_5$ consists of five rotations and five reflections,

$$D_5 = \{(1),\ (12345),\ (13524),\ (14253),\ (15432),\ (25)(34),\\ (13)(45),\ (15)(24),\ (12)(35),\ (14)(23)\}.$$

It is easy to see that we always obtain $n$ reflections if $n$ is odd, so $|D_n| = 2n$ for every $n \geq 1$.

### The Alternating Group

Every permutation can be expressed as a product of *transpositions*, which are cycles of length 2. For example, the cycle $(123)$ can be written as the product $(12)(23)$, and the permutation $(1234)(567)$ can be expressed as the product of six transpositions: $(12)(23)(34)(56)(67)$. Such a decomposition is not unique; for instance, $(123)$ may also be written as $(23)(13)$, or $(12)(23)(13)(13)$. However, the number of transpositions in any representation of one permutation is either always an even number, or always an odd number. Exercise 6 outlines a proof of this fact. If a permutation $\pi$ always decomposes into an even number of transpositions, we say that $\pi$ is an *even* permutation; otherwise, it is an *odd* permutation. Notice that the identity permutation is even, since it is represented by a product of zero transpositions.

The *alternating group* $A_n$ consists of the even permutations of $S_n$. For example, $A_3 = \{(1), (123), (132)\} = C_3$, and

$$A_4 = \{(1),\ (123),\ (132),\ (124),\ (142),\ (134),\ (143),\\ (234),\ (243),\ (12)(34),\ (13)(24),\ (14)(23)\}.$$

Exercises 7 and 8 ask you to verify that $A_n$ is a group of size $|A_n| = n!/2$ for $n \geq 2$, and that $A_n$ is not abelian for $n \geq 5$.

**Exercises**

1. Show that the identity element of a group is unique.

2. (a) Suppose that $M$ is a finite set and $f : M \to M$ is an injective function. Show that $f$ is a bijection.

   (b) Suppose that $M$ is a finite set and $f : M \to M$ is a surjective function. Show that $f$ is a bijection.

   (c) Show that neither of these statements is necessarily true if $M$ is an infinite set.

3. In each part, determine all values of $n$ that satisfy the statement.

   (a) $C_n$ is a subgroup of $A_n$.

   (b) $D_n$ is a subgroup of $A_n$.

   (c) $C_n$ is a subgroup of $D_{n+1}$.

   (d) $C_n$ is a subgroup of $S_{n+1}$.

4. Verify that $S_n$ forms a group under composition of functions by checking that each of the required properties is satisfied.

   (a) Closure. If $\pi_1$ and $\pi_2$ are bijections on $\{1, 2, \ldots, n\}$, show that $\pi_1 \circ \pi_2$ is also a bijection on $\{1, 2, \ldots, n\}$.

   (b) Associativity. If $\pi_1$, $\pi_2$, and $\pi_3$ are in $S_n$, show that $\pi_1 \circ (\pi_2 \circ \pi_3)$ and $(\pi_1 \circ \pi_2) \circ \pi_3$ represent the same function in $S_n$.

   (c) Identity. Check that $\pi_0 \circ \pi = \pi \circ \pi_0 = \pi$, for every $\pi$ in $S_n$. Here, $\pi_0$ is the identity map on $\{1, 2, \ldots, n\}$.

   (d) Inverses. Given a bijection $\pi$ in $S_n$, construct a bijection $\pi^{-1}$ in $S_n$ satisfying $\pi \circ \pi^{-1} = \pi^{-1} \circ \pi = \pi_0$.

5. Suppose that $G$ is a group and $g$ is an element of $G$.

   (a) Show that $\langle g \rangle$ is a subgroup of $G$.

   (b) Show that $\langle g \rangle$ is abelian.

6. Let $\mathbf{x}$ denote the vector of $n$ variables $(x_1, x_2, \ldots, x_n)$. Define

$$P(\mathbf{x}) = \prod_{1 \leq i < j \leq n} (x_i - x_j),$$

and if $\pi \in S_n$, let

$$P_\pi(\mathbf{x}) = \prod_{1 \leq i < j \leq n} (x_{\pi(i)} - x_{\pi(j)}).$$

   (a) Show that $P_\pi(\mathbf{x}) = \pm P(\mathbf{x})$.
   (b) Show that $P_\pi(\mathbf{x}) = -P(\mathbf{x})$ if $\pi$ is a transposition.
   (c) Conclude that no permutation $\pi$ in $S_n$ can be represented both as a product of an even number of transpositions and as a product of an odd number of transpositions.

7. (a) Prove that $A_n$ is a group.
   (b) Show that $A_n$ is not abelian for $n \geq 5$.

8. Let $n \geq 2$, let $B_n$ denote the set of odd permutations in $S_n$, and let $\tau$ be a transposition in $S_n$.
   (a) Show that the map $T : S_n \to S_n$ defined by $T(\pi) = \tau \circ \pi$ is a bijection.
   (b) Show that $T$ maps $A_n$ to $B_n$, and $B_n$ to $A_n$.
   (c) Conclude that $|A_n| = n!/2$.

9. Determine the group of symmetries of each of the following objects.
   (a) The vertices of a regular tetrahedron.
   (b) The vertices of a cube.
   (c) The vertices of a regular octahedron.

### 2.7.2 Burnside's Lemma

*Burnside had submitted the scheme to Meade and myself, and we both approved of it, as a means of keeping the men occupied.*
— *Personal Memoirs of U. S. Grant*

Armed with our knowledge of permutation groups, we now develop a general method for counting combinations in the presence of symmetry. In general, we are given a set of objects $S$, a set of colorings of these objects $C$, and a group of permutations $G$ representing symmetries possessed by configurations of the objects. We consider two colorings in $C$ to be equivalent if one of the permutations in $G$ transforms one coloring to the other, and we would like to determine the number of nonequivalent colorings in $C$.

For example, suppose $S = \{1, 2, 3, 4\}$ is the set of vertices of the square in Figure 2.5, and $C$ is the set of all possible colorings of these vertices using two colors, red and green. Let $rrgr$ denote the coloring where vertices 1, 2, and 4 are red and vertex 3 is green. Then

$$\begin{aligned} C = \{&gggg,\ gggr,\ ggrg,\ ggrr,\ grgg,\ grgr,\ grrg,\ grrr,\\ &rggg,\ rggr,\ rgrg,\ rgrr,\ rrgg,\ rrgr,\ rrrg,\ rrrr\}. \end{aligned} \tag{2.51}$$

We consider two colorings in $C$ to be equivalent if one can be transformed to the other by a rotation of the square. For example, rotating the coloring $rrgr$ yields the equivalent colorings $rrrg$, $grrr$, and $rgrr$. So we choose $G$ to be the group of rotations, $C_4$. A permutation $\pi$ in $C_4$ is a function defined on the set $\{1, 2, 3, 4\}$, but $\pi$ induces a map $\pi^*$ defined on the set of colorings $C$ in a natural way. For example, if $\pi$ is the 180-degree rotation $(13)(24)$, then the induced map $\pi^*$ rotates a coloring by the same amount, so $\pi^*(rrgr) = grrr$, and $\pi^*(grgr) = grgr$.

If $c_1$ and $c_2$ are two equivalent colorings in $C$, so $\pi^*(c_1) = c_2$ for some $\pi \in G$, we write $c_1 \sim c_2$. Using the fact that $G$ is a group, it is easy to verify (Exercise 1) that the relation $\sim$ on the set of colorings is

- reflexive: $c \sim c$ for all colorings $c$,
- symmetric: $c_1 \sim c_2$ implies $c_2 \sim c_1$, and
- transitive: $c_1 \sim c_2$ and $c_2 \sim c_3$ implies $c_1 \sim c_3$.

A relation possessing these three properties is called an *equivalence relation*. By grouping together collections of mutually equivalent elements, an equivalence relation on a set partitions the set into a number of disjoint subsets, called *equivalence classes*. Our goal then is to determine the number of equivalence classes of $C$ under the relation $\sim$.

In our example, the group $C_4$ partitions our set of colorings (2.51) into six equivalence classes:

$$\begin{gathered} \{gggg\},\\ \{gggr,\ ggrg,\ grgg,\ rggg\},\\ \{ggrr,\ grrg,\ rggr,\ rrgg\},\\ \{grgr,\ rgrg\},\\ \{grrr,\ rgrr,\ rrgr,\ rrrg\},\\ \{rrrr\}. \end{gathered}$$

Therefore, there are just six ways to color the vertices of a square using two colors, after discounting rotational symmetries.

We can now translate the problems from the introduction to this section into this more abstract setting. In the round table problem, $S$ is the set of $n$ places at the table, $G$ is $C_n$, and $C$ is the collection of the $n!$ seating assignments. In the necklace problem, $S$ is the set of $n$ bead positions, $G$ is $D_n$, and $C$ is the collection of the $m^n$ possible arrangements of the $m$ kinds of beads on the necklace.

Before presenting a general method to solve problems like these, we introduce three sets that will be useful in our analysis. Given a permutation $\pi$ in $G$, define $C_\pi$ to be the set of colorings that are invariant under action by the induced map $\pi^*$,

$$C_\pi = \{c \in C \ : \ \pi^*(c) = c\}. \tag{2.52}$$

This set is called the *invariant set* of $\pi$ in $C$. Similarly, given a coloring $c$ in $C$, define $G_c$ to be the set of permutations $\pi$ in $G$ for which $c$ is a fixed coloring,

$$G_c = \{\pi \in G \ : \ \pi^*(c) = c\}. \tag{2.53}$$

This set is called the *stabilizer* of $c$ in $G$. It is always a subgroup of $G$. Finally, let $\overline{c}$ be the set of colorings in $C$ that are equivalent to $c$ under the action of the group $G$,

$$\overline{c} = \{\pi^*(c) \ : \ \pi \in G\}. \tag{2.54}$$

The set $\overline{c}$ is thus the equivalence class of $c$ under the relation $\sim$. It is also called the *orbit* of $c$ under the action of $G$.

For example, if $C$ is given by (2.51) and $G$ is the dihedral group $D_4$, we have

$$\overline{gggr} = \{gggr, \ ggrg, \ grgg, \ rggg\}$$

and

$$G_{gggr} = \{(1), \ (13)\}.$$

Also,

$$\overline{grgr} = \{grgr, \ rgrg\}$$

and

$$G_{grgr} = \{(1), \ (13)(24), \ (13), \ (24)\}.$$

Notice that in both cases, the product of the size of the stabilizer of a coloring with the size of the equivalence class of the same coloring equals the number of elements in the group. The following lemma proves that this is always the case.

**Lemma 2.8.** *Suppose a group $G$ acts on a set of colorings $C$. For any coloring $c$ in $C$, we have $|G_c|\,|\,\overline{c}\,| = |G|$.*

*Proof.* We prove this by showing that every permutation in $G$ may be represented in a unique way as a composition of a permutation in $G_c$ with a permutation in a particular set $P$, where $|P| = |\,\overline{c}\,|$. Suppose there are $m$ colorings in the equivalence class of $c$, $\overline{c} = \{c_1, c_2, \ldots, c_m\}$. For each $i$ between 1 and $m$, select a permutation $\pi_i \in G$ such that $\pi_i^*(c) = c_i$, and let $P = \{\pi_1, \pi_2, \ldots, \pi_m\}$.

Now let $\pi$ be an arbitrary permutation in $G$. Then $\pi^*(c) = c_i$ for some $i$, so $\pi^*(c) = \pi_i^*(c)$. Thus $(\pi_i^{-1} \circ \pi)^*(c) = c$, and so $\pi_i^{-1} \circ \pi \in G_c$. Since $\pi_i \circ (\pi_i^{-1} \circ \pi) = \pi$, we see that $\pi$ has at least one representation in the desired form. Suppose now that $\pi = \pi_i \circ \sigma = \pi_j \circ \tau$, for some $\pi_i$ and $\pi_j$ in $P$ and some $\sigma$ and $\tau$ in $G_c$. Then $\pi_i(\sigma(c)) = \pi_i(c) = c_i$ and $\pi_j(\tau(c)) = c_j$, so $c_i = c_j$, and hence $i = j$. Therefore, $\sigma = \tau$, so the representation of $\pi$ is unique. $\square$

The following formula for the number of equivalence classes of $C$ under the action of a group $G$ is usually named for Burnside (the English mathematician, not the American Civil War general), as it was popularized by his book [45]. This result was first proved by Frobenius [115], however, and Burnside even attributes the formula to Frobenius in the first edition of his textbook [45]. Further details on the history of this result appear in Neumann [213] and Wright [288].

Briefly, Burnside's Lemma states that the number of equivalence classes of colorings is the average size of the invariant sets.

**Theorem 2.9** (Burnside's Lemma)**.** *The number of equivalence classes $N$ of the set $C$ in the presence of the group of symmetries $G$ is given by*

$$N = \frac{1}{|G|} \sum_{\pi \in G} |C_\pi| . \tag{2.55}$$

*Proof.* If $P$ is a logical expression, let $[P]$ be 1 if $P$ is true and 0 if $P$ is false. Then

$$\begin{aligned}
\frac{1}{|G|} \sum_{\pi \in G} |C_\pi| &= \frac{1}{|G|} \sum_{\pi \in G} \sum_{c \in C} [\pi^*(c) = c] \\
&= \frac{1}{|G|} \sum_{c \in C} \sum_{\pi \in G} [\pi^*(c) = c] \\
&= \frac{1}{|G|} \sum_{c \in C} |G_c| \\
&= \sum_{c \in C} \frac{1}{|\overline{c}|} \\
&= \sum_{\overline{c}} \sum_{c \in \overline{c}} \frac{1}{|\overline{c}|} \\
&= \sum_{\overline{c}} 1 \\
&= N.
\end{aligned}$$

We applied Lemma 2.8 to obtain the fourth line. □

We may apply Burnside's Lemma to solve the problems we described earlier. In the round table problem, $|G| = n$. The invariant set of the identity permutation is the entire set of colorings, $C_{(1)} = C$, and the invariant set of any nontrivial rotation $\pi$ is empty, $C_\pi = \{\ \}$. Therefore, the number of nonequivalent seating arrangements is $|C|/n = (n-1)!$.

To determine the number of nonequivalent necklaces with four beads using two different kinds of beads, we calculate $|C_{(1)}| = 16$, $|C_{(13)}| = |C_{(24)}| = 8$, $|C_{(12)(34)}| = |C_{(13)(24)}| = |C_{(14)(23)}| = 4$, and $|C_{(1234)}| = |C_{(1432)}| = 2$. Therefore, $N = (16 + 2 \cdot 8 + 3 \cdot 4 + 2 \cdot 2)/8 = 6$. Last, we calculate the number of nonequivalent three-bead necklaces using three different kinds of beads. Here, $|C_{(1)}| = 27$, $|C_{(12)}| = |C_{(13)}| = |C_{(23)}| = 9$, and $|C_{(123)}| = |C_{(132)}| = 3$, so $N = 60/6 = 10$.

**Exercises**

1. Show that $\sim$ is an equivalence relation on $C$.

2. Prove that $G_c$ is a subgroup of $G$.

3. How many different necklaces having five beads can be formed using three different kinds of beads if we discount:

   (a) Both flips and rotations?

   (b) Rotations only?

   (c) Just one flip?

4. The commander of a space cruiser wishes to post four sentry ships arrayed around the cruiser at the vertices of a tetrahedron for defensive purposes, since an attack can come from any direction.

   (a) How many ways are there to deploy the ships if there are two different kinds of sentry ships available, and we discount all symmetries of the tetrahedral formation?

   (b) How many ways are there if there are three different kinds of sentry ships available?

5. (a) How many ways are there to label the faces of a cube with the numbers 1 through 6 if each number may be used more than once?

   (b) What if each number may only be used once?

### 2.7.3 The Cycle Index

*Lance Armstrong (7), Jacques Anquetil (5), Bernard Hinault (5), Miguel Indurain (5), Eddy Merckx (5), Louision Bobet (3), Greg LeMond (3), Philippe Thys (3).*

— Multiple Tour de France winners

To use Burnside's Lemma to count the number of equivalence classes of a set of colorings $C$, we must compute the size of the invariant set $C_\pi$ associated with every permutation $\pi$ in a group of symmetries $G$. A simple observation allows us to compute the size of this set easily in many situations.

Suppose we wish to determine the number of ways to color $n$ objects using up to $m$ colors, discounting symmetries on the objects described by a group $G$. If a coloring is invariant under the action of a permutation $\pi$ in $G$, then every object permuted by one cycle of $\pi$ must have the same color. Therefore, if $\pi$

has $k$ disjoint cycles, the number of colorings invariant under the action of $\pi$ is $|C_\pi| = m^k$. For example, if $S$ is the set of vertices of a square and $G = D_4$, then $\left|C_{(1234)}\right| = m$, $\left|C_{(12)(34)}\right| = m^2$, $\left|C_{(13)(2)(4)}\right| = m^3$, and $\left|C_{(1)(2)(3)(4)}\right| = m^4$. Notice that it is essential to include the cycles of length 1 in these calculations.

With this in mind, we define the *cycle index* of a group $G$ of permutations on $n$ objects. For a permutation $\pi$ in $G$, define a monomial $M_\pi$ associated with $\pi$ in the following way. If $\pi$ is a product of $k$ cycles, and the $i$th cycle has length $\ell_i$, let

$$M_\pi = M_\pi(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{k} x_{\ell_i}. \tag{2.56}$$

Here, $x_1$, $x_2$, …, $x_n$ are indeterminates. The cycle index of $G$ is defined by

$$P_G(\mathbf{x}) = \frac{1}{|G|} \sum_{\pi \in G} M_\pi(\mathbf{x}), \tag{2.57}$$

where $\mathbf{x}$ denotes the vector $(x_1, x_2, \ldots, x_n)$.

For example, for $G = D_4$, we find that

$$\begin{gathered} M_{(1)(2)(3)(4)} = x_1^4, \\ M_{(13)(2)(4)} = M_{(1)(24)(3)} = x_1^2 x_2, \\ M_{(12)(34)} = M_{(13)(24)} = M_{(14)(23)} = x_2^2, \\ M_{(1234)} = M_{(1432)} = x_4. \end{gathered}$$

Therefore,

$$P_{D_4}(x_1, x_2, x_3, x_4) = \tfrac{1}{8}\left(x_1^4 + 2x_1^2 x_2 + 3x_2^2 + 2x_4\right), \tag{2.58}$$

and

$$P_{C_4}(x_1, x_2, x_3, x_4) = \tfrac{1}{4}\left(x_1^4 + x_2^2 + 2x_4\right). \tag{2.59}$$

By Burnside's Lemma, the number of ways to color $n$ objects using up to $m$ colors, discounting the symmetries of $G$, is $P_G(m, m, \ldots, m)$. For example, the number of equivalence classes of four-bead necklaces composed using $m$ different kinds of beads is

$$P_{D_4}(m, m, m, m) = \tfrac{1}{8}\left(m^4 + 2m^3 + 3m^2 + 2m\right).$$

Substituting $m = 2$, we find there are six different colorings, as before.

Finally, let us compute the number of twenty-bead necklaces composed of rhodonite, rose quartz, and lapis lazuli beads. We must determine the cycle index for the group $D_{20}$. We find that eight of the rotations, those by $18k$ degrees with $k = 1$, 3, 7, 9, 11, 13, 17, or 19, are a single cycle of length 20, yielding the term $8x_{20}$ in the cycle index. Four rotations, $k = 2$, 6, 14, and 18, make two cycles of length 10, contributing $4x_{10}^2$. Rotations with $k = 4$, 8, 12, or 16 make four cycles of length $5$, adding $4x_5^4$, and $k = 5$ or $15$ contributes $2x_4^5$. The rotation with $k = 10$ yields $x_2^{10}$, and the identity adds $x_1^{20}$. Ten of the reflections, the ones about axes of symmetry that pass through midpoints of edges, are each represented by ten transpositions, contributing $10x_2^{10}$. The other ten reflections, flipping about opposite vertices, yield $10x_1^2x_2^9$. Therefore,

$$P_{D_{20}}(x_1, \ldots, x_{20}) = \tfrac{1}{40}\left(x_1^{20} + 10x_1^2x_2^9 + 11x_2^{10} + 2x_4^5 + 4x_5^4 + 4x_{10}^2 + 8x_{20}\right), \tag{2.60}$$

and the number of different twenty-bead necklaces that can be made using three kinds of beads is $P_{D_{20}}(3, \ldots, 3) = 87\,230\,157$.

**Exercises**

1. Show that the monomial $M_\pi$ defined in (2.57) has the property that the sum $\sum_{i=1}^{k} \ell_i = n$.
2. (a) Determine the cycle index for $S_4$ and for $A_4$.
   (b) Show that $P_{S_4}(m, m, m, m)$ may be written as a binomial coefficient.
   (c) Determine the smallest value of $m$ for which $P_{A_4}(m, m, m, m) > P_{S_4}(m, m, m, m)$.
3. Determine the number of different necklaces with 21 beads that can be made using four kinds of beads. Your equivalence classes should account for both rotations and flips.
4. Determine the number of eight-bead necklaces that can be made using red, green, blue, and white beads under each of the following groups of symmetries.
   (a) $D_8$.
   (b) A subgroup of $D_8$ having four elements. How does the answer depend on the subgroup you choose?
5. Determine the cycle index for the group of symmetries of the faces of a cube, and use this to determine the number of different six-sided dice that can be manufactured using $m$ different labels for the faces of the dice. Assume that each label may be used any number of times.

### 2.7.4 Pólya's Enumeration Formula

*I have yet to see any problem, however complicated, which, when looked at in the right way, did not become still more complicated.*
— Poul Anderson

We can use the cycle index to solve more complicated problems on arrangements in the presence of symmetry. Suppose we need to determine the number of equivalence classes of colorings of $n$ objects using the $m$ colors $y_1, y_2, \ldots, y_m$, where each color $y_i$ occurs a prescribed number of times. For example, how many different necklaces can be made using exactly two rhodonite, nine rose quartz, and nine lapis lazuli beads?

Let us define the *pattern inventory* of the different ways to color $n$ objects using $m$ colors with respect to a symmetry group $G$ as a generating function in $m$ variables,

$$F_G(y_1, y_2, \ldots, y_m) = \sum_{\mathbf{v}} a_{\mathbf{v}} y_1^{n_1} y_2^{n_2} \cdots y_m^{n_m}, \tag{2.61}$$

where the sum runs over all vectors $\mathbf{v} = (n_1, n_2, \ldots, n_m)$ of nonnegative integers satisfying $n_1 + n_2 + \cdots + n_m = n$, and $a_{\mathbf{v}}$ represents the number of nonequivalent colorings of the $n$ objects where the color $y_i$ occurs precisely $n_i$ times. For example, if we denote a rhodonite bead by $r$, a rose quartz bead by $q$, and a lapis lazuli bead by $l$, we see that the answer to our question above is the coefficient of $r^2q^9l^9$ in the generating function

$$F_{D_{20}}(r, q, l) = \sum_{\substack{i+j+k=20\\ i,j,k\geq 0}} a_{(i,j,k)}\, r^i q^j l^k.$$

In his influential paper [224] (translated into English by Read [226]), Pólya found that the cycle index can be used to compute the pattern inventory in a simple way. Recall that each occurrence of $x_k$ in the cycle index arises from a permutation having a cycle of length $k$, and if a coloring is invariant under this permutation, then these $k$ elements must have the same color. So either each of the $k$ objects permuted by this cycle has color $y_1$, or each one has color $y_2$, etc. In the spirit of generating functions, this choice can be represented by the formal sum $y_1^k + y_2^k + \cdots + y_m^k$. Pólya found that substituting this expression for $x_k$ for each $k$ in the cycle index yields the pattern inventory for the coloring.

**Theorem 2.10** (Pólya's Enumeration Formula)**.** *Suppose $S$ is a set of $n$ objects and $G$ is a subgroup of the symmetric group $S_n$. Let $P_G(\mathbf{x})$ be the cycle index of $G$. Then the pattern inventory for the nonequivalent colorings of $S$ under the action of $G$ using colors $y_1$, $y_2$, …, $y_m$ is*

$$F_G(\mathbf{y}) = P_G\left(\sum_{i=1}^{m} y_i,\ \sum_{i=1}^{m} y_i^2,\ \ldots,\ \sum_{i=1}^{m} y_i^n\right). \tag{2.62}$$

The proof we present follows Stanley [262, sec. 7.24].

*Proof.* Let $\mathbf{v} = (n_1, n_2, \ldots, n_m)$ be a vector of nonnegative integers of length $m$ whose components sum to $n$, and let $C_{\mathbf{v}}$ denote the set of colorings of $S$ where exactly $n_i$ of the objects have the color $y_i$, for each $i$. Let $C_{\mathbf{v},\pi}$ denote the invariant set of $C_{\mathbf{v}}$ under the action of a permutation $\pi$.

If a permutation $\pi$ in $G$ does not disturb a particular coloring, then every object permuted by one cycle of $\pi$ must have the same color. Therefore, $|C_{\mathbf{v},\pi}|$ is the coefficient of $y_1^{n_1} y_2^{n_2} \cdots y_m^{n_m}$ in $M_\pi(\sum y_i, \sum y_i^2, \ldots, \sum y_i^n)$, where $M_\pi$ is the monomial defined by (2.56). Let $\mathbf{y}^{\mathbf{v}}$ denote the term $y_1^{n_1} y_2^{n_2} \cdots y_m^{n_m}$. Then, summing over all permissible vectors $\mathbf{v}$, we obtain

$$\sum_{\mathbf{v}} |C_{\mathbf{v},\pi}| \, \mathbf{y}^{\mathbf{v}} = M_\pi \left( \sum_{i=1}^m y_i, \; \sum_{i=1}^m y_i^2, \; \ldots, \; \sum_{i=1}^m y_i^n \right).$$

Now we sum both expressions over all $\pi \in G$ and divide by $|G|$. On the left side, we have

$$\begin{aligned} \frac{1}{|G|} \sum_{\pi \in G} \sum_{\mathbf{v}} |C_{\mathbf{v},\pi}| \, \mathbf{y}^{\mathbf{v}} &= \sum_{\mathbf{v}} \left( \frac{1}{|G|} \sum_{\pi \in G} |C_{\mathbf{v},\pi}| \right) \mathbf{y}^{\mathbf{v}} \\ &= \sum_{\mathbf{v}} a_{\mathbf{v}} \mathbf{y}^{\mathbf{v}} \end{aligned}$$

by Burnside's Lemma, and this is the pattern inventory (2.61). On the right side, using (2.57), we obtain (2.62), the cycle index of $G$ evaluated at $x_k = \sum_i y_i^k$:

$$\begin{aligned} F_G(\mathbf{y}) &= \frac{1}{|G|} \sum_{\pi \in G} M_\pi \left( \sum_{i=1}^m y_i, \; \sum_{i=1}^m y_i^2, \; \ldots, \; \sum_{i=1}^m y_i^n \right) \\ &= P_G \left( \sum_{i=1}^m y_i, \; \sum_{i=1}^m y_i^2, \; \ldots, \; \sum_{i=1}^m y_i^n \right). \end{aligned}$$ □

For example, the pattern inventory for nonequivalent four-bead necklaces under $D_4$ using colors red ($r$), green ($g$), and blue ($b$) is

$$\begin{aligned} F_{D_4}(r,g,b) &= P_{D_4} \left( r+g+b, \; r^2+g^2+b^2, \; r^3+g^3+b^3, \; r^4+g^4+b^4 \right) \\ &= r^4+g^4+b^4+r^3g+rg^3+r^3b+rb^3+g^3b+gb^3 \\ &\quad + 2r^2g^2+2r^2b^2+2g^2b^2+2r^2gb+2rg^2b+2rgb^2. \end{aligned}$$

The pattern inventory for nonequivalent four-bead necklaces under $C_4$ using the same three colors is

$$\begin{aligned} F_{C_4}(r,g,b) &= P_{C_4} \left( r+g+b, \; r^2+g^2+b^2, \; r^3+g^3+b^3, \; r^4+g^4+b^4 \right) \\ &= r^4+g^4+b^4+r^3g+rg^3+r^3b+rb^3+g^3b+gb^3 \\ &\quad + 2r^2g^2+2r^2b^2+2g^2b^2+3r^2gb+3rg^2b+3rgb^2. \end{aligned}$$

Notice that there are three nonequivalent necklaces with two red beads, one green bead, and one blue bead under $C_4$, but only two under $D_4$. Can you explain this?

Using (2.60) and Theorem 2.10, we may compute the pattern inventory for twenty-bead necklaces composed of rhodonite ($r$), rose quartz ($q$), and lapis lazuli ($l$) beads. This pattern inventory is shown in Figure 2.7, where we see that there are exactly 231 260 different necklaces with two rhodonite, nine rose quartz, and nine lapis lazuli beads.

$$
\begin{aligned}
F_{D_{20}}(r,q,l) = {} & r^{20} + r^{19}q + r^{19}l + 10r^{18}q^2 + 10r^{18}ql + 10r^{18}l^2 + 33r^{17}q^3 + 90r^{17}q^2l \\
& + 90r^{17}ql^2 + 33r^{17}l^3 + 145r^{16}q^4 + 489r^{16}q^3l + 774r^{16}q^2l^2 + 489r^{16}ql^3 + 145r^{16}l^4 \\
& + 406r^{15}q^5 + 1956r^{15}q^4l + 3912r^{15}q^3l^2 + 3912r^{15}q^2l^3 + 1956r^{15}ql^4 + 406r^{15}l^5 \\
& + 1032r^{14}q^6 + 5832r^{14}q^5l + 14724r^{14}q^4l^2 + 19416r^{14}q^3l^3 + 14724r^{14}q^2l^4 \\
& + 5832r^{14}ql^5 + 1032r^{14}l^6 + 1980r^{13}q^7 + 13608r^{13}q^6l + 40824r^{13}q^5l^2 \\
& + 67956r^{13}q^4l^3 + 67956r^{13}q^3l^4 + 40824r^{13}q^2l^5 + 13608r^{13}ql^6 + 1980r^{13}l^7 \\
& + 3260r^{12}q^8 + 25236r^{12}q^7l + 88620r^{12}q^6l^2 + 176484r^{12}q^5l^3 + 221110r^{12}q^4l^4 \\
& + 176484r^{12}q^3l^5 + 88620r^{12}q^2l^6 + 25236r^{12}ql^7 + 3260r^{12}l^8 + 4262r^{11}q^9 \\
& + 37854r^{11}q^8l + 151416r^{11}q^7l^2 + 352968r^{11}q^6l^3 + 529452r^{11}q^5l^4 + 529452r^{11}q^4l^5 \\
& + 352968r^{11}q^3l^6 + 151416r^{11}q^2l^7 + 37854r^{11}ql^8 + 4262r^{11}l^9 + 4752r^{10}q^{10} \\
& + 46252r^{10}q^9l + 208512r^{10}q^8l^2 + 554520r^{10}q^7l^3 + 971292r^{10}q^6l^4 + 1164342r^{10}q^5l^5 \\
& + 971292r^{10}q^4l^6 + 554520r^{10}q^3l^7 + 208512r^{10}q^2l^8 + 46252r^{10}ql^9 + 4752r^{10}l^{10} \\
& + 4262r^9q^{11} + 46252r^9q^{10}l + 231260r^9q^9l^2 + 693150r^9q^8l^3 + 1386300r^9q^7l^4 \\
& + 1940568r^9q^6l^5 + 1940568r^9q^5l^6 + 1386300r^9q^4l^7 + 693150r^9q^3l^8 + 231260r^9q^2l^9 \\
& + 46252r^9ql^{10} + 4262r^9l^{11} + 3260r^8q^{12} + 37854r^8q^{11}l + 208512r^8q^{10}l^2 \\
& + 693150r^8q^9l^3 + 1560534r^8q^8l^4 + 2494836r^8q^7l^5 + 2912112r^8q^6l^6 + 2494836r^8q^5l^7 \\
& + 1560534r^8q^4l^8 + 693150r^8q^3l^9 + 208512r^8q^2l^{10} + 37854r^8ql^{11} + 3260r^8l^{12} \\
& + 1980r^7q^{13} + 25236r^7q^{12}l + 151416r^7q^{11}l^2 + 554520r^7q^{10}l^3 + 1386300r^7q^9l^4 \\
& + 2494836r^7q^8l^5 + 3326448r^7q^7l^6 + 3326448r^7q^6l^7 + 2494836r^7q^5l^8 + 1386300r^7q^4l^9 \\
& + 554520r^7q^3l^{10} + 151416r^7q^2l^{11} + 25236r^7ql^{12} + 1980r^7l^{13} + 1032r^6q^{14} \\
& + 13608r^6q^{13}l + 88620r^6q^{12}l^2 + 352968r^6q^{11}l^3 + 971292r^6q^{10}l^4 + 1940568r^6q^9l^5 \\
& + 2912112r^6q^8l^6 + 3326448r^6q^7l^7 + 2912112r^6q^6l^8 + 1940568r^6q^5l^9 + 971292r^6q^4l^{10} \\
& + 352968r^6q^3l^{11} + 88620r^6q^2l^{12} + 13608r^6ql^{13} + 1032r^6l^{14} + 406r^5q^{15} + 5832r^5q^{14}l \\
& + 40824r^5q^{13}l^2 + 176484r^5q^{12}l^3 + 529452r^5q^{11}l^4 + 1164342r^5q^{10}l^5 + 1940568r^5q^9l^6 \\
& + 2494836r^5q^8l^7 + 2494836r^5q^7l^8 + 1940568r^5q^6l^9 + 1164342r^5q^5l^{10} + 529452r^5q^4l^{11} \\
& + 176484r^5q^3l^{12} + 40824r^5q^2l^{13} + 5832r^5ql^{14} + 406r^5l^{15} + 145r^4q^{16} + 1956r^4q^{15}l \\
& + 14724r^4q^{14}l^2 + 67956r^4q^{13}l^3 + 221110r^4q^{12}l^4 + 529452r^4q^{11}l^5 + 971292r^4q^{10}l^6 \\
& + 1386300r^4q^9l^7 + 1560534r^4q^8l^8 + 1386300r^4q^7l^9 + 971292r^4q^6l^{10} + 529452r^4q^5l^{11} \\
& + 221110r^4q^4l^{12} + 67956r^4q^3l^{13} + 14724r^4q^2l^{14} + 1956r^4ql^{15} + 145r^4l^{16} + 33r^3q^{17} \\
& + 489r^3q^{16}l + 3912r^3q^{15}l^2 + 19416r^3q^{14}l^3 + 67956r^3q^{13}l^4 + 176484r^3q^{12}l^5 \\
& + 352968r^3q^{11}l^6 + 554520r^3q^{10}l^7 + 693150r^3q^9l^8 + 693150r^3q^8l^9 + 554520r^3q^7l^{10} \\
& + 352968r^3q^6l^{11} + 176484r^3q^5l^{12} + 67956r^3q^4l^{13} + 19416r^3q^3l^{14} + 3912r^3q^2l^{15} \\
& + 489r^3ql^{16} + 33r^3l^{17} + 10r^2q^{18} + 90r^2q^{17}l + 774r^2q^{16}l^2 + 3912r^2q^{15}l^3 \\
& + 14724r^2q^{14}l^4 + 40824r^2q^{13}l^5 + 88620r^2q^{12}l^6 + 151416r^2q^{11}l^7 + 208512r^2q^{10}l^8 \\
& + \mathbf{231260r^2q^9l^9} + 208512r^2q^8l^{10} + 151416r^2q^7l^{11} + 88620r^2q^6l^{12} + 40824r^2q^5l^{13} \\
& + 14724r^2q^4l^{14} + 3912r^2q^3l^{15} + 774r^2q^2l^{16} + 90r^2ql^{17} + 10r^2l^{18} + rq^{19} + 10rq^{18}l \\
& + 90rq^{17}l^2 + 489rq^{16}l^3 + 1956rq^{15}l^4 + 5832rq^{14}l^5 + 13608rq^{13}l^6 + 25236rq^{12}l^7 \\
& + 37854rq^{11}l^8 + 46252rq^{10}l^9 + 46252rq^9l^{10} + 37854rq^8l^{11} + 25236rq^7l^{12} \\
& + 13608rq^6l^{13} + 5832rq^5l^{14} + 1956rq^4l^{15} + 489rq^3l^{16} + 90rq^2l^{17} + 10rql^{18} + rl^{19} \\
& + q^{20} + q^{19}l + 10q^{18}l^2 + 33q^{17}l^3 + 145q^{16}l^4 + 406q^{15}l^5 + 1032q^{14}l^6 + 1980q^{13}l^7 \\
& + 3260q^{12}l^8 + 4262q^{11}l^9 + 4752q^{10}l^{10} + 4262q^9l^{11} + 3260q^8l^{12} + 1980q^7l^{13} \\
& + 1032q^6l^{14} + 406q^5l^{15} + 145q^4l^{16} + 33q^3l^{17} + 10q^2l^{18} + ql^{19} + l^{20}
\end{aligned}
$$

FIGURE 2.7. Pattern inventory for necklaces with twenty beads formed using three kinds of beads.

Pólya's enumeration formula has many applications in several fields, including chemistry, physics, and computer science. Pólya devotes a large portion of his paper [224] to applications involving enumeration of graphs, trees, and chemical isomers.

**Exercises**

1. What is the pattern inventory for coloring $n$ objects using the $m$ colors $y_1$, $y_2, \ldots, y_m$ if the group of symmetries is $S_n$?

2. Use Pólya's enumeration formula to determine the number of six-sided dice that can be manufactured if each of three different labels must be placed on two of the faces.

3. The hydrocarbon benzene has six carbon atoms arranged at the vertices of a regular hexagon, and six hydrogen atoms, with one bonded to each carbon atom. Two molecules are said to be *isomers* if they are composed of the same number and types of atoms, but have different structure.

   (a) Show that exactly three isomers (ortho-dichlorobenzene, meta-dichlorobenzene, and para-dichlorobenzene) may be constructed by replacing two of the hydrogen atoms of benzene with chlorine atoms.

   (b) How many isomers may be obtained by replacing two of the hydrogen atoms with chlorine atoms, and two others with bromine atoms?

4. The hydrocarbon naphthalene has ten carbon atoms arranged in a double hexagon as in Figure 2.8, and eight hydrogen atoms attached at each of the positions labeled 1 through 8.

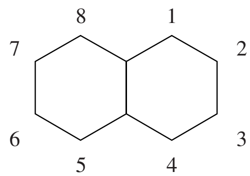



FIGURE 2.8. Naphthalene.

   (a) Naphthol is obtained by replacing one of the hydrogen atoms of naphthalene with a hydroxyl group (OH). How many isomers of naphthol are there?

   (b) Tetramethylnaphthalene is obtained by replacing four of the hydrogen atoms of naphthalene with methyl groups ($CH_3$). How many isomers of tetramethylnaphthalene are there?

(c) How many isomers may be constructed by replacing three of the hydrogen molecules of naphthalene with hydroxyl groups, and another three with methyl groups?

(d) How many isomers may be constructed by replacing two of the hydrogen molecules of naphthalene with hydroxyl groups, two with methyl groups, and two with carboxyl groups (COOH)?

5. The hydrocarbon anthracene has fourteen carbon atoms arranged in a triple hexagon as in Figure 2.9, with ten hydrogen atoms bonded at the numbered positions.

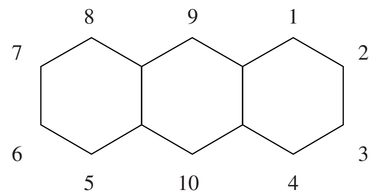



FIGURE 2.9. Anthracene.

(a) How many isomers of trimethylanthracene can be formed by replacing three hydrogen atoms with methyl groups?

(b) How many isomers can be formed by replacing four of the hydrogen atoms with chlorine, and two others with hydroxyl groups?

6. The molecule triphenylamine has three rings of six carbon atoms attached to a central nitrogen atom, as in Figure 2.10, and fifteen hydrogen atoms, with one attached to each carbon atom except the three carbons attached to the central nitrogen atom.

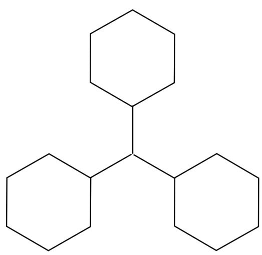


FIGURE 2.10. Triphenylamine.

(a) How many isomers can be formed by replacing six hydrogen atoms with hydroxyl groups?

(b) How many isomers can be formed by replacing five hydrogen atoms with methyl groups, and five with fluorine atoms?

7. The hydrocarbon tetraphenylmethane consists of four rings of six carbon atoms, each bonded to a central carbon atom, as in Figure 2.11, together with twenty hydrogen atoms, with one hydrogen atom attached to each carbon atom in the rings except for those attached to the carbon at the center.

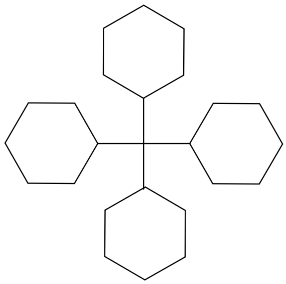


FIGURE 2.11. Tetraphenylmethane.

(a) How many isomers can be formed by replacing five hydrogen atoms of tetraphenylmethane with chlorine?

(b) How many isomers can be formed by replacing five hydrogen atoms with bromine, and six others with hydroxyl groups?

8. Suppose a medical relief agency plans to design a symbol for their organization in the shape of a regular cross, as in Figure 2.12. To symbolize the purpose of the organization and emphasize its international constituency, its board of directors decides that the cross should be white in color, with each of the twelve line segments outlining the cross colored red, green, blue, or yellow, with an equal number of lines of each color. If we discount rotations and flips, how many different ways are there to design the symbol?

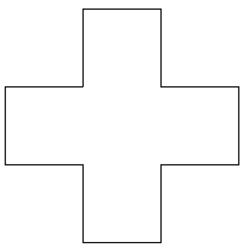


FIGURE 2.12. Symbol of a relief agency.

### 2.7.5 de Bruijn's Generalization

*It doesn't matter what color, well that gets a nope!*
*Be it pink, purple, or heliotrope!*

— *Boundin'*, Pixar Films

Suppose a jewelry company plans to market a new line of unisex bracelets under the brand name OPPOSITES ATTRACT. The bracelets are sold in pairs, for a couple to share. Each bracelet consists of $n$ beads, some gold and some silver, and the two bracelets in a pair are opposites, in the sense that one can be obtained from the other by changing each silver bead to a gold one and each gold to a silver. For example, if one bracelet has two adjacent gold beads and $n-2$ silver beads, then its mate has two adjacent silver beads and $n-2$ gold beads. The companion then of the all-gold bracelet is the all-silver one. How many different pairs of $n$-bead bracelets are possible in the OPPOSITES ATTRACT product line?

We have seen that there are exactly six different bracelets for the case $n=4$, if we discount both rotations and flips. These are represented by the configurations $gggg$, $gggs$, $ggss$, $gsgs$, $gsss$, and $ssss$ of gold and silver beads. This produces just four different (unordered) pairs of bracelets for the product line when $n=4$:

$$gggg + ssss,\ gggs + gsss,\ ggss + ggss,\ gsgs + gsgs. \tag{2.63}$$

Recall that each of the configurations we listed for $n=4$ in fact represent an equivalence class of the set of two-colorings of the vertices of a square, where we consider two colorings to be equivalent if one can be obtained from the other by the action of some element of the group of symmetries of the square, $D_4$. In the same way, we may consider each of the pairs of bracelets in our product line as representing a single set of two-colorings of the square—the union of the equivalence classes of the two bracelets in the set. For example, the four pairs listed in (2.63) correspond to the following partition of the sixteen ways to color the vertices of a square using at most two colors:

$$\begin{aligned}
&\{gggg, ssss\},\\
&\{gggs, ggsg, gsgg, sggg, sssg, ssgs, sgss, gsss\},\\
&\{ggss, gssg, ssgg, sggs\},\\
&\{gsgs, sgsg\}.
\end{aligned}$$

This partition is precisely the collection of equivalence classes of two-colorings under a different equivalence relation. Now we consider two colorings to be equivalent if one can be obtained from the other by first performing some geometric transformation corresponding to a symmetry of the bracelet, then possibly inverting all the colors. It is easy to check that this is indeed an equivalence relation.

We can generalize this problem in the following way. Given a set of objects $S$, a set of colors $R$, a group $G$ acting on $S$, and a group $H$ acting on $R$. Let $C$ denote the set of colorings of $S$ using the colors in $R$, so this is the set of all functions from $S$ into $R$. We consider two colorings in $C$ to be equivalent if one can be obtained from the other by first applying a permutation from $G$ on the objects, then applying a permutation from $H$ on the colors. Exercise 1 asks you to verify that this does in fact form an equivalence relation on $C$. We would like to know the number of equivalence classes of $C$ with respect to $G$ and $H$.

In our example with four-bead bracelets, we have that $S$ is the set of vertices of a square, $R = \{g, s\}$ for gold and silver beads, $G = D_4$, and $H = S_2$, since we may either leave the beads unchanged, or swap them. A permutation $\pi \in G$ induces a map $\pi^*$ on $C$ in the usual way. For instance, if $\pi$ is the 90-degree rotation $(1234)$, then $\pi^*(gggs) = sggg$. Similarly, a permutation $\rho \in H$ induces a map $\rho^*$ on $C$. For example, if $\rho = (12)$ then $\rho^*(gggs) = sssg$.

Of course, if $H$ is the trivial group consisting only of the identity permutation $(1)$, then we can use the cycle index and the enumeration formula of Pólya to determine the answer. The Dutch mathematician Nicolaas Govert de Bruijn generalized the method of Pólya for arbitrary color groups $H$, and we describe this theory here. The first step is computing the set of equivalence classes of $C$ with respect to the object group $G$ which are invariant with respect to a given permutation of the colors. Our proof follows de Bruijn's paper [69].

**Theorem 2.11.** *Suppose $S$ is a set of $n$ objects, $R = \{y_1, \ldots, y_m\}$ is a set of $m$ colors, $G$ is a subgroup of the symmetric group $S_n$, and $\rho \in S_m$. Let $P_G(\mathbf{x})$ denote the cycle index of $G$. Then the pattern inventory for the colorings of $S$ which are nonequivalent with respect to the action of $G$ on $S$, but invariant with respect to the action of $\rho$ on $R$, is*

$$F_{G,\rho}(\mathbf{y}) = P_G(\alpha_1(\rho), \alpha_2(\rho), \ldots, \alpha_n(\rho)), \tag{2.64}$$

*where*

$$\alpha_k(\rho) = \sum_{\rho^k(j)=j} \prod_{i=0}^{k-1} y_{\rho^i(j)}$$

*for $1 \le k \le n$.*

*Proof.* Let $C$ denote the set of all colorings of $S$, so $C$ is the set of maps from $S$ into $R$. For a particular coloring $c \in C$, let $\overline{c}$ denote its orbit with respect to the group $G$, so $\overline{c} = \{\pi^*(c) : \pi \in G\}$. Also, let $\mathbf{v}(c) = (n_1, n_2, \ldots, n_m)$, where for each $i$ the integer $n_i$ records the number of elements of $S$ assigned the color $y_i$ in $c$, and let $\mathbf{y}^{\mathbf{v}(c)}$ denote the monomial $y_1^{n_1} y_2^{n_2} \cdots y_m^{n_m}$. Since $\mathbf{v}(\pi^*(c)) = \mathbf{v}(c)$ for any $\pi \in G$, we may define $\mathbf{y}^{\mathbf{v}(\overline{c})}$ by $\mathbf{y}^{\mathbf{v}(\overline{c})} = \mathbf{y}^{\mathbf{v}(c)}$.

Suppose that $c$ is invariant under the action of $\rho$, that is $\rho^*(c) \in \overline{c}$. Since we want to find the pattern inventory for the classes of colorings of $S$ that are nonequivalent with respect to $G$, but invariant with respect to $\rho$, we need to study the generating function

$$F_{G,\rho}(\mathbf{y}) = \sum_{\rho(\overline{c})=\overline{c}} \mathbf{y}^{\mathbf{v}(\overline{c})}.$$

Since $G$ is a group, it is straightforward to show that the set of all colorings that are invariant under $\rho$ is the union of all the orbits $\overline{c}$ where $\rho(\overline{c}) = \overline{c}$. Thus, using

Lemma 2.8 we find that

$$F_{G,\rho}(\mathbf{y}) = \sum_{\rho^*(c)\in\overline{c}} \frac{\mathbf{y}^{\mathbf{v}(c)}}{|\overline{c}|} = \frac{1}{|G|} \sum_{\rho^*(c)\in\overline{c}} |G_c|\, \mathbf{y}^{\mathbf{v}(c)},$$

where $G_c$ is the stabilizer of $c$ in $G$. Now since $\rho^*(c) \in \overline{c}$, there exists a permutation $\pi_c \in G$ such that $\rho^*(c) = \pi_c^*(c)$. Also, the set of permutations $\{\pi_c \circ \pi : \pi \in G_c\}$ is exactly the same as the set $\{\pi \in G : \pi^*(c) = \pi_c^*(c)\}$, so $|G_c|$ equals the number of permutations in $G$ which have the same effect as $\rho$ on $c$:

$$|G_c| = |\{\pi \in G : \pi^*(c) = \rho^*(c)\}|\,.$$

Let $U_\pi$ denote the set of colorings $c$ for which $\pi$ and $\rho$ have the same effect,

$$U_\pi = \{c \in C : \pi^*(c) = \rho^*(c)\}.$$

Note that if $c \in U_\pi$ then automatically $\rho^*(c) \in \overline{c}$. Thus, we find that

$$F_{G,\rho}(\mathbf{y}) = \frac{1}{|G|} \sum_{\pi\in G} \sum_{c\in U_\pi} \mathbf{y}^{\mathbf{v}(c)}. \tag{2.65}$$

Now suppose $\pi \in G$, and $\pi$ has $\lambda_i$ cycles of length $i$, for each $i$ with $1 \le i \le n$. Let $\ell_i$ denote the length of the $i$th cycle (when $\pi$ is written in cycle notation in the canonical way), and let $s_i$ denote the smallest element of the $i$th cycle. For example, if $n = 7$ and $\pi = (1245)(37)(6)$, then $\lambda_1 = \lambda_2 = \lambda_4 = 1$, $\ell_1 = 4$, $\ell_2 = 2$, $\ell_3 = 1$, $s_1 = 1$, $s_2 = 3$, and $s_3 = 6$. Also, let $M_\pi(x_1, \ldots, x_n)$ denote the monomial obtained from $\pi$ as in (2.56), so in the example we have $M_\pi(x_1, \ldots, x_7) = x_1x_2x_4$.

Suppose that $c \in U_\pi$, so that applying $\pi$ to $c$ has the same effect as applying $\rho$. If position $s_i$ has color $y_j$ in $c$, it follows that position $\pi^{-1}(s_i)$ has color $y_{\rho(j)}$, position $\pi^{-2}(s_i)$ has color $y_{\rho^2(j)}, \ldots,$ position $\pi^{-(\ell_i-1)}(s_i)$ has color $y_{\rho^{(\ell_i-1)}(j)}$, and we require that $\rho^{\ell_i}(j) = j$. It therefore follows that

$$\sum_{c\in U_\pi} \mathbf{y}^{\mathbf{v}(c)} = M_\pi(\alpha_1(\rho), \alpha_2(\rho), \ldots, \alpha_n(\rho)),$$

and the theorem follows by combining this with (2.57) and (2.65). □

We can apply this theorem to our original example on bracelets, where $n = 4$, $m = 2$, $G = D_4$, and $\rho = (12)$. Write $y_1 = g$ for a gold bead, and $y_2 = s$ for a silver one. Then $\alpha_1(\rho) = 0$, since no color is left unchanged by $\rho$. Next, $\alpha_2(\rho) = y_1y_2 + y_2y_1 = 2gs$, since $\rho^2(j) = j$ for both $j = 1$ and $j = 2$. We then find that $\alpha_3(\rho) = 0$ since $\rho^3(1) = 2$ and $\rho^3(2) = 1$, and $\alpha_4(\rho) = y_1y_2y_1y_2 + y_2y_1y_2y_1 = 2g^2s^2$. Using (2.58), we obtain then that

$$F_{D_4,(12)}(g, s) = P_{D_4}(0, 2gs, 0, 2g^2s^2) = 2g^2s^2,$$

and we verify that there are indeed just two four-bead bracelets which are invariant under bead swapping, discounting rotations and flips: $gsgs$ and $ggss$.

If we introduce another type of bead in this example, say $y_3 = b$ for bronze, and keep $\rho = (12)$, then we obtain $\alpha_1(\rho) = b$, $\alpha_2(\rho) = 2gs + b^2$, $\alpha_3(\rho) = b^3$, and $\alpha_4(\rho) = 2g^2s^2 + b^4$, and we calculate that

$$F_{D_4,(12)}(g, s, b) = P_{D_4}(b, 2gs + b^2, b^3, 2g^2s^2 + b^4) = 2g^2s^2 + 2gsb^2 + b^4.$$

The five different configurations in this case are represented by $ggss$, $gsgs$, $gbsb$, $gsbb$, and $bbbb$.

We may now use Theorem 2.11 to solve our original problem. We would like to obtain the pattern inventory for a set of colorings $C$ when we account for both a group of symmetries $G$ on the objects, and a group $H$ of symmetries on the colors. We compute this pattern inventory by averaging the patterns $F_{G,\rho}(\mathbf{y})$ over all permutations $\rho$ in $H$, then combining the terms that correspond to equivalent patterns of colors.

**Theorem 2.12** (de Bruijn's Enumeration Formula)**.** *Suppose $S$ is a set of $n$ objects, $R = \{y_1, \ldots, y_m\}$ is a set of $m$ colors, $G$ is a subgroup of the symmetric group $S_n$, and $H$ is a subgroup of $S_m$. Then the pattern inventory $\hat{F}_{G,H}(\mathbf{y})$ for the colorings of $S$ which are nonequivalent with respect to both the action of $G$ on $S$ and the action of $H$ on $R$ is obtained by identifying equivalent color patterns in the polynomial*

$$F_{G,H}(\mathbf{y}) = \frac{1}{|H|} \sum_{\rho \in H} F_{G,\rho}(\mathbf{y}), \tag{2.66}$$

*where $F_{G,\rho}(\mathbf{y})$ is given by* (2.64)*.*

We describe one example before providing the proof. With $n = 4$, $m = 2$, $R = \{g, s\}$, $G = D_4$, and $H = S_2$, we compute

$$\begin{aligned} F_{D_4,S_2}(g, s) &= \tfrac{1}{2}\big(P_{D_4}(g + s, g^2 + s^2, g^3 + s^3, g^4 + s^4) \\ &\qquad + P_{D_4}(0, 2gs, 0, 2g^2s^2)\big) \\ &= \tfrac{1}{2}(g^4 + s^4) + \tfrac{1}{2}(g^3s + gs^3) + 2g^2s^2. \end{aligned} \tag{2.67}$$

The color patterns $g^4$ and $s^4$ are equivalent under the color group $H = S_2$, so we let $[g^4]$ denote either one of these patterns. Likewise, we let $[g^3s]$ denote either of the equivalent patterns $g^3s$ or $gs^3$. The last pattern, $g^2s^2$, is not equivalent to any of the others, so we let $[g^2s^2]$ designate this single pattern. We obtain the pattern inventory by combining the equivalent terms:

$$\hat{F}_{D_4,S_2}(g, s) = [g^4] + [g^3s] + 2[g^2s^2].$$

*Proof of Theorem 2.12.* Let $C$ denote the set of all colorings of $S$ using the colors of $R$, and let $\overline{c}$ denote the orbit of the coloring $c$ under the action of $G$, so $\overline{c} = \{\pi^*(c) : \pi \in G\}$. The group $H$ acts on the set of equivalence classes $\{\overline{c} : c \in C\}$, and we let $\overline{\overline{c}}$ denote the orbit of $\overline{c}$ under this action, so

$$\overline{\overline{c}} = \{\rho^*(\overline{c}) : \rho \in H\}.$$

In the example above, if $c = gggs$, then

$$\overline{c} = \{gggs, ggsg, gsgg, sggg\}$$

and

$$\overline{\overline{c}} = \{\{gggs, ggsg, gsgg, sggg\}, \{sssg, ssgs, sgss, gsss\}\};$$

if $c = gsgs$, then $\overline{c} = \{gsgs, sgsg\}$ and $\overline{\overline{c}} = \{\{gsgs, sgsg\}\}$.

Employing the notation we introduced in the proof of Theorem 2.9, and using Lemma 2.8, we compute

$$\begin{aligned}
\frac{1}{|H|}\sum_{\rho\in H} F_{G,\rho}(\mathbf{y}) &= \frac{1}{|H|}\sum_{\rho\in H}\sum_{\overline{c}}[\rho^*(\overline{c}) = \overline{c}]\,\mathbf{y}^{\mathbf{v}(\overline{c})} \\
&= \frac{1}{|H|}\sum_{\overline{c}}\mathbf{y}^{\mathbf{v}(\overline{c})}\sum_{\rho\in H}[\rho^*(\overline{c}) = \overline{c}] \\
&= \frac{1}{|H|}\sum_{\overline{c}}|H_{\overline{c}}|\,\mathbf{y}^{\mathbf{v}(\overline{c})} \qquad (2.68)\\
&= \sum_{\overline{c}}\left|\overline{\overline{c}}\right|^{-1}\mathbf{y}^{\mathbf{v}(\overline{c})} \\
&= \sum_{\overline{\overline{c}}}\left|\overline{\overline{c}}\right|^{-1}\sum_{\overline{c}\in\overline{\overline{c}}}\mathbf{y}^{\mathbf{v}(\overline{c})}.
\end{aligned}$$

Since the color patterns in the set $\{\mathbf{y}^{\mathbf{v}(\overline{c})} : \overline{c} \in \overline{\overline{c}}\}$ are equivalent under $H$, we select one pattern from this set to represent the class $\overline{\overline{c}}$, and denote this equivalence class of patterns by $[\mathbf{y}^{\mathbf{v}(\overline{c})}]$. By replacing each term $\mathbf{y}^{\mathbf{v}(\overline{c})}$ in the last line of (2.68) by its representative class $[\mathbf{y}^{\mathbf{v}(\overline{c})}]$, we obtain the pattern inventory,

$$\hat{F}_{G,H}(\mathbf{y}) = \sum_{\overline{\overline{c}}}[\mathbf{y}^{\mathbf{v}(\overline{c})}]. \qquad \square$$

We can use Theorem 2.12 to determine the number different ten-bead pairs of bracelets in the OPPOSITES ATTRACT product line having a given configuration of colors. Since

$$P_{D_{10}}(\mathbf{x}) = \tfrac{1}{2}\left(x_1^{10} + x_2^5 + 4x_5^2 + 4x_{10} + 5x_2^5 + 5x_1^2x_2^4\right),$$

we compute

$$\begin{aligned}
F_{D_{10},S_2}(g,s) &= \tfrac{1}{2}\big(P_{D_{10}}(g+s, g^2+s^2, \ldots, g^{10}+s^{10}) \\
&\qquad + P_{D_{10}}(0, 2gs, \ldots, 0, 2g^5s^5)\big) \\
&= \tfrac{1}{2}(g^{10}+s^{10}) + \tfrac{1}{2}(g^9s+gs^9) + \tfrac{5}{2}(g^8s^2+g^2s^8) \\
&\qquad + 4(g^7s^3+g^3s^7) + 8(g^6s^4+g^4s^6) + 13g^5s^5,
\end{aligned}$$

so the pattern inventory for these pairs of bracelet is

$$\hat{F}_{D_{10},S_2}(g,s) = [g^{10}] + [g^9s] + 5[g^8s^2] + 8[g^7s^3] + 16[g^6s^4] + 13[g^5s^5].$$

The situation is much simpler if we need only compute the total number of distinct colorings with respect to $G$ and $H$, and we do not need the finer information provided by the pattern inventory. For this case, we need only set each $y_i = 1$ in $F_{G,H}(\mathbf{y})$, and there is no need to compute $\hat{F}_{G,H}(\mathbf{y})$. Since this case is so common, we describe its solution as a corollary to Theorem 2.12. Its proof is left as an exercise.

**Corollary 2.13.** *Suppose $S$ is a set of $n$ objects, $R$ is a set of $m$ colors, $G$ is a subgroup of the symmetric group $S_n$, and $H$ is a subgroup of $S_m$. Then the number of colorings of $S$ using the colors in $R$ which are nonequivalent with respect to both the action of $G$ on $S$ and the action of $H$ on $R$ is*

$$N_{G,H}(n,m) = \frac{1}{|H|}\sum_{\rho\in H} P_G(\beta_1(\rho), \beta_2(\rho), \ldots, \beta_n(\rho)), \tag{2.69}$$

*where $\beta_k(\rho) = \sum_{j|k} j\lambda_j(\rho)$, with the sum extending over all the positive divisors $j$ of $k$, and $\lambda_j(\rho)$ is the number of cycles of $\rho$ of length $j$.*

For example, for our ten-bead bracelet problem with $m = 2$ and $H = S_2$, we find that the only nonzero values of the $\lambda_j(\rho)$ are $\lambda_1(\,(1)(2)\,) = 2$ and $\lambda_2(\,(12)\,) = 1$. It follows that $\beta_k(\,(1)(2)\,) = 2$ for $1 \le k \le 10$, and

$$\beta_k(\,(12)\,) = \begin{cases} 2 & \text{if } k \text{ is even,} \\ 0 & \text{if } k \text{ is odd.} \end{cases}$$

Therefore,

$$N_{D_{10},S_2}(10,2) = \tfrac{1}{2}\big(P_{D_{10}}(2,2,\ldots,2) + P_{D_{10}}(0,2,\ldots,0,2)\big) = 44.$$

Last, we return to the problem from earlier sections concerning twenty-bead necklaces using rhodonite, rose quartz, and lapis lazuli beads. Using $H = \langle(123)\rangle$, we find that

$$\beta_k(\,(123)\,) = \beta_k(\,(132)\,) = \begin{cases} 3 & \text{if } 3 \mid k, \\ 0 & \text{if } 3 \nmid k, \end{cases}$$

and $\beta_k(\,(1)\,) = 3$ for each $k$. Thus,

$$\begin{aligned} N_{D_{20},C_3}(20,3) &= \tfrac{1}{3}\big(P_{D_{20}}(3,\ldots,3) + 2P_{D_{20}}(0,0,3,\ldots,0,0,3,0,0)\big) \\ &= \tfrac{1}{3}P_{D_{20}}(3,\ldots,3) = 29\,076\,719, \end{aligned}$$

since none of the variables $x_{3k}$ appears in $P_{D_{20}}(\mathbf{x})$. This then is the number of different 20-bead necklaces if we discount rotations, flips, and the bead substitutions rhodonite $\to$ rose quartz $\to$ lapis lazuli $\to$ rhodonite, or rhodonite $\to$ lapis lazuli $\to$ rose quartz $\to$ rhodonite.

Using $H = S_3$ instead, we obtain

$$\beta_k(\,(12)\,) = \beta_k(\,(13)\,) = \beta_k(\,(23)\,) = \begin{cases} 1 & \text{if } k \text{ is odd,} \\ 3 & \text{if } k \text{ is even,} \end{cases}$$

and so

$$\begin{aligned} N_{D_{20},S_3}(20,3) &= \tfrac{1}{6}\big(P_{D_{20}}(3,\ldots,3) + 2P_{D_{20}}(0,0,3,\ldots,0,0,3,0,0) \\ &\qquad + 3P_{D_{20}}(1,3,\ldots,1,3)\big) \\ &= \tfrac{1}{6}(87\,230\,157 + 63\,519) = 14\,548\,946. \end{aligned}$$

This is therefore the number of different 20-bead necklaces if we discount rotations, flips, and any permutation of the bead types.

**Exercises**

1. Suppose that a group $G$ acts on a set $S$ of objects, and a group $H$ acts on a set $R$ of colors. Let $C$ denote the set of functions from $S$ into $R$, that is, the number of colorings of $S$ using the colors in $R$. If $c_1$ and $c_2$ are two colorings in $C$, write $c_1 \sim c_2$ if there exists an element $g \in G$ and an element $h \in H$ such that applying $g$ to the underlying objects of $c_1$, then $h$ to its colors, produces $c_2$. Show that $\sim$ induces an equivalence relation on $C$.

2. Suppose that $G$ is a group acting on a set of objects $S$, and that $C$ is the set of colorings of elements of $S$ using the colors in a set $R$. Let $\overline{c}$ denote the orbit of $c$ in $C$ with respect to the action of $G$. Let $\rho$ be a permutation acting on $R$. Prove that $\{c \in C : \rho^*(c) \in \overline{c}\} = \bigcup_{\rho^*(\overline{c})=\overline{c}} \overline{c}$.

3. Compute the number of different pairs of bracelets in the OPPOSITES ATTRACT product line for $n = 6$, $n = 7$, and $n = 8$.

4. Our jewelry company plans to extend their line of bracelets by introducing sets of $m$ bracelets formed using $m$ different colors of beads, so that a set may be shared among a group of $m$ people. If one bracelet in a package has the coloring $c$, then the others in the package have the coloring $\rho^*(c)$, $(\rho^*)^2(c)$, …, $(\rho^*)^{m-1}(c)$, where $\rho$ is the cyclic permutation $(1\,2\cdots m)$. Use $D_n$ for the object group $G$ in each of the following problems.

   (a) Compute the number of different packages of bracelets for $m = 3$ when $n = 6$, then $n = 7$, then $n = 9$.

   (b) Compute the number of different packages of bracelets for $m = 4$ when $n = 10$, then when $n = 12$.

   (c) Determine the pattern inventory $\hat{F}_{D_n,C_m}(\mathbf{y})$ for the case $m = 3$ and $n = 6$, then $n = 9$.

5. Compute the pattern inventory $\hat{F}_{D_n,S_2}(x,y)$ for $n = 6$, $n = 7$, and $n = 8$.

6. Compute the pattern inventory $\hat{F}_{C_n,S_2}(x,y)$ for $n = 6$, $n = 7$, and $n = 8$.

7. Verify that the pattern inventory $\hat{F}_{D_{20},S_3}(r,q,l)$ for 20-bead necklaces with three kinds of beads, using the full symmetric group $S_3$ for $H$, is

$$\begin{aligned}
\hat{F}_{D_{20},S_3}(r,q,l) &= [r^{20}] + [r^{19}q] + 10[r^{18}q^2] + 10[r^{18}ql] + 33[r^{17}q^3] \\
&\quad + 90[r^{17}q^2l] + 145[r^{16}q^4] + 489[r^{16}q^3l] + 430[r^{16}q^2l^2] \\
&\quad + 406[r^{15}q^5] + 1956[r^{15}q^4l] + 3912[r^{15}q^3l^2] + 1032[r^{14}q^6] \\
&\quad + 5832[r^{14}q^5l] + 14724[r^{14}q^4l^2] + 9924[r^{14}q^3l^3] + 1980[r^{13}q^7] \\
&\quad + 13608[r^{13}q^6l] + 40824[r^{13}q^5l^2] + 67956[r^{13}q^4l^3] + 3260[r^{12}q^8] \\
&\quad + 25236[r^{12}q^7l] + 88620[r^{12}q^6l^2] + 176484[r^{12}q^5l^3] \\
&\quad + 111270[r^{12}q^4l^4] + 4262[r^{11}q^9] + 37854[r^{11}q^8l] \\
&\quad + 151416[r^{11}q^7l^2] + 352968[r^{11}q^6l^3] + 529452[r^{11}q^5l^4] \\
&\quad + 2518[r^{10}q^{10}] + 46252[r^{10}q^9l] + 208512[r^{10}q^8l^2] \\
&\quad + 554520[r^{10}q^7l^3] + 971292[r^{10}q^6l^4] + 583784[r^{10}q^5l^5] \\
&\quad + 116398[r^9q^9l^2] + 693150[r^9q^8l^3] + 1386300[r^9q^7l^4] \\
&\quad + 1940568[r^9q^6l^5] + 782141[r^8q^8l^4] + 2494836[r^8q^7l^5] \\
&\quad + 1458578[r^8q^6l^6] + 1665912[r^7q^7l^6].
\end{aligned}$$

8. Prove Corollary 2.13.

9. Consider the symbol of the medical relief agency shown in Figure 2.12. Each of the twelve line segments outlining the cross shape must be colored red, green, blue, or yellow.

   (a) How many ways are there to design the symbol, if we consider two configurations equivalent if one can be obtained from the other by some combination of a rotation, flip, and color reversal? A color reversal exchanges red and green, and exchanges blue and yellow.

   (b) How many of these configurations have the same number of edges of each color?

   (c) Repeat the first two problems, but this time consider two colorings to be equivalent if one can be obtained from the other by either exchanging red and green, or exchanging blue and yellow, or both.

   (d) Repeat the first two problems, but now consider two colorings to be equivalent if one can be obtained from the other by an iterate of the cyclic permutation red $\rightarrow$ green $\rightarrow$ blue $\rightarrow$ yellow $\rightarrow$ red.

   (e) Suppose now that black is added as a possible color for a segment of the border. How many ways are there to design the symbol, if we consider two configurations equivalent if one can be obtained from the other by some combination of a rotation, flip, and color reversal?

A color reversal exchanges red and green, exchanges blue and yellow, and leaves black fixed.

(f) Repeat the previous problem, but this time consider two colorings to be equivalent if one can be obtained from the other by either exchanging red and green, or exchanging blue and yellow, or both.

10. Determine the number of ways to color the faces of a cube using the three colors maroon, cardinal, and burnt orange, if two colorings are considered to be equivalent if one can be obtained from the other by rotating the cube in some way in three-dimensional space, and possibly exchanging maroon and burnt orange. Then determine the number of such colorings in which maroon and burnt orange appear the same number of times.

11. Determine the number of ways to color the faces of an octahedron using the four colors heliotrope, lavender, thistle, and wisteria, if two colorings are considered to be equivalent if one can be obtained from the other by rotating the octahedron in some way, and possibly exchanging heliotrope and lavender, or thistle and wisteria, or both. Then determine the number of such colorings in which the number of faces colored heliotrope matches the number colored lavender, and at the same time the number of faces colored thistle matches the number colored wisteria.

## 2.8 More Numbers

*Truly, I thought there had been one number more…*

— William Shakespeare, *The Merry Wives of Windsor*, Act IV, Scene I

Many questions in combinatorics can be answered by analyzing the number of ways to arrange a particular collection of objects into a number of bins, without regard to the order of placement. There are four basic kinds of problems of this form: The objects may be identical or distinguishable, and similarly for the bins. Problems of this form in combinatorics are called *occupancy problems*.

We have already studied occupancy problems for the case of distinguishable bins. If the objects are identical, then we saw in Section 2.6.2 that the number of ways to distribute $n$ objects among $k$ bins is the binomial coefficient $\binom{n+k-1}{n}$. This is the same as the number of ways to select $n$ objects from a set of $k$ different objects with repetition allowed, and we described the correspondence between these two problems in the earlier section. On the other hand, if the objects are distinguishable, then the number of ways to distribute $n$ objects among $k$ bins is simply $k^n$ by the product rule, since each object can be placed in any of the bins. For example, consider the problem of determining the number of $n$-letter words that can be formed using an $k$-letter alphabet. We can model this as an occupancy problem by taking the integers between 1 and $n$ as our objects, and the $k$ letters of the alphabet as our bins. Each placement of the objects in the bins corresponds to an $n$-letter word: The placement of 1 indicates the first letter, 2 the second letter, etc. Furthermore, it is clear that every possible $n$-letter word can be obtained in this way.

In subsequent sections, we consider some occupancy problems where the bins are indistinguishable. We call the bins *groups* or *piles* in this case, with the understanding that they are always unlabeled. The remaining two basic types of occupancy problems each produce important sequences of numbers in combinatorics. The problem of arranging a number *identical* objects into piles gives rise to *partitions*, which are studied in Section 2.8.1. The case of distributing a collection of *distinguishable* objects into groups produces the Stirling set numbers, discussed in Section 2.8.3, and the Bell numbers of Section 2.8.4. We also study two other important combinatorial sequences here: the Stirling cycle numbers in Section 2.8.2, and the Eulerian numbers in Section 2.8.5. Both of these are connected to the structure of permutations.

We study some important properties of each of these classes numbers, aided by generating functions. We also introduce some different kinds of generating functions to assist with our derivations. Some analysis illuminates for instance some interesting connections between ordinary powers, rising and falling factorial powers, and binomial coefficients.

### 2.8.1 Partitions

*Whew! Don't try to eat these so-called chips!*

— Homer Simpson, after choking during a poker game,
*The Simpsons*, episode 103, *Secrets of a Successful Marriage*

Suppose a winning hand in poker nets you a pot of $n$ identical poker chips, and you want to organize your winnings into a number of neat stacks, in order to intimidate your opponents. Individual stacks are not labeled or distinguishable in any way, except for the number of chips they contain, so an arrangement of chips simply corresponds to a collection of positive numbers that sums to $n$. How many ways are there to organize your winnings?

An arrangement of $n$ identical objects into a number of (unlabeled) piles is called a *partition* of the objects, so we want to know the number of partitions of the $n$ objects, or, for short, the number of partitions of $n$. Let $p_n$ denote this number. We might also investigate the number of ways to divide $n$ identical objects into a specific number $k$ of piles. Let $p_{n,k}$ denote this number. Since the piles are unlabeled, we can discount the possibility of an empty pile, so it follows that $p_n = p_{n,1} + p_{n,2} + \cdots + p_{n,n}$ for $n \geq 1$. For example, Figure 2.13 exhibits the fifteen ways to divide $n = 7$ poker chips into stacks. Thus $p_7 = 15$, and we see for instance that $p_{7,3} = 4$ and $p_{7,4} = 3$. Each configuration here is also displayed with a list showing the size of the stacks in descending order. We will always denote partitions in this way. It follows that we can define $p_n$ as the number of ways to write $n$ as a sum of positive integers, with the summands listed in descending

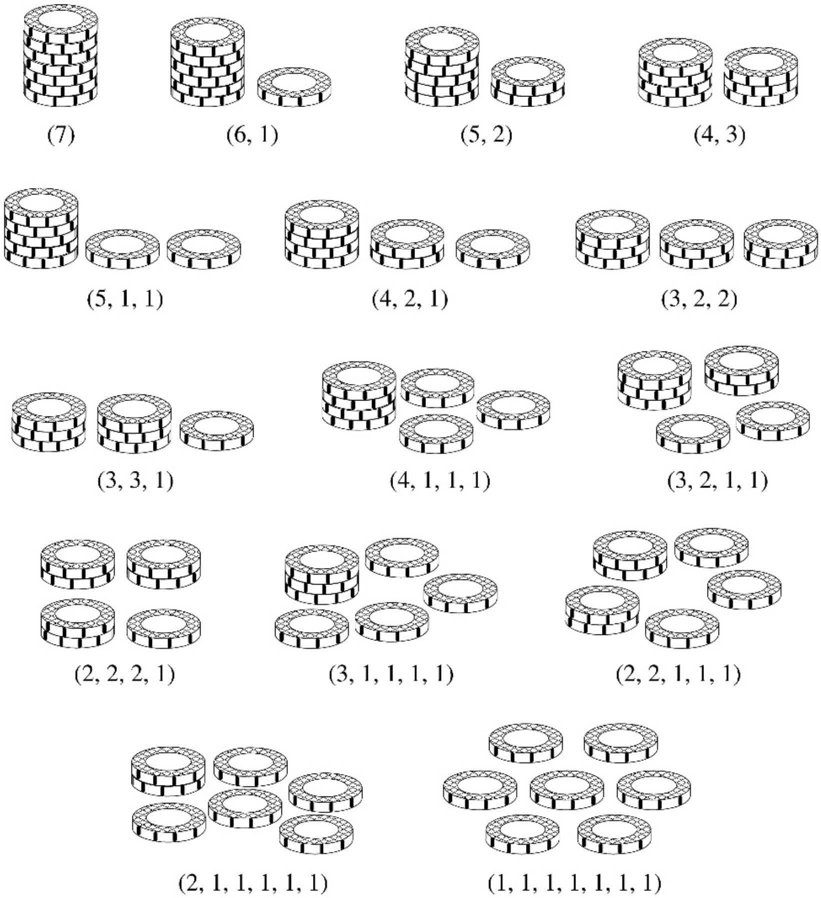



FIGURE 2.13. The fifteen ways to stack seven poker chips.

order. For example, the partitions of $n = 4$ are $(4)$, $(3, 1)$, $(2, 2)$, $(2, 1, 1)$, and $(1, 1, 1, 1)$.

We first note some particular values for the $p_{n,k}$. As a special case, we set

$$p_{0,k} = \begin{cases} 1 & \text{if } k = 0, \\ 0 & \text{if } k \neq 0, \end{cases} \tag{2.70}$$

so $p_0 = 1$. Also, we set $p_{n,k} = 0$ for all $k$ if $n < 0$, so $p_n = 0$ if $n$ is negative. For $n \geq 1$, certainly there is just one way to write $n$ using a single summand, and just one way using $n$ summands, so $p_{n,1} = p_{n,n} = 1$ for $n \geq 1$. Further, it is impossible to express a positive integer as a sum with zero terms, or more than $n$ terms, or a negative number of terms, so we set

$$p_{n,k} = 0, \quad \text{if } k \leq 0 \text{ or } k > n. \tag{2.71}$$

Thus, for each integer $n$ we have the identity

$$p_n = \sum_k p_{n,k}. \tag{2.72}$$

We can now derive a recurrence relation for $p_{n,k}$. Suppose that $(a_1, \ldots, a_k)$ is a partition of $n$, with the summands in descending order. If $a_k = 1$, then $(a_1, \ldots, a_{k-1})$ is a partition of $n-1$, and every partition of $n-1$ into $k-1$ parts can be obtained in this way. Thus, the number of partitions of $n$ into $k$ parts, where the smallest part is 1, is precisely $p_{n-1,k-1}$. Suppose then that $a_k \geq 2$. In this case, we see that $(a_1 - 1, \ldots, a_k - 1)$ is a partition of $n-k$ into exactly $k$ parts, and every partition of $n-k$ can be obtained in this way. It follows that the number of partitions of $n$ into $k$ parts, where the smallest part is at least 2, is $p_{n-k,k}$. Therefore, we find that

$$p_{n,k} = p_{n-1,k-1} + p_{n-k,k} \tag{2.73}$$

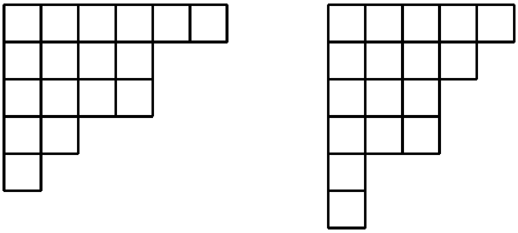
for $n \geq 1$. This recurrence relation, together with the initial condition $p_{0,0} = 1$, allows us to compute the value of $p_{n,k}$, for any $n$ and $k$. A table of these values for $n \leq 10$ appears in Table 2.3.

| $p_{n,k}$ | $k=0$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | $p_n$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n=0$ | 1 | | | | | | | | | | | 1 |
| 1 | 0 | 1 | | | | | | | | | | 1 |
| 2 | 0 | 1 | 1 | | | | | | | | | 2 |
| 3 | 0 | 1 | 1 | 1 | | | | | | | | 3 |
| 4 | 0 | 1 | 2 | 1 | 1 | | | | | | | 5 |
| 5 | 0 | 1 | 2 | 2 | 1 | 1 | | | | | | 7 |
| 6 | 0 | 1 | 3 | 3 | 2 | 1 | 1 | | | | | 11 |
| 7 | 0 | 1 | 3 | 4 | 3 | 2 | 1 | 1 | | | | 15 |
| 8 | 0 | 1 | 4 | 5 | 5 | 3 | 2 | 1 | 1 | | | 22 |
| 9 | 0 | 1 | 4 | 7 | 6 | 5 | 3 | 2 | 1 | 1 | | 30 |
| 10 | 0 | 1 | 5 | 8 | 9 | 7 | 5 | 3 | 2 | 1 | 1 | 42 |

TABLE 2.3. Number of partitions $p_{n,k}$ of $n$ into $k$ parts, and the number of partitions $p_n$ of $n$.

We would like to determine a more efficient way of computing $p_n$, without using (2.73) to determine all of the $p_{n,k}$. In order to do this, we first introduce a useful way to visualize a partition known as a *Young diagram*. The Young diagram of a partition $(a_1, \ldots, a_k)$ of $n$ consists of $n$ boxes arranged in $k$ rows, with $a_1$ boxes in the top row, $a_2$ boxes in the second row, and so on, and each row is aligned on the left. For example, Figure 2.14(a) illustrates the Young diagram for the partition $(6, 4, 4, 2, 1)$ of $n = 17$. This is then much like our stacks of poker chips of Figure 2.13, only turned sideways.

Many texts use arrays of dots instead of arrays of boxes for illustrating partitions, and in this case the diagrams are known as *Ferrers diagrams*. We find the Young diagrams more convenient to use. (Young diagrams earned a distinct name due to their use in visualizing more complicated structures known as *Young tableaux*, where the boxes are filled with integers according to particular rules.)



(a) $\lambda = (6, 4, 4, 2, 1)$. (b) $\lambda' = (5, 4, 3, 3, 1, 1)$.

FIGURE 2.14. The Young diagram for a partition $\lambda$, and its conjugate $\lambda'$.

We now define the *conjugate* $\lambda'$ of a given partition $\lambda$ of $n$ as the partition of $n$ obtained by counting the stacks of boxes in the columns of the Young diagram for $\lambda$. For example, that the conjugate partition of $\lambda = (6, 4, 4, 2, 1)$ in Figure 2.14(a) is $\lambda' = (5, 4, 3, 3, 1, 1)$. The diagram for $\lambda'$ is displayed in Figure 2.14(b). Also, the conjugate of the partition of $n$ that consists of all 1s is the trivial partition $(n)$.

Clearly, different partitions cannot have the same conjugate, and every partition is the conjugate of some partition, so the conjugation mapping is a permutation on the set of partitions of $n$. This fact is very useful in establishing properties of the numbers $p_{n,k}$ and $p_n$. For example, it is immediate that the number of partitions of $n$ which have largest summand $a_1 = k$ is simply $p_{n,k}$, since conjugating the partitions with this property yields precisely the set of partitions of $n$ into exactly $k$ parts.

Next, we consider some generating functions. From our work on the money-changing problems of Section 2.6.3, we know that the generating function $P(x)$ for the sequence $p_n$ is given by an infinite product,

$$P(x) = \prod_{k\geq 1} \frac{1}{1-x^k}. \tag{2.74}$$

Let $\Phi(x) = 1/P(x)$, so

$$\Phi(x) = \prod_{k\geq 1} (1-x^k). \tag{2.75}$$

Then $\Phi(x)$ is itself the generating function for some sequence $\{c_n\}$. If we imagine expanding enough terms of this product to determine $c_n$, we see that each partition $(a_1, \ldots, a_k)$ of $n$ into *distinct* parts $a_1 > \cdots > a_k$ contributes $(-1)^k$ to $c_n$, and these terms determine $c_n$. Define $q_e(n)$ to be the number of partitions of $n$ into an even number of distinct parts, and let $q_o(n)$ be the number of partitions of $n$ into an odd number of distinct parts. It follows that $c_n = q_e(n) - q_o(n)$, and so

$$\Phi(x) = \sum_{n\geq 0} (q_e(n) - q_o(n))x^n, \tag{2.76}$$

with the understanding that $q_e(0) = 1$ and $q_o(0) = 0$.

By expanding a number of terms of the product for $\Phi(x)$, we can compute the values of these coefficients up to $n = 100$:

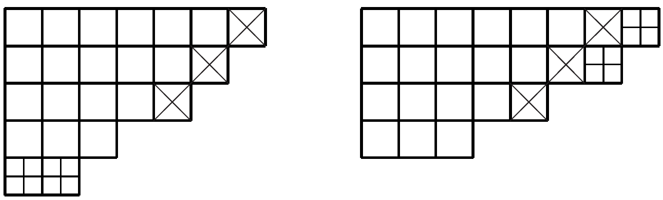
$$\Phi(x) = 1 - x - x^2 + x^5 + x^7 - x^{12} - x^{15} + x^{22} + x^{26}$$
$$- x^{35} - x^{40} + x^{51} + x^{57} - x^{70} - x^{77} + x^{92} + x^{100} - \cdots$$

Thus, it appears that $q_e(n)$ and $q_o(n)$ are often equal, and moreover differ by at most 1. Euler first established this fact; the proof we exhibit here employs Young diagrams and is due to Franklin in 1881 [111]. The reason for the curious name of this theorem is explored in Exercise 7.

**Theorem 2.14** (Euler's Pentagonal Number Theorem)**.** *Let $n$ be a nonnegative integer, and let $q_e(n)$ and $q_o(n)$ be defined as above. Then*

$$q_e(n) - q_o(n) = \begin{cases} (-1)^k & \text{if } n = \frac{k(3k\pm1)}{2}, \\ 0 & \text{otherwise.} \end{cases}$$

*Proof.* Let $\lambda$ be a partition of $n$ into distinct parts. Let $s(\lambda)$ denote the smallest part of $\lambda$, and let $r(\lambda)$ be the number of consecutive integers in $\lambda$, starting with its largest part. In the Young diagram for $\lambda$, the number of squares on the bottom row is $s(\lambda)$, and $r(\lambda)$ is the number of boxes in the diagram that lie on a 45° line anchored at the rightmost box. For example, Figure 2.15(a) exhibits a partition of $n = 23$ into five distinct parts. Here $r(\lambda) = 3$ and $s(\lambda) = 2$, and the relevant boxes for these quantities are marked respectively with x's and +'s.

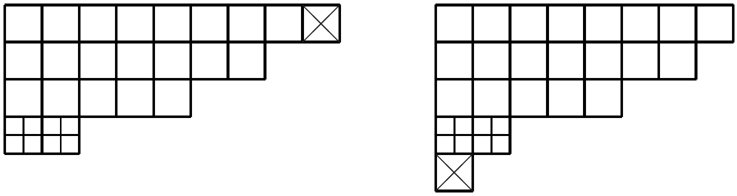


(a) $\lambda = (7, 6, 5, 3, 2)$. (b) $\mu = (8, 7, 5, 3)$.

FIGURE 2.15. Constructing $\mu$ when $s(\lambda) \le r(\lambda)$.

We aim to transform $\lambda$ into another partition $\mu$ of $n$ with distinct parts. The number of parts of $\mu$ will be either one more or one less than the number of parts of $\lambda$, so one of these two partitions will have an even number of parts, and the other will have an odd number. The transformation is described in terms of the Young diagram for $\lambda$, and depends on the relative sizes of $r(\lambda)$ and $s(\lambda)$.

If $s(\lambda) \le r(\lambda)$, then we move the boxes in the bottom row of the Young diagram for $\lambda$ to the ends of the top $s(\lambda)$ rows of the diagram. Figure 2.15(b) shows the resulting partition $\mu$ obtained from the partition $\lambda$ of Figure 2.15(a). On the other hand, if $s(\lambda) > r(\lambda)$, then we move the rightmost boxes of the top $r(\lambda)$ rows of the diagram for $\lambda$ to make a new row at the bottom of the diagram. Figure 2.16 shows this procedure for $\lambda = (9, 7, 5, 2)$, yielding $\mu = (8, 7, 5, 2, 1)$.



(a) $\lambda = (9, 7, 5, 2)$. (b) $\mu = (8, 7, 5, 2, 1)$.

FIGURE 2.16. Constructing $\mu$ when $s(\lambda) > r(\lambda)$.

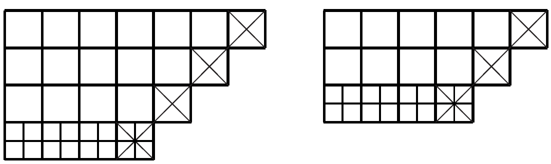
The procedure for creating $\mu$ from $\lambda$ fails in some special cases. The first case breaks down precisely when $s(\lambda) = r(\lambda)$ and the corresponding boxes in the Young diagram overlap, as in Figure 2.17(a). In this case, writing $k$ for $r(\lambda)$, we compute that the total number of boxes in the diagram is

$$n = \sum_{j=k}^{2k-1} j = \frac{k(3k-1)}{2}.$$

The second case fails precisely when $s(\lambda) = r(\lambda) + 1$ and the boxes overlap, as in Figure 2.17(b). Again writing $k$ for $r(\lambda)$, we find that

$$n = \sum_{j=k+1}^{2k} j = \frac{k(3k+1)}{2}$$

in this case.



(a) $\lambda = (7, 6, 5, 4)$. (b) $\lambda = (6, 5, 4)$.

FIGURE 2.17. Exceptional partitions.

Since our mapping on Young diagrams is its own inverse (see Exercise 6), it follows that it defines a bijection between the set of partitions of $n$ into a distinct odd number of parts, and the set of partitions of $n$ into a distinct even number of parts, provided that $n \neq k(3k \pm 1)/2$. When $n$ is one of these exceptional values, there is exactly one additional partition with an even number of parts if $k$

is even, and exactly one extra partition into an odd number of parts if $k$ is odd. The statement then follows. □

By combining (2.76) with Theorem 2.14, we see that

$$\Phi(x) = 1 + \sum_{k\geq 1}(-1)^k \left(x^{k(3k-1)/2} + x^{k(3k+1)/2}\right), \tag{2.77}$$

and so

$$\begin{aligned} 1 &= P(x)\Phi(x) \\ &= \left(\sum_{k\geq 0} p_k x^k\right)\left(1 + \sum_{k\geq 1}(-1)^k \left(x^{k(3k-1)/2} + x^{k(3k+1)/2}\right)\right). \end{aligned} \tag{2.78}$$

It follows that the coefficient of $x^n$ on the right side of (2.78) is 0 for $n \geq 1$. We therefore immediately obtain the following result.

**Theorem 2.15.** *Let $n$ be a positive integer. Then*

$$p_n + \sum_{k\geq 1}(-1)^k \left(p_{n-k(3k-1)/2} + p_{n-k(3k+1)/2}\right) = 0,$$

*that is,*

$$p_n = p_{n-1} + p_{n-2} - p_{n-5} - p_{n-7} + p_{n-12} + p_{n-15} - \cdots. \tag{2.79}$$

We can now use (2.79) to determine values of $p_n$ without using the recurrence (2.73) for the $p_{n,k}$. For example, using the values of $p_n$ computed in Table 2.3, we compute

$$p_{11} = p_{10} + p_9 - p_6 - p_4 = 42 + 30 - 11 - 5 = 56,$$

then

$$p_{12} = p_{11} + p_{10} - p_7 - p_5 + p_0 = 56 + 42 - 15 - 7 + 1 = 77,$$

and so on. Table 2.4 displays the values of $p_n$ computed in this way up to $n = 50$, where

$$p_{50} = p_{49} + p_{48} - p_{45} - p_{43} + p_{38} + p_{35} - p_{28} - p_{24} + p_{15} + p_{10} = 204\,226.$$

We close this section with another interesting fact about the partition sequence. In 1918, Hardy and Ramanujan [152] established a remarkable nonrecursive formula for $p_n$ as the value of a certain convergent series. Their formula was refined by Rademacher in 1937 [230]. We do not reproduce this formula here, but we mention only that it involves the number $\pi$, a certain complex root of the polynomial $x^{24} - 1$, and the hyperbolic sine function. From this formula, however, one

| $n$ | $p_n$ | $n$ | $p_n$ | $n$ | $p_n$ | $n$ | $p_n$ |
|---|---|---|---|---|---|---|---|
| 11 | 56 | 21 | 792 | 31 | 6842 | 41 | 44583 |
| 12 | 77 | 22 | 1002 | 32 | 8349 | 42 | 53174 |
| 13 | 101 | 23 | 1255 | 33 | 10143 | 43 | 63261 |
| 14 | 135 | 24 | 1575 | 34 | 12310 | 44 | 75175 |
| 15 | 176 | 25 | 1958 | 35 | 14883 | 45 | 89134 |
| 16 | 231 | 26 | 2436 | 36 | 17977 | 46 | 105558 |
| 17 | 297 | 27 | 3010 | 37 | 21637 | 47 | 124754 |
| 18 | 385 | 28 | 3718 | 38 | 26015 | 48 | 147273 |
| 19 | 490 | 39 | 4565 | 39 | 31185 | 49 | 173525 |
| 20 | 627 | 30 | 5604 | 40 | 37338 | 50 | 204226 |

TABLE 2.4. The number of partitions of $n$.

can obtain information on the rate of growth of the sequence $p_n$. Asymptotically, the number of partitions of $n$ satisfies

$$p_n \sim \frac{e^{\pi\sqrt{2n/3}}}{4n\sqrt{3}}, \tag{2.80}$$

where $a_n \sim b_n$ means that $\lim_{n\to\infty} a_n/b_n = 1$. See the book by Andrews [9] for the details and a proof, as well as much more information on this rich topic.

**Exercises**

1. Establish formulas for $p_{n,2}$, $p_{n,n-1}$, and $p_{n,n-2}$.

2. Use (2.73) and Table 2.3 to compute the values of $p_{11,k}$, $p_{12,k}$, and $p_{13,k}$ for each $k$.

3. Use (2.79) and Table 2.4 to compute the value of $p_{51}$, then $p_{52}$.

4. Use Young diagrams to prove that $q_0(n)$ equals the number of partitions $\lambda$ of $n$ which are invariant under conjugation, that is, for which $\lambda = \lambda'$.

5. Use generating functions to prove that the number of partitions of $n$ into distinct parts equals the number of partitions of $n$ where each part is odd.

6. Suppose that $\lambda$ is a partition of $n$, and that $\lambda$ is not an exceptional partition like those shown in Figure 2.17. Let $\mu$ be the partition obtained by applying the procedure described in the proof of Theorem 2.14 on $\lambda$. Show that $r(\mu) < s(\mu)$ if and only if $r(\lambda) \ge s(\lambda)$. Then conclude that this map defines a permutation on the set of non-exceptional partitions of $n$ into distinct parts, and that this permutation is its own inverse.

7. (a) Show that (2.77) may be written more simply as

$$\Phi(x) = \sum_k (-1)^k x^{k(3k-1)/2}.$$

(b) The $k$th *pentagonal number* $\alpha_k$ is the number of disks in a pentagonal shape formed by stacking a triangular arrangement of $1 + 2 + \cdots + (k-1)$ disks atop a square arrangement of $k \times k$ disks, as shown in Figure 2.18. Determine a closed fomula for the $k$th pentagonal number. Why is Theorem 2.14 called the Pentagonal Number Theorem?

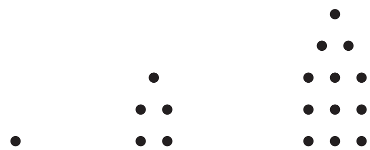


FIGURE 2.18. Pentagonal numbers: $\alpha_1 = 1$, $\alpha_2 = 5$, and $\alpha_3 = 12$.

8. Let $s_{n,k}$ denote the number of partitions of $n$ whose smallest element is $k$, so $p_n = s_{n,1} + s_{n,2} + \cdots + s_{n,n}$. Prove that

$$s_{n,k} = \begin{cases} p_{n-1} & \text{if } k = 1, \\ s_{n-1,k-1} - s_{n-k,k-1} & \text{if } k \geq 2. \end{cases}$$

Then use this recurrence, together with the base values $s_{n,k} = 0$ for $k > n$ and $p_{0,0} = 0$, to produce a table of values for the $s_{n,k}$ for $1 \leq n \leq 10$, similar to Table 2.3.

9. Prove that $p_n \leq p_{n-1} + p_{n-2}$ for $n \geq 1$ by considering first the number of partitions of $n$ that have at least two parts equal to 1, then the other partitions. Then use this to establish that $p_n \leq F_{n+1}$ for $n \geq 0$, where $F_k$ denotes the $k$th Fibonacci number.

10. A *composition* of $n$ is a list of positive integers $\langle a_1, a_2, \ldots, a_k \rangle$ whose sum is $n$, where the order of the integers matters. For example, there are four different compositions of $n = 3$: $\langle 3 \rangle$, $\langle 2, 1 \rangle$, $\langle 1, 2 \rangle$, and $\langle 1, 1, 1 \rangle$. Let $c_n$ denote the number of compositions of $n$, and let $c_{n,k}$ denote the number of compositions of $n$ into exactly $k$ parts.

(a) Compute the value of $c_{n,k}$ for each $k$ and $n$ with $1 \leq k \leq n$ and $1 \leq n \leq 5$ by listing all the compositions, and then calculate the value of $c_n$ for $1 \leq n \leq 5$.

(b) Using these examples, conjecture formulas for $c_{n,k}$ and $c_n$, for arbitrary positive integers $n$ and $k$. Then prove that your formulas are correct.

### 2.8.2 Stirling Cycle Numbers

*The Round Table soon heard of the challenge, and of course it was a good deal discussed...*
— Mark Twain, *A Connecticut Yankee in King Arthur's Court*

Suppose King Arthur decides to divide his knights into committees in order to better govern Britain. True to his egalitarian nature, he crafts $k$ identical round tables for this purpose. How many ways are there to seat $n$ knights at these tables, if each table can seat any number of knights, and no table can be empty? Here, we count two seating arrangements as different only if some knight has a different neighbor on his left side (or his right) in each one. Since the tables are identical, the particular table occupied by a group of knights is immaterial. Thus, once a group of knights is assigned to a table, we must account for all the possible seating arrangements there. From Section 2.7.2, we know that there are $(m-1)!$ different ways to seat $m$ people at one round table.

Let us represent the $n$ knights by the integers 1 through $n$, and denote the seating of knights $K_1$, $K_2$, ..., $K_m$ in clockwise order around one table by $(K_1K_2 \ldots K_m)$. Of course, $(K_2K_3 \ldots K_mK_1)$ denotes the same arrangement of knights around the table, so to make our notation unique we demand that the knight represented by the smallest number appear first in the list. An arrangement of knights at the $k$ tables is then uniquely represented by a list of $k$ strings of integers in parentheses, where each integer between 1 and $n$ appears exactly once. For example, with six knights and three tables, we might seat knights 1, 3, and 5 in clockwise order around one table, knights 2 and 6 at another table, and knight 4 alone at the third table. This arrangement is denoted by $(135)(26)(4)$. This is precisely the cycle notation we used to describe a permutation on six objects. We see that each seating arrangement of $n$ knights at $k$ tables corresponds to a unique permutation $\pi \in S_n$ having exactly $k$ cycles, and every such permutation corresponds to a unique seating arrangement.

We define the *Stirling cycle number*, denoted by $\left[{n \atop k}\right]$, to be the number of ways to seat $n$ knights at $k$ identical tables, or, equivalently, the number of permutations $\pi \in S_n$ having exactly $k$ cycles. These numbers are also known as the *signless Stirling numbers of the first kind*. A signed version of these numbers is also often defined by

$$s(n,k) = (-1)^{n-k} \begin{bmatrix} n \\ k \end{bmatrix}, \tag{2.81}$$

but we will employ only the signless numbers here.

We derive a few properties of the Stirling cycle numbers. First, it is impossible to seat $n$ knights at zero tables, unless there are no knights, so

$$\begin{bmatrix} n \\ 0 \end{bmatrix} = \begin{cases} 1 & \text{if } n = 0, \\ 0 & \text{if } n > 0. \end{cases} \tag{2.82}$$

Second, if there is only one table, then

$$\begin{bmatrix} n \\ 1 \end{bmatrix} = (n-1)!, \quad n \geq 1. \tag{2.83}$$

Next, if there are $n$ tables, then each knight must sit at his own table, and if there are $n-1$ tables, then one pair of knights must sit at one table, and the others must each sit alone. Thus

$$\begin{bmatrix} n \\ n \end{bmatrix} = 1, \tag{2.84}$$

and

$$\begin{bmatrix} n \\ n-1 \end{bmatrix} = \binom{n}{2}. \tag{2.85}$$

There are no arrangements possible if there are more tables than knights, or a negative number of tables, so

$$\begin{bmatrix} n \\ k \end{bmatrix} = 0 \quad \text{if } k < 0 \text{ or } k > n. \tag{2.86}$$

Further, because of the correspondence between seating arrangements and permutations, we have

$$\sum_k \begin{bmatrix} n \\ k \end{bmatrix} = n!. \tag{2.87}$$

Consider now the case $n = 4$ and $k = 2$. Suppose one knight, delayed by an armor adjustment, picks his place after the first three knights are already seated. If the first three knights are seated at one table, then the last knight must sit at the second table by himself. The number of arrangements in this case is the number of ways to seat the first three knights at one table, so $\left[{3 \atop 1}\right] = 2$. On the other hand, if two of the first three knights occupy one table, and the third sits at the second table, then the last knight may then either join the single knight, or the table with two knights. There are two possibilities in the latter case, since the fourth knight may sit on the left side of either of the knights already at the table. Thus, there are $3\left[{3 \atop 2}\right] = 9$ possibilities in this case, and we find that $\left[{4 \atop 2}\right] = 3\left[{3 \atop 2}\right] + \left[{3 \atop 1}\right] = 11$. Figure 2.19 shows these eleven arrangements when Tristram joins Bedivere, Lancelot, and Percival at two tables.

This technique generalizes to produce a recurrence relation for these numbers. To seat $n$ knights at $k$ tables, we can first seat $n-1$ knights at $k-1$ tables, then seat the last knight alone at the $k$th table. Alternatively, we can seat the first $n-1$ knights at $k$ tables, then insert the last knight at one of these tables. This knight must sit on the left side of one of the other $n-1$ knights, so there are $n-1$ different places to seat the last knight. Therefore,

$$\begin{bmatrix} n \\ k \end{bmatrix} = (n-1)\begin{bmatrix} n-1 \\ k \end{bmatrix} + \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}, \quad n \geq 1. \tag{2.88}$$

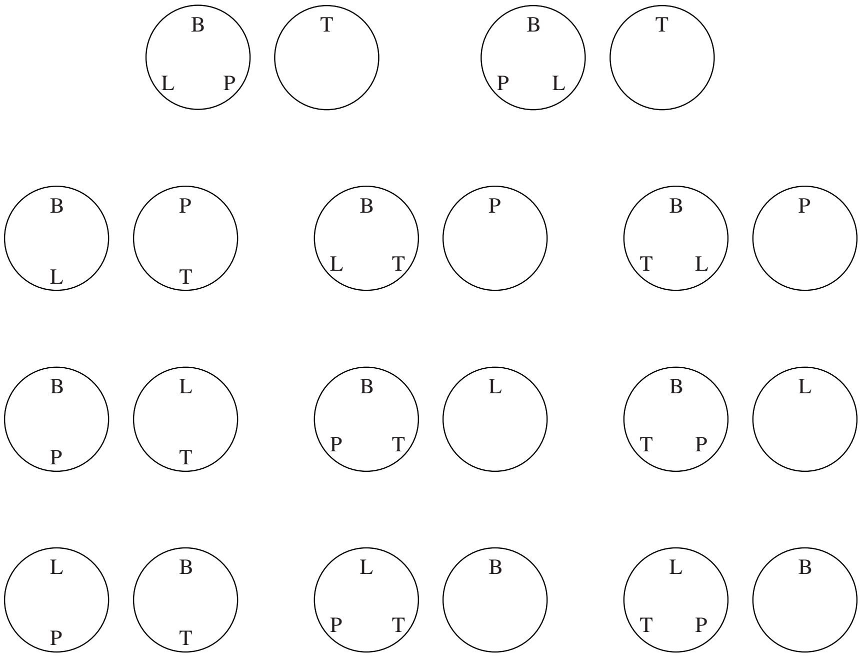



FIGURE 2.19. Seating Bedivere, Lancelot, Percival, and Tristram at two tables.

We can use this formula to compute a triangle of Stirling cycle numbers, just as we used the addition identity for binomial coefficients to obtain Pascal's triangle. These computations appear in Table 2.5.

Recall that for fixed $n$ the generating function for the sequence of binomial coefficients has a particularly nice form: $\sum_k \binom{n}{k} x^k = (x+1)^n$. We can use the identity (2.88) to obtain an analogous representation for the sequence of Stirling cycle numbers. Let $G_n(x) = \sum_k \left[{n \atop k}\right] x^k$. Clearly, $G_0(x) = 1$, and for $n \geq 1$,

$$\begin{aligned} G_n(x) &= \sum_k \left[{n \atop k}\right] x^k \\ &= (n-1)\sum_k \left[{n-1 \atop k}\right] x^k + \sum_k \left[{n-1 \atop k-1}\right] x^k \\ &= (n-1)G_{n-1}(x) + xG_{n-1}(x), \end{aligned}$$

so $G_n(x) = (x+n-1)G_{n-1}(x)$. It is easy to verify by induction that this implies that $G_n(x) = x(x+1)(x+2)\cdots(x+n-1) = x^{\overline{n}}$. Thus,

$$x^{\overline{n}} = \sum_k \left[{n \atop k}\right] x^k. \tag{2.89}$$

| $\left[{n \atop k}\right]$ | $k=0$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | $n!$ |
|---|---|---|---|---|---|---|---|---|---|---|
| $n=0$ | 1 | | | | | | | | | 1 |
| 1 | 0 | 1 | | | | | | | | 1 |
| 2 | 0 | 1 | 1 | | | | | | | 2 |
| 3 | 0 | 2 | 3 | 1 | | | | | | 6 |
| 4 | 0 | 6 | 11 | 6 | 1 | | | | | 24 |
| 5 | 0 | 24 | 50 | 35 | 10 | 1 | | | | 120 |
| 6 | 0 | 120 | 274 | 225 | 85 | 15 | 1 | | | 720 |
| 7 | 0 | 720 | 1764 | 1624 | 735 | 175 | 21 | 1 | | 5040 |
| 8 | 0 | 5040 | 13068 | 13132 | 6769 | 1960 | 322 | 28 | 1 | 40320 |

TABLE 2.5. Stirling cycle numbers, $\left[{n \atop k}\right]$.

for $n \geq 0$. Therefore, the Stirling cycle numbers allow us to express rising factorial powers as linear combinations of ordinary powers. Exercise 7 establishes a similar connection for the falling factorial powers.

**Exercises**

1. Use (2.88) and Table 2.5 to compute the values of $\left[{9 \atop k}\right]$ and $\left[{10 \atop k}\right]$ for each $k$.

2. Prove that

$$\sum_k (-1)^k \begin{bmatrix} n \\ k \end{bmatrix} = \begin{cases} 1 & \text{if } n = 0, \\ -1 & \text{if } n = 1, \\ 0 & \text{if } n \geq 2. \end{cases}$$

3. Use a combinatorial argument to show that

$$\begin{bmatrix} n \\ 2 \end{bmatrix} = \frac{n!}{2} \sum_{m=1}^{n-1} \frac{1}{m(n-m)}.$$

4. Use a combinatorial argument to determine a simple formula for $\left[{n \atop n-2}\right]$.

5. Use a combinatorial argument to show that

$$\begin{bmatrix} n+1 \\ m \end{bmatrix} = \sum_{k=0}^{n} \begin{bmatrix} n-k \\ m-1 \end{bmatrix} n^{\underline{k}}$$

for nonnegative integers $n$ and $m$.

6. Prove that if $n$ and $m$ are nonnegative integers then

$$\begin{bmatrix} n+1 \\ m+1 \end{bmatrix} = \sum_k \begin{bmatrix} n \\ k \end{bmatrix} \binom{k}{m}.$$

7. Prove that if $n \geq 0$ then

$$x^{\underline{n}} = \sum_k (-1)^{n-k} \begin{bmatrix} n \\ k \end{bmatrix} x^k = \sum_k s(n,k) x^k. \tag{2.90}$$

8. Use (2.89) to prove that if $n \geq 0$ then

$$\sum_k \begin{bmatrix} n \\ k \end{bmatrix} y^{n-k} = \prod_{k=1}^{n} (1 + ky).$$

Then use this to prove that $\left[{n \atop k}\right]$ equals the sum of all products of $n - k$ distinct integers selected from $\{1, \ldots, n-1\}$. For example, $\left[{6 \atop 3}\right] = 1 \cdot 2 \cdot 3 + 1{\cdot}2{\cdot}4 + 1{\cdot}2{\cdot}5 + 1{\cdot}3{\cdot}4 + 1{\cdot}3{\cdot}5 + 1{\cdot}4{\cdot}5 + 2{\cdot}3{\cdot}4 + 2{\cdot}3{\cdot}5 + 2{\cdot}4{\cdot}5 + 3{\cdot}4{\cdot}5 = 225$.

9. Let $d(n, k)$ denote the number of ways to place $n$ knights at $k$ identical tables, with at least *two* knights at each table. For example, Figure 2.19 shows that $d(4, 2) = 3$. Set $d(0, 0) = 1$.

   (a) Use a combinatorial argument to show that $d(n, k)$ satisfies the recurrence relation

   $$d(n,k) = (n-1)\big(d(n-1,k) + d(n-2,k-1)\big)$$

   for $n \geq 1$.

   (b) Compute the table of values of $d(n, k)$ for $0 \leq n \leq 8$, similar to Table 2.5.

   (c) Prove that if $n \geq 0$ then

   $$\sum_k d(n,k) = n¡,$$

   where $n¡$ denotes the number of derangements of $n$.

### 2.8.3 Stirling Set Numbers

36 *(Roger Federer, 2006–07),* 35 *(John McEnroe, 1984),*
26 *(Stefan Edberg, 1991–92),* 25 *(Ilie Nastase, 1972–73).*

— Most consecutive sets won in Grand Slam matches in men's tennis

How many ways are there to divide $n$ guests at a party into exactly $k$ groups, if we disregard the arrangement of people within each group? Rephrased, this problem asks for the number of ways to partition a set of $n$ objects into exactly $k$ nonempty subsets, so that each element in the original set appears exactly once among the $k$ subsets. For example, there are three ways to partition the set $\{a, b, c\}$ into two nonempty subsets: $\{a, b\}$, $\{c\}$; $\{a, c\}$, $\{b\}$; and $\{b, c\}$, $\{a\}$. There is just one way to partition $\{a, b, c\}$ into one subset: $\{a, b, c\}$, and just one way to partition $\{a, b, c\}$ into three subsets: $\{a\}$, $\{b\}$, $\{c\}$.

The number of ways to divide $n$ objects into exactly $k$ groups is denoted by $\left\{ n \atop k \right\}$. Thus, $\left\{ 3 \atop 2 \right\} = 3$, and $\left\{ 3 \atop 1 \right\} = \left\{ 3 \atop 3 \right\} = 1$. These numbers are called the *Stirling set numbers*, or the *Stirling numbers of the second kind*. The notation $S(n, k)$ is also often used to denote these numbers.

We begin by listing a few properties of these numbers. First, for $n \geq 1$ we have

$$\left\{ n \atop 1 \right\} = \left\{ n \atop n \right\} = 1, \tag{2.91}$$

since there is only one way to place $n$ people into a single group, and only one way to split them into $n$ groups. Second,

$$\left\{ n \atop 0 \right\} = \begin{cases} 1 & \text{if } n = 0, \\ 0 & \text{if } n > 0, \end{cases} \tag{2.92}$$

since one cannot divide $n$ people into zero groups, unless there are no people. Third, to divide $n$ people into $n - 1$ groups, we must pick two people to be in one group, then place the rest of the people in groups by themselves, so

$$\left\{ n \atop n-1 \right\} = \binom{n}{2}. \tag{2.93}$$

Next, we set

$$\left\{ n \atop k \right\} = 0, \quad \text{if } k < 0 \text{ or } k > n. \tag{2.94}$$

Also, the Stirling cycle number $\left[ n \atop k \right]$ distinguishes among the different ways to arrange $n$ people within $k$ groups, and the Stirling set number $\left\{ n \atop k \right\}$ does not, so

$$\left\{ n \atop k \right\} \leq \left[ n \atop k \right] \tag{2.95}$$

for all $n \geq 0$ and all $k$.

We now derive a recurrence relation for $\left\{ n \atop k \right\}$. Suppose we plan to divide $n \geq 1$ people into $k$ groups for a party, and we know that one person will arrive late. We could divide the first $n - 1$ people into $k - 1$ groups, then place the last person in her own group when she arrives, or we can arrange the first $n - 1$ people into $k$ groups, then pick a group for the last person to join. There are $\left\{ n-1 \atop k-1 \right\}$ different ways to arrange the guests in the first case, and $k\left\{ n-1 \atop k \right\}$ different possibilities in the second. Therefore,

$$\left\{ n \atop k \right\} = k\left\{ n-1 \atop k \right\} + \left\{ n-1 \atop k-1 \right\}, \quad n \geq 1. \tag{2.96}$$

For example, to partition the set $\{a, b, c, d\}$ into two subsets, we can place $d$ in its own set, yielding $\{a, b, c\}$, $\{d\}$, or we can split $\{a, b, c\}$ into two sets, then add $d$

to one of these sets. The latter possibility yields the six different partitions

$$\begin{array}{l}\{a,b,d\},\{c\};\ \{a,c,d\},\{b\};\ \{b,c,d\},\{a\};\\ \{a,b\},\{c,d\};\ \{a,c\},\{b,d\};\ \{b,c\},\{a,d\};\end{array} \tag{2.97}$$

and $\left\{ {4 \atop 2} \right\} = 2\left\{ {3 \atop 2} \right\} + \left\{ {3 \atop 1} \right\} = 7$.

Using identity (2.96), we can generate the triangle of Stirling set numbers shown in Table 2.6. The sequence $\{b_n\}$ that appears in this table as the sum across the rows of the triangle is studied in the next section.

| $\left\{ {n \atop k} \right\}$ | $k=0$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | $b_n$ |
|---|---|---|---|---|---|---|---|---|---|---|
| $n=0$ | 1 | | | | | | | | | 1 |
| 1 | 0 | 1 | | | | | | | | 1 |
| 2 | 0 | 1 | 1 | | | | | | | 2 |
| 3 | 0 | 1 | 3 | 1 | | | | | | 5 |
| 4 | 0 | 1 | 7 | 6 | 1 | | | | | 15 |
| 5 | 0 | 1 | 15 | 25 | 10 | 1 | | | | 52 |
| 6 | 0 | 1 | 31 | 90 | 65 | 15 | 1 | | | 203 |
| 7 | 0 | 1 | 63 | 301 | 350 | 140 | 21 | 1 | | 877 |
| 8 | 0 | 1 | 127 | 966 | 1701 | 1050 | 266 | 28 | 1 | 4140 |

TABLE 2.6. Stirling set numbers, $\left\{ {n \atop k} \right\}$, and Bell numbers, $b_n$.

Exercise 8 analyzes the generating function for the sequence of Stirling set numbers $\left\{ {n \atop k} \right\}$ with $n$ fixed. We can obtain a more useful relation, however, if we replace the ordinary powers of $x$ in this generating function with falling factorial powers. For fixed $n$, let

$$F_n(x) = \sum_k \left\{ {n \atop k} \right\} x^{\underline{k}},$$

so $F_0(x) = 1$. If $n \geq 1$, then

$$\begin{aligned} F_n(x) &= \sum_k \left( k\left\{ {n-1 \atop k} \right\} + \left\{ {n-1 \atop k-1} \right\} \right) x^{\underline{k}} \\ &= \sum_k k\left\{ {n-1 \atop k} \right\} x^{\underline{k}} + \sum_k \left\{ {n-1 \atop k} \right\} x^{\underline{k+1}} \\ &= \sum_k k\left\{ {n-1 \atop k} \right\} x^{\underline{k}} + \sum_k (x-k)\left\{ {n-1 \atop k} \right\} x^{\underline{k}} \\ &= xF_{n-1}(x), \end{aligned}$$

so by induction we obtain

$$x^n = \sum_k \left\{ {n \atop k} \right\} x^{\underline{k}}, \quad n \geq 0. \tag{2.98}$$

Therefore, the Stirling set numbers allow us to express ordinary powers as combinations of falling factorial powers.

We can derive another useful formula by considering the generating function for the numbers $\left\{ {n \atop k} \right\}$ with $k$ fixed. Let

$$H_k(x) = \sum_{n \geq 0} \left\{ {n \atop k} \right\} x^n,$$

so $H_0(x) = 1$. For $k \geq 1$, we obtain

$$\begin{aligned} H_k(x) &= \sum_{n \geq 1} \left\{ {n \atop k} \right\} x^n \\ &= \sum_{n \geq 1} \left( k \left\{ {n-1 \atop k} \right\} + \left\{ {n-1 \atop k-1} \right\} \right) x^n \\ &= kx \sum_{n \geq 0} \left\{ {n \atop k} \right\} x^n + x \sum_{n \geq 0} \left\{ {n \atop k-1} \right\} x^n \\ &= kxH_k(x) + xH_{k-1}(x), \end{aligned}$$

so

$$H_k(x) = \frac{x}{1 - kx} H_{k-1}(x),$$

and therefore

$$H_k(x) = \frac{x^k}{(1-x)(1-2x)\cdots(1-kx)}. \tag{2.99}$$

Next, we use partial fractions to expand this rational function. Our calculations are somewhat simpler if we multiply by $k!$ first, so we wish to find constants $A_1$, $A_2, \ldots, A_k$ such that

$$\frac{k!x^k}{\prod_{m=1}^{k}(1 - mx)} = \sum_{m=1}^{k} \frac{A_m}{1 - mx}.$$

Clearing denominators, we have

$$k!x^k = \sum_{m=1}^{k} A_m \prod_{j=1}^{m-1}(1 - jx) \prod_{j=m+1}^{k}(1 - jx),$$

and setting $x = 1/m$, we obtain

$$\frac{k!}{m^k} = A_m \prod_{j=1}^{m-1}\left(1 - \frac{j}{m}\right) \prod_{j=m+1}^{k}\left(1 - \frac{j}{m}\right),$$

so

$$\begin{aligned} k! &= mA_m \prod_{j=1}^{m-1} (m-j) \prod_{j=m+1}^{k} (m-j) \\ &= mA_m (m-1)!(-1)^{k-m} \prod_{j=m+1}^{k} (j-m) \\ &= (-1)^{k-m} m!(k-m)! A_m, \end{aligned}$$

and

$$A_m = (-1)^{k-m} \binom{k}{m}.$$

Thus

$$\begin{aligned} H_k(x) &= \frac{1}{k!} \sum_{m=1}^{k} (-1)^{k-m} \frac{\binom{k}{m}}{1-mx} \\ &= \frac{1}{k!} \sum_{m=1}^{k} (-1)^{k-m} \binom{k}{m} \sum_{n \ge 0} (mx)^n \\ &= \sum_{n \ge 0} \left( \frac{1}{k!} \sum_{m=1}^{k} (-1)^{k-m} \binom{k}{m} m^n \right) x^n, \end{aligned}$$

and therefore

$$\left\{ {n \atop k} \right\} = \frac{1}{k!} \sum_{m=0}^{k} (-1)^{k-m} \binom{k}{m} m^n, \tag{2.100}$$

for any nonnegative integers $n$ and $k$. This produces a formula for the Stirling set numbers. For example, we may compute $\left\{ {6 \atop 3} \right\} = \frac{1}{3!}(3 \cdot 1^6 - 3 \cdot 2^6 + 1 \cdot 3^6) = 90$.

**Exercises**

1. Use (2.96) and Table 2.6 to compute the values of $\left\{ {9 \atop k} \right\}$ and $\left\{ {10 \atop k} \right\}$ for each $k$.
2. A hungry fraternity brother stops at the drive-through window of a fast-food restaurant and orders twelve different items. The server plans to convey the items using either three or four identical cardboard trays, and empty trays are never given to a customer. Use (2.96) and your augmented table from Exercise 1 to determine the number of ways that the server can arrange the items on the trays.
3. Use combinatorial arguments to determine simple formulas for $\left\{ {n \atop 2} \right\}$ and $\left\{ {n \atop n-2} \right\}$.
4. A new casino game takes ten ping-pong balls, each labeled with a different number between 1 and 10, and drops each one at random into one of three identical buckets. A bucket may be empty after the ten balls are dropped.

(a) Suppose a bet consists of identifying which balls have landed together in each bucket. For example, a bet may state that one bucket is empty, another has just the balls numbered 2, 3, and 7, and the rest are in the other bucket. How many bets are possible?

(b) Suppose instead that a bet consists of identifying only the number of balls that land in the buckets. For example, a bet might state that one bucket is empty, another has three balls, and the other has seven. The numbers on the balls have no role in the bet. How many bets are possible?

5. How many different fifty-character sequences use every character of the 26-letter alphabet at least once? More generally, how many ways can one place $n$ distinguishable objects into $k$ distinguishable bins, if no bin may be empty?

6. Use (2.99) to prove that $\left\{ {n \atop k} \right\}$ equals the sum of all products of $n-k$ integers selected from $\{1, \ldots, k\}$. For example, $\left\{ {6 \atop 3} \right\} = 1\cdot 1\cdot 1 + 1\cdot 1\cdot 2 + 1\cdot 1\cdot 3 + 1\cdot 2\cdot 2 + 1\cdot 2\cdot 3 + 1\cdot 3\cdot 3 + 2\cdot 2\cdot 2 + 2\cdot 2\cdot 3 + 2\cdot 3\cdot 3 + 3\cdot 3\cdot 3 = 90$.

7. Let $r_{n,k}$ denote the number of ways to divide $n$ people into $k$ groups, with at least *two* people in each group. For example, the list (2.97) shows that $r_{4,2} = 3$. Set $r_{0,0} = 1$.

(a) Use a combinatorial argument to show that $r_{n,k}$ satisfies the recurrence relation

$$r_{n,k} = kr_{n-1,k} + (n-1)r_{n-2,k-1}$$

for $n \geq 1$.

(b) Define $r_n$ for $n \geq 0$ by $r_n = \sum_k r_{n,k}$. Compute the table of values of $r_{n,k}$ and $r_n$ for $0 \leq n \leq 8$, similar to Table 2.6.

(c) Determine a formula for $r_{2n,n}$, for a positive integer $n$.

(d) A *rhyming scheme* describes the pattern of rhymes in a poem. For example, the rhyming scheme of a limerick is $(a, a, b, b, a)$, since a limerick has five lines, with the first, second, and last line exhibiting one rhyme, and the third and fourth showing a different rhyme. Also, a sonnet is a poem with fourteen lines. Shakespearean sonnets have the rhyming scheme $(a, b, a, b, c, d, c, d, e, f, e, f, g, g)$; many Petrarchan sonnets exhibit the scheme $(a, b, b, a, a, b, b, a, c, d, e, c, d, e)$. Argue that $r_n$ counts the number of possible rhyming schemes for a poem with $n$ lines, if each line must rhyme with at least one other line.

8. Let $G_n(x) = \sum_k \left\{ {n \atop k} \right\} x^k$, so $G_0(x) = 1$. Show that $G_n(x) = x(G_{n-1}(x) + G'_{n-1}(x))$ for $n \geq 1$, and use this recurrence to compute $G_4(x)$.

9. Show that
$$x^n = \sum_k \left\{ {n \atop k} \right\} (-1)^{n-k} x^{\overline{k}}. \tag{2.101}$$

10. Use (2.90) and (2.98), or (2.89) and (2.101), to prove the following identities.
$$\sum_k \left[ {n \atop k} \right] \left\{ {k \atop m} \right\} (-1)^{(n-k)} = \begin{cases} 1 & \text{if } n = m, \\ 0 & \text{otherwise.} \end{cases} \tag{2.102}$$
$$\sum_k \left\{ {n \atop k} \right\} \left[ {k \atop m} \right] (-1)^{(n-k)} = \begin{cases} 1 & \text{if } n = m, \\ 0 & \text{otherwise.} \end{cases} \tag{2.103}$$

11. Prove that
$$\sum_{k \geq 0} k^n x^k = \sum_k \left\{ {n \atop k} \right\} \frac{k! x^k}{(1-x)^{k+1}}$$
for any nonnegative integer $n$.

12. Suppose $\{r_1, \ldots, r_\ell\}$ and $\{s_1, \ldots, s_\ell\}$ are two sets of positive integers, $f(x) = \sum_{j=1}^{\ell} (x^{r_j} - x^{s_j})$, and $N$ is a positive integer. Prove that
$$\sum_{j=1}^{\ell} r_j^n = \sum_{j=1}^{\ell} s_j^n$$
for every $n$ with $1 \leq n \leq N$ if and only if $f^{(n)}(1) = 0$ for every $n$ with $1 \leq n \leq N$. Here, $f^{(n)}(x)$ denotes the $n$th derivative of $f(x)$.

For example, select $\{1, 5, 9, 17, 18\}$ and $\{2, 3, 11, 15, 19\}$ as the two sets, and select $N = 4$. Then $1 + 5 + 9 + 17 + 18 = 2 + 3 + 11 + 15 + 19 = 50$, $1^2 + 5^2 + 9^2 + 17^2 + 18^2 = 2^2 + 3^2 + 11^2 + 15^2 + 19^2 = 720$, $1^3 + 5^3 + 9^3 + 17^3 + 18^3 = 2^3 + 3^3 + 11^3 + 15^3 + 19^3 = 11600$, and $1^4 + 5^4 + 9^4 + 17^4 + 18^4 = 2^4 + 3^4 + 11^4 + 15^4 + 19^4 = 195684$; and $f(x) = x - x^2 + x^5 - x^3 + x^9 - x^{11} + x^{17} - x^{15} + x^{18} - x^{19}$ has $f^{(n)}(1) = 0$ for $1 \leq n \leq 4$.

### 2.8.4 Bell Numbers

*Silence that dreadful bell: it frights the isle…*

— William Shakespeare, *Othello*, Act II, Scene III

The *Bell number* $b_n$ is the number of ways to divide $n$ people into any number of groups. It is therefore a sum of Stirling set numbers,
$$b_n = \sum_k \left\{ {n \atop k} \right\}. \tag{2.104}$$

The first few values of this sequence appear in Table 2.6.

We can derive a recurrence relation for the Bell numbers. To divide $n$ people into groups, consider the different ways to form a group containing one particular person. We must choose some number $k$ of the other $n-1$ people to join this person in one group, then divide the other $n-1-k$ people into groups. It follows that

$$b_n = \sum_k \binom{n-1}{k} b_{n-1-k}.$$

Reindexing the sum by replacing $k$ with $n-1-k$, then applying the symmetry identity for binomial coefficients, we find the somewhat simpler relation

$$b_n = \sum_k \binom{n-1}{k} b_k, \quad n \geq 1. \tag{2.105}$$

Rather than analyze the ordinary generating function for the sequence of Bell numbers, we introduce another kind of generating function that is often useful in combinatorial analysis. The *exponential generating function* for the sequence $\{a_n\}$ is defined as the ordinary generating function for the sequence $\{a_n/n!\}$. For example, the exponential generating function for the constant sequence $a_n = c$ is $\sum_{n\geq 0} cx^n/n! = ce^x$, and for the sequence $a_n = (-1)^n n!$, it is $1/(1+x)$. The exponential generating function for the sequence of Bell numbers is therefore

$$E(x) = \sum_{n\geq 0} \frac{b_n}{n!} x^n. \tag{2.106}$$

We can compute a closed form for this series. Differentiating, we find

$$\begin{aligned}
E'(x) &= \sum_{n\geq 1} \frac{b_n}{(n-1)!} x^{n-1} \\
&= \sum_{n\geq 1} \frac{1}{(n-1)!} \left( \sum_k \binom{n-1}{k} b_k \right) x^{n-1} \\
&= \sum_{n\geq 1} \sum_{k=0}^{n-1} \frac{b_k}{k!(n-1-k)!} x^{n-1} \\
&= \sum_{k\geq 0} \sum_{n\geq k+1} \frac{b_k}{k!(n-1-k)!} x^{n-1} \\
&= \sum_{k\geq 0} \sum_{n\geq 0} \frac{b_k}{k!n!} x^{n+k} \\
&= \left( \sum_{n\geq 0} \frac{x^n}{n!} \right) \left( \sum_{k\geq 0} \frac{b_k}{k!} x^k \right) \\
&= e^x E(x).
\end{aligned}$$

Therefore,

$$(\ln E(x))' = e^x,$$

and so

$$\ln E(x) = e^x + c$$

for some constant $c$. Since $E(0) = b_0 = 1$, we must have $c = -1$. Thus,

$$E(x) = e^{e^x - 1}. \tag{2.107}$$

We can use this closed form to determine a formula for $b_n$. Using the Maclaurin series for the exponential function twice, we find that

$$\begin{aligned} E(x) &= \frac{1}{e} e^{e^x} \\ &= \frac{1}{e} \sum_{k \geq 0} \frac{(e^x)^k}{k!} \\ &= \frac{1}{e} \sum_{k \geq 0} \frac{1}{k!} \sum_{n \geq 0} \frac{(kx)^n}{n!} \\ &= \frac{1}{e} \sum_{n \geq 0} \left( \sum_{k \geq 0} \frac{k^n}{k!} \right) \frac{x^n}{n!}. \end{aligned}$$

Therefore,

$$b_n = \frac{1}{e} \sum_{k \geq 0} \frac{k^n}{k!}. \tag{2.108}$$

This formula is sometimes called *Dobiński's formula* [79].

**Exercises**

1. How many ways are there to put ten different dogs into pens, if each pen can hold any number of dogs, and every pen is exactly the same?

2. Determine a closed form for the exponential generating function for each of the following sequences.

   (a) $a_k = c^k$, with $c$ a constant.

   (b) $a_k = 1$ if $k$ is even and $0$ if $k$ is odd.

   (c) $a_k = k$.

   (d) $a_k = k^n$, for a fixed nonnegative integer $n$. The number of terms in the answer may depend on $n$.

3. Verify that equation (2.108) for $b_n$ produces the correct value for $b_0$, $b_1$, and $b_2$.

4. Show that the series in equation (2.108) converges for every $n \geq 0$.

5. Use a combinatorial argument to show that

$$\begin{aligned} \left\{ {n \atop m} \right\} &= \sum_k \binom{n-1}{k} \left\{ {n-k-1 \atop m-1} \right\} \\ &= \sum_k \binom{n-1}{k} \left\{ {k \atop m-1} \right\}, \end{aligned} \tag{2.109}$$

for $n \geq 1$, and use this to derive the recurrence (2.105) for Bell numbers.

6. Define the *complementary Bell number* $\widetilde{b}_n$ for $n \geq 0$ by

$$\widetilde{b}_n = \sum_k (-1)^k \left\{ {n \atop k} \right\}.$$

Wilf asked if $\widetilde{b}_n = 0$ for infinitely many $n$, or if there even exists an integer $n > 2$ where $\widetilde{b}_n = 0$. The first few complementary Bell numbers are $1$, $-1$, $0$, $1$, $1$, $-2$, $-9$, $-9$, $50$, $267$, $413$, $-2180$, $-17731$, $-50533$, and $110176$.

(a) Describe a combinatorial interpretation of $\widetilde{b}_n$.

(b) Use (2.109) to determine a recurrence for the complementary Bell numbers. Then determine a closed form for their exponential generating function, $\widetilde{E}(x)$. How is this function related to the function $E(x)$ of this section?

(c) Use $\widetilde{E}(x)$ to determine a formula for $\widetilde{b}_n$, similar to the expression (2.108) for $b_n$.

It is known that the sequence $\widetilde{b}_n$ changes sign infinitely often, and that $\widetilde{b}_n \neq 0$ for almost all values of $n$. See Yang [289] and de Wannemacker, Laffey, and Osburn [71] for more information on this problem.

7. Suppose $P(x)$ is the exponential generating function for the sequence $\{p_n\}$, and $Q(x)$ is the exponential generating function for $\{q_n\}$. Prove that the product $P(x)Q(x)$ is the exponential generating function for the sequence $\{\sum_k \binom{n}{k} p_k q_{n-k}\}$.

8. Let $r_n$ denote the number of rhyming schemes for a poem with $n$ lines, if each line must rhyme with at least one other line, as in Exercise 7d of Section 2.8.3. Recall that $r_0 = 1$.

(a) Prove that

$$r_n = \sum_{k=0}^{n-2} \binom{n-1}{k} r_k.$$

(b) Determine a closed form similar to (2.107) for the exponential generating function $R(x)$ for the sequence $\{r_n\}$.

(c) Use this generating function, together with Exercise 7, to show that

$$r_n = \sum_k \binom{n}{k} (-1)^{n-k} b_k.$$

(d) Prove that the number of rhyming schemes for $n+1$ lines in which each line rhymes with at least one other line equals the number of rhyming schemes for $n$ lines in which at least one line rhymes with no other line. Note that $b_n$ is the total number of rhyming schemes on $n$ lines, including schemes where some lines rhyme with no others.

9. Let $E_k(x)$ denote the exponential generating function for the sequence of Stirling cycle numbers with $k$ fixed,

$$E_k(x) = \sum_{n\geq 0} \begin{bmatrix} n \\ k \end{bmatrix} \frac{x^n}{n!}.$$

Prove that

$$E'_k(x) = \frac{E_{k-1}(x)}{1-x},$$

for $k \geq 1$, and use this to derive a closed form for $E_k(x)$,

$$\sum_{n\geq 0} \begin{bmatrix} n \\ k \end{bmatrix} \frac{x^n}{n!} = \frac{(-1)^k}{k!} \left(\ln(1-x)\right)^k. \tag{2.110}$$

Comtet [60] uses this identity, together with (2.100) and (2.113), to derive a complicated formula due to Schlömilch for the Stirling cycle numbers. We include it here without proof:

$$\begin{bmatrix} n \\ k \end{bmatrix} = \sum_{m=0}^{n-k} (-1)^{n-k-m} \binom{n-1+m}{k-1} \binom{2n-k}{n-k-m} \begin{Bmatrix} n-k+m \\ m \end{Bmatrix} \tag{2.111}$$

$$= \sum_{m=0}^{n-k} \sum_{j=0}^{m} (-1)^{n-k-j} \binom{n-1+m}{k-1} \binom{2n-k}{n-k-m} \binom{m}{j} \frac{j^{n-k+m}}{m!}. \tag{2.112}$$

10. Use an argument similar to that of Exercise 9 to prove that

$$\sum_{n\geq 0} \begin{Bmatrix} n \\ k \end{Bmatrix} \frac{x^n}{n!} = \frac{1}{k!} \left(e^x - 1\right)^k \tag{2.113}$$

for every $k \geq 0$.

### 2.8.5 Eulerian Numbers

> 3 *(Al Hamilton),* 7 *(Paul Coffey),* 11 *(Mark Messier),* 17 *(Jari Kurri),* 31 *(Grant Fuhr),* 99 *(Wayne Gretzky).*
>
> — Retired jersey numbers, Edmonton Oilers

Suppose that a pipe organ having $n$ pipes needs to be installed at a concert hall. Each pipe has a different length, and the pipes must be arranged in a single row. Let us say that two adjacent pipes in an arrangement form an *ascent* if the one on the left is shorter than the one on the right, and a *descent* otherwise. Arranging the pipes from shortest to tallest yields an arrangement with $n-1$ ascents and no descents; arranging them from tallest to shortest results in no ascents and $n-1$ descents.

Whether for aesthetic or acoustical reasons, the eccentric director of the concert hall demands that there be exactly $k$ ascents in the arrangement of the $n$ pipes. How many ways are there to install the organ? The answer is the *Eulerian number* $\left\langle {n \atop k} \right\rangle$. Stated in more abstract terms, $\left\langle {n \atop k} \right\rangle$ is the number of permutations $\pi$ of the integers $\{1, \ldots, n\}$ having $\pi(i) < \pi(i+1)$ for exactly $k$ numbers $i$ between 1 and $n-1$.

We list a few properties of these numbers. It is easy to see that there is only one arrangement of $n$ pipes with no ascents, and only one with $n-1$ ascents, so

$$\left\langle {n \atop 0} \right\rangle = 1, \quad n \geq 0, \tag{2.114}$$

and

$$\left\langle {n \atop n-1} \right\rangle = 1, \quad n \geq 1. \tag{2.115}$$

The Eulerian numbers have a symmetry property similar to that of the binomial coefficients. An arrangement of $n$ pipes with $k$ ascents has $n-1-k$ descents, so reversing this arrangement yields a complementary configuration with $n-1-k$ ascents and $k$ descents. Thus,

$$\left\langle {n \atop k} \right\rangle = \left\langle {n \atop n-1-k} \right\rangle. \tag{2.116}$$

Next, by summing over $k$ we count every possible arrangement of pipes precisely once, so

$$\sum_k \left\langle {n \atop k} \right\rangle = n!. \tag{2.117}$$

We also note the degenerate cases

$$\left\langle {n \atop k} \right\rangle = 0, \quad \text{if } n > 0, \text{ and } k < 0 \text{ or } k \geq n, \tag{2.118}$$

and

$$\left\langle {0 \atop k} \right\rangle = 0, \quad \text{if } k \neq 0. \tag{2.119}$$

We can derive a recurrence relation for the Eulerian numbers. To arrange $n$ pipes with exactly $k$ ascents, suppose we first place every pipe except the tallest into a configuration with exactly $k$ ascents. Then the tallest pipe can be inserted either in the first position, or between two pipes forming any ascent. Any other position would yield an additional ascent. There are therefore $k+1$ different places to insert the tallest pipe in this case. Alternatively, we can line up the $n-1$ shorter pipes so that there are $k-1$ ascents, then insert the last pipe either at the end of the row, or between two pipes forming any descent. There are $n-2-(k-1)=n-k-1$ descents, so there are $n-k$ different places to insert the tallest pipe in this case. It is impossible to create a permissible configuration by inserting the tallest pipe into any other arrangement of the $n-1$ shorter pipes, so

$$\left\langle {n \atop k} \right\rangle = (k+1)\left\langle {n-1 \atop k} \right\rangle + (n-k)\left\langle {n-1 \atop k-1} \right\rangle, \quad n \geq 1. \tag{2.120}$$

For example, $\left\langle {3 \atop 1} \right\rangle = 2\left\langle {2 \atop 1} \right\rangle + 2\left\langle {2 \atop 0} \right\rangle = 4$, and $\left\langle {4 \atop 2} \right\rangle = 3\left\langle {3 \atop 2} \right\rangle + 2\left\langle {3 \atop 1} \right\rangle = 3+8=11$. Figure 2.20 shows these eleven arrangements of four pipes with two ascents.

We can use the recurrence (2.120) to compute the triangle of Eulerian numbers, shown in Table 2.7.

| $\left\langle {n \atop k} \right\rangle$ | $k=0$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | $n!$ |
|---|---|---|---|---|---|---|---|---|---|
| $n=0$ | 1 | | | | | | | | 1 |
| 1 | 1 | | | | | | | | 1 |
| 2 | 1 | 1 | | | | | | | 2 |
| 3 | 1 | 4 | 1 | | | | | | 6 |
| 4 | 1 | 11 | 11 | 1 | | | | | 24 |
| 5 | 1 | 26 | 66 | 26 | 1 | | | | 120 |
| 6 | 1 | 57 | 302 | 302 | 57 | 1 | | | 720 |
| 7 | 1 | 120 | 1191 | 2416 | 1191 | 120 | 1 | | 5040 |
| 8 | 1 | 247 | 4293 | 15619 | 15619 | 4293 | 247 | 1 | 40320 |

TABLE 2.7. Eulerian numbers, $\left\langle {n \atop k} \right\rangle$.

Next, we study some generating functions involving the Eulerian numbers. Recall that in Section 2.6.5 we computed the generating function for the sequence $\{0,1,2,3,\ldots\}$ by differentiating both sides of the identity $\sum_{k\geq 0} x^k = \frac{1}{1-x}$, then multiplying by $x$:

$$\sum_{k\geq 0} kx^k = x \cdot \frac{d}{dx}\left(\sum_{k\geq 0} x^k\right) = x \cdot \frac{d}{dx}\left(\frac{1}{1-x}\right) = \frac{x}{(1-x)^2}. \tag{2.121}$$

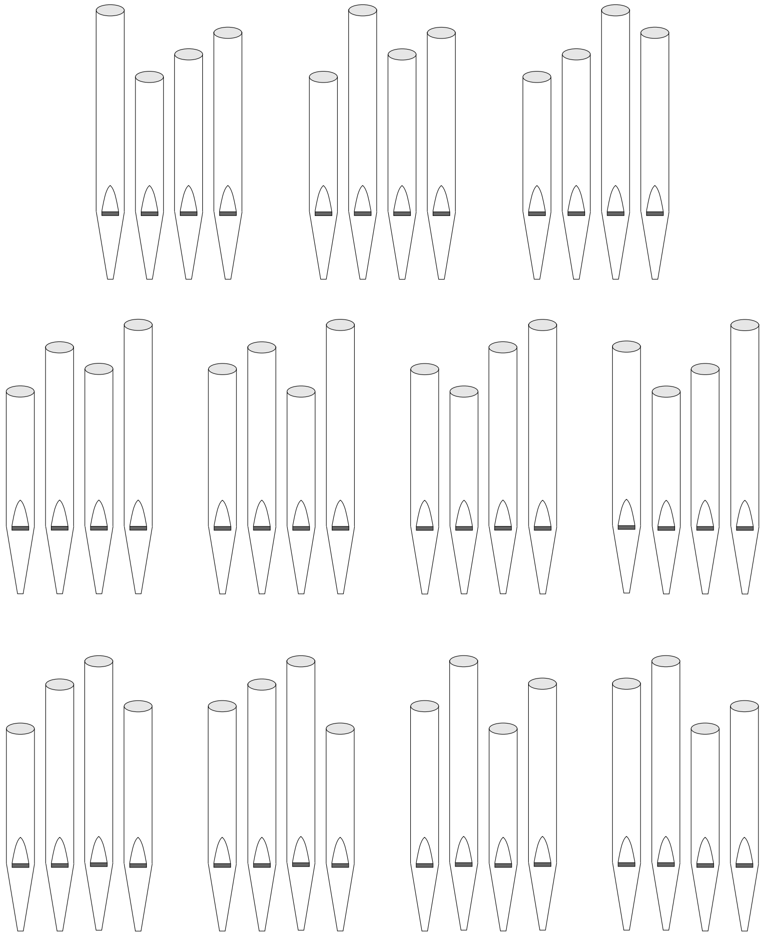


FIGURE 2.20. Four organ pipes with two ascents.

Clearly, we can obtain a generating function for the sequence of squares $\{k^2\}$ by applying the same differentiate-and-multiply operator to (2.121). We find that

$$\begin{aligned}\sum_{k\geq 0} k^2 x^k &= x \cdot \frac{d}{dx}\left(\frac{x}{(1-x)^2}\right)\\ &= x\left(\frac{2x}{(1-x)^3} + \frac{1}{(1-x)^2}\right)\\ &= \frac{x(1+x)}{(1-x)^3}.\end{aligned} \tag{2.122}$$

In the same way, we may use this operator to calculate the generating function for the sequence of cubes, then fourth powers and fifth powers. After a bit of simplifying, we find that

$$\sum_{k\geq 0} k^3 x^k = \frac{x(1+4x+x^2)}{(1-x)^4}, \tag{2.123}$$

$$\sum_{k\geq 0} k^4 x^k = \frac{x(1+11x+11x^2+x^3)}{(1-x)^5}, \tag{2.124}$$

$$\sum_{k\geq 0} k^5 x^k = \frac{x(1+26x+66x^2+26x^3+x^4)}{(1-x)^6}. \tag{2.125}$$

A glance at Table 2.7 shows that the coefficients appearing on the right side of these formulas are all Eulerian numbers, and we would suspect that the numbers $\left\langle {n \atop k} \right\rangle$ will appear in the generating function for the sequence of $n$th powers of integers. This is in fact the case.

**Theorem 2.16.** *If* $n \geq 0$ *then*

$$\sum_{k\geq 1} k^n x^k = \frac{x}{(1-x)^{n+1}} \sum_k \left\langle {n \atop k} \right\rangle x^k. \tag{2.126}$$

*Proof.* We use induction on $n$. The formula is easy to verify when $n = 0$, so we assume it holds for a nonnegative integer $n$. We calculate

$$\begin{aligned}
\sum_{k\geq 1} k^{n+1} x^k &= x \cdot \frac{d}{dx}\left(\sum_{k\geq 1} k^n x^k\right) \\
&= x \cdot \frac{d}{dx}\left(\frac{x}{(1-x)^{n+1}} \sum_k \left\langle {n \atop k} \right\rangle x^k\right) \\
&= x\left(\frac{1}{(1-x)^{n+1}} \sum_k \left\langle {n \atop k} \right\rangle (k+1)x^k + \frac{n+1}{(1-x)^{n+2}} \sum_k \left\langle {n \atop k} \right\rangle x^{k+1}\right) \\
&= \frac{x}{(1-x)^{n+2}}\left((1-x)\sum_k \left\langle {n \atop k} \right\rangle (k+1)x^k + (n+1)\sum_k \left\langle {n \atop k-1} \right\rangle x^k\right) \\
&= \frac{x}{(1-x)^{n+2}}\left(\sum_k (k+1)\left\langle {n \atop k} \right\rangle x^k + \sum_k (n+1-k)\left\langle {n \atop k-1} \right\rangle x^k\right) \\
&= \frac{x}{(1-x)^{n+2}} \sum_k \left\langle {n+1 \atop k} \right\rangle x^k.
\end{aligned}$$

The last step follows from the recurrence relation (2.120). $\square$

We can use (2.126) to obtain a formula for $\left\langle {n \atop k} \right\rangle$ in terms of binomial coefficients and powers. We calculate

$$\begin{aligned}
\sum_k \left\langle {n \atop k} \right\rangle x^k &= \frac{(1-x)^{m+1}}{x} \sum_{m \geq 1} m^n x^m \\
&= \sum_{m \geq 0} (m+1)^n x^m \sum_j \binom{n+1}{j} (-1)^j x^j \\
&= \sum_{m \geq 0} \sum_{j \geq 0} (-1)^j \binom{n+1}{j} (m+1)^n x^{j+m} \\
&= \sum_{k \geq 0} \sum_{j=0}^{k} (-1)^j \binom{n+1}{j} (k+1-j)^n x^k.
\end{aligned} \tag{2.127}$$

Now the first and last expressions in (2.127) are power series in $x$, so we can equate coefficients to obtain a formula for the Eulerian number $\left\langle {n \atop k} \right\rangle$. We find that

$$\left\langle {n \atop k} \right\rangle = \sum_{j=0}^{k} (-1)^j \binom{n+1}{j} (k+1-j)^n. \tag{2.128}$$

Last, we derive one more interesting identity involving Eulerian numbers, binomial coefficients, and ordinary powers. Consider a sort of generating function for the sequence $\{\left\langle {n \atop k} \right\rangle\}$ with $n$ fixed, where we use the binomial coefficient $\binom{x+k}{n}$ in place of $x^k$. Let

$$F_n(x) = \sum_k \left\langle {n \atop k} \right\rangle \binom{x+k}{n},$$

so that $F_0(x) = 1$. For $n \geq 1$, we calculate

$$\begin{aligned}
F_n(x) &= \sum_k \left( (k+1) \left\langle {n-1 \atop k} \right\rangle + (n-k) \left\langle {n-1 \atop k-1} \right\rangle \right) \binom{x+k}{n} \\
&= \sum_k (k+1) \left\langle {n-1 \atop k} \right\rangle \binom{x+k}{n} + \sum_k (n-k) \left\langle {n-1 \atop k-1} \right\rangle \binom{x+k}{n} \\
&= \sum_k (k+1) \left\langle {n-1 \atop k} \right\rangle \binom{x+k}{n} + \sum_k (n-k-1) \left\langle {n-1 \atop k} \right\rangle \binom{x+k+1}{n}.
\end{aligned}$$

Combining the two sums on the right, and replacing the term $\binom{x+k+1}{n}$ by the sum $\binom{x+k}{n} + \binom{x+k}{n-1}$, we find that

$$\begin{aligned}
F_n(x) &= \sum_k \left\langle {n-1 \atop k} \right\rangle \left( n \binom{x+k}{n} + (n-k-1) \binom{x+k}{n-1} \right) \\
&= \sum_k \left\langle {n-1 \atop k} \right\rangle \frac{(x+k)^{\underline{n-1}}}{(n-1)!} \big( (x+k-n+1) + (n-k-1) \big)
\end{aligned}$$

$$= x \sum_k \left\langle {n-1 \atop k} \right\rangle \binom{x+k}{n-1}$$
$$= xF_{n-1}(x).$$

Therefore, $F_n(x) = x^n$, so we obtain

$$x^n = \sum_k \left\langle {n \atop k} \right\rangle \binom{x+k}{n}, \quad n \geq 0. \tag{2.129}$$

This is known as *Worpitzky's identity* [287]. Thus, Eulerian numbers allow us to write ordinary powers as linear combinations of certain generalized binomial coefficients. For example, $x^4 = \binom{x}{4} + 11\binom{x+1}{4} + 11\binom{x+2}{4} + \binom{x+3}{4}$.

**Exercises**

1. Use an ordinary generating function to find a simple formula for $\left\langle {n \atop 1} \right\rangle$, and verify your formula using (2.128).

2. Let $E_n(x)$ denote the polynomial
$$E_n(x) = \sum_k \left\langle {n \atop k} \right\rangle x^k.$$
Use (2.126) to show that the exponential generating function for the sequence of polynomials $\{E_n(x)\}_{n\geq 0}$ is
$$E(x,t) = \frac{1-x}{e^{t(x-1)} - x}.$$
That is, show that
$$E(x,t) = \sum_{n\geq 0} \frac{E_n(x)t^n}{n!}.$$

3. (From [282].) Use (2.126) and Exercise 11 of Section 2.8.3 to prove that
$$\sum_k \left\langle {n \atop k} \right\rangle 2^k = \sum_k \left\{ {n \atop k} \right\} k!$$
for any nonnegative integer $n$.

4. Use (2.128) to establish the following identity for $n \geq 1$:
$$\sum_{j=0}^{n} (-1)^j \binom{n}{j} (j+1)^{n-1} = 0.$$

5. A neurotic running back for an American football team will run between two offensive linemen only if the jersey number of the player on the left is less than the jersey number of the player on the right. The player will not run outside the last player on either end of the offensive line. The coach wants to be sure that the running back has at least three options on every play. If the coach always puts seven players on the offensive line, and there are fifteen players on the team capable of playing any position on the offensive line, each of whom has a different jersey number, how many formations of linemen are possible?

## 2.9 Stable Marriage

*How do I love thee? Let me count the ways.*
— Elizabeth Barrett Browning, Sonnet 43,
*Sonnets from the Portuguese*

Most of the problems we have considered in this chapter are questions in enumerative combinatorics, concerned with counting arrangements of objects subject to various constraints. In this section we consider a very different kind of combinatorial problem.

Suppose we must arrange $n$ marriages between $n$ men and $n$ women. Each man supplies us with a list of the women ranked according to his preference; each woman does the same for the men. Is there always a way to arrange the marriages so that no unmatched man and woman prefer each other to their assigned spouses? Such a pairing is called a *stable matching*.

Consider a simple example with $n = 2$. Suppose Aaron prefers Yvonne over Zoë, and Björn prefers Zoë over Yvonne. We denote these preferences by

$$A : Y > Z,$$
$$B : Z > Y.$$

Suppose also that Yvonne and Zoë both prefer Aaron over Björn, so

$$Y : A > B,$$
$$Z : A > B.$$

Then the matching of Aaron with Zoë and Björn with Yvonne is unstable, since Aaron and Yvonne prefer each other over their partners. The preferences of Björn and Zoë are irrelevant: Indeed, Zoë would prefer to remain with Aaron in this case. On the other hand, the matching of Aaron with Yvonne and Björn with Zoë is stable, for no unmatched pair prefers to be together over their assigned partners.

The stable marriage problem is a question of existential combinatorics, since it asks whether a particular kind of arrangement exists. We might also consider it as a problem in constructive combinatorics, if we ask for an efficient algorithm for finding a stable matching whenever one does exist. In fact, we develop just such an algorithm in Section 2.9.1.

The stable marriage problem and its variations have many applications in problems involving scheduling and assignments. We mention three examples.

1. Stable Roommates.

   Suppose $2n$ students at a university must be paired off and assigned to $n$ dorm rooms. Each student ranks all of the others in order of preference. A pairing is stable if no two unmatched students prefer to room with each other over their assigned partners. Must a stable pairing always exist? This variation of the stable marriage problem, known as *stable roommates*, is considered in Exercise 1.

2. College Admissions.

   Suppose a number of students apply for admission to a number of universities. Each student ranks the universities, and each university ranks the students. Is there a way to assign the students to universities in such a way that no student and university prefer each other over their assignment? This problem is similar to the original stable marriage question, since we are matching elements from two sets using information on preferences. However, there are some significant differences—probably not every student applies to every university, and each university needs to admit a number of students. Some variations on the stable marriage problem that cover extensions like these are considered in Section 2.9.2.

3. Hospitals and Residents.

   The problem of assigning medical students to hospitals for residencies is similar to the problem of matching students and universities: Each medical student ranks hospital residency programs in order of preference, and each hospital ranks the candidates. In this case, however, a program has been used to make most of the assignments in the U.S. since 1952. The National Resident Matching Program was developed by a group of hospitals to try to ensure a fair method of hiring residents. Since medical students are not obligated to accept the position produced by the matching program, it is important that the algorithm produce a stable matching. (Since the program's inception, a large majority of the medical students have accepted their offer.) We describe this matching algorithm in the next section.

## Exercises

1. Suppose that four fraternity brothers, Austin, Bryan, Conroe, and Dallas, need to pair off as roommates. Each of the four brothers ranks the other three brothers in order of preference. Prove that there is a set of rankings for which no stable matching of roommates exists.

2. Suppose $M_1$ and $M_2$ are two stable matchings between $n$ men and $n$ women, and we allow each woman to choose between the man she is paired with in $M_1$ and the partner she receives in $M_2$. Each woman always chooses the man she prefers. Show that the result is a stable matching between the men and the women.

3. Suppose that in the previous problem we assign each woman the man she likes less between her partners in the two matchings $M_1$ and $M_2$. Show that the result is again a stable matching.

4. The following preference lists for four men, $\{A, B, C, D\}$, and four women, $\{W, X, Y, Z\}$, admit exactly ten different stable matchings.

| | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $A$ | $W$ | $X$ | $Y$ | $Z$ |
| $B$ | $X$ | $W$ | $Z$ | $Y$ |
| $C$ | $Y$ | $Z$ | $W$ | $X$ |
| $D$ | $Z$ | $Y$ | $X$ | $W$ |

| | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $W$ | $D$ | $C$ | $B$ | $A$ |
| $X$ | $C$ | $D$ | $A$ | $B$ |
| $Y$ | $B$ | $A$ | $D$ | $C$ |
| $Z$ | $A$ | $B$ | $C$ | $D$ |

(a) Prove that the matching $\{(A, X), (B, Z), (C, W), (D, Y)\}$ is stable.

(b) Determine the remaining nine stable matchings.

### 2.9.1 The Gale–Shapley Algorithm

*Matchmaker, matchmaker, make me a match!*

— Chava and Hodel, *Fiddler on the Roof*

In 1962, Gale and Shapley [117] proved that a stable matching between $n$ men and $n$ women always exists by describing an algorithm for constructing such a matching. Their algorithm is essentially the same as the one used by the hospitals to select residents, although apparently no one realized this for several years [143, chap. 1].

In the algorithm, we first choose either the men or the women to be the *proposers*. Suppose we select the men; the women will have their chance soon. Then the men take turns proposing to the women, and the women weigh the offers that they receive. More precisely, the Gale–Shapley algorithm has three principal steps.

**Algorithm 2.17** (Gale–Shapley)**.** *Construct a stable matching.*

*Input.* A set of $n$ men, a set of $n$ women, a ranked list of the $n$ women for each man, and a ranked list of the $n$ men for each woman.

*Output.* A stable matching that pairs the $n$ men and $n$ women.

*Description.*

*Step 1.* Label every man and woman as free.

*Step 2.* While some man $m$ is free, do the following.

Let $w$ be the highest-ranked woman on the preference list of $m$ to whom $m$ has not yet proposed. If $w$ is free, then label $m$ and $w$ as engaged to each other. If $w$ is engaged to $m'$ and $w$ prefers $m$ over $m'$, then label $m'$ as free and label $m$ and $w$ as engaged to one another. Otherwise, if $w$ prefers $m'$ over $m$, then $w$ remains engaged to $m'$ and $m$ remains free.

*Step 3.* Match all of the engaged couples.

For example, consider the problem of arranging marriages between five men, Mack, Mark, Marv, Milt, and Mort, and five women, Walda, Wanda, Wendy, Wilma, and Winny. The men's and women's preferences are listed in Table 2.8.

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Mack | Winny | Wilma | Wanda | Walda | Wendy |
| Mark | Wanda | Winny | Wendy | Wilma | Walda |
| Marv | Winny | Walda | Wanda | Wilma | Wendy |
| Milt | Winny | Wilma | Wanda | Wendy | Walda |
| Mort | Wanda | Winny | Walda | Wilma | Wendy |
| | | | | | |
| Walda | Milt | Mort | Mack | Mark | Marv |
| Wanda | Milt | Marv | Mort | Mark | Mack |
| Wendy | Mort | Mack | Milt | Mark | Marv |
| Wilma | Mark | Mort | Milt | Mack | Marv |
| Winny | Marv | Mort | Mark | Milt | Mack |

TABLE 2.8. Preferences for five men and women.

First, Mack proposes to Winny, who accepts, and Mark proposes to Wanda, who also accepts. Then Marv proposes to Winny. Winny likes Marv much better than her current fiancé, Mack, so Winny rejects Mack and becomes engaged to Marv. This leaves Mack without a partner, so he proceeds to the second name on his list, Wilma. Wilma currently has no partner, so she accepts. Our engaged couples are now

(Mack, Wilma), (Mark, Wanda), and (Marv, Winny).

Next, Milt proposes to his first choice, Winny. Winny prefers her current partner, Marv, so she rejects Milt. Milt proceeds to his second choice, Wilma. Wilma rejects Mack in favor of Milt, and Mack proposes to his third choice, Wanda. Wanda prefers to remain with Mark, so Mack asks Walda, who accepts. Our engaged couples are now

(Mack, Walda), (Mark, Wanda), (Marv, Winny), and (Milt, Wilma).

Now our last unmatched man, Mort, asks his first choice, Wanda. Wanda accepts Mort over Mark, then Mark asks his second choice, Winny. Winny rejects Mark in favor of her current partner, Marv, so Mark proposes to his third choice, Wendy. Wendy is not engaged, so she accepts. Now all the men and women are engaged, so we have our matching:

(Mack, Walda), (Mark, Wendy), (Marv, Winny),
(Milt, Wilma), and (Mort, Wanda).

We prove that this is in fact a stable matching.

**Theorem 2.18.** *The Gale–Shapley algorithm produces a stable matching.*

*Proof.* First, each man proposes at most $n$ times, so the procedure must terminate after at most $n^2$ proposals. Thus, the procedure is an algorithm. Second, the algorithm always produces a matching. This follows from the observations that a woman, once engaged, is thereafter engaged to exactly one man, and every man ranks every woman, so the last unmatched man must eventually propose to the last unmatched woman. Third, we prove that the matching is stable. Suppose $m$ prefers $w$ to his partner in the matching. Then $m$ proposed to $w$, and was rejected in favor of another suitor. This suitor is ranked higher than $m$ by $w$, so $w$ must prefer her partner in the matching to $m$. Therefore, the matching is stable. □

We remark that the Gale–Shapley algorithm is quite efficient: A stable matching is always found after at most $n^2$ proposals. (Exercise 8 establishes a better upper bound.)

Suppose that we choose the women as the proposers. Does the algorithm produce the same stable matching? We test this by using the lists of preferences in Table 2.8. First, Walda proposes to Milt, who accepts. Next, Wanda proposes to Milt, and Milt prefers Wanda over Walda, so he accepts. Walda must ask her second choice, Mort, who accepts. Then Wendy proposes to Mort, who declines, so she asks Mack, and Mack accepts. Last, Wilma asks Mark, and Winny proposes to Marv, and both accept. We therefore obtain a different stable matching:

(Walda, Mort), (Wanda, Milt), (Wendy, Mack),
(Wilma, Mark), and (Winny, Marv).

Only Winny and Marv are paired together in both matchings; everyone else receives a higher-ranked partner precisely when he or she is among the proposers. Table 2.9 illustrates this for the two different matchings. The pairing obtained with the men as proposers is in boldface; the matching resulting from the women as proposers is underlined.

The next theorem shows that this is no accident. The proposers always obtain the best possible stable matching, and those in the other group, which we call the *proposees*, always receive the worst possible stable matching. We define two terms before stating this theorem. We say a stable matching is *optimal* for a person $p$ if $p$ can do no better in any stable matching. Thus, if $p$ is matched with $q$ in an optimal matching for $p$, and $p$ prefers $r$ over $q$, then there is no stable matching

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Mack | Winny | Wilma | Wanda | **Walda** | <u>Wendy</u> |
| Mark | Wanda | Winny | **Wendy** | <u>Wilma</u> | Walda |
| Marv | <u>**Winny**</u> | Walda | Wanda | Wilma | Wendy |
| Milt | Winny | **Wilma** | <u>Wanda</u> | Wendy | Walda |
| Mort | **Wanda** | Winny | <u>Walda</u> | Wilma | Wendy |
| | | | | | |
| Walda | Milt | <u>Mort</u> | **Mack** | Mark | Marv |
| Wanda | <u>Milt</u> | Marv | **Mort** | Mark | Mack |
| Wendy | Mort | <u>Mack</u> | Milt | **Mark** | Marv |
| Wilma | <u>Mark</u> | Mort | **Milt** | Mack | Marv |
| Winny | <u>**Marv**</u> | Mort | Mark | Milt | Mack |

TABLE 2.9. Two stable matchings.

where $p$ is paired with $r$. Similarly, a stable matching is *pessimal* for $p$ if $p$ can do no worse in any stable matching. So if $p$ is matched with $q$ in a pessimal matching for $p$, and $p$ prefers $q$ over $r$, then there is no stable matching where $p$ is paired with $r$. Finally, a stable matching is optimal for a set of people $P$ if it is optimal for every person $p$ in $P$, and likewise for a pessimal matching.

**Theorem 2.19.** *The stable matching produced by the Gale–Shapley algorithm is independent of the order of proposers, optimal for the proposers, and pessimal for the proposees.*

*Proof.* Suppose the men are the proposers. We first prove that the matching produced by the Gale–Shapley algorithm is optimal for the men, regardless of the order of the proposers. Order the men in an arbitrary manner, and suppose that a man $m$ and woman $w$ are matched by the algorithm. Suppose also that $m$ prefers a woman $w'$ over $w$, denoted by $m : w' > w$, and assume that there exists a stable matching $M$ with $m$ paired with $w'$. Then $m$ was rejected by $w'$ at some time during the execution of the algorithm. We may assume that this was the first time a potentially stable couple was rejected by the algorithm. Say $w'$ rejected $m$ in favor of another man $m'$, so $w' : m' > m$. Then $m'$ has no stable partner he prefers over $w'$, by our assumption. Let $w''$ be the partner of $m'$ in the matching $M$. Then $w'' \neq w'$, since $m$ is matched with $w'$ in $M$, and so $m' : w' > w''$. But then $m'$ and $w'$ prefer each other to their partners in $M$, and this contradicts the stability of $M$.

The optimality of the matching for the proposers is independent of the order of the proposers, so the first statement in the theorem follows immediately.

Finally, we show that the algorithm is pessimal for the proposees. Suppose again that the men are the proposers. Assume that $m$ and $w$ are matched by the algorithm, and that there exists a stable matching $M$ where $w$ is matched with a man $m'$ and $w : m > m'$. Let $w'$ be the partner of $m$ in $M$. Since the Gale–Shapley algorithm produces a matching that is optimal for the men, we have $m :$

$w > w'$. Therefore, $m$ and $w$ prefer each other over their partners in $M$, and this contradicts the stability of $M$. □

**Exercises**

1. Our four fraternity brothers, Austin, Bryan, Conroe, and Dallas, plan to ask four women from the neighboring sorority, Willa, Xena, Yvette, and Zelda, to a dance on Friday night. Each person's preferences are listed in the following table.

| | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Austin | Yvette | Xena | Zelda | Willa |
| Bryan | Willa | Yvette | Xena | Zelda |
| Conroe | Yvette | Xena | Zelda | Willa |
| Dallas | Willa | Zelda | Yvette | Xena |
| | | | | |
| Willa | Austin | Dallas | Conroe | Bryan |
| Xena | Dallas | Bryan | Austin | Conroe |
| Yvette | Dallas | Bryan | Conroe | Austin |
| Zelda | Austin | Dallas | Conroe | Bryan |

   (a) What couples attend the dance, if each man asks the women in his order of preference, and each woman accepts the best offer she receives?

   (b) Suppose the sorority hosts a "Sadie Hawkins" dance the following weekend, where the women ask the men out. Which couples attend this dance?

2. Determine the total number of stable matchings that pair the four men Axel, Buzz, Clay, and Drew with the four women Willow, Xuxa, Yetty, and Zizi, given the following preference lists.

| | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Axel | Yetty | Willow | Zizi | Xuxa |
| Buzz | Yetty | Xuxa | Zizi | Willow |
| Clay | Zizi | Yetty | Xuxa | Willow |
| Drew | Xuxa | Zizi | Willow | Yetty |
| | | | | |
| Willow | Buzz | Drew | Axel | Clay |
| Xuxa | Buzz | Axel | Clay | Drew |
| Yetty | Drew | Clay | Axel | Buzz |
| Zizi | Axel | Drew | Buzz | Clay |

3. Determine a list of preferences for four men and four women where no one obtains his or her first choice, regardless of who proposes.

4. Determine a list of preferences for four men and four women where one proposer receives his or her lowest-ranked choice.

5. Determine a list of preferences for four men and four women where one proposer receives his or her lowest-ranked choice, and the rest of the proposers receive their penultimate choice.

6. Suppose that all the men have identical preference lists in an instance of the stable marriage problem. Show that there exists exactly one stable matching by completing the following argument. Let $M$ be the matching obtained by the Gale-Shapley algorithm using the men as proposers, and suppose another stable matching $M'$ exists. Among all women who change partners between $M$ and $M'$, let $w$ be the woman who ranks lowest on the men's common preference list. Suppose $m$ and $w$ are matched in $M$, and $m$ and $w'$ in $M'$. Determine a contradiction.

7. Suppose that the preference lists of the men $m_1, \ldots, m_n$ and the women $w_1, \ldots, w_n$ have the property that $m_i$ ranks $w_i$ ahead of each of the women $w_{i+1}, \ldots, w_n$, and $w_i$ ranks $m_i$ ahead of each of the men $m_{i+1}, \ldots, m_n$, for each $i$.

   (a) Show that the matching $(m_1, w_1), \ldots, (m_n, w_n)$ is stable.

   (b) (Eeckhout [86].) Show that this is the unique stable matching in this case.

   (c) Prove that there are $(n!)^{n-1}$ different sets of preference lists for $m_1$, $\ldots$, $m_n$ that have the property that $m_i$ ranks $w_i$ ahead of each of the women $w_{i+1}, \ldots, w_n$, for each $i$.

   (d) Prove that at least $1/n!$ of the possible instances of the stable marriage problem for $n$ couples admits a unique solution.

8. (Knuth [178].) Prove that the Gale–Shapley algorithm terminates after at most $n^2 - n + 1$ proposals by showing that at most one proposer receives his or her lowest-ranked choice.

9. Suppose that more than one woman receives her lowest-ranked choice when the men propose. Prove that there exist at least two stable matchings between the men and the women.

### 2.9.2 Variations on Stable Marriage

> *I want what any princess wants—to live happily ever after, with the ogre I married.*
>
> — Princess Fiona, *Shrek 2*

The stable marriage problem solves matching problems of a rather special sort. Each member of one set must rank all the members of the other set, and the two sets must have the same number of elements. In this section, we consider several variations of the stable marriage problem, in order to apply this theory much more broadly. In each case, we study two main questions. First, how does the change affect the existence and structure of the stable pairings? Second, can we amend the Gale-Shapley algorithm to construct a stable matching in the new setting?

### Unacceptable Partners

Suppose each of $n$ men and $n$ women ranks only a subset of their potential mates. Potential partners omitted from a person's list are deemed *unacceptable* to that person, and we do not allow any pairing in which either party is unacceptable to the other. Clearly, we cannot in general guarantee even a complete matching, since for instance a confirmed bachelor could mark all women as unacceptable. This suggests a modification of our notion of a stable matching for this problem. We say a matching (or partial matching) $M$ is *unstable* if there exists a man $m$ and woman $w$ who are unmatched in $M$, each of whom is acceptable to the other, and each is either single in $M$, or prefers the other to their partner in $M$. We will show that every such problem admits a matching that is stable in this sense, and further that every stable matching pairs the same subcollection of men and women. We first require a preliminary observation. We say a person $p$ prefers a matching $M_1$ over a matching $M_2$ if $p$ strictly prefers his or her partner in $M_1$ to $p$'s match in $M_2$.

**Lemma 2.20.** *Suppose $M_1$ and $M_2$ are stable matchings of $n$ men and $n$ women, whose preference lists may include unacceptable partners. If $m$ and $w$ are matched in $M_1$ but not in $M_2$, then one of $m$ or $w$ prefers $M_1$ over $M_2$, and the other prefers $M_2$ over $M_1$.*

*Proof.* Suppose $m_0$ and $w_0$ are paired in $M_1$ but not $M_2$. Then $m_0$ and $w_0$ cannot both prefer $M_1$, since otherwise $M_2$ would not be stable. Suppose that both prefer $M_2$. Then both have partners in $M_2$, so suppose $(m_0, w_1)$ and $(m_1, w_0)$ are in $M_2$. Both $m_0$ and $w_1$ cannot prefer $M_2$, since $M_1$ is stable, so $w_1$ must prefer $M_1$, and likewise $m_1$ must prefer $M_1$. These two cannot be paired in $M_1$, so denote their partners in $M_1$ by $m_2$ and $w_2$. By the same reasoning, both of these people must prefer $M_2$, but cannot be matched together in $M_2$, so we obtain $m_3$ and $w_3$, who prefer $M_1$, but are not paired to each other in $M_1$. We can continue this process indefinitely, obtaining a sequence $m_0, w_0, m_2, w_2, m_4, w_4, \ldots$ of distinct men and women who prefer $M_2$ over $M_1$, and another sequence $m_1, w_1, m_3, w_3, \ldots$ of different people who prefer $M_1$ over $M_2$. This is impossible, since there are only finitely many men and women. □

We can now establish an important property of stable matchings when some unacceptable partners may be included: For a given set of preferences, every stable matching leaves the same group of men and women single.

**Theorem 2.21.** *Suppose each of $n$ women ranks a subset of $n$ men as potential partners, with the remaining men deemed unacceptable, and suppose each of the* *men rank the women in the same way. Then there exists a subset $X_0$ of the women and a subset $Y_0$ of the men such that every stable matching of the $n$ men and $n$ women leaves precisely the members of $X_0$ and $Y_0$ unassigned.*

*Proof.* Suppose $M_1$ and $M_2$ are distinct stable matchings, and suppose $m_1$ is matched in $M_1$ but not in $M_2$. Let $w_1$ be the partner of $m_1$ in $M_1$. Since $m_1$ clearly prefers $M_1$ over $M_2$, by Lemma 2.20 $w_1$ must prefer $M_2$ over $M_1$. Let $m_2$ be the partner of $w_1$ in $M_2$. Then $m_2$ prefers $M_1$, and so his partner $w_2$ in $M_1$ must prefer $M_2$ over $M_1$. Continuing in this way, we obtain an infinite sequence $(m_1, w_1)$, $(m_2, w_2)$, $(m_3, w_3)$, ... of distinct couples in $M_1$ (and another sequence $(m_2, w_1)$, $(m_3, w_2)$, $(m_4, w_3)$, ... in $M_2$), which is impossible. □

We still need to show that at least one stable matching exists, and we can do this by altering the Gale-Shapley algorithm for preference lists that may include unacceptable partners. We require just two modifications. First, we terminate the loop either when all proposers are engaged, or when no free proposer has any remaining acceptable partners to ask. Second, proposals from unacceptable partners are always rejected. It is straightforward to show that this amended procedure always produces a stable matching (see Exercise 1). We can illustrate it with an example. Suppose the four men Iago, Julius, Kent, and Laertes each rank a subset of the four women Silvia, Thaisa, Ursula, and Viola, and each of the women ranks a subset of the men, as shown in Figure 2.21. Potential partners omitted from a person's list are deemed unacceptable to that person, so for example Iago would not consider marrying Thaisa or Ursula.

| | 1 | 2 | 3 |
|---|---|---|---|
| $I$ | $V$ | $S$ | |
| $J$ | $S$ | $V$ | |
| $K$ | $U$ | $T$ | $S$ |
| $L$ | $S$ | $T$ | $V$ |

| | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $S$ | $I$ | $K$ | $L$ | $J$ |
| $T$ | $J$ | $K$ | | |
| $U$ | $L$ | $I$ | $J$ | |
| $V$ | $K$ | $J$ | | |

FIGURE 2.21. Preferences with unacceptable partners.

Suppose the men propose. Iago first asks Viola, but she rejects him as an unacceptable partner, so he asks Silvia, who happily accepts. Next, Julius asks Silvia, who rejects him in favor of Iago, so he proposes to Viola, who now accepts. Ursula then rejects Kent, then Thaisa accepts his proposal. Finally, Laertes proposes to Silvia, then Thaisa, then Viola, but each rejects him. Our stable matching is then (Iago, Silvia), (Julius, Viola), and (Kent, Thaisa). The set $X_0$ of unmatchable bachelorettes contains only Ursula, and $Y_0 = \{\text{Laertes}\}$.

We have shown how to adapt the Gale-Shapley algorithm to handle incomplete preference lists, but we can also describe a way to alter the data in such a way that we can apply the Gale-Shapley algorithm without any modifications. To do this, we introduce a fictitious man to mark the boundary between the acceptable and unacceptable partners on each woman's list, and similarly introduce a fictitious woman for the men's lists. We'll call our invented man the *ogre*, and our fictitious woman, the *ogress*. Append the ogre to each woman's ranked list of acceptable partners, then add her unacceptable partners afterwards in an arbitrary order. Thus, each woman would sooner marry an ogre than one of her unacceptable partners. Do the same for the men with the ogress. The ogre prefers any woman over the ogress, and the ogress prefers any man over the ogre (people are tastier!), but the rankings of the humans on the ogre's and ogress' lists are immaterial. For example, we can augment the preference lists of Figure 2.21 to obtain the $5 \times 5$ system of Figure 2.22, using $\mathbf{M}$ to denote the ogre and $\mathbf{W}$ for the ogress.

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $I$ | $V$ | $S$ | $\mathbf{W}$ | $T$ | $U$ |
| $J$ | $S$ | $V$ | $\mathbf{W}$ | $T$ | $U$ |
| $K$ | $U$ | $T$ | $S$ | $\mathbf{W}$ | $V$ |
| $L$ | $S$ | $T$ | $V$ | $\mathbf{W}$ | $U$ |
| $\mathbf{M}$ | $S$ | $T$ | $U$ | $V$ | $\mathbf{W}$ |

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $S$ | $I$ | $K$ | $L$ | $J$ | $\mathbf{M}$ |
| $T$ | $J$ | $K$ | $\mathbf{M}$ | $I$ | $L$ |
| $U$ | $L$ | $I$ | $J$ | $\mathbf{M}$ | $K$ |
| $V$ | $K$ | $J$ | $\mathbf{M}$ | $I$ | $L$ |
| $\mathbf{W}$ | $I$ | $J$ | $K$ | $L$ | $\mathbf{M}$ |

FIGURE 2.22. Augmented preference lists.

We can now characterize when the original configuration has a complete stable matching, that is, a stable pairing where no one is left single.

**Theorem 2.22.** *Suppose each of $n$ men ranks some subset of $n$ women as acceptable partners, and each of the women does the same for the men. Suppose further that we obtain an instance of the standard stable marriage problem on $n+1$ men and women by adding an ogre $\mathbf{M}$ and ogress $\mathbf{W}$, and augmenting the preference lists in the manner described above. Then the original system has a complete stable matching if and only if the augmented system has a stable matching where $\mathbf{M}$ is paired with $\mathbf{W}$.*

*Proof.* Suppose the original system has a complete stable matching. Then each woman prefers her partner in this matching to the ogre under the augmented preferences, and likewise no man would leave his partner for the ogress. Thus, adding $(\mathbf{M}, \mathbf{W})$ to this pairing produces a stable matching for the augmented system. Next, suppose the augmented system has a stable matching $P'$ that includes $(\mathbf{M}, \mathbf{W})$, and let $P = P' \setminus \{(\mathbf{M}, \mathbf{W})\}$. Suppose $(m, w) \in P$. If $m$ is unacceptable to $w$, then $w$ would prefer the ogre $\mathbf{M}$ over $m$, and certainly $\mathbf{M}$ prefers $w$ over $\mathbf{W}$. This contradicts the stability of $P'$. Similarly, $w$ must be acceptable to $m$. Thus, $P$ is a complete matching of mutually acceptable partners, and stability follows at once from the stability of $P'$. $\square$

Exercise 2 asks you to show that $\mathbf{M}$ and $\mathbf{W}$ must be paired together in all stable matchings of the augmented system, if they are paired in any particular stable matching. Thus, we can determine if a complete stable matching exists by running the original Gale-Shapley algorithm on the augmented preference lists, choosing either set as the proposers.

While applying the Gale-Shapley algorithm in this way always produces a matching that is stable with respect to the augmented preferences, it is important to note that restricting such a pairing back to the original preferences might not produce a stable matching! For example, when the men propose using the augmented lists of Figure 2.22, we obtain the stable matching

$$\begin{gathered}(\text{Iago, Silvia}),\ (\text{Julius, Viola}),\ (\text{Kent, Ursula}),\\ (\text{Laertes, Ogress}),\ (\text{Ogre, Thaisa}).\end{gathered} \tag{2.130}$$

However, Kent is not acceptable to Ursula, so we must disband this pair when we restrict to the original preference lists. The surviving pairs are (Iago, Silvia) and (Julius, Viola), and now Kent and Thaisa are unmatched but mutually acceptable.

### Indifference

In the original stable marriage problem, we required that all preferences be *strictly ordered*, since each person needed to assign each potential partner a different rank. However, rankings often contain items that are valued equally. What happens if we allow *weakly ordered* rankings, that is, rankings that may contain some elements of the same rank? Suppose that each of $n$ men supplies a weak ordering of a set of $n$ women, and each of the women does the same for the men. We'll assume for now that all rankings are complete, so there are no unacceptable partners. Must a stable ranking exist? Can we construct one?

We first require a clarification of our notion of stability for this situation. We say a matching $M$ of the men and women is *unstable* if there exists an unmatched couple $m$ and $w$, each of whom strictly prefers the other to his or her partner in $M$. For example, if $m$ strictly prefers $w$ to his partner, but $w$ ranks $m$ equal to her partner, then the pair $m$ and $w$ do not violate stability under this definition.

One can certainly study this problem with other notions of stability. For instance, one could demand that no unmatched man and woman weakly prefer each other to their assigned partners. A matching with no such couples is called *superstable*. Or one could require that no unmatched couple prefer each other, one in a strict sense and the other in a weak manner. Such a matching is said to be *strongly stable*. Since the notion that we employ is the least restrictive, matchings with this property are often called *weakly stable*.

Given a collection of weakly ordered preference lists for $n$ men and $n$ women, we can certainly create a corresponding set of strongly ordered preference lists by breaking each tie in an arbitrary way. We call the strongly ordered preferences a *refinement* of the original weak preferences. A stable matching for the refined lists certainly exists, and it is easy to see that this matching is also a (weakly) stable matching for the original, weakly ordered lists. Furthermore, every stable matching for the original preferences can be obtained in this way. We can summarize these facts in the following theorem.

**Theorem 2.23.** *Suppose each of $n$ men ranks a collection of $n$ women, with tied rankings allowed, and each woman does the same for the men. Then a stable* *matching for these preferences exists, and further every such stable matching is a stable matching for some refinement of these weakly ordered rankings.*

*Proof.* For the first part, let $P'$ be a refinement of the given list of preferences $P$, and let $M$ be a stable matching for $P'$. If $m$ and $w$ are unmatched in $M$, and according to $P$ strictly prefer each other to their partners in this matching, then they also strictly prefer each other according to $P'$. This is impossible, since $M$ is stable with respect to $P'$. Thus, $M$ is stable with respect to $P$.

For the second part, suppose $M$ is a stable matching with respect to $P$. We need to construct a refinement $P'$ of $P$ where $M$ is stable. If $(m, w) \in M$, and $m$ ranks $w'$ equal to $w$ in $P$, then let $m$ rank $w$ ahead of $w'$ in $P'$. Likewise, if $w$ ranks $m'$ equal to $m$ in $P$, then $w$ ranks $m$ ahead of $m'$ in $P'$. Any remaining tied rankings are broken arbitrarily to complete $P'$. Suppose then that $m_0$ and $w_0$ are unmatched in $P'$, but prefer each other (according to $P'$) to their partners in $M$. Since $M$ is stable with respect to $P$, then either $m_0$ ranks $w_0$ equal to his partner in $M$, or $w_0$ ranks $m_0$ equal to her partner in $M$ (or both). We obtain a contradiction in either case, by the construction of $P'$. $\square$

$$
\begin{array}{rl}
G: & D > A = C > F \\
H: & A = F > C = D \\
I: & F > C > D > A \\
K: & D > A = C = F
\end{array}
\qquad\qquad
\begin{array}{rl}
A: & I > G = H = K \\
C: & H > G = I > K \\
D: & I > K > H > G \\
F: & H = I > G = K
\end{array}
$$

FIGURE 2.23. Preference lists with indifference.

The Gale-Shapley algorithm requires no modifications for this variation, once a refinement is selected. Of course, the algorithm may produce different matchings for different refinements, even when the same group proposes. For example, suppose the four men Gatsby, Hawkeye, Ishmael, and Kino, and four women Apolonia, Cora, Daisy, and Fayaway, submit the preference lists shown in Figure 2.23. Using the refinement obtained by replacing each $=$ in these lists with $>$, the Gale-Shapley algorithm produces the following matching when the men propose:

$$
\begin{gathered}
\text{(Gatsby, Apolonia)}, \ \text{(Hawkeye, Fayaway)}, \\
\text{(Ishmael, Cora)}, \ \text{(Kino, Daisy)}.
\end{gathered}
\tag{2.131}
$$

However, if we reverse the order of Apolonia and Cora in the refinement of Gatsby's list, and the order of Apolonia and Fayaway in Hawkeye's, we then obtain a very different stable matching:

$$
\begin{gathered}
\text{(Gatsby, Cora)}, \ \text{(Hawkeye, Fayaway)}, \\
\text{(Ishmael, Daisy)}, \ \text{(Kino, Apolonia)}.
\end{gathered}
\tag{2.132}
$$

Finally, we may also ask about combining this extension of the stable marriage problem with the prior one. Suppose the men and women supply weakly ordered rankings, and may also declare some potential partners as unacceptable. The stable matching problem becomes much more complicated in this case. Even the size of a stable matching may vary, in contrast to the case of unacceptable partners with strict rankings, where Theorem 2.21 guarantees that all stable matchings have not only the same size, but match exactly the same men and women. For example, consider the following $2 \times 2$ system from [196], where $A$ finds $Y$ acceptable but not $Z$, and $Z$ finds $B$ acceptable but not $A$.

$$\begin{array}{lll} A : Y & \qquad & Y : A = B \\ B : Y > Z & \qquad & Z : B \end{array}$$

These preferences admit exactly two stable matchings, which have different sizes: $\{(A,Y),(B,Z)\}$ and $\{(B,Y)\}$.

We might ask if we could determine a stable matching of maximal size in a problem like this, since this would often be desirable. However, no fast algorithm is known for computing this in the general $n \times n$ case. (Here, a "fast" algorithm would have its running time bounded by a polynomial in $n$.) In fact, it is known [196] that this problem belongs to a family of difficult questions known as NP-complete problems. The problem remains hard even if ties are allowed in only the men's or only the women's preferences, and all ties occur at the end of each list, even if each person is allowed at most one tied ranking.

**Sets of Different Sizes**

Every stable marriage problem we have considered so far required an equal number of men and women. Suppose now that one group is larger than the other. Of course, we could not possibly match everyone with a partner now, but can we find a stable matching that pairs everyone in the smaller set? Here, we say a matching (or partial matching) $M$ is *unstable* if there exists a man $m$ and woman $w$, unmatched in $M$, such that each is either single in $M$, or prefers the other to his or her partner in $M$.

We can solve this variation by considering it to be a special case of the problem with unacceptable partners. Suppose we have $k$ men and $n$ women, with $n > k$. Suppose also that each of the men rank each of the women in strict order, and each of the women reciprocate for the men. We introduce $n-k$ *ghosts* to the set of men. Each ghost finds no woman to be an acceptable partner, and each women would not accept any ghost. Then a stable matching exists by the modified Gale-Shapley algorithm for unacceptable partners, and by Theorem 2.21 there exists a set $X_0$ of women and $Y_0$ of ghosts and men such that the members of $X_0$ and $Y_0$ are precisely the unassigned parties in any stable matching. Certainly $Y_0$ includes all the ghosts, since they have no acceptable partners. But no man can be unassigned in a stable matching, since each man is acceptable to all the women. Thus, $X_0$ is empty and $Y_0$ is precisely the set of ghosts, and we obtain the following theorem.

**Theorem 2.24.** *Suppose each of $k$ men ranks each of $n$ women in a strict ordering, and each of the women ranks the men in the same way. Then*

*(i) a stable matching exists,*

*(ii) every stable matching pairs every member of the smaller set, and*

*(iii) there exists a subset $X$ of the larger set such that every stable matching leaves the members of $X$ unassigned, and the others all matched.*

An example with groups of different sizes appears in Exercise 6. Some other interesting variations (and combinations of variations) on the stable marriage problem are introduced in the exercises too. We will study marriage problems further in Chapter 3, where in Section 3.8 we investigate matchings for various *infinite* sets.

**Exercises**

1. Prove that the Gale-Shapley algorithm, amended to handle unacceptable partners, always produces a stable matching.

2. Prove that if the ogre and ogress are paired in some stable matching for an augmented system of preferences as in Theorem 2.22, then they must be paired in every such stable matching.

3. (a) Verify the stable matching (2.130) produced by the Gale-Shapley algorithm when the men propose using the preferences in Figure 2.22.

   (b) Compute the stable matching obtained when the women propose using these preferences. Does this pairing restrict to a stable matching for Figure 2.21?

   (c) In the augmentation procedure for the case of unacceptable partners, we can list the unacceptable partners for each person in any order after the ogre or ogress, and we can list the humans in any order in the lists for the ogre and ogress. Show that one can select orderings when augmenting the preferences of Figure 2.21 so that when the men propose in the Gale-Shapley algorithm, one obtains a pairing that restricts to a stable matching of Figure 2.21.

4. The following problems all refer to the weakly ordered preference lists of Figure 2.23.

   (a) Verify the matching (2.131) obtained from the refinement obtained by replacing each $=$ with $>$, when the men propose in the Gale-Shapley algorithm. Then determine the matching obtained when the women propose.

   (b) Verify (2.132) using the refinement obtained from the previous one by reversing the order of Apolonia and Cora in Gatsby's list, and Apolonia and Fayaway in Hawkeye's. Then determine the matching obtained when the women propose.

   (c) Construct another refinement by ranking any tied names in reverse alphabetical order. Compute the stable matchings constructed by the

Gale-Shapley algorithm when the men propose, then when the women propose.

5. Construct three refinements of the following preference lists so that the Gale-Shapley algorithm, amended for unacceptable partners, produces a stable matching of a different size in each case.

$$\begin{array}{ll} A : W & W : A = B = C = D \\ B : W > X & X : B = C = D \\ C : W > X > Y & Y : C = D \\ D : W > X > Y > Z & Z : D \end{array}$$

6. Suppose the five men Arceneaux, Boudreaux, Comeaux, Duriaux, and Gautreaux, each rank the three women Marteaux, Robichaux, and Thibodeaux in order of preference, and the women each rank the men, as shown in the following tables.

| | 1 | 2 | 3 |
|---|---|---|---|
| $A$ | $R$ | $T$ | $M$ |
| $B$ | $T$ | $R$ | $M$ |
| $C$ | $M$ | $T$ | $R$ |
| $D$ | $T$ | $M$ | $R$ |
| $G$ | $R$ | $T$ | $M$ |

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $M$ | $A$ | $D$ | $B$ | $C$ | $G$ |
| $R$ | $D$ | $G$ | $A$ | $C$ | $B$ |
| $T$ | $G$ | $A$ | $D$ | $C$ | $B$ |

Determine the stable matching obtained when the men propose, then the matching found when the women propose. What is the set $X$ of Theorem 2.24 for these preferences?

7. Suppose we allow weakly ordered rankings in the hypothesis of Theorem 2.24. Determine which of the conclusions still hold, and which do not necessarily follow. Supply a proof for any parts that do hold, and supply a counterexample for any parts that do not.

8. Suppose that each of $n$ students, denoted $S_1, S_2, \ldots, S_n$, ranks each of $m$ universities, $U_1, U_2, \ldots, U_m$, and each university does the same for the students. Suppose also that university $U_k$ has $p_k$ open positions. We say an assignment of students to universities is *unstable* if there exists an unpaired student $S_i$ and university $U_j$ such that $S_i$ is either unassigned, or prefers $U_j$ to his assignment, and $U_j$ either has an unfilled position, or prefers $S_i$ to some student in the new class.

   (a) Assume that $\sum_{k=1}^{m} p_k = n$. Explain how to amend the preference lists so that the Gale-Shapley algorithm may be used to compute a stable assignment of students to universities, with no university exceeding its capacity.

(b) Repeat this problem without assuming that the number of students matches the total number of open positions.

(c) Suppose each student ranks only a subset of the universities, and each university ranks only a subset of the students who apply to that school. Assume that unranked possibilities are unacceptable choices. Modify the definition of stability for this case, then describe how to use the Gale-Shapley algorithm to determine a stable assignment.

9. Suppose that each of $n$ students, denoted $S_1, S_2, \ldots, S_n$, needs to enroll in a number of courses from among $m$ possible offerings, denoted $C_1, C_2, \ldots, C_m$. Assume that student $S_i$ can register for up to $q_i$ courses, and course $C_j$ can admit up to $r_j$ students. An *enrollment* is a set of pairs $(S_i, C_j)$ where each student $S_i$ appears in at most $q_i$ such pairs, and each course $C_j$ appears in at most $r_j$ pairs. Suppose each student ranks a subset of acceptable courses in order of preference, and the supervising professor of each course ranks a subset of acceptable students. Define a *stable enrollment* in an appropriate way.

## 2.10 Combinatorial Geometry

*We should expose the student to some material that has strong intuitive appeal, is currently of research interest to professional mathematicians, and in which the student himself may discover interesting problems that even the experts are unable to solve.*

— Victor Klee, from the translator's preface to *Combinatorial Geometry in the Plane* [144]

The subject of *combinatorial geometry* studies combinatorial problems regarding arrangements of points in space, and the geometric figures obtained from them. Such figures include lines and polygons in two dimensions, planes and polyhedra in three, and hyperplanes and polytopes in $n$-dimensional space. This subject has much in common with the somewhat broader subject of *discrete geometry*, which treats all sorts of geometric problems on discrete sets of points in Euclidean space, especially extremal problems concerning quantities such as distance, direction, area, volume, perimeter, intersection counts, and packing density.

In this section, we provide an introduction to the field of combinatorial geometry by describing two famous problems regarding points in the plane: a question of Sylvester concerning the collection of lines determined by a set of points, and a problem of Erdős, Klein, and Szekeres on the existence of certain polygons that can be formed from large collections of points in the plane. The latter problem leads us again to Ramsey's theorem, and we prove this statement in a more general form than what we described in Section 1.8. (Ramsey theory is developed further in Chapter 3.) In particular, we establish some of the bounds on the Ramsey numbers $R(p, q)$ that were cited in Section 1.8.

### 2.10.1 Sylvester's Problem

*Thufferin' thuccotash!*

— Sylvester the cat, *Looney Tunes*

James Joseph Sylvester, a British-born mathematician, spent the latter part of his career at Johns Hopkins University, where he founded the first research school in mathematics in America, and established the first American research journal in the subject, *The American Journal of Mathematics*. Toward the end of his career, Sylvester posed the following problem in 1893, in the "Mathematical Questions" column of the British journal, *Educational Times* [265].

**Sylvester's Problem.** *Given $n \geq 3$ points in the plane which do not all lie on the same line, must there exist a line that passes through exactly two of them?*

Given a collection of points in the plane, we say a line is *ordinary* if it passes through exactly two of the points. Thus, Sylvester's problem asks if an ordinary line always exists, as long as the points are not all on the same line.

This problem remained unsolved for many years, and seemed to have been largely forgotten until Erdős rediscovered it in 1933. Tibor Gallai, a friend of Erdős' who is also known as T. Grünwald, found the first proof in the same year. Erdős helped to revive the problem by posing it in the "Problems" section of the *American Mathematical Monthly* in 1933 [89], and Gallai's solution was published in the solution the following year [264].

Kelly also produced a clever solution, which was published in a short article by Coxeter in 1948 [62], along with a version of Gallai's argument. Forty years later, the computer scientist Edsger Dijkstra derived a similar proof, but with a more algorithmic viewpoint [76]. The proof we present here is based on Dijkstra's algorithm. Given any collection of three or more points which do not all lie on the same line, it constructs a line with the required property.

In this method, we start with an arbitrary line $\ell_1$ connecting at least two points of the set, and some point $S_1$ from the set that does not lie on $\ell_1$. If $\ell_1$ contains just two of the points, we are done, so suppose that at least three of the points lie on $\ell_1$. The main idea of the method is to construct from the current line $\ell_1$ and point $S_1$ another line $\ell_2$ and point $S_2$, with $S_2$ not on $\ell_2$. Then we iterate this process, constructing $\ell_3$ and $S_3$, then $\ell_4$ and $S_4$, etc., until one is assured of obtaining a line that connects exactly two of the points of the original collection. In order to ensure that the procedure does not cycle endlessly, we introduce a *termination argument*: a strictly monotone function of the state of the algorithm. A natural candidate is the distance $d_k$ from the point $S_k$ to the line $\ell_k$, so $d_k = d(S_k, \ell_k)$. We therefore aim to construct $\ell_{k+1}$ and $S_{k+1}$ from $\ell_k$ and $S_k$ in such a way that $d_{k+1} < d_k$. Since there are only finitely many points, there are only finitely many possible values for $d_k$, so if we can achieve this monotonicity, then it would follow that the procedure must terminate.

We derive a procedure that produces a strictly decreasing sequence $\{d_k\}$. Suppose the line $\ell_k$ contains the points $P_k$, $Q_k$, and $R_k$ from our original collection, and $S_k$ is a point from the set that does not lie on $\ell_k$. We need to choose $\ell_{k+1}$ and

$S_{k+1}$ so that $d_{k+1} < d_k$. Suppose we set $S_{k+1}$ to be one of the points that we labeled on $\ell_k$, say $S_{k+1} = Q_k$. Certainly $Q_k$ does not lie on either of the lines $\overline{P_k S_k}$ or $\overline{R_k S_k}$, so we might choose one of these two lines for our $\ell_{k+1}$. Can we guarantee that one of these choices will produce a good value for $d_{k+1}$? To test this, let

$$p_k = d(Q_k, \overline{P_k S_k})$$

and

$$r_k = d(Q_k, \overline{R_k S_k}).$$

We require then that

$$\min(p_k, r_k) < d_k. \tag{2.133}$$

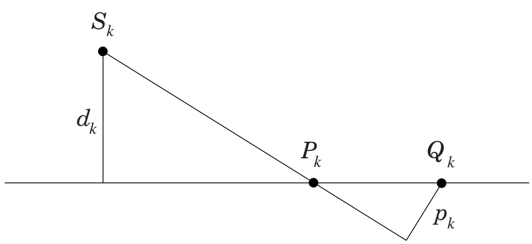



FIGURE 2.24. Similar triangles in the construction.

Using similar triangles in Figure 2.24, we see that the inequality $p_k < d_k$ is equivalent to the statement

$$d(P_k, Q_k) < d(P_k, S_k), \tag{2.134}$$

and likewise $r_k < d_k$ is equivalent to the inequality

$$d(Q_k, R_k) < d(S_k, R_k). \tag{2.135}$$

Now at least one of (2.134) or (2.135) must hold if

$$d(P_k, Q_k) + d(Q_k, R_k) < d(P_k, S_k) + d(S_k, R_k).$$

Further, since $S_k$ does not lie on $\ell_k$, by the triangle inequality we know that

$$d(P_k, R_k) < d(P_k, S_k) + d(S_k, R_k).$$

Therefore, inequality (2.133) follows from the statement

$$d(P_k, Q_k) + d(Q_k, R_k) \le d(P_k, R_k).$$

However, by the triangle inequality, we know that

$$d(P_k, Q_k) + d(Q_k, R_k) \ge d(P_k, R_k).$$

Thus, we require that

$$d(P_k, Q_k) + d(Q_k, R_k) = d(P_k, R_k).$$

Clearly, this latter condition holds if and only if $Q_k$ lies between $P_k$ and $R_k$ on $\ell_k$. We therefore obtain the following algorithm for solving Sylvester's problem.

**Algorithm 2.25.** *Construct an ordinary line.*

*Input.* A set of $n \geq 3$ points in the plane, not all on the same line.

*Output.* A line connecting exactly two of the points.

*Description.*

*Step 1.* Let $\ell_1$ be a line connecting at least two of the points in the given set, and let $S_1$ be a point from the collection that does not lie on $\ell_1$. Set $k = 1$, then perform Step 2.

*Step 2.* If $\ell_k$ contains exactly two points from the original collection, then output $\ell_k$ and stop. Otherwise, perform Step 3.

*Step 3.* Let $P_k$, $Q_k$, and $R_k$ be three points from the given set that lie on $\ell_k$, with $Q_k$ lying between $P_k$ and $R_k$. Set $S_{k+1} = Q_k$, and set $\ell_{k+1} = \overline{P_k S_k}$ if $d(Q_k, \overline{P_k S_k}) < d(Q_k, \overline{P_k R_k})$; otherwise set $\ell_{k+1} = \overline{R_k S_k}$. Then increment $k$ by 1 and repeat Step 2.

Now Sylvester's problem is readily solved: The monotonicity of the sequence $\{d_k\}$ guarantees that the algorithm must terminate, so it must produce a line connecting just two points of the given set. An ordinary line must therefore always exist.

We can illustrate Dijkstra's algorithm with an example. Figure 2.25 shows a collection of thirteen points that produce just six ordinary lines (shown in bold), along with 21 lines that connect at least three of the points. Figure 2.26 illustrates the action of Algorithm 2.25 on these points, using a particular initial configuration. Each successive diagram shows the line $\ell_k$, the point $S_k$ off the line, and the points $P_k$, $Q_k$, and $R_k$ on the line.

Much more is now known about Sylvester's problem. For example, Csima and Sawyer [64, 65] proved that every arrangement of $n \geq 3$ points in the plane, not all on the same line, must produce at least $6n/13$ ordinary lines, except for certain arrangements of $n = 7$ points. Figure 2.25 shows that this bound is best possible, and Exercise 2 asks you to determine an exceptional configuration for $n = 7$. Also, it has long been conjectured that there are always at least $\lceil n/2 \rceil$ ordinary lines for a set of $n$ non-colinear points, except for $n = 7$ and $n = 13$, but this remains unresolved. For additional information on Sylvester's problem and several of its generalizations, see the survey article by Borwein and Moser [34], or the book by Brass, Moser, and Pach [37, sec. 7.2].

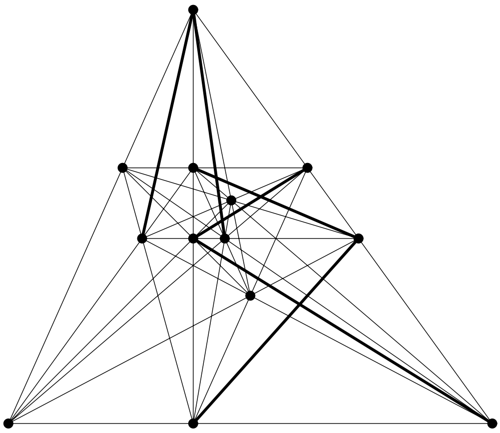


FIGURE 2.25. A collection of thirteen points with just six ordinary lines.

## Exercises

1. Exhibit an arrangement of six points in the plane that produce exactly three ordinary lines.

2. Exhibit an arrangement of seven points in the plane that produce exactly three ordinary lines.

3. Exhibit an arrangement of eight points in the plane that produce exactly four ordinary lines.

4. Exhibit an arrangement of nine points in the plane that produce exactly six ordinary lines.

5. Suppose $n \geq 3$ points in the plane do not all lie on the same line. Show that if one joins each pair of points with a straight line, then one must obtain at least $n$ distinct lines.

6. We say a set of points $B$ is *separated* if there exists a positive number $\delta$ such that the distance $d(P, Q) \geq \delta$ for every pair of points $P$ and $Q$ in $B$. Describe an infinite, separated set of points in the plane, not all on the same line, for which no ordinary line exists. What happens if you apply Dijkstra's algorithm to this set of points?

7. Repeat problem 6, if each of the points $(x, y)$ must in addition satisfy $|y| \leq 1$.

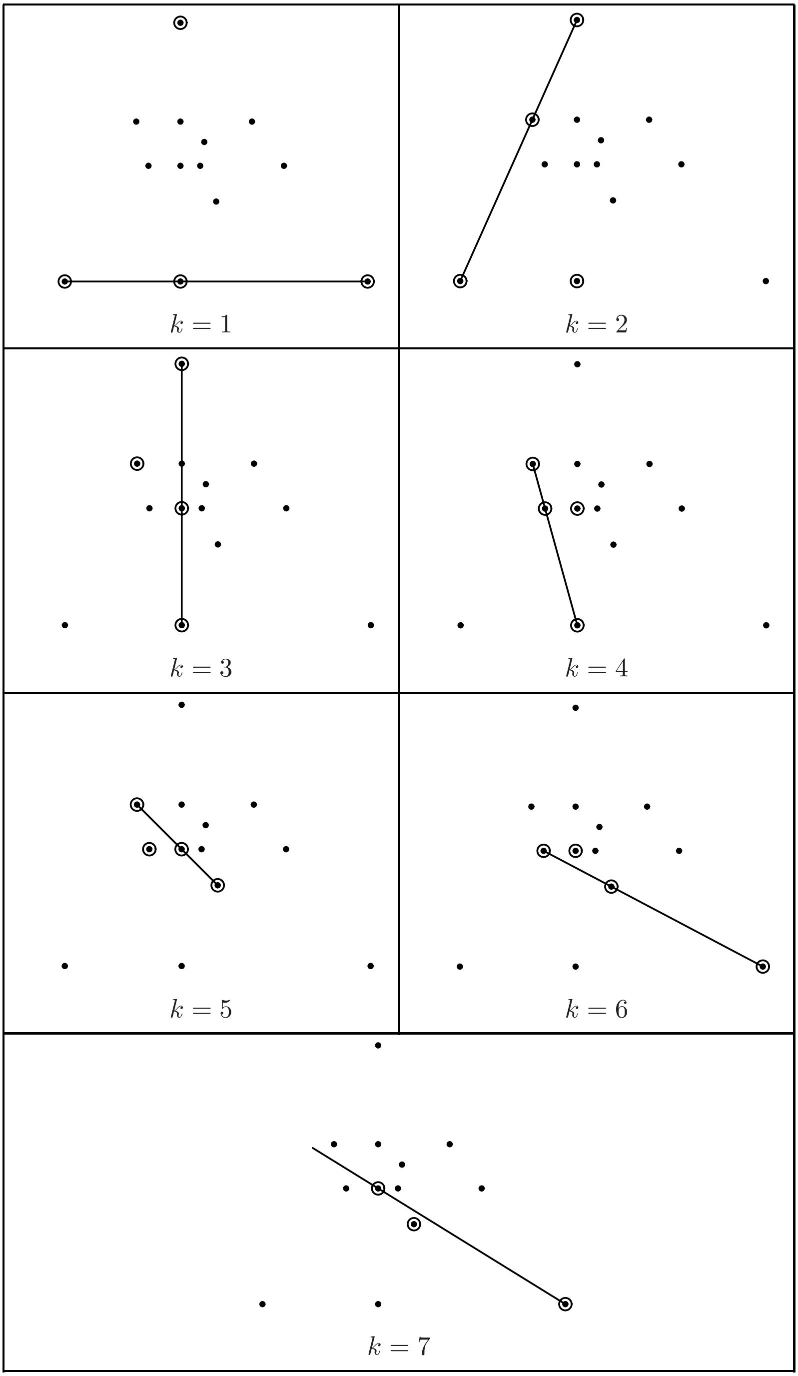



FIGURE 2.26. Dijkstra's algorithm.

8. Let the set $S$ consist of the point $(0,0)$, together with all the points in the plane of the form $(\frac{1}{3k-1}, \frac{1}{3k-1})$, $(\frac{-1}{3k-1}, \frac{1}{3k-1})$, or $(0, \frac{2}{3k-2})$, where $k$ is an arbitrary integer. Show that every line connecting two points of $S$ must intersect a third point of $S$.

9. Consider the following collection $T$ of three-element subsets of the seven-element set $S = \{a, b, c, d, e, f, g\}$:

$$T = \{\{a, b, c\}, \{a, d, e\}, \{a, f, g\}, \{b, d, f\}, \{b, e, g\}, \{c, d, g\}, \{c, e, f\}\}.$$

   (a) Verify that each two-element subset of $S$ is in fact a subset of one of the members of $T$, and that any two distinct sets in $T$ have at most one element in common.

   (b) Explain how this example is germane to Sylvester's problem. Hint: Try thinking of the elements of $S$ as points, and the elements of $T$ as lines.

### 2.10.2 Convex Polygons

*I would certainly pay \$500 for a proof of Szekeres' conjecture.*

— Paul Erdős, [92, p. 66]

A set of points $S$ in the plane is said to be *convex* if for each pair of points $a$ and $b$ in $S$, the line segment joining $a$ to $b$ lies entirely in $S$. Loosely, then, a convex set has no "holes" in its interior, and no "dents" in its boundary. Line segments, triangles, rectangles, and ellipses are thus all examples of convex sets.

The *convex hull* of a finite collection of points $T$ in the plane is defined as the intersection of all closed convex sets which contain $T$. Less formally, if one imagines $T$ represented by a set of pushpins in a bulletin board, then the convex hull of $T$ is the shape enclosed by a rubber band when it is snapped around all the pushpins. The convex hull of a set of three points then is either a triangle or a line segment, and for four points we may obtain one of these shapes, or a convex quadrilateral.

In order to avoid degenerate cases, we will assume in this section that our given collection of points is in *general position*, which means that no three points lie on the same line, or, using the term from the previous section, that each line connecting two of the points is ordinary. Thus, the convex hull of a set of four points in general position forms either a quadrilateral, or a triangle whose interior contains the fourth point of the collection. In the early 1930s, Esther Klein observed that one can always find a convex quadrilateral in a collection of five points in general position.

**Theorem 2.26.** *Any collection of five points in the plane in general position contains a four-element subset whose convex hull is a quadrilateral.*

*Proof.* Suppose we are given a collection of five points in the plane, with no three on the same line. If their convex hull is a pentagon or a quadrilateral, then the statement follows, so suppose that it forms a triangle. Let $a$ and $b$ be the two points of the collection lying inside the triangle, and let $\ell$ be the line connecting $a$ and $b$. Since the points are in general position, two of the vertices of the triangle lie on one side of $\ell$. Label them $c$ and $d$. Then the convex hull of $\{a, b, c, d\}$ is a quadrilateral. See Figure 2.27. □

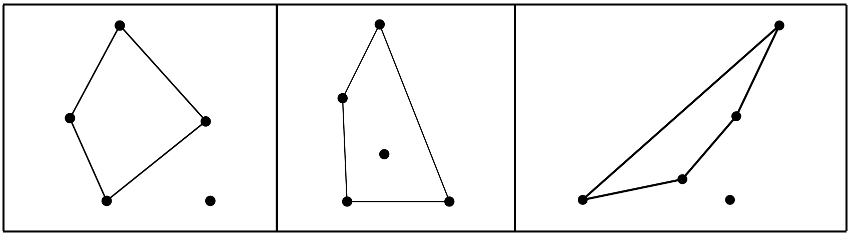


FIGURE 2.27. A convex quadrilateral may always be found among five points in general position.

Klein then asked about a natural generalization. How many points in the plane (in general position) are required in order to be certain that some subset forms the convex hull of a polygon with $n$ sides? Does such a number exist for each $n$? For example, Figure 2.28 illustrates a collection of eight points, no five of which produce a convex pentagon, and a set of sixteen points, no six of which forms a convex hexagon. Thus, at least nine points are needed for $n = 5$, and at least seventeen for $n = 6$.

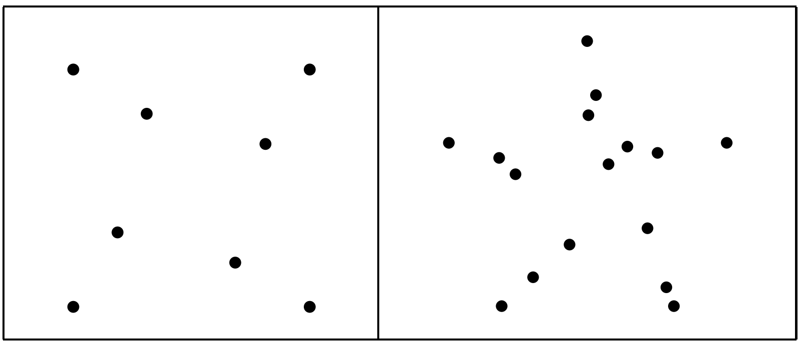


FIGURE 2.28. Eight points with no convex pentagon, and sixteen points with no convex hexagon.

Erdős and Szekeres studied this problem in their first joint paper, in 1935 [94]. There they independently developed a version of Ramsey's theorem, and the proof we describe in this section is based on their argument. The statement we develop here is much more general than the special case of Ramsey's theorem that we described in Section 1.8, although Ramsey in fact established a still more general result in his seminal paper of 1930 [232] (see Exercise 7). We will also derive the bounds on the ordinary Ramsey numbers $R(m, n)$ stated in Theorems 1.63 and 1.64 of Section 1.8 as special cases.

Let $\mathrm{ES}(n)$ denote the minimal number of points in the plane in general position that are required so that there must exist a subcollection of $n$ points whose convex hull is a polygon with $n$ sides (an $n$-gon). Thus, we have seen that $\mathrm{ES}(3) = 3$, $\mathrm{ES}(4) = 5$, and, from Figure 2.28, that $\mathrm{ES}(5) \geq 9$ and $\mathrm{ES}(6) \geq 17$. We aim to show that $\mathrm{ES}(n)$ exists for each $n$ by obtaining an upper bound on its value, in terms of $n$. As a first step, we show that it is enough to find a collection of $n$ points, each of whose four-element subsets forms a convex quadrilateral.

**Theorem 2.27.** *Suppose $S$ is a set of $n$ points in the plane in general position with the property that each four-element subset of $S$ is the vertex set of a convex quadrilateral. Then $S$ is the set of vertices of a convex $n$-gon.*

*Proof.* Let $H$ denote the convex hull of $S$, and suppose $a \in S$ lies in the interior of $H$. Let $b \in S$ with $a \neq b$. Divide $H$ into triangles by joining $b$ to each vertex of $H$. Then $a$ lies in the interior of one of these triangles, and we label its vertices $b$, $c$, and $d$. But then $\{a, b, c, d\}$ is a four-element subset of $S$ whose convex hull is a triangle, contradicting our assumption. $\square$

Next, we develop the more general version of Ramsey's theorem. Recall that in Section 1.8 we defined $R(m, n)$ to be the smallest positive integer $N$ such that any 2-coloring of the edges of the complete graph $K_N$ (using the colors red and blue) must produce either a red $K_m$ or a blue $K_n$ as a subgraph. Coloring each edge of $K_N$ is certainly equivalent to assigning a color to each of the $\binom{N}{2}$ subsets of size 2 of the set $\{1, 2, \ldots, N\}$, and so we might consider what happens more generally when we assign a color to each of the $\binom{N}{k}$ subsets of size $k$, for a fixed positive integer $k$. We call such a subset a *$k$-subset* of the original set. Ramsey's theorem extends in a natural way to this setting. For convenience, we let $[N]$ denote the set $\{1, 2, \ldots, N\}$, and we define the generalized Ramsey numbers in the following way.

**Definition.** For positive integers $k$, $m$, and $n$, with $m \geq k$ and $n \geq k$, the *Ramsey number* $R_k(m, n)$ is defined as the smallest positive integer $N$ such that in any 2-coloring of the $k$-subsets of $[N]$ (using the colors red and blue) there must exist either a subset of $m$ elements, each of whose $k$-subsets is red, or a subset of $n$ elements, each of whose $k$-subsets is blue.

Thus, the Ramsey numbers $R(m, n)$ of Section 1.8 are denoted by $R_2(m, n)$ here. Also, just as the ordinary Ramsey numbers can be described in terms of coloring edges of complete graphs, so too can we describe $R_k(m, n)$ in terms of coloring edges of certain hypergraphs (see Exercise 1).

The next theorem establishes that the Ramsey numbers $R_k(m, n)$ always exist, and provides an upper bound on their values.

**Theorem 2.28** (Ramsey's Theorem)**.** *Let $k$, $m$, and $n$ be positive integers, with $\min\{m, n\} \geq k$. Then the Ramsey number $R_k(m, n)$ exists. Furthermore, for each such $k$, $m$, and $n$, we have*

$$R_1(m, n) = m + n - 1, \tag{2.136}$$

$$R_k(k, n) = n, \tag{2.137}$$

$$R_k(m, k) = m, \tag{2.138}$$

*and, if* $\min\{m, n\} > k \geq 2$*, then*

$$R_k(m, n) \leq R_{k-1}\big(R_k(m-1, n) + R_k(m, n-1)\big) + 1. \tag{2.139}$$

*Proof.* First, consider the case $k = 1$. If the elements of $[N]$ are each colored red or blue, and there are fewer than $m$ red elements and fewer than $n$ blue elements, then certainly $N \leq m + n - 2$, and (2.136) follows.

Second, suppose $k = m$, and suppose that each $k$-subset of $[N]$ is colored red or blue. If any is red then we have a qualifying $m$-subset, so suppose all are blue. Then we have a qualifying $n$-subset precisely when $N \geq n$. Thus, the formula (2.137) follows, and by symmetry so does (2.138).

To establish (2.139), suppose $\min\{m, n\} > k \geq 2$. Using induction on $k$, we may assume that $R_{k-1}(a, b)$ exists for all integers $a$ and $b$ with $\min\{a, b\} \geq k - 1$, and further by induction on $m + n$ we may assume that $R_k(m - 1, n)$ and $R_k(m, n - 1)$ both exist. Let $m' = R_k(m - 1, n)$, $n' = R_k(m, n - 1)$, and $N = R_{k-1}(m', n') + 1$, and consider an arbitrary 2-coloring $C$ of the $k$-subsets of $[N]$ using the colors red and blue. Create a coloring $C'$ of the $(k - 1)$-subsets of $[N - 1]$ by assigning a subset $X$ of size $k - 1$ the color of the set $X \cup \{N\}$ in $C$. Since $N - 1 = R_{k-1}(m', n')$, the coloring $C'$ must produce either a subset of $[N - 1]$ of cardinality $m'$, each of whose $(k - 1)$-subsets is red, or a subset of $[N - 1]$ of cardinality $n'$, each of whose $(k - 1)$-subsets is blue. Suppose the first possibility occurs (the argument for the second case is symmetric), and let $S$ be a qualifying subset of $[N - 1]$. Since $S$ has $m' = R_k(m - 1, n)$ elements, there must exist either a subset of size $m - 1$ of $S$, each of whose $k$-subsets is red in the original coloring $C$, or a subset of size $n$ of $S$, each of whose $k$-subsets is blue in $C$. In the latter case, we are done, so suppose the former case occurs, and let $T$ be such a subset of $[N - 1]$. Let $T' = T \cup \{N\}$, and suppose $X$ is a $k$-subset of $T'$. If $N \notin X$, then $X \subseteq S$, so $X$ is red in $C$. If $N \in X$, then $X \setminus \{N\}$ is a $(k - 1)$-subset of $S$ and so is red in $C'$, and thus $X$ is red in $C$. □

Using this result, we can now establish the upper bound for the original Ramsey numbers $R_2(m, n)$ that was cited in Section 1.8.

**Corollary 2.29.** *Suppose $m$ and $n$ are integers with* $\min\{m, n\} \geq 2$*. Then*

$$R_2(m, n) \leq R_2(m - 1, n) + R_2(m, n - 1) \tag{2.140}$$

*and*

$$R_2(m, n) \leq \binom{m + n - 2}{m - 1}. \tag{2.141}$$

*Proof.* The inequality (2.140) follows at once from (2.136) and (2.139). The formulas (2.137) and (2.138) produce equality in (2.141) for the cases $m = 2$ and $n = 2$ respectively, and the general inequality follows by induction on $m+n$ (see Exercise 3). □

Armed with Ramsey's theorem, we may now prove that a sufficiently large collection of points in the plane in general position must contain a subset that forms the vertices of a convex $n$-gon, for any positive integer $n$.

**Theorem 2.30.** *If* $n \geq 3$ *is an integer, then* $\mathrm{ES}(n) \leq R_4(5, n)$.

*Proof.* Let $S$ be a collection of $N = R_4(5, n)$ points in the plane in general position. For each four-element subset $T$ of $S$, assign $T$ the color red if its convex hull is a triangle, and assign it the color blue if it is a quadrilateral. By Ramsey's Theorem, there must exist either a five-element subset of $S$ whose $4$-subsets are all red, or an $n$-element subset of $S$ whose $4$-subsets are all blue. The former case is impossible by Theorem 2.26, so the latter case must occur, and this implies that the $n$ points form the vertex set of a convex $n$-gon by Theorem 2.27. □

Much more is known about the quantity $\mathrm{ES}(n)$. In the same article [94], Erdős and Szekeres employ a separate geometric argument to show that in fact

$$\mathrm{ES}(n) \leq \binom{2n-4}{n-2} + 1.$$

Since then, this bound has been improved several times. For example, in 2005 Tóth and Valtr [268] proved that

$$\mathrm{ES}(n) \leq \binom{2n-5}{n-2} + 1$$

for $n \geq 5$.

Few exact values of $\mathrm{ES}(n)$ have been determined. In [94], Erdős and Szekeres noted that Makai first proved that $\mathrm{ES}(5) = 9$, so Figure 2.28 exhibits an extremal configuration. Proofs of this statement were published later in [171] and [30]. In 2006, Szekeres and Peters [266] employed a computational strategy to establish that $\mathrm{ES}(6) = 17$. Thus, again Figure 2.28 illustrates an optimal arrangement. Erdős and Szekeres conjectured that in fact $\mathrm{ES}(n) = 2^{n-2} + 1$ for all $n \geq 3$, and this problem remains open. This is the \$500 conjecture that Erdős was referring to in the quote that opens this section.

It is known that $\mathrm{ES}(n)$ cannot be any smaller than the conjectured value. In 1961, Erdős and Szekeres [95] described a method for placing $2^{n-2}$ points in the plane in general position so that no convex $n$-gon appears. Their construction was later corrected by Kalbfleisch and Stanton [172]. Thus, certainly

$$\mathrm{ES}(n) \geq 2^{n-2} + 1$$

for $n \geq 7$. For additional information on this problem and many of its generalizations, see for instance the books by Brass, Moser, and Pach [37, sec. 8.2] and

Matoušek [200, chap. 3], the survey article by Morris and Soltan [208], or the note by Dumitrescu [82].

**Exercises**

1. State Ramsey's theorem in terms of coloring edges of certain hypergraphs.

2. Exhibit a collection of eight points in general position in the plane whose convex hull is a triangle, so that no subset of four points forms the vertex set of a convex quadrilateral.

3. Complete the proof of Corollary 2.29.

4. (Johnson [169].) If $S$ is a finite set of points in the plane in general position, and $T$ is a subset of $S$ of size 3, let $\psi_S(T)$ denote the number of points of $S$ that lie in the interior of the triangle determined by $T$. Complete the following argument to establish a different upper bound on $\mathrm{ES}(n)$.

   (a) Let $n \geq 3$ be an integer. Prove that if $S$ is sufficiently large, then there exists a subset $U$ of $S$ of size $n$ such that either every 3-subset $T$ of $U$ has $\psi_S(T)$ even, or every such subset has $\psi_S(T)$ odd.

   (b) If $U$ does not form the vertex set of a convex $n$-gon, then by Theorem 2.27 there exist four points $a$, $b$, $c$, and $d$ of $U$, with $d$ lying inside the triangle determined by $a$, $b$, and $c$. Show that

   $$\psi_S(\{a, b, c\}) = \psi_S(\{a, b, d\}) + \psi_S(\{b, c, d\}) + \psi_S(\{a, c, d\}) + 1.$$

   (c) Establish a contradiction and conclude that $\mathrm{ES}(n) \leq R_3(n, n)$.

5. (Tarsy [188].) If $a$, $b$, and $c$ form the vertices of a triangle in the plane, let $\theta(a, b, c) = 1$ if the path $a \to b \to c \to a$ induces a clockwise orientation of the boundary, and let $\theta(a, b, c) = -1$ if it is counterclockwise. Thus, for example, $\theta(a, b, c) = -\theta(a, c, b)$. Complete the following argument to establish an upper bound on $\mathrm{ES}(n)$.

   (a) Let $n \geq 3$ be an integer, and let $S = \{v_1, v_2, \ldots, v_N\}$ be a set of labeled points in the plane in general position. Prove that if $N$ is sufficiently large, then there exists a subset $U$ of $S$ of size $n$ such that either every 3-subset $\{v_i, v_j, v_k\}$ of $U$ with $i < j < k$ has $\theta(v_i, v_j, v_k) = 1$, or every such subset has $\theta(v_i, v_j, v_k) = -1$.

   (b) Prove that if $S$ contains a 4-subset whose convex hull is a triangle, then this subset must contain triangles of both orientations with respect to the ordering of the vertices.

   (c) Conclude that $\mathrm{ES}(n) \leq R_3(n, n)$.

6. Complete the proof of Theorem 1.64 by proving that if $m$ and $n$ are positive integers with $\min\{m, n\} \geq 2$, and $R_2(m-1, n)$ and $R_2(m, n-1)$ are both even, then

$$R_2(m, n) \leq R_2(m-1, n) + R_2(m, n-1) - 1.$$

Use the following strategy. Let $r_1 = R_2(m-1, n)$, $r_2 = R_2(m, n-1)$, and $N = r_1 + r_2 - 1$. Suppose that the edges of $K_N$ are 2-colored, using the colors red and blue, in such a way that no red $K_m$ nor blue $K_n$ appears.

(a) Show that the red degree of any vertex in the graph must be less than $r_1$.

(b) Show that the red degree of any vertex in the graph must equal $r_1 - 1$.

(c) Compute the number of red edges in the graph, and establish a contradiction.

7. Prove the following more general version of Ramsey's theorem. Let $k$, $n_1$, $n_2$, ..., $n_r$ be positive integers, with $\min\{n_1, \ldots, n_r\} \geq k$, and let $c_1$, $c_2$, ..., $c_r$ denote $r$ different colors. Then there exists a positive integer $R_k(n_1, \ldots, n_r)$ such that in any $r$-coloring of the $k$-subsets of a set with $N \geq R_k(n_1, \ldots, n_r)$ elements, there must exist a subset of $n_i$ elements, each of whose $k$-subsets has color $c_i$, for some $i$ with $1 \leq i \leq r$.

8. (Schur [251].) If $C$ is an $r$-coloring of the elements of $[N]$, then let $C'$ be the $r$-coloring of 2-subsets of $[N] \cup \{0\}$ obtained by assigning the pair $\{a, b\}$ the color of $|b - a|$ in $C$.

(a) Use the generalized Ramsey's Theorem of Exercise 7 to assert that if $N$ is sufficiently large then in $[N] \cup \{0\}$ there must exist a set of three nonnegative integers, each of whose 2-subsets has the same color in $C'$.

(b) Conclude that if $N$ is sufficiently large then there exist integers $a$ and $b$ in $[N]$, with $a + b \leq N$, such that $a$, $b$, and $a + b$ all have the same color in $C$.

9. Let $S$ be a finite set of points in the plane, and let $P$ be a convex polygon whose vertices are all selected from $S$. We say $P$ is *empty* (with respect to $S$) if its interior contains no points of $S$. Erdős asked if for each integer $n \geq 3$ there exists a positive integer $\mathrm{ES}_0(n)$ such that any set of at least $\mathrm{ES}_0(n)$ points in general position in the plane must contain an empty $n$-gon, but this need not be the case for sets with fewer than $\mathrm{ES}_0(n)$ points.

(a) Compute $\mathrm{ES}_0(3)$ and $\mathrm{ES}_0(4)$.

(b) (Ehrenfeucht [91].) Prove that $\mathrm{ES}_0(5)$ exists by completing the following argument. Let $S$ be a set of $\mathrm{ES}(6)$ points in general position in the plane, and let $P$ be a convex hexagon whose vertices lie in $S$, selected so that its interior contains a minimal number of points of $S$. Denote this number by $m$.

i. Complete the proof if $m = 0$ or $m = 1$.
ii. If $m \geq 2$, let $H$ be the convex hull of the points of $S$ lying inside $P$, and let $\ell$ be a line determined by two points on the boundary of $H$. Finish the proof for this case.

The argument above establishes that $\mathrm{ES}_0(5) \leq 17$; in 1978 Harborth [151] showed that in fact $\mathrm{ES}_0(5) = 10$. Horton [164] in 1983 proved the surprising result that $\mathrm{ES}_0(n)$ does not exist for $n \geq 7$. More recently, Gerken [121] and Nicolás [215] solved the problem for $n = 6$: A sufficiently large set of points in the plane in general position must contain an empty convex hexagon. The precise value of $\mathrm{ES}_0(6)$ remains unknown, though it must satisfy $30 \leq \mathrm{ES}_0(6) \leq \mathrm{ES}(9) \leq 1717$. (An example by Overmars [219] establishes the lower bound; additional information on the upper bound can be found in [182, 271].)

## 2.11 References

*You may talk too much on the best of subjects.*
— Benjamin Franklin, *Poor Richard's Almanack*

We list several additional references for the reader who wishes to embark on further study.

### General References

The text by van Lint and Wilson [273] is a broad and thorough introduction to the field of combinatorics, covering many additional topics. Classical introductions to combinatorial analysis include Riordan [235] and Ryser [246], and many topics in discrete mathematics and enumerative combinatorics are developed extensively in Graham, Knuth, and Patashnik [133]. The text by Pólya, Tarjan, and Woods [227] is a set of notes from a course in enumerative and constructive combinatorics. A problems-oriented introduction to many topics in combinatorics and graph theory can be found in Lovász [191]. The book by Nijenhuis and Wilf [216] describes efficient algorithms for solving a number of problems in combinatorics and graph theory, and a constructive view of the subject is developed in Stanton and White [263]. Texts by Aigner [4, 5], Berge [24], Comtet [60], Hall [146], and Stanley [261, 262] present more advanced treatments of many aspects of combinatorics.

### Combinatorial Identities

The history of binomial coefficients and Pascal's triangle is studied in Edwards [85], and some interesting patterns in the rows of Pascal's triangle are observed by Granville [138]. Combinatorial identities are studied in Riordan [236], and automated techniques for deriving and proving identities involving binomial coefficients and other quantities are developed in Petkovšek, Wilf, and Zeilberger

[222]. Combinatorial proofs for many identities are also developed in the book by Benjamin and Quinn [22].

### Pigeonhole Principle

More nice applications of the pigeonhole principle, together with many other succinct proofs in combinatorics and other subjects, are described in Aigner and Ziegler [6]. An interesting card trick based in part on a special case of Theorem 2.4 is described by Mulcahy [210]. Polynomials with $\{-1, 0, 1\}$ coefficients and a root of prescribed order $m$ at $x = 1$, as in Exercise 14 of Section 2.4, are studied by Borwein and Mossinghoff [35].

### Generating Functions

More details on generating functions and their applications can be found for instance in the texts by Wilf [284] and Graham, Knuth, and Patashnik [133], and in the survey article by Stanley [260]. The problem of determining the minimal degree $d_k$ of a polynomial with $\{0, 1\}$ coefficients that is divisible by $(x+1)^k$, as in Exercise 5 of Section 2.6.5, is studied by Borwein and Mossinghoff [36]. Some properties of the generalized Fibonacci numbers (Exercise 8b of Section 2.6.5) are investigated by Miles [203].

### Pólya's Theory of Counting

Pólya's seminal paper on enumeration in the presence of symmetry is translated into English by Read in [226]. Redfield [233] independently devised the notion of a cycle index for a group, which he termed the *group reduction formula*, ten years before Pólya's paper. As a result, many texts call this topic *Pólya-Redfield theory*. This theory, along with the generalization incorporating a color group, is also described in the expository article by de Bruijn [68], and his research article [69]. Further generalizations of this theory are explored by de Bruijn in [70], culminating in a "monster theorem." Another view of de Bruijn's theorem is developed by Harary and Palmer in [149; 150, chap. 6].

Applications of this theory in chemistry are described in the text by Fujita [116], and additional references for enumeration problems in this field are collected in the survey article [13]. Some applications of Pólya's and de Bruijn's theorems in computer graphics appear for example in articles by Banks, Linton, and Stockmeyer [15, 16].

### More Numbers

The book [10] by Andrews and Eriksson is an introduction to the theory of partitions of integers, directed toward undergraduates. A more advanced treatment is developed by Andrews [9]. Euler's original proof of the pentagonal number theorem, along with some of its additional ramifications, is described by Andrews in [8].

The history of Stirling numbers, the notations developed for them, and many interesting identities they satisfy are discussed by Knuth in [177]. Rhyming schemes, as in Exercise 7d of Section 2.8.3 and Exercise 8 of Section 2.8.4, are analyzed by Riordan [237]. Stirling set numbers arise in a natural way in an interesting problem on juggling in an article by Warrington [280]. Some identities involving the complementary Bell numbers (Exercise 6 of Section 2.8.4) are established in the article by Uppuluri and Carpenter [270].

Eulerian numbers appear in the computation of the volume of certain slabs of $n$-dimensional cubes in articles by Chakerian and Logothetti [51] and Marichal and Mossinghoff [197], and in the solution to a problem concerning a novel graduation ceremony in an article by Gessel [122].

The reference book by Sloane and Plouffe [258] and website by Sloane [257] catalog thousands of integer sequences, many of which arise in combinatorics and graph theory, and list references to the literature for almost all of these sequences. The book by Conway and Guy [61] is an informal discussion of several kinds of numbers, including many common combinatorial sequences.

### Stable Marriage

The important results of Gale and Shapley appeared in [117]. A fast algorithm that solves the "stable roommates" problem whenever a solution exists was first described by Irving in [166]. Stable matching problems are studied in Knuth [178] as motivation for the mathematical analysis of algorithms, and the structure of stable matchings in marriage and roommate problems is described in detail by Gusfield and Irving [143], along with algorithms for their computation. A matching algorithm for the "many-to-many" variation of the stable marriage problem, as in Exercise 9 of Section 2.9.2, is developed by Baïou and Balinski [14]. The monograph by Feder [103] studies extensions of the stable matching problem to more general settings.

### Combinatorial Geometry

A survey on Sylvester's problem regarding ordinary lines for collections of points, as well as related problems, appears in Borwein and Moser [34]. A variation of Sylvester's theorem for an infinite sequence of points lying within a bounded region in the plane is investigated by Borwein [33]. The influential paper of Erdős and Szekeres on convex polygons, first published in [94], also appears in the collection by Gessel and Rota [123]. The survey by Morris and Soltan [208] summarizes work on this problem and several of its variations. Dozens of problems in combinatorial geometry, both solved and unsolved, are described in the books by Brass, Moser, and Pach [37], Hadwiger, Debrunner, and Klee [144], Herman, Kučera, and Šimša [158], and Matoušek [200], as well as the survey article by Erdős and Purdy [93].

## Collected Papers

The collection [123] by Gessel and Rota contains many influential papers in combinatorics and graph theory, including the important articles by Erdős and Szekeres [94], Pólya [225], and Ramsey [232]. The two-volume set edited by Graham and Nešetřil [134, 135] is a collection of articles on the mathematics of Paul Erdős, including many contributions regarding his work in combinatorics and graph theory. The *Handbook of Combinatorics* [131, 132] provides an overview of dozens of different areas of combinatorics and graph theory for mathematicians and computer scientists.

Springer

http://www.springer.com/978-0-387-79710-6

Combinatorics and Graph Theory
Harris, J.; Hirst, J.L.; Mossinghoff, M.
2008, XV, 381 p., Hardcover
ISBN: 978-0-387-79710-6