Combinatorics Through Guided Discovery

Kenneth P. Bogart

Preface

This book is an introduction to combinatorial mathematics, also known as combinatorics. The book focuses especially but not exclusively on the part of combinatorics that mathematicians refer to as "counting." The book consist almost entirely of problems. Some of the problems are designed to lead you to think about a concept, others are designed to help you figure out a concept and state a theorem about it, while still others ask you to prove the theorem. Other problems give you a chance to use a theorem you have proved. From time to time there is a discussion that pulls together some of the things you have learned or introduces a new idea for you to work with. Many of the problems are designed to build up your intuition for how combinatorial mathematics works. There are problems that some people will solve quickly, and there are problems that will take days of thought for everyone. Probably the best way to use this book is to work on a problem until you feel you are not making progress and then go on to the next one. Think about the problem you couldn't get as you do other things. The next chance you get, discuss the problem you are stymied on with other members of the class. Often you will all feel you've hit dead ends, but when you begin comparing notes and listening *carefully* to each other, you will see more than one approach to the problem and be able to make some progress. In fact, after comparing notes you may realize that there is more than one way to interpret the problem. In this case your first step should be to think together about what the problem is actually asking you to do. You may have learned in school that for every problem you are given, there is a method that has already been taught to you, and you are supposed to figure out which method applies and apply it. That is not the case here. Based on some simplified examples, you will discover the method for yourself. Later on, you may recognize a pattern that suggests you should try to use this method again.

The point of learning from this book is that you are learning how to discover ideas and methods for yourself, not that you are learning to apply methods that someone else has told you about. The problems in this book are designed to lead you to discover for yourself and prove for yourself the main ideas of combinatorial mathematics. There is considerable evidence that this leads to deeper learning and more understanding.

You will see that some of the problems are marked with bullets. Those are the problems that I feel are essential to having an understanding of what

comes later, whether or not it is marked by a bullet. The problems with bullets are the problems in which the main ideas of the book are developed. Your instructor may leave out some of these problems because he or she plans not to cover future problems that rely on them. Many problems, in fact entire sections, are not marked in this way, because they use an important idea rather than developing one. Some other special symbols are described in what follows; a summary appears in Table 0.0.1.

●	essential
○	motivational material
+	summary
⇒	especially interesting
*	difficult
·	essential for this section or the next

Table 0.0.1: The meaning of the symbols to the left of problem numbers.

Some problems are marked with open circles. This indicates that they are designed to provide motivation for, or an introduction to, the important concepts, motivation with which some students may already be familiar. You will also see that some problems are marked with arrows. These point to problems that I think are particularly interesting. Some of them are also difficult, but not all are. A few problems that summarize ideas that have come before but aren't really essential are marked with a plus, and problems that are essential if you want to cover the section they are in or, perhaps, the next section, are marked with a dot (a small bullet). If a problem is relevant to a much later section in an essential way, I've marked it with a dot and a parenthetical note that explains where it will be essential. Finally, problems that seem unusually hard to me are marked with an asterisk. Some I've marked as hard only because I think they are difficult in light of what has come before, not because they are intrinsically difficult. In particular, some of the problems marked as hard will not seem so hard if you come back to them after you have finished more of the problems.

If you are taking a course, your instructor will choose problems for you to work on based on the prerequisites for and goals of the course. If you are reading the book on your own, I recommend that you try all the problems in a section you want to cover. Try to do the problems with bullets, but by all means don't restrict yourself to them. Often a bulleted problem makes more sense if you have done some of the easier motivational problems that come before it. If, after you've tried it, you want to skip over a problem without a bullet or circle, you should not miss out on much by not doing that problem. Also, if you don't find the problems in a section with no bullets interesting, you can skip them, understanding that you may be skipping an entire branch of combinatorial mathematics! And no matter what, read the textual material

that comes before, between, and immediately after problems you are working on!

One of the downsides of how we learn math in high school is that many of us come to believe that if we can't solve a problem in ten or twenty minutes, then we can't solve it at all. There will be problems in this book that take hours of hard thought. Many of these problems were first conceived and solved by professional mathematicians, and *they* spent days or weeks on them. How can you be expected to solve them at all then? You have a context in which to work, and even though some of the problems are so open ended that you go into them without any idea of the answer, the context and the leading examples that preceded them give you a structure to work with. That doesn't mean you'll get them right away, but you will find a real sense of satisfaction when you see what you can figure out with concentrated thought. Besides, you can get hints!

Some of the questions will appear to be trick questions, especially when you get the answer. They are not intended as trick questions at all. Instead they are designed so that they don't tell you the answer in advance. For example the answer to a question that begins "How many..." might be "none." Or there might be just one example (or even no examples) for a problem that asks you to find all examples of something. So when you read a question, unless it directly tells you what the answer is and asks you to show it is true, don't expect the wording of the problem to suggest the answer. The book isn't designed this way to be cruel. Rather, there is evidence that the more open-ended a question is, the more deeply you learn from working on it. If you do go on to do mathematics later in life, the problems that come to you from the real world or from exploring a mathematical topic are going to be open-ended problems because nobody will have done them before. Thus working on open-ended problems now should help to prepare you to do mathematics later on.

You should try to write up answers to all the problems that you work on. If you claim something is true, you should explain why it is true; that is you should prove it. In some cases an idea is introduced before you have the tools to prove it, or the proof of something will add nothing to your understanding. In such problems there is a remark telling you not to bother with a proof. When you write up a problem, remember that the instructor has to be able to "get" your ideas and understand exactly what you are saying. Your instructor is going to choose some of your solutions to read carefully and give you detailed feedback on. When you get this feedback, you should think it over carefully and then write the solution again! You may be asked not to have someone else read your solutions to some of these problems until your instructor has. This is so that the instructor can offer help which is aimed at your needs. On other problems it is a good idea to seek feedback from other students. One of the best ways of learning to write clearly is to have someone point out to you where it is hard to figure out what you mean. The crucial thing is to make it clear to your reader that you really want to know where you may have left something out, made an unclear statement, or failed to support a

statement with a proof. It is often very helpful to choose people who have not yet become an expert with the problems, as long as they realize it will help you most for them to tell you about places in your solutions they do not understand, even if they think it is their problem and not yours!

As you work on a problem, think about why you are doing what you are doing. Is it helping you? If your current approach doesn't feel right, try to see why. Is this a problem you can decompose into simpler problems? Can you see a way to make up a simple example, even a silly one, of what the problem is asking you to do? If a problem is asking you to do something for every value of an integer n, then what happens with simple values of n like 0, 1, and 2? Don't worry about making mistakes; it is often finding mistakes that leads mathematicians to their best insights. Above all, don't worry if you can't do a problem. Some problems are given as soon as there is one technique you've learned that might help do that problem. Later on there may be other techniques that you can bring back to that problem to try again. The notes have been designed this way on purpose. If you happen to get a hard problem with the bare minimum of tools, you will have accomplished much. As you go along, you will see your ideas appearing again later in other problems. On the other hand, if you don't get the problem the first time through, it will be nagging at you as you work on other things, and when you see the idea for an old problem in new work, you will know you are learning.

There are quite a few concepts that are developed in this book. Since most of the intellectual content is in the problems, it is natural that definitions of concepts will often be within problems. When you come across an unfamiliar term in a problem, it is likely it was defined earlier. Look it up in the index, and with luck (hopefully no luck will really be needed!) you will be able to find the definition.

Above all, this book is dedicated to the principle that doing mathematics is fun. As long as you know that some of the problems are going to require more than one attempt before you hit on the main idea, you can relax and enjoy your successes, knowing that as you work more and more problems and share more and more ideas, problems that seemed intractable at first become a source of satisfaction later on.

Contents

Chapter 1

What is Combinatorics?

Combinatorial mathematics arises from studying how we can *combine* objects into arrangements. For example, we might be combining sports teams into a tournament, samples of tires into plans to mount them on cars for testing, students into classes to compare approaches to teaching a subject, or members of a tennis club into pairs to play tennis. There are many questions one can ask about such arrangements of objects. Here we will focus on questions about *how many ways* we may combine the objects into arrangements of the desired type. These are called *counting problems*. Sometimes, though, combinatorial mathematicians ask if an arrangement is possible (if we have ten baseball teams, and each team has to play each other team once, can we schedule all the games if we only have the fields available at enough times for forty games?). Sometimes they ask if all the arrangements we might be able to make have a certain desirable property (Do all ways of testing 5 brands of tires on 5 different cars [with certain additional properties] compare each brand with each other brand on at least one common car?). Problems of these sorts come up throughout physics, biology, computer science, statistics, and many other subjects. However, to demonstrate all these relationships, we would have to take detours into all these subjects. While we will give some important applications, we will usually phrase our discussions around everyday experience and mathematical experience so that the student does not have to learn a new context before learning mathematics in context!

1.1   About These Notes

These notes are based on the philosophy that you learn the most about a subject when you are figuring it out directly for yourself, and learn the least when you are trying to figure out what someone else is saying about it. On the other hand, there is a subject called combinatorial mathematics, and that is what we are going to be studying, so we will have to tell you some basic facts. What we are going to try to do is to give you a chance to discover many of the interesting examples that usually appear as textbook examples and dis-

cover the principles that appear as textbook theorems. Your main activity will be solving problems designed to lead you to discover the basic principles of combinatorial mathematics. Some of the problems lead you through a new idea, some give you a chance to describe what you have learned in a sequence of problems, and some are quite challenging. When you find a problem challenging, don't give up on it, but don't let it stop you from going on with other problems. Frequently you will find an idea in a later problem that you can take back to the one you skipped over or only partly finished in order to finish it off. With that in mind, let's get started. In the problems that follow, you will see some problems marked on the left with various symbols. The preface gives a full explanation of these symbols and discusses in greater detail why the book is organized as it is! Table 1.1.1, which is repeated from the preface, summarizes the meaning of the symbols.

•	essential
○	motivational material
+	summary
⇒	especially interesting
*	difficult
·	essential for this section or the next

Table 1.1.1: The meaning of the symbols to the left of problem numbers.

1.2   Basic Counting Principles

Activity 1. Five schools are going to send their baseball teams to a tournament, in which each team must play each other team exactly once. How many games are required?

Hint.   Answer the questions in Problem 2 for the case of five schools.

Solution.   Think of numbering the five teams. The first team must play all four others, the second team will be in one of these games but must play in three more games, with the third, fourth and fifth team. The third team is in two of the games we've already mentioned, and must still play the fourth and fifth team for two more games, and the fourth team must play the fifth team in addition to playing in three of the games already mentioned. Thus there are $4+3+2+1 = 10$ games. Alternatively, there are five teams, each of which must play in four games, giving us 20 pairings of two teams each. However each game involves two of these pairings, so there are $20/2 = 10$ games.

Activity 2. Now some number n of schools are going to send their baseball teams to a tournament, and each team must play each other team exactly once. Let us think of the teams as numbered 1 through n.

(a) How many games does team 1 have to play in?

   Solution. $n - 1$

(b) How many games, other than the one with team 1, does team two have to play in?

   Solution. $n - 2$

(c) How many games, other than those with the first $i - 1$ teams, does team i have to play in?

   Solution. $n - i$

(d) In terms of your answers to the previous parts of this problem, what is the total number of games that must be played?

   Solution. $1 + 2 + \cdots + n - 1$. Although this need not be part of the answer, a formula that we usually use in both algebra and calculus courses tells us this sum is $n(n - 1)/2$.

Activity 3. One of the schools sending its team to the tournament has to send its players from some distance, and so it is making sandwiches for team members to eat along the way. There are three choices for the kind of bread and five choices for the kind of filling. How many different kinds of sandwiches are available?

Hint. For each kind of bread, how many sandwiches are possible?

Solution. $3 \cdot 5 = 15$, or $5 + 5 + 5 = 15$.

Activity 4. An *ordered pair* (a, b) consists of two things we call a and b. We say a is the first member of the pair and b is the second member of the pair. If M is an m element set and N is an n-element set, how many ordered pairs are there whose first member is in M and whose second member is in N? Does this problem have anything to do with any of the previous problems?

Solution. $m \cdot n$. This is because if $M = \{x_1, x_2, \ldots, x_m\}$, then we have n ordered pairs starting with x_1, n ordered pairs starting with x_2, and so on, so the total number of ordered pairs is a sum of m terms, all equal to n. In problem 3 we were looking at ordered pairs of bread and filling. Less directly, and so not required for the answer, In Problem 1 we have 20 ordered pairs, and each baseball game involved two of the ordered pairs so we had 10 baseball games. The same argument applies to Problem 2; namely we have n teams each of which is in an ordered pair with $n - 1$ other teams, so we have $n(n-1)$ ordered pairs, and each game corresponds to two ordered pairs so we have $n(n - 1)/2$ games. This gives us one proof of the formula we mentioned in the solution to that problem.

Activity 5. Since a sandwich by itself is pretty boring, students from the school in Problem 3 are offered a choice of a drink (from among five different kinds), a sandwich, and a fruit (from among four different kinds). In how many ways may a student make a choice of the three items now?

Solution. $5 \cdot 15 \cdot 4 = 300$. Why do we multiply? Multiplying five by 15 is equivalent to adding 15, the number of sandwiches, once for each drink, giving us 75 combinations of drink and sandwich. For each such pair we have 4 choices of fruit, and we can either think of adding 75 fours or adding four 75s to get three hundred. Thus we multiply because multiplication is repeated addition.

Activity 6. The coach of the team in Problem 3 knows of an ice cream parlor along the way where she plans to stop to buy each team member a triple decker cone. There are 12 different flavors of ice cream, and triple decker cones are made in homemade waffle cones. Having chocolate ice cream as the bottom scoop is different from having chocolate ice cream as the top scoop. How many possible ice cream cones are going to be available to the team members? How many cones with three different kinds of ice cream will be available?

Hint. Try to solve the problem first with a two-scoop cone. (Look for an earlier problem that is analogous.) Then, for each two scoop cone, in how many ways can you put on a top scoop?

Solution. $12 \cdot 12 \cdot 12 = 1728$ possible cones. If the flavors must be different, we have $12 \cdot 11 \cdot 10 = 1320$ possible cones. In both cases, the reason we are multiplying is as a shortcut for repeated addition.

Activity 7. The idea of a function is ubiquitous in mathematics. A function f from a set S to a set T is a relationship between the two sets that associates exactly one member $f(x)$ of T with each element x in S. We will come back to the ideas of functions and relationships in more detail and from different points of view from time to time. However, the quick review above should probably let you answer these questions. If you have difficulty with them, it would be a good idea to go now to Appendix A and work through ⟨⟨Unresolved xref, reference "relationsandfunctions"; check spelling or use "provisional" attribute⟩⟩ which covers this definition in more detail. You might also want to study Section to learn to visualize the properties of functions. We will take up the topic of this section later in this chapter as well, but in less detail than is in the appendix.

(a) Using f, g, ..., to stand for the various functions, write down all the different functions you can from the set $\{1, 2\}$ to the set $\{a, b\}$. For example, you might start with $f(1) = a$, $f(2) = b$. How many functions are there from the set $\{1, 2\}$ to the set $\{a, b\}$?

   Hint. Ask yourself "how many choices do we have for $f(1)$?" Then ask how many choices we have for $f(2)$.

Solution. $f(1) = a$, $f(2) = b$. Or, $h(1) = a$, $h(2) = a$. Or, $g(1) = b$, $g(2) = a$. Or $j(1) = b$, $j(2) = b$. We have exhausted all the possibilities for functions that associate a with 1 and all possibilities for functions that associate b with 1, so we have exhausted all possibilities. There are four such functions.

(b) How many functions are there from the three element set $\{1, 2, 3\}$ to the two element set $\{a, b\}$?

Hint. It may not be practical to write down rules for all the functions for this problem. But you could ask yourself how many choices we have for $f(1)$, how many we have for $f(2)$ and how many we have for $f(3)$.

Solution. $2 \cdot 2 \cdot 2 = 8$

(c) How many functions are there from the two element set $\{a, b\}$ to the three element set $\{1, 2, 3\}$?

Hint. If you are choosing a function f, how many choices do you have for $f(a)$? Then how many choices do you have for $f(b)$?

Solution. $3 \cdot 3 = 9$

(d) How many functions are there from a three element set to a 12 element set?

Solution. $12 \cdot 12 \cdot 12 = 1728$

(e) The function f is called **one-to-one** or an **injection** if whenever x is different from y, $f(x)$ is different from $f(y)$. How many one-to-one functions are there from a three element set to a 12 element set?

Solution. $12 \cdot 11 \cdot 10 = 1320$

(f) Explain the relationship between this problem and Problem 6.

Solution. When we counted the number of possible ice cream cones we were counting functions from the three places in the cone where ice cream would sit to the 12 flavors. When we counted the number of possible ice cream cones with different flavors, we were counting the number of one-to-one functions from the three places in the cone where ice cream would sit to the 12 flavors.

Activity 8. A group of hungry team members in Problem 6 notices it would be cheaper to buy three pints of ice cream for them to split than to buy a triple decker cone for each of them, and that way they would get more ice cream. They ask their coach if they can buy three pints of ice cream.

(a) In how many ways can they choose three pints of different flavors out of the 12 flavors?

   Hint.   You know how to figure out in how many ways they could make a list of three flavors out of the twelve. But each set of three flavors can be listed in a number of different ways. Try to figure out in how many ways a set of three flavors can be listed, and then try to see how this helps you.

   Solution.   There are $12 \cdot 11 \cdot 10 = 1320$ ways to make a list of three flavors. But a choice of three flavors accounts for $3 \cdot 2 \cdot 1 = 6$ of those lists. Therefore there are $1320/6 = 220$ ways to choose the pints if the flavors are different.

(b) In how many ways may they choose three pints if the flavors don't have to be different?

   Hint.   Try to break the problem up into cases you can solve by previous methods; then figure out how to get the answer to the problem by using these answers for the cases.

   Solution.   If the flavors need not be different, we must add in the number of ways to choose two pints of one flavor and one of a second and also the number fo ways to choose three pints of one flavor. the first of these is $12 \cdot 11 = 132$ and the third is 12, so we have $220+132+12 = 364$ ways to choose three pints. We can do a more elegant solution after we learn about multisets in Problem 115.

Activity 9. Two sets are said to be **disjoint** if they have no elements in common. For example, $\{1, 3, 12\}$ and $\{6, 4, 8, 2\}$ are disjoint, but $\{1, 3, 12\}$ and $\{3, 5, 7\}$ are not. Three or more sets are said to be **mutually disjoint** if no two of them have any elements in common. What can you say about the size of the union of a finite number of finite (mutually) disjoint sets? Does this have anything to do with any of the previous problems?

Solution.   The size of a union of disjoint sets is the sum of their sizes. We used this principle in Problems 1 and 2 directly, and indirectly in every other problem when we multiplied the number of ways of doing one thing times the number of ways of doing another. Note that we used this principle informally in the explanation in the solution of Problem 4.

Activity 10. Disjoint subsets are defined in Problem 9. What can you say about the size of the union of m (mutually) disjoint sets, each of size n? Does this have anything to do with any of the previous problems?

Solution. The size of the union is $m \cdot n$. This is because the size of a union of disjoint sets is the sum of their sizes, and a sum of m terms each equal to n is $m \cdot n$.

The sum and product principles

These problems contain among them the kernels of many of the fundamental ideas of combinatorics. For example, with luck, you just stated the sum principle (illustrated in Figure 1.2.1), and product principle (illustrated in Figure 1.2.2) in Problems 9 and Activity 10. These two counting principles are the basis on which we will develop many other counting principles.
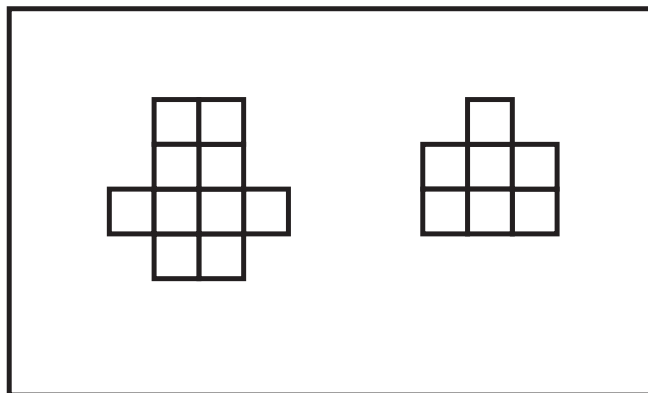
Figure 1.2.1: The union of these two disjoint sets has size 17.
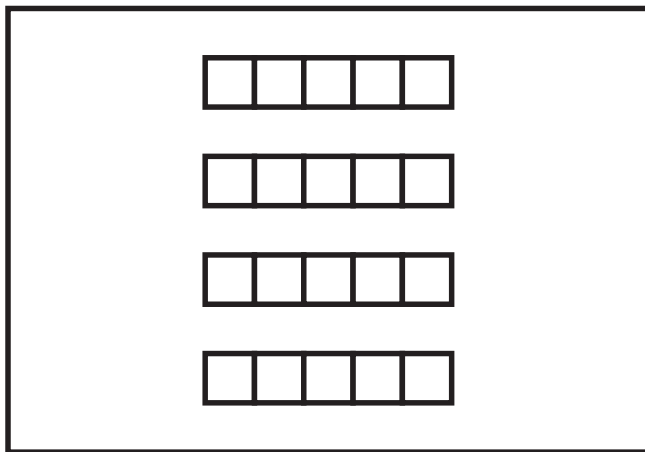
Figure 1.2.2: The union of four disjoint sets of size five.

You may have noticed some standard mathematical words and phrases such as *set*, *ordered pair*, *function* and so on creeping into the problems. One of our goals in these notes is to show how most counting problems can be recognized as counting all or some of the elements of a set of standard mathematical objects. For example Problem 4 is meant to suggest that the question we asked in Problem 3 was really a problem of counting all the ordered pairs consisting of a bread choice and a filling choice. We use $A \times B$ to stand for the set of all ordered pairs whose first element is in A and whose second element is in B and we call $A \times B$ the **Cartesian product** of A and B, so you can think of Problem 4 as asking you for the size of the Cartesian product of M and N, that is, asking you to count the number of elements of this Cartesian product.

When a set S is a union of disjoint sets $B_1, B_2, \ldots, B_m$ we say that the sets $B_1, B_2, \ldots, B_m$ are a **partition** of the set S. Thus a partition of S is a (special kind of) set of sets. So that we don't find ourselves getting confused between the set S and the sets B_i into which we have divided it, we often call the sets $B_1, B_2, \ldots, B_m$ the *blocks* of the partition. In this language, the **sum principle**says that

> if we have a partition of a set S, then the size of S is the sum of the sizes of the blocks of the partition.

The **product principle** says that

> if we have a partition of a set S into m blocks, each of size n, then S has size mn.

You'll notice that in our formal statement of the sum and product pinciple we talked about a partition of a finite set. We could modify our language a bit to cover infinite sizes, but whenever we talk about sizes of sets in what follows, we will be working with finite sets. So as to avoid possible complications in

the future, let us agree that when we refer to the size of a set, we are implicitly assuming the set is finite. There is another version of the product principle that applies directly in problems like Problem 5 and Problem 6, where we were not just taking a union of m disjoint sets of size n, but rather m disjoint sets of size n, each of which was a union of m' disjoint sets of size n'. This is an inconvenient way to have to think about a counting problem, so we may rephrase the product principle in terms of a sequence of decisions:

Activity 11. If we make a sequence of m choices for which

- there are k_1 possible first choices, and

- for each way of making the first $i - 1$ choices, there are k_i ways to make the ith choice,

then in how many ways may we make our sequence of choices? (You need not prove your answer correct at this time.)

The counting principle you gave in Problem 11 is called the *general product principle*. We will outline a proof of the general product pinciple from the original product principle in Problem 78. Until then, let us simply accept it as another counting principle. For now, notice how much easier it makes it to explain why we multiplied the things we did in Problem 5 and Problem 6.

Activity 12. A tennis club has $2n$ members. We want to pair up the members by twos for singles matches.

(a) In how many ways may we pair up all the members of the club? (Hint: consider the cases of 2, 4, and 6 members.)

Hint. Suppose you have a list in alphabetical order of names of the members of the club. In how many ways can you pair up the first person on the list? In how many ways can you pair up the next person who isn't already paired up?

Solution. Suppose we list the people in the club in some way, and keep that list for the remainder of the problem. Take the first person from the list and pair that person with any of the $2n - 1$ remaining people. Now take the next *unpaired* person from the list and pair that person with any of the remaining $2n - 3$ unpaired people. Continuing in this way, once k pairs have been selected, take the next unpaired person from the list and pair that person with any of the remaining $2n - 2k - 1$ unpaired people. Every pairing can arise in this way, and no pairing can arise twice in this process. Thus the number of outcomes is $\prod_{i=0}^{n-1} 2n - 2i - 1$.

(b) Suppose that in addition to specifying who plays whom, for each pairing we say who serves first. Now in how many ways may we specify our pairs?

Activity 13. Let us now return to Problem 7 and justify—or perhaps finish—our answer to the question about the number of functions from a three-element set to a 12-element set.

(a) How can you justify your answer in Problem 7 to the question "How many functions are there from a three element set (say $[3] = \{1, 2, 3\}$) to a twelve element set (say $[12]$)?"

Solution. The number of functions from $[3]$ to $[12]$ with $f(3) = 1$ is the number of functions from $[2]$ to $[12]$, namely 144. The same is true for the number of functions with $f(3) = 2$ and $f(3) = 3$, and the size of the set S_i. The union of the sets S_i is the set of all functions from $[3]$ to $[12]$, and since there are twelve sets S_i, this union has size $12 \cdot 144 = 1728$.

(b) Based on the examples you've seen so far, make a conjecture about how many functions there are from the set

$$[m] = \{1, 2, 3, \dots, m\}$$

to $[n] = \{1, 2, 3, \dots, n\}$ and prove it.

Solution. n^m.

(c) A common notation for the set of all functions from a set M to a set N is N^M. Why is this a good notation?

Solution. Because there are n^m such functions, at least according to our conjecture.

Activity 14. Now suppose we are thinking about a set S of functions f from $[m]$ to some set X. (For example, in Problem 6 we were thinking of the set of functions from the three possible places for scoops in an ice-cream cone to 12 flavors of ice cream.) Suppose there are k_1 choices for $f(1)$. (In Problem 6, k_1 was 12, because there were 12 ways to choose the first scoop.) Suppose that for each choice of $f(1)$ there are k_2 choices for $f(2)$. (For example, in Problem 6 k_2 was 12 if the second flavor could be the same as the first, but k_2 was 11 if the flavors had to be different.) In general, suppose that for each choice of $f(1)$, $f(2)$, ... $f(i-1)$, there are k_i choices for $f(i)$. (For example, in Problem 6, if the flavors have to be different, then for each coice of $f(1)$ and $f(2)$, there are 10 choices for $f(3)$.)

What we have assumed so far about the functions in S may be summarized as

- There are k_1 choices for $f(1)$.

- For each choice of $f(1)$, $f(2)$, ..., $f(i-1)$, there are k_i choices for $f(i)$.

Solution. $\prod_{i=1}^{m} k_i$.

The point of Problem 14 is that the general product principle can be stated informally, as we did originally, or as a statement about counting sets of standard concrete mathematical objects, namely functions.

Activity 15. A roller coaster car has n rows of seats, each of which has room for two people. If n men and n women get into the car with a man and a woman in each row, in how many ways may they choose their seats?

Hint. In how many ways may you assign the men to their rows? The women? Once a woman and a man have a row to share, in how many ways may they choose their seats?

Solution. $(n!)^2 2^n$

Activity 16. How does the general product principle apply to Problem 6?

Solution. By the general product principle, there are $12 \cdot 11 \cdot 10$ triple decker cones.

Activity 17. In how many ways can we pass out k distinct pieces of fruit to n children (with no restriction on how many pieces of fruit a child may get)?

Solution. Either by the formula for the number of functions from an m-element set to an n-element set or the general product principle, there are k^n ways. (Each distribution is a function from the set of fruit to the set of children, because each piece of fruit goes to one and only one child.)

Activity 18. How many subsets does a set S with n elements have?

Hint. Try applying the product principle in the case n = 2 and n = 3. How might you apply it in general?

Solution. For each of the n elements of S, we have two options: either we put the element into the subset or we do not. Thus, the general product principle tells us that there are 2^n subsets of S.

Activity 19. Assuming $k \leq n$, in how many ways can we pass out k distinct pieces of fruit to n children if each child may get at most one? What is the number if $k > n$? Assume for both questions that we pass out all the fruit.

Hint. Ask yourself if either the sum principle or product principle applies.

Hint. Remember that zero is a number.

Solution. We are asking for the number of k-element permutations of n children, which is $\prod_{i=1}^{k} n - i + 1$, and is zero if $k > n$.

Activity 20. Another name for a list, in a specific order, of k distinct things chosen from a set S is a k-**element permutation** of S. We can also think of a k-element permutation of S as a one-to-one function (or, in other words, injection) from $[k] = \{1, 2, \ldots, k\}$ to S. How many k-element permutations does an n-element set have? (For this problem it is natural to assume $k \leq n$.

However the question makes sense even if $k > n$. What is the number of k-element permutations of an n-element set if $k > n$?

Hint.   Do you see an analogy between this problem and any of the previous problems?

Solution.   By the general product principle, the number is

$$\prod_{i=1}^{k}(n - i + 1).$$

In the case that $k > n$, there are no such lists with distinct entries, and that is what the formula gives us, because $n - (n + 1) + 1 = 0$.

There are a number of different notations for the number of k-element permutations of an n-element set. The one we shall use was introduced by Don Knuth; namely $n^{\underline{k}}$, read "n to the k falling" or "n to the k down". In Problem 20 you may have shown that

$$n^{\underline{k}} = n(n - 1) \cdots (n - k + 1) = \prod_{i=1}^{k}(n - i + 1). \qquad (1.1)$$

It is standard to call $n^{\underline{k}}$ the **k-th falling factorial power of** n, which explains why we use exponential notation. Of course we call it a *factorial* power since $n^{\underline{n}} = n(n - 1) \cdots 1$ which we call *n-factorial* and denote by $n!$. If you are unfamiliar with the Π notation, or **product notation** we introduced for products in Equation (1.1), it works just like the Σ notation works for summations.

Activity 21. Express $n^{\underline{k}}$ as a quotient of factorials.

Solution. $n^{\underline{k}} = n!/(n - k)!$

Activity 22. How should we define $n^{\underline{0}}$?

Solution.   We define $n^{\underline{0}}$ to be 1.

Functions and directed graphs

As another example how standard mathematical language relates to counting problems, Problem 7 explicitly asked you to relate the idea of counting functions to the question of Problem 6. You have probably learned in algebra or calculus how to draw graphs in the Cartesian plane of functions from a set of numbers to a set of numbers. You may recall how we can determine whether a graph is a graph of a function by examining whether each vertical straight line crosses the graph at most one time. You might also recall how we can determine whether such a function is one-to-one by examining whether each horizontal straight line crosses the graph at most one time. The functions we deal with will often involve objects which are not numbers, and will often be

functions from one finite set to another. Thus graphs in the cartesian plane will often not be available to us for visualizing functions.

However, there is another kind of graph called a *directed graph* or *digraph* that is especially useful when dealing with functions between finite sets. We take up this topic in more detail in Appendix A, particularly Section and Section . In Figure 1.2.3 we show several examples of digraphs of functions.

(a) The function given by $f(x) = x^2$ on the domain $\{1,2,3,4,5\}$.

(b) The function from the set $\{0,1,2,3,4,5,6,7\}$ to the set of triples of zeros and ones given by $f(x)$ = the binary representation of x.

(c) The function from the set $\{-2,-1,0,1,2\}$ to the set $\{0,1,2,3,4\}$ given by $f(x) = x^2$.

(d) Not the digraph of a function.
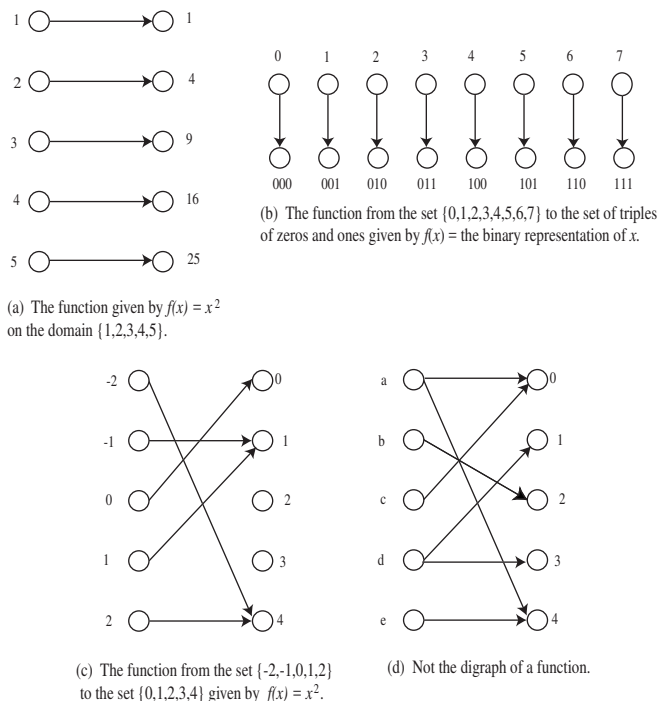
Figure 1.2.3: What is a digraph of a function?

If we have a function f from a set S to a set T, we draw a line of dots or circles, called *vertices* to represent the elements of S and another (usually parallel) line of circles or dots to represent the elements of T. We then draw an arrow from the circle for x to the circle for y if $f(x) = y$. Sometimes, as in part (e) of the figure, if we have a function from a set S to itself, we draw only one set of vertices representing the elements of S, in which case we can have arrows both entering and leaving a given vertex. As you see, the digraph can be more enlightening in this case if we experiment with the function to find a nice placement of the vertices rather than putting them in a row.

The figure needs to have part (e) added to it.

Notice that there is a simple test for whether a digraph whose vertices represent the elements of the sets S and T is the digraph of a function from S to T. There must be one and only one arrow leaving each vertex of the digraph representing an element of S. The fact that there is one arrow means that $f(x)$ is defined for each x in S. The fact that there is only one arrow

means that each x in S is related to exactly one element of T. (Note that these remarks hold as well if we have a function from S to S and draw only one set of vertices representing the elements of S.) For further discussion of functions and digraphs see Sections and Subsection of Appendix A.

Activity 23. Draw the digraph of the function from the set {Alice, Bob, Dawn, Bill} to the set {A, B, C, D, E} given by

$$f(X) = \text{the first letter of the name } X.$$

Activity 24. A function $f : S \to T$ is called an *onto function* or *surjection* if each element of T is $f(x)$ for some $x \in S$. Choose a set S and a set T so that you can draw the digraph of a function from S to T that is one-to-one but not onto, and draw the digraph of such a function.

Solution. The digraph of one such function follows
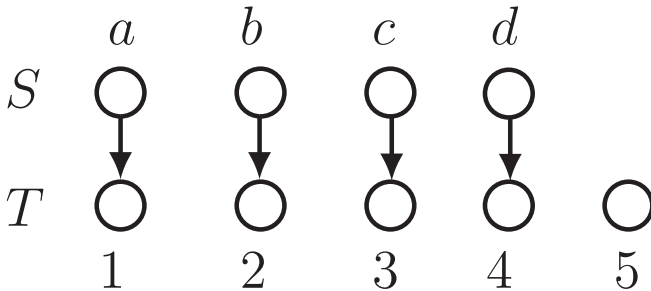
Activity 25. Choose a set S and a set T so that you can draw the digraph of a function from S to T that is onto but not one-to-one, and draw the digraph of such a function.

Solution. The digraph of one such function follows
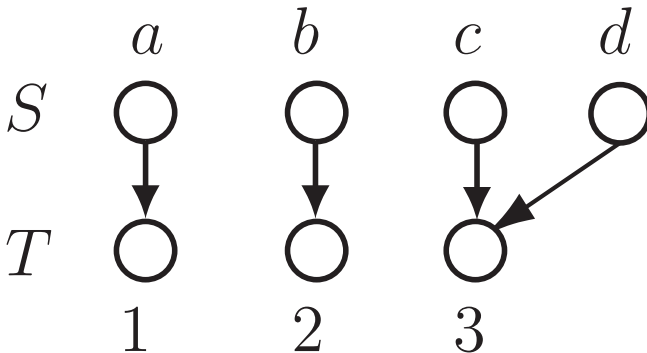
Activity 26.

Activity 27. The word *permutation* is actually used in two different ways in mathematics. A **permutation** of a set S is one-to-one from S onto S. How many permutations does an n-element set have?

Solution. $n!$.

Notice that there is a great deal of consistency between the use of the word permutation in Problem 27 and the use in Problem 20. If we have some way $a_1, a_2, \ldots, a_n$ of listing our set, then any other list $b_1, b_2, \ldots, b_n$ gives us the permutation of S whose rule is $f(a_i) = b_i$, and any permutation of S, say the one given by $g(a_i) = c_i$ gives us a list $c_1, c_2, \ldots, c_n$ of S. Thus there is really very little difference between the idea of a permutation of S and an n-element permutation of S when n is the size of S.

The bijection principle

Another name for a one-to-one and onto function is **bijection**. The digraphs marked (a), (b), and (e) in Figure 1.2.3 are digraphs of bijections. The description in Problem of the digraph of a bijection from X to Y illustrates one of the fundamental principles of combinatorial mathematics, the **bijection principle**:

> Two sets have the same size if and only if there is a bijection between them.

It is surprising how this innocent sounding principle guides us into finding insight into some otherwise very complicated proofs.

Counting subsets of a set

Activity 28. The **binary representation** of a number m is a list, or string, $a_1 a_2 \ldots a_k$ of zeros and ones such that $m = a_1 2^{k-1} + a_2 2^{k-2} + \cdots + a_k 2^0$. Describe a bijection between the binary representations of the integers between 0 and $2^n - 1$ and the subsets of an n-element set. What does this tell you about the number of subsets of an n-element set?

Hint. The problem is asking you to describe a one-to-one function from the set of binary representations of numbers between 0 and $2^n - 1$ onto the set of subsets of the set $[n]$. Write down these two sets for $n = 2$. They should both have four elements. The set of binary representations should contain the string 00. You could think of this as the instruction "take no ones and take no twos." In that context, what could you think of the string 11 as standing for? This should help you describe a function. Of course now you have to figure out how to show it is one-to-one and onto.

Solution. The sequence $a_1 a_2 \ldots a_k$ corresponds to the set of i such that $a_i = 1$. This is a bijection because each sequence gives a set, and each set can be the set of places where a sequence is 1. Since there are 2^n integers which are between 0 and $2^n - 1$, and they correspond to sequences of length n (notice, we have another bijection, the one between a number and its binary representation), there are 2^n subsets of the n-element set $[n]$.

Notice that the first question in Problem 8 asked you for the number of ways to choose a three element subset from a 12 element subset. You may have seen a notation like $\binom{n}{k}$, $C(n, k)$, or $_nC_k$ which stands for the number of

ways to choose a k-element subset from an n-element set. The number $\binom{n}{k}$ is read as "n choose k" and is called a **binomial coefficient** for reasons we will see later on. Another frequently used way to read the binomial coefficient notation is "the number of combinations of n things taken k at a time." You are going to be asked to construct two bijections that relate to these numbers and figure out what famous formula they prove. We are going to think about subsets of the n-element set $[n] = \{1, 2, 3, \ldots, n\}$. As an example, the set of two-element subsets of $[4]$ is

$$\{\{1, 2\}, \{1, 3\}, \{1, 4\}, \{2, 3\}, \{2, 4\}, \{3, 4\}\}.$$

This example tells us that $\binom{4}{2} = 6$.

Activity 29. Let C be the set of k-element subsets of $[n]$ that contain the number n, and let D be the set of k-element subsets of $[n]$ that don't contain n.

(a) Let C' be the set of $(k-1)$-element subsets of $[n-1]$. Describe a bijection from C to C'. (A verbal description is fine.)

   Solution.   Let $f(X) = X - \{n\}$, the set X with n removed. This is a bijection because two different sets containing n must yield different sets when n is removed (one-to-one), and each $(k-1)$-element subset X of $[n-1]$ may be obtained from the k-element subset $X \cup \{n\}$ of $[n]$ by removing n (onto).

(b) Let D' be the set of k-element subsets of $[n - 1] = \{1, 2, \ldots n - 1\}$. Describe a bijection from D to D'. (A verbal description is fine.)

   Solution.   Simply let $f(X) = X$. This is one-to-one by definition, and onto because the subsets of $[n - 1]$ are identical with the subsets of $[n]$ not containing n.

(c) Based on the two previous parts, express the sizes of C and D in terms of binomial coefficients involving $n - 1$ instead of n.

   Solution. $|C| = \binom{n-1}{k-1}$; $|D| = \binom{n-1}{k}$

(d) Apply the sum principle to C and D and obtain a formula that expresses $\binom{n}{k}$ in terms of two binomial coefficients involving $n - 1$. You have just derived the Pascal Equation that is the basis for the famous Pascal's Triangle.

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

Pascal's Triangle

The Pascal Equation that you derived in Problem 29 gives us the triangle in Figure 1.2.4. This figure has the number of k-element subsets of an n-element set as the kth number over in the nth row (we call the top row the zeroth row and the beginning entry of a row the zeroth number over). You'll see that your formula doesn't say anything about $\binom{n}{k}$ if $k = 0$ or $k = n$, but otherwise it says that each entry is the sum of the two that are above it and just to the left or right.

```
                    1
                 1     1
              1     2     1
           1     3     3     1
        1     4     6     4     1
     1     5    10    10     5     1
  1     6    15    20    15     6     1
1     7    21    35    35    21     7     1
```
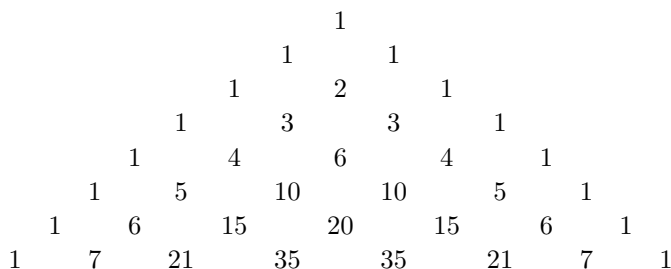
Figure 1.2.4: Pascal's Triangle

Activity 30. Just for practice, what is the next row of Pascal's triangle?

Solution.

$$1 \quad 8 \quad 28 \quad 56 \quad 70 \quad 56 \quad 28 \quad 8 \quad 1$$

Activity 31. Without writing out the rows completely, write out enough of Pascal's triangle to get a numerical answer for the first question in Problem 8.

Hint. Starting with the row 1 8 28 56 70 56 28 8 1, put dots below it where the elements of row 9 should be. Then put dots below that where the elements of row 10 should be. Do the same for rows 11 and 12. Mark the dot where row 12 should appear. Now mark the dots you need in row 11 to compute the entry in column 3 of row 12. Now mark the dots you need in row 10 to compute the marked entries in row 11. Do the same for rows 9 and 8. Now you should be able to see what you need to do.

Solution. Starting with row 9, we get

```
        1       9        36        84
     1      10       45       120
   1     11      55      165
 1    12     66     220
```

so the answer is 220.

It is less common to see Pascal's triangle as a right triangle, but it actually makes your formula easier to interpret. In Pascal's Right Triangle, the element in row n and column k (with the convention that the first row is row zero and the first column is column zero) is $\binom{n}{k}$. In this case your formula says each entry in a row is the sum of the one above and the one above and to the left, except for the leftmost and rightmost entries of a row, for which that doesn't make sense. Since the leftmost entry is $\binom{n}{0}$ and the rightmost entry is $\binom{n}{n}$, these entries are both one (to see why, ask yourself how many 0-element subsets and how many n-element subsets an n-element set has), and your formula then tells how to fill in the rest of the table.

	$k=0$	1	2	3	4	5	6	7
$n=0$	1							
1	1	1						
2	1	2	1					
3	1	3	3	1				
4	1	4	6	4	1			
5	1	5	10	10	5	1		
6	1	6	15	20	15	6	1	
7	1	7	21	35	35	21	7	1

Table 1.2.5: Pascal's Right Triangle

Seeing this right triangle leads us to ask whether there is some natural way to extend the right triangle to a rectangle. If we did have a rectangular table of binomial coefficients, counting the first row as row zero (i.e., $n = 0$) and the first column as column zero (i.e., $k = 0$), the entries we don't yet have are values of $\binom{n}{k}$ for $k > n$. But how many k-element subsets does an n-element set have if $k > n$? The answer, of course, is zero, so all the other entries we would fill in would be zero, giving us the rectangular array in Figure 1.2.6. It is straightforward to check that Pascal's equation now works for all the entries in the rectangle that have an entry above them and an entry above and to the left.

	k = 0	1	2	3	4	5	6	7
n = 0	1	0	0	0	0	0	0	0
1	1	1	0	0	0	0	0	0
2	1	2	1	0	0	0	0	0
3	1	3	3	1	0	0	0	0
4	1	4	6	4	1	0	0	0
5	1	5	10	10	5	1	0	0
6	1	6	15	20	15	6	1	0
7	1	7	21	35	35	21	7	1

Table 1.2.6: Pascal's Rectangle

Activity 32. Because our definition told us that $\binom{n}{k}$ is 0 when $k > n$, we got a rectangular table of numbers that satisfies the Pascal Equation.

(a) Is there any other way to define $\binom{n}{k}$ when $k > n$ in order to get a rectangular table that agrees with Pascal's Right Triangle for $k \leq n$ and satisfies the Pascal Equation?

Hint. Begin by trying to figure out what the entries just above the diagonal of the rectangle are. After that, what other entries can you figure out?

Solution. No, because there must be a zero above each one not in column zero. Then above each zero not in column zero or one, there must be yet another zero and so on.

(b) Suppose we want to extend Pascal's Rectangle to the left and define $\binom{n}{-k}$ for $n \geq 0$ and $k > 0$ so that $-k < 0$. What should we put into row n and column $-k$ of Pascal's Rectangle in order for the Pascal Equation to hold true?

Hint. See if you can figure out what the entries in column -1 have to be.

Solution. To the left of all the ones in column zero, we must have zeros for the Pascal Equation to hold. To the left of those zeros, we must again have zeros, and so on.

(c) What should we put into row $-n$ and column k or column $-k$ in order for the Pascal Equation to continue to hold? Do we have any freedom of choice?

Hint. What does the sum of two consecutive values in row -1 have to be? Could this sum depend on which two consecutive values we take?

Is there some value of row -1 that we could choose arbitrarily? Now what about row -2? Can we make arbitrary choices there? If so, how many can we make, and is their position arbitrary?

Solution.   Above row zero, we have some freedom. The $-1, -1$ and the $-1, 0$ entry must add to one, so they can be $-x$ and $x + 1$ for any number x. To the right of the $-1, 0$ entry they must alternate between $(-x-1)$ and $x+1$ while to the left of the $-1, 1$ entry they must alternate between $-x$ and x, ending with the $-x$ in position $-1, 1$. Now if we know the entry in row -2 and column 0, we can use the Pascal equation (in the form $\binom{n-1}{k-1} = \binom{n}{k} - \binom{n-1}{k}$) to compute all the entries to the left of it, and (in a different form) to compute all the entries to the right of it. Thus we may be arbitrary about the entries in column 0 (or, in fact, one entry in each row) and then the Pascal Equation tells us how to fill in the rest of each row. We shall see later on that there is one very natural choice for how to fill in all the rows above row zero.

Activity 33. There is yet another bijection that lets us prove that a set of size n has 2^n subsets. Namely, for each subset S of $[n] = \{1, 2, \ldots, n\}$, define a function (traditionally denoted by χ_S) as follows.[1]

$$\chi_S(i) = \begin{cases} 1 & \text{if } i \in S \\ 0 & \text{if } i \notin S \end{cases}$$

The function χ_S is called the **characteristic function** of S. Notice that the characteristic function is a function from $[n]$ to $\{0, 1\}$.

(a) For practice, consider the function $\chi_{\{1,3\}}$ for the subset $\{1, 3\}$ of the set $\{1, 2, 3, 4\}$. What are

    (i) $\chi_{\{1,3\}}(1)$?
    (ii) $\chi_{\{1,3\}}(2)$?
    (iii) $\chi_{\{1,3\}}(3)$?
    (iv) $\chi_{\{1,3\}}(4)$?

(b) We define a function f from the set of subsets of $[n] = \{1, 2, \ldots, n\}$ to the set of functions from $[n]$ to $\{0, 1\}$ by $f(S) = \chi_S$. Explain why f is a bijection.

Solution.   If $i \in S$ but $i \notin T$, then $\chi_S(i) = 1$ but $\chi_T(i) = 0$. Thus if $S \neq T$, the $\chi_S \neq \chi_T$. Therefore f is one-to-one. Given a function g from $[n]$ to $\{0, 1\}$, let $S = \{i | g(i) = 1\}$. Then by definition, $g = \chi_S = f(S)$. Therefore f is onto, so it is a bijection.

---

[1]The symbol χ is the Greek letter chi that is pronounced Ki, with the i sounding like "eye."

(c) Why does the fact that f is a bijection prove that $[n]$ has 2^n subsets?

Solution.   We have seen that there are 2^n functions from $[n]$ to a two-element set, and we have a bijection between the set of all such functions and the subsets of $[n]$.

In Problems 18, Activity 28, and Activity 33 you gave three proofs of the following theorem.

Theorem 1.2.7. *The number of subsets of an n-element set is 2^n.*

The proofs in Problem 28 and Activity 33 use essentially the same bijection, but they interpret sequences of zeros and ones differently, and so end up being different proofs. We will give yet another proof, using bijections similar to those we used in proving the Pascal Equation, at the beginning of Chapter 2.

The quotient principle

Activity 34.

Activity 35. Based on what you observed in Problem , how many k-element subsets does an n-element set have?

Solution.   Following the reasoning of Problem 34, there are $n^{\underline{k}}$ k-element permutations of an n-element set, and $k!$ of these permutations list the same set of k elements, so the number of k-element sets is $\frac{n^{\underline{k}}}{k!} = \frac{n!}{k!(n-k)!}$.

Activity 36. The formula you proved in Problem 35 is symmetric in k and $n-k$; that is, it gives the same number for $\binom{n}{k}$ as it gives for $\binom{n}{n-k}$. Whenever two quantities are counted by the same formula it is good for our insight to find a bijection that demonstrates the two sets being counted have the same size. In fact this is a guiding principle of research in combinatorial mathematics. Find a bijection that proves that $\binom{n}{k}$ equals $\binom{n}{n-k}$.

Hint.   The first thing you need to decide is "What are the two sets whose elements we are counting?" Then it will be easier to think of a bijection between these two sets. It turns out that these two sets are sets of sets!

Solution.   For each k-element subset K of the n-element set N, define $f(K)$ to be the set of all elements of N *not* in K. Then f is the desired bijection.

Activity 37. In how many ways can we pass out k (identical) ping-pong balls to n children if each child may get at most one?

Hint.   Ask yourself "What is a problem like this doing in the middle of a bunch of problems about counting subsets of a set? Is it related, or is it supposed to gives us a break from sets?"

Solution. $\binom{n}{k}$, because we choose the k children to whom we give ping-pong balls.

Activity 38. In how many ways may n people sit around a round table? (Assume that when people are sitting around a round table, all that really matters is who is to each person's right. For example, if we can get one arrangement of people around the table from another by having everyone get up and move to the right one place and sit back down, we get an equivalent arrangement of people. Notice that you can get a list from a seating arrangement by marking a place at the table, and then listing the people at the table, starting at that place and moving around to the right.) There are at least two different ways of doing this problem. Try to find them both.

Hint. The problem suggests that you think about how to get a list from a seating arrangement. Could every list of n distinct people come from a seating chart? How many lists of n distinct people are there? How many lists could we get from a given seating chart by taking different starting places?

Hint. Second Hint for Problem 38: For a different way of doing the problem, suppose that you have chosen one person, say the first one in a list of the people in alphabetical order by name. Now seat that person. Does it matter where they sit? In ways can you seat the remaining people? Does it matter where the second person in alphabetical order sits?

Solution. The total number of ways to list how the n people sit around the table is $n!$. However, two lists are the same if we get one from the other by shifting everyone right the same number of places. This divides the set of lists up into sets of n mutually equivalent lists. The number m of such sets is the number of seating arrangements. However by the product principle, $mn = n!$, because we have partitioned up the set of $n!$ lists into m sets of size n. Therefore $m = (n-1)!$ A second solution may be obtained by choosing one of the n people and letting this person sit anywhere. Since all that matters is who is to the right of each person, it doesn't matter where this person sits. Once this person is seated, let everybody else sit down. If they sit down first in one order clockwise around the table and then in some other order, the person to the right of somebody has changed. Thus there are $(n-1)!$ ways (the number of ways to seat everybody else) to seat the people around the table.

We are now going to analyze the result of Problem 35 in more detail in order to tease out another counting principle that we can use in a wide variety of situations.

abc	*acb*	*bac*	*bca*	*cab*	*cab*
abd	*adb*	*bad*	*bda*	*dab*	*dba*
abe	*aeb*	*bae*	*bea*	*eab*	*eba*
acd	*adc*	*cad*	*cda*	*dac*	*dca*
ace	*aec*	*cae*	*cea*	*eac*	*eca*
ade	*aed*	*dae*	*dea*	*ead*	*eda*
bcd	*bdc*	*cbd*	*cdb*	*dbc*	*dcb*
bce	*bec*	*cbe*	*ceb*	*ebc*	*ecb*
bde	*bed*	*dbe*	*deb*	*ebd*	*edb*
cde	*ced*	*dce*	*dec*	*ecd*	*edc*

Table 1.2.8: The 3-element permutations of $\{a, b, c, d, e\}$ organized by which 3-element set they permute.

In Table reftab:permsof3 we list all three-element permutations of the 5-element set $\{a, b, c, d, e\}$. Each row consists of all 3-element permutations of some subset of $\{a, b, c, d, e\}$. Because a given k-element subset can be listed as a k-element permutation in $k!$ ways, there are $3! = 6$ permutations in each row. Because each 3-element permutation appears exactly once in the table, each row is a block of a partition of the set of 3-element permutations of $\{a, b, c, d, e\}$. Each block has size six. Each block consists of all 3-element permutations of some three-element subset of $\{a, b, c, d, e\}$. Since there are ten rows, we see that there are ten 3-element subsets of $\{a, b, c, d, e\}$. An alternate way to see this is to observe that we partitioned the set of all 60 three-element permutations of $\{a, b, c, d, e\}$ into some number q of blocks, each of size six. Thus by the product principle, $q \cdot 6 = 60$, so $q = 10$.

Activity 39.

Activity 40.

Activity 41. In Problem 38, describe a way to partition the n-element permutations of the n people into blocks so that there is a bijection between the set of blocks of the partition and the set of arrangements of the n people around a round table. What method of solution for Problem 38 does this correspond to?

Solution. Put two permutations in the same block if we can get one from the other by moving everyone (circularly) some number r places to the right. This corresponds to the method that gives $n!/n$ as the answer. Many students should be able to answer this question by saying "See the answer to Problem 38."

Activity 42. In Problems and Activity 41, you have been using the product principle in a new way. One of the ways in which we previously stated the product principle was "If we partition a set into m blocks each of size n, then the set has size $m \cdot n$." In problems and Activity 41 we knew the size p of a set P of permutations of a set, and we knew we had partitioned P into some

unknown number of blocks, each of a certain known size r. If we let q stand for the number of blocks, what does the product principle tell us about p, q, and r? What do we get when we solve for q?

Solution. $p = qr$, so that $q = p/r$.

   The formula you found in the Problem 42 is so useful that we are going to single it out as another principle. The *quotient principle* says:

   If we partition a set P into q blocks, each of size r, then $q = p/r$.

The quotient principle is really just a restatement of the product principle, but thinking about it as a principle in its own right often leads us to find solutions to problems. Notice that it does not always give us a formula for the number of blocks of a partition; it only works when all the blocks have the same size. In Chapter 6, we develop a way to solve problems with different block sizes in cases where there is a good deal of symmetry in the problem. (The roundness of the table was a symmetry in the problem of people at a table; the fact that we can order the sets in any order is the symmetry in the problem of counting k-element subsets.)

   In Section of Appendix A we introduce the idea of an equivalence relation, see what equivalence relations have to do with partitions, and discuss the quotient principle from that point of view. While that appendix is not required for what we are doing here, if you want a more thorough discussion of the quotient principle, this would be a good time to work through that appendix.

Activity 43. In how many ways may we string n distinct beads on a necklace without a clasp? (Perhaps we make the necklace by stringing the beads on a strong, and then carefully gluing the two ends of the string together so that the joint can't be seen. Assume someone can pick up the necklace, move it around in space and put it back down, giving an apparently different way of stringing the beads that is equivalent to the first.)

Hint.   How could we get a list of beads from a necklace?

Hint.   When we cut the necklace and string it out on a table, there are 2n lists of beads we could get. Why is it $2n$ rather than n?

Solution.   We can obtain a permutation of the beads by cutting the necklace and stretching it out in a straight line. We can partition the permutations according to which necklace they come from in this process. Two permutations are in the same block if we get one either by circularly permuting the other or by reversing the other (this corresponds to flipping the necklace over in space). Thus each necklace corresponds to $2n$ permutations so by the quotient principle we have $n!/2n = (n-1)!/2$ ways to string n distinct beads on a necklace.

Activity 44.

Activity 45. (This becomes especially relevant in Chapter 6, though it makes an important point here.) In how many ways may we attach two identical red

beads and two identical blue beads to the corners of a square (with one bead per corner) free to move around in (three-dimensional) space?

Hint. It might be helpful to just draw some pictures of the possible configurations. There aren't that many.

Solution. Two ways; either the red beads are side-by-side or diagonally opposite. If we think about partitioning lists of 2 Rs and 2 Bs so that two are in the same block if we get one from the other by moving the square, we get two blocks, $\{RRBB, BRRB, BBRR, RBBR\}$ and $\{RBRB, BRBR\}$.

Activity 46. While the formula you proved in Problems 35 and Task   is very useful, it doesn't give us a sense of how big the binomial coefficients are. We can get a very rough idea, for example, of the size of $\binom{2n}{n}$ by recognizing that we can write $(2n)^{\underline{n}}/n!$ as $\frac{2n}{n} \cdot \frac{2n-1}{n-1} \cdots \frac{n+1}{1}$, and each quotient is at least 2, so the product is at least 2^n. If this were an accurate estimate, it would mean the fraction of n-element subsets of a $2n$-element set would be about $2^n/2^{2n} = 1/2^n$, which becomes very small as n becomes large. However it is pretty clear the approximation will not be a very good one, because some of the terms in that product are much larger than 2. In fact, if $\binom{2n}{k}$ were the same for every k, then each would be the fraction $\frac{1}{2n+1}$ of 2^{2n}. This is much larger than the fraction $\frac{1}{2^n}$. But our intuition suggests that $\binom{2n}{n}$ is much larger than $\binom{2n}{1}$ and is likely larger than $\binom{2n}{n-1}$ so we can be sure our approximation is a bad one. For estimates like this, James Stirling developed a formula to approximate $n!$ when n is large, namely $n!$ is about $\left(\sqrt{2\pi n}\right) n^n/e^n$. In fact the ratio of $n!$ to this expression approaches 1 as n becomes infinite.[2] We write this as
$$n! \sim \sqrt{2\pi n}\frac{n^n}{e^n}.$$

We read this notation as $n!$ is asymptotic to $\sqrt{2\pi n}\frac{n^n}{e^n}$. Use Stirling's formula to show that the fraction of subsets of size n in an $2n$-element set is approximately $1/\sqrt{\pi n}$. This is a much bigger fraction than $\frac{1}{2^n}$!

1.3   Some Applications of the Basic Principles

Lattice paths and Catalan Numbers

Activity 47. In a part of a city, all streets run either north-south or east-west, and there are no dead ends. Suppose we are standing on a street corner. In how many ways may we walk to a corner that is four blocks north and six blocks east, using as few blocks as possible?

Hint. Note that we must walk at least ten blocks, so ten is the smallest number of blocks possible. In how many of those ten blocks must we walk north?

---

[2] Proving this takes more of a detour than is advisable here; however there is an elementary proof which you can work through in the problems of the end of Section 1 of Chapter 1 of *Introductory Combinatorics* by Kenneth P. Bogart, Harcourt Academic Press, (2000).

Solution.   The shortest possible walk is going to be ten blocks. To plan a walk, we must choose which four of those ten blocks go north; the other six blocks we will have to go east. There are $\binom{10}{4}$ ways to make this selection.

Activity 48. Problem 47 has a geometric interpretation in a coordinate plane. A *lattice path* in the plane is a "curve" made up of line segments that either go from a point (i, j) to the point $(i + 1, j)$ or from a point (i, j) to the point $(i, j + 1)$, where i and j are integers. (Thus lattice paths always move either up or to the right.) The length of the path is the number of such line segments.

(a) What is the length of a lattice path from $(0, 0)$ to (m, n)?

  Solution.   The length of a lattice path from $(0, 0)$ to (m, n) is $m + n$.

(b) How many such lattice paths of that length are there?

  Hint.   In Problem 47 you saw that we had to make ten choices of north or east, choosing north four times.

  Solution.   The number of such paths is $\binom{m+n}{n}$.

(c) How many lattice paths are there from (i, j) to (m, n), assuming i, j, m, and n are integers?

  Hint.   This problem is actually a bit tricky. What happens to the answer if $i > m$ or $j > n$? Remember that paths go up or to the right.

  Solution.   Since lattice paths move up and to the right, there are no paths from (i, j) to (m, n) unless $i \leq m$ and $j \leq n$. In that case, the number of paths is $\binom{m+n-i-j}{n-j}$ which is the same as $\binom{m+n-i-j}{m-i}$.

Activity 49. Another kind of geometric path in the plane is a *diagonal lattice path*. Such a path is a path made up of line segments that go from a point (i, j) to $(i + 1, j + 1)$ (this is often called an *upstep*) or $(i + 1, j - 1)$ (this is often called a *downstep*), again where i and j are integers. (Thus diagonal lattice paths always move towards the right but may move up or down.)

(a) Describe which points are connected to $(0, 0)$ by diagonal lattice paths.

  Hint.   Where can you go from (0,0) in one step? In two steps? In any of these cases, what can you say about the sum of the coordinates of a point you can get to? Can you find any other relationship between the x- and y-coordinates of a point you can get to? For example, can you get to the point $(1, 3)$? The points (m, n) connected to $(0, 0)$ by diagonal lattice paths will have $m + n$ is even, because each upstep adds two to the sum of i and j while each downstep does not change the sum.

Further, since we go one step to the right each time we go up or down, we cannot get above the line $y = x$ or below the line $y = -x$. However for any point (m, n) with m and n nonnegative integers such that $m + n$ even and $-m \leq n \leq m$, we can get to (m, n) by making $\frac{m-n}{2}$ downsteps and $\frac{m+n}{2}$ upsteps.

(b) What is the length of a diagonal lattice path from $(0, 0)$ to (m, n)?

Solution. From the previous part of this problem, we will make a total of $\frac{m-n}{2} + \frac{m+n}{2} = m$ steps, and our total motion parallel to the y axis will be $\frac{m+n}{2} - \frac{m-n}{2} = n$. The length of such a path is $m\sqrt{2}$; we might informally just call it m steps.

(c) Assuming that (m, n) is such a point, how many diagonal lattice paths are there from $(0, 0)$ to (m, n)?

Hint. How many choices do you have to make in order to choose a path?

Solution. The number of possible paths is the number of ways we can choose which of the m steps are upsteps (or equivalently downsteps) this number is $\binom{m}{\frac{m+n}{2}}$.

Activity 50. A school play requires a ten dollar donation per person; the donation goes into the student activity fund. Assume that each person who comes to the play pays with a ten dollar bill or a twenty dollar bill. The teacher who is collecting the money forgot to get change before the event. If there are always at least as many people who have paid with a ten as a twenty as they arrive the teacher won't have to give anyone an IOU for change. Suppose $2n$ people come to the play, and exactly half of them pay with ten dollar bills.

(a) Describe a bijection between the set of sequences of tens and twenties people give the teacher and the set of lattice paths from $(0, 0)$ to (n, n).

Solution. For each ten dollar bill take a rightstep and for each twenty dollar bill take an upstep (where rightstep and upstep have the hopefully natural meaning). The assumption that there are an equal number of ten and twenty dollar bills means that the path will end up at (n, n). Each sequence of tens and twenties gives a lattice path and each lattice path corresponds to such a sequence, so we have a bijection.

(b) Describe a bijection between the set of sequences of tens and twenties that people give the teacher and the set of diagonal lattice paths from $(0, 0)$ and $(0, 2n)$.

Solution.   For each ten dollar bill take an upstep and for each twenty dollar bill take a downstep. Each sequence of tens and twenties will give us a diagonal lattice path from $(0,0)$, and each diagonal lattice path from $(0,0)$ to $(0,2n)$ will give us a sequence of tens and twenties with an equal number of tens and twenties, so we have a bijection.

(c) In each case, what is the geometric interpretation of a sequence that does not require the teacher to give any IOUs?

Hint.   In each part, each such sequence corresponds to a path that can't cross over (but may touch) a certain line.

Solution.   In the first case a sequence that does not require the teacher to give any IOUs will correspond to a lattice path that stays on or below the line $y = x$, and in the second case such a sequence will correspond to a diagonal lattice path that stays on or above the x-axis.

Activity 51. Notice that a lattice path from $(0,0)$ to (n,n) stays inside (or on the edges of) the square whose sides are the x-axis, the y-axis, the line $x = n$ and the line $y = n$. In this problem we will compute the number of lattice paths from $(0,0)$ to (n,n) that stay inside (or on the edges of) the triangle whose sides are the x-axis, the line $x = n$ and the line $y = x$. For example, in Figure 1.3.1 we show the grid of points with integer coordinates for the triangle whose sides are the x-axis, the line $x = 4$ and the line $y = x$.
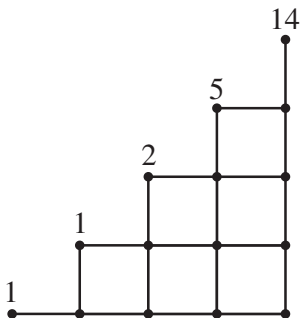
Figure 1.3.1: The lattice paths from $(0,0)$ to (i,i) for $i = 0,1,2,3,4$. The number of paths to the point (i,i) is shown just above that point.

(a) Explain why the number of lattice paths from $(0,0)$ to (n,n) that go outside the triangle is the number of lattice paths from $(0,0)$ to (n,n) that either touch or cross the line $y = x + 1$.

Solution.   If a lattice path between $(0,0)$ and (n,n) goes outside the triangle, it can only do so on an upstep. (A step from (i,j) to $(i,j+1)$.)

And an upstep must originate at a point with integer coordinates. If $j < i$ an upstep from (i, j)) cannot leave the triangle. Thus to leave the triangle, the upstep must leave from a point of the form (i, i), and go to $(i, i + 1)$, which is on the line $y = x + 1$.

(b) Find a bijection between lattice paths from $(0, 0)$ to (n, n) that touch (or cross) the line $y = x + 1$ and lattice paths from $(-1, 1)$ to (n, n).

Hint. Given a path from $(0, 0)$ to (n, n) which touches or crosses the line $y = x + 1$, how can you modify the part of the path from $(0, 0)$ to the first touch of $y = x + 1$ so that the modified path starts instead at $(-1, 1)$? The trick is to do this in a systematic way that will give you your bijection.

Solution. Suppose we have a lattice path form $(0, 0)$ to (n, n) which touches or crosses the line $y = x + 1$. Let $(k, k + 1)$ be the first point on the line $y = x + 1$ that the lattice path touches. From that point, work backwards, replacing every upstep with a step one unit to the left and every rightstep with a step one unit down. The segment of the path you just changed will have moved left $k + 1$ times, so its leftmost x coordinate will be -1, and it will have moved down k times, so its lowest y coordinate will be 1. Thus we now have a lattice path from $(-1, 1)$ to (n, n). Further, given a lattice path from $(-1, 1)$ to (n, n), it must cross the line $y = x + 1$ at least once, because it starts above the line and ends below it. At the first point where such a path touches the line $y = x + 1$, say $(k', k' + 1)$, work backwards replacing every upstep with a step to the left and every rightstep with a step downwards. The leftmost point on this path will have x coordinate 0, and the lowest point will have y coordinate 0, so the new path will be a lattice path from $(0, 0)$ to (n, n) that touches the line $y = x + 1$. Clearly these two processes reverse each other, and so they give us a bijection between paths form $(0, 0)$ to (n, n) that touch the line $y = x + 1$ and lattice lattice paths from $(-1, 1)$ to (n, n). Notice that geometrically what we are doing to get the bijection is to take the portion of a lattice path that goes from the initial point till the first touch of the line $y = x + 1$ and reflecting it around that line. This idea of reflection was introduced by Feller, and is called Feller's reflection principle.

(c) Find a formula for the number of lattice paths from $(0, 0)$ to (n, n) that do not cross the line $y = x$. The number of such paths is called a *Catalan Number* and is usually denoted by C_n.

Hint. A path either touches the line $y = x + 1$ or it doesn't. This partitions the set of paths into two blocks.

Solution. $C_n = \binom{2n}{n} - \binom{2n}{n+1} = \frac{1}{n+1}\binom{2n}{n}$.

Activity 52. Your formula for the Catalan Number can be expressed as a binomial coefficient divided by an integer. Whenever we have a formula that calls for division by an integer, an ideal combinatorial explanation of the formula is one that uses the quotient principle. The purpose of this problem is to find such an explanation using diagonal lattice paths.[3] A diagonal lattice path that never goes below the y-coordinate of its first point is called a **Dyck Path**. We will call a Dyck Path from $(0,0)$ to $(2n,0)$ a **Catalan Path** of length $2n$. Thus the number of Catalan Paths of length $2n$ is the Catalan Number C_n.

(a) If a Dyck Path has n steps (each an upstep or downstep), why do the first k steps form a Dyck Path for each nonnegative $k \leq n$?

  Solution.   If no points on the path are lower than the first point, then no points among the first k steps are lower than the first point.

(b) Thought of as a curve in the plane, a diagonal lattice path can have many local maxima and minima, and can have several absolute maxima and minima, that is, several highest points and several lowest points. What is the y-coordinate of an absolute minimum point of a Dyck Path starting at $(0,0)$? Explain why a Dyck Path whose rightmost absolute minimum point is its last point is a Catalan Path.

  Hint.   Look back at the definition of a Dyck Path and a Catalan Path.

  Solution.   Since the path starts at $(0,0)$ and can't go below it, the y coordinate of an absolute minimum must be zero. If the last point is an absolute minimum, then (because it ends with the same y coordinate with which it starts) the path has an even number $2k$ of steps and ends at $(2k,0)$.

(c) Let D be the set of all diagonal lattice paths from $(0,0)$ to $(2n,0)$. (Thus these paths can go below the x-axis.) Suppose we partition D by letting B_i be the set of lattice paths in D that have i upsteps (perhaps mixed with some downsteps) following the last absolute minimum. How many blocks does this partition have? Give a succinct description of the block B_0.

  Hint.   What makes this part difficult is understanding how we are partitioning the paths. As an example, B_0 is the set of all paths that have no upsteps following the last absolute minimum. Can such a path have downsteps after the last absolute minimum? (The description we gave of B_0 is not succinct enough to be the answer to the second question

───────────

[3]The result we will derive is called the Chung-Feller Theorem; this approach is based of a paper of Wen-jin Woan "Uniform Partitions of Lattice Paths and Chung-Feller Generalizations," *American Mathematics Monthly 58* June/July 2001, p556.

of this part.) As another example B_1 is the set of all paths that have exactly one upstep and perhaps some downsteps after the last absolute minimum. Is it possible, though, for a path in B_1 to have any downsteps after the last absolute minimum? A path in B_2 has exactly two upsteps after its last absolute minimum. If is possible to have one downstep after the last absolute minimum, but it has to be in a special place. What place is that? Now to figure out how many parts our partition has, we need to know the maximum number of upsteps a path can have following its last absolute minimum. What is this maximum? It might help to draw some pictures with $n = 5$ or 6. In particular, is it possible that all upsteps occur after the last absolute minimum?

Solution. The path must have n upsteps total, and so can have any number between 0 and n upsteps after the rightmost absolute minimum. Thus the partition has $n + 1$ blocks. Block B_0 consists of the Catalan Paths.

(d) How many upsteps are in a Catalan Path?

Solution. n.

(e) We are going to give a bijection between the set of Catalan Paths and the block B_i for each i between 1 and n. For now, suppose the value of i, while unknown, is fixed. We take a Catalan path and break it into three pieces. The piece F (for "front") consists of all steps before the ith upstep in the Catalan path. The piece U (for "up") consists of the ith upstep. The piece B (for "back") is the portion of the path that follows the ith upstep. Thus we can think of the path as FUB. Show that the function that takes FUB to BUF is a bijection from the set of Catalan Paths onto the block B_i of the partition. (Notice that BUF can go below the x axis.)

Hint. Using d for down and u for up, we could have $uudduuddudud$ as our Catalan path. Suppose that $i = 5$. The fifth upstep is the u in position 9. Thus $F = uudduudd$, $U = u$, and $B = dud$. Now BUF is $duduuudduudd$. This is a Dyck path that begins by going below the x-axis. The d's in positions 1 and 3 take the path to the y-coordinate -1. Then the y coordinate climbs to 2, goes back to 0, up to 2 again, and finally down to 0. So the absolute minimum is -1, and it occurs in the first and third position. There are five u's after the third positon. So this Dyck path is in the block B_5 of our partition. Now comes the crucial question. Why were there five u's after that last absolute minimum in position 3? Try with the same path and $i = 3$. Figure out why there are three u's after the last absolute minimum in the resulting path. All this discussion should explain why when $i = 5$, the set of all Catalan

paths is mapped into the set B_5. Keeping $i = 5$ for a while, try to see why this correspondence between Catalan paths and B_5 is a bijection. Then, if you need to, do the same thing with $i = 3$. This should give you enough insight to do the general case.

Solution. Since we are starting with a Catalan path, the point on the path at the beginning of the ith upstep must have y coordinate greater or equal to than zero. Thus wherever we start the sequence F of upsteps and downsteps, a path constructed by this sequence never goes lower than its starting point. Thus in BUF the last absolute minimum is either right before the U or earlier. But B is the final segment of a Catalan Path, so its final point is at least as low as its starting point. Thus the point at the beginning of the U in BUF is an absolute minimum, and there are i upsteps after that local minimum. If we take two different sequences and rearrange them in the same way, we get two different sequences, so the function we just described is a one-to-one function. If we take an arbitrary diagonal lattice path from $(0,0)$ to $(2n, 0)$, let U' be the first upstep after the last absolute minimum, F' be the portion of the path that follows U', and B' be the portion that precedes U', then $F'U'B'$ is a Catalan Path, and U' is its ith upstep if and only if in $B'U'F'$ there are i upsteps after the last absolute minimum. Thus the mapping from FUB to BUF is a bijection.

(f) Explain how you have just given another proof of the formula for the Catalan Numbers.

Solution. We have taken the set of all $\binom{2n}{n}$ diagonal lattice paths of length $2n$ from $(0,0)$ to $(2n,0)$ and partitioned it into $n + 1$ blocks all of size C_n. Thus by the quotient principle, $C_n = \frac{1}{n+1}\binom{2n}{n}$.

The Binomial Theorem

Activity 53. We know that $(x+y)^2 = x^2 + 2xy + y^2$. Multiply both sides by $(x+y)$ to get a formula for $(x+y)^3$ and repeat to get a formula for $(x+y)^4$. Do you see a pattern? If so, what is it? If not, repeat the process to get a formula for $(x+y)^5$ and look back at Figure 1.2.4 to see the pattern. Conjecture a formula for $(x+y)^n$.

Solution. $(x+y)^3 = x^3 + 2x^2y + xy^2 + x^2y + +2xy^2 + y^3 = x^3 + 3x^2y + +3xy^2 + y^3$.

Similarly, $(x+4)^4 = x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4$,

and $(x+y)^5 = x^5 + 5x^4y + 10x^3y^2 + 10x^2y^3 + 5xy^4 + y^5$. The pattern is that the coefficient of x^iy^j is $\binom{i+j}{i}$ which is the same as $\binom{i+j}{j}$. Said differently, the coefficient of $x^{n-i}y^i$ is $\binom{n}{i}$ or the coefficient of x^iy^{n-i} is $\binom{n}{i}$. We conjecture that

$$(x+y)^n = \sum_{i=0}^{n} \binom{n}{i} x^{n-i}y^i.$$

(The reason for putting $x^{n-i}y^i$ into the sum is so that as i goes from 0 to n, the powers of x decrease from n to 0.)

Activity 54. When we apply the distributive law n times to $(x+y)^n$, we get a sum of terms of the form $x^i y^{n-i}$ for various values of the integer i.

(a) If it is clear to you that each term of the form $x^i y^{n-i}$ that we get comes from choosing an x from i of the $(x+y)$ factors and a y from the remaining $n-i$ of the factors and multiplying these choices together, then answer this part of the problem and skip the next part. Otherwise, do the next part instead of this one. In how many ways can we choose an x from i terms and a y from $n-i$ terms?

Solution. The number of ways to choose an x from i of the factors and a y from the remaining ones is the way to choose the i factors from the n factors; that is, $\binom{n}{i}$.

(b) Expand the product $(x_1 + y_1)(x_2 + y_2)(x_3 + y_3)$.

Solution.

$$(x_1 + y_1)(x_2 + y_2)(x_3 + y_3) = x_1 x_2 x_3 + \quad x_1 x_2 y_3 + x_1 y_2 x_3 +$$
$$y_1 x_2 x_3 + x_1 y_2 y_3 + y_1 x_2 y_3 + \quad y_1 y_2 x_3 + y_1 y_2 y_3.$$

(c) What do you get when you substitute x for each x_i and y for each y_i?

Solution. When you substitute x for each x_i and y for each y_i, you get $(x+y)^3 = x^3 + 3x^2 y + 3xy^2 + y^3$.

(d) Now imagine expanding

$$(x_1 + y_1)(x_2 + y_2) \cdots (x_n + y_n).$$

Once you apply the commutative law to the individual terms you get, you will have a sum of terms of the form

$$x_{k_1} x_{k_2} \cdots x_{k_i} \cdot y_{j_1} y_{j_2} \cdots y_{j_{n-i}}.$$

What is the set $\{k_1, k_2, \ldots, k_i\} \cup \{j_1, j_2, \ldots, j_{n-i}\}$?

Solution. $\{k_1, k_2, \ldots, k_i\} \cup \{j_1, j_2, \ldots, j_{n-i}\} = \{1, 2, \ldots, n\}$.

(e) In how many ways can you choose the set $\{k_1, k_2, \ldots, k_i\}$?

Solution. You can choose the set $\{k_1, k_2, \ldots k_i\}$ in $\binom{n}{i}$ ways.

(f) Once you have chosen this set, how many choices do you have for
$\{j_1, j_2, \ldots, j_{n-i}\}$?

Solution.   Once you have chosen the set of ks, there is just one way
to choose the set of js.

(g) If you substitute x for each x_i and y for each y_i, how many terms of the
form $x^i y^{n-i}$ will you have in the expanded product

$$(x_1 + y_1)(x_2 + y_2) \cdots (x_n + y_n) = (x + y)^n?$$

Solution.   If you substitute x for x_i and substitute y for y_i, you will
get $\binom{n}{i}$ terms of the form $x^i y^{n-i}$.

(h) How many terms of the form $x^{n-i} y^i$ will you have?

Solution.   You will also get $\binom{n}{i}$ terms of the form $x^{n-i} y^i$.

(i) Explain how you have just proved your conjecture from Problem 53.
The theorem you have proved is called the *Binomial Theorem*.

Solution.   We have proved that the coefficient of $x^i y^{n-i}$ in $(x + y)^n$ is
$\binom{n}{i}$, or equivalently that the coefficient of $x^{n-i} y^i$ in $(x + y)^n$ is $\binom{n}{i}$.

Activity 55. What is $\sum_{i=1}^{n} \binom{10}{i} 3^i$?

Hint.   What would the lower limit of the sum have to be for this problem
to be a routine application of the binomial theorem?

Solution. $\sum_{i=1}^{n} \binom{10}{i} 3^i = \sum_{i=0}^{n} \binom{10}{i} 3^i - \binom{10}{0} 3^0 = (1 + 3)^{10} - 1 = 4^{10} - 1$

Activity 56. What is $\binom{n}{0} - \binom{n}{1} + \binom{n}{2} - \cdots \pm \binom{n}{n}$ if n is an integer bigger
than 0?

Hint.   What does the binomial theorem give you for $(x - y)n$?

Solution.   The sum is 0 because it is $(-1 + 1)^n$.

Activity 57. Explain why

$$\sum_{i=0}^{m} \binom{m}{i} \binom{n}{k-i} = \binom{m+n}{k}.$$

Find two different explanations.

Hint.   Consider $(x + y)m(x + y)n$.

Hint.   What does $\binom{m+n}{k}$ count? What does $\binom{m}{i} \binom{m}{k-1}$ count?

Solution. When we expand both sides of $(x+y)^m(x+y)^n = (x+y)^{m+n}$ by the binomial theorem we get $\sum_{i=0}^{m} \binom{m}{i}\binom{n}{k-i}$ as the coefficient of $x^{m+n-k}y^k$ on the left hand side and $\binom{m+n}{k}$ on the right hand side.

For a second explanation, to choose k elements out of the union of an m-element set and a disjoint n-element set, chose some number $i \le m$ of them from the m-element set and the remaining $k - i$ of them from the n-element set. The sum on the left hand side of the equation simply sums the number of such choices over all possible i, and the binomial coefficient on the right hand side of the equation says we will end up choosing k elements from among our $m + n$ elements.

Activity 58. From the symmetry of the binomial coefficients, it is not too hard to see that when n is an odd number, the number of subsets of $\{1, 2, \ldots, n\}$ of odd size equals the number of subsets of $\{1, 2, \ldots, n\}$ of even size. Is it true that when n is even the number of subsets of $\{1, 2, \ldots, n\}$ of even size equals the number of subsets of odd size? Why or why not?

Hint. For example when $n = 3$, we have $\binom{3}{0} = \binom{3}{3}$ and $\binom{3}{1} = \binom{3}{2}$. The number of subsets of even size is $\binom{3}{0} + \binom{3}{2}$ and the number of subsets of odd size is $\binom{3}{1} + \binom{3}{3}$, and the two sums can be paired off into equal terms. When we subtract the number of subsets of odd size from the number of subsets of even size, the pairing also gives us $\binom{3}{0} - \binom{3}{1} + \binom{3}{2} - \binom{3}{3} = 0$.

Solution. It is true, because if $n > 0$, when you expand $(1 - 1)^n$ by the binomial theorem, you get an alternating sum of binomial coefficients equal to 0, and so the sum of the binomial coefficients $\binom{n}{i}$ with i even must equal the sum of the binomial coefficients $\binom{n}{i}$ with i odd.

Activity 59. What is $\sum_{i=0}^{n} i\binom{n}{i}$? (Hint: think about how you might use calculus.)

Hint. Take the derivative of something interesting.

Solution. $\sum_{i=0}^{n}(\binom{n}{i})x^i = (1 + x)^n$. Taking derivatives of both sides gives us $\sum_{i=0}^{n} i\binom{n}{i}x^{i-1} = n(1 + x)^{n-1}$.. Now substitute 1 for x and you get $\sum_{i=0}^{n} i\binom{n}{i} = n2^{n-1}$.

Notice how the proof you gave of the binomial theorem was a counting argument. It is interesting that an apparently algebraic theorem that tells us how to expand a power of a binomial is proved by an argument that amounts to counting the individual terms of the expansion. Part of the reason that combinatorial mathematics turns out to be so useful is that counting arguments often underlie important results of algebra. As the algebra becomes more sophisticated, so do the families of objects we have to count, but nonetheless we can develop a great deal of algebra on the basis of counting.

The pigeonhole principle

Activity 60. American coins are all marked with the year in which they were made. How many coins do you need to have in your hand to guarantee that

on two (at least) of them, the date has the same last digit? (When we say "to guarantee that on two (at least) of them,…" we mean that you can find two with the same last digit. You might be able to find three with that last digit, or you might be able to find one pair with the last digit 1 and one pair with the last digit 9, or any combination of equal last digits, as long as there is at least one pair with the same last digit.)

Solution.   Since there are ten possible last digits, you need at least 11 coins, and with 11 coins, at least two last digits must be the same.

There are many ways in which you might explain your answer to Activity 60. For example, you can partition the coins according to the last digit of their date; that is, you put all the coins with a given last digit in a block together, and put no other coins in that block; repeating until all coins are in some block. Then you have a partition of your set of coins. If no two coins have the same last digit, then each block has exactly one coin. Since there are only ten digits, there are at most ten blocks and so by the sum principle there are at most ten coins. In fact with ten coins it is possible to have no two with the same last digit, but with 11 coins some block must have at least two coins in order for the sum of the sizes of at most ten blocks to be 11. This is one explanation of why we need 11 coins in Activity 60. This kind of situation arises often in combinatorial situations, and so rather than always using the sum principle to explain our reasoning , we enunciate another principle which we can think of as yet another variant of the sum principle. The **pigeonhole principle** states that

> If we partition a set with more than n elements into n parts, then at least one part has more than one element.

The pigeonhole principle gets its name from the idea of a grid of little boxes that might be used, for example, to sort mail, or as mailboxes for a group of people in an office. The boxes in such grids are sometimes called pigeonholes in analogy with stacks of boxes used to house homing pigeons when homing pigeons were used to carry messages. People will sometimes state the principle in a more colorful way as "if we put more than n pigeons into n pigeonholes, then some pigeonhole has more than one pigeon."

Activity 61. Show that if we have a function from a set of size n to a set of size less than n, then f is not one-to-one.

Hint.   To prove that each function from a set S of size n to a set of size less than n is not one-to-one, we must prove that regardless of the function f that we choose, there are always two elements, say x and y, such that $f(x) = f(y)$.

Solution.   Let T be the set of size less than n, and S be the set of size n. Let $B_j = \{i | f(i) = j\}$ for each j in T. Then the nonempty sets among the B_js form a partition of S and the number of blocks is less than the size of S. Therefore by the pigeonhole principle, there is at least one block with at least two elements, so there are two elements i_1 and i_2 such that $f(i_1) = f(i_2)$.

Activity 62. Show that if S and T are finite sets of the same size, then a function f from S to T is one-to-one if and only if it is onto.

Hint. The previous exercise could help you prove that if f is one-to-one, then it is onto.

Hint. The sum principle can help you show that if f is an onto function, then f is one-to-one.

Solution. First suppose that f is a one-to-one function from S to T, sets which have the same size. Let $B_j = \{i|f(i) = j\}$ for each j in T.. If f is not onto, then the number of nonempty sets B_j is smaller than the number of elements of T and thus is smaller than the size of S. The nonempty sets B_j are a partition of S. But then by the pigeonhole principle, some nonempty B_j has two or more elements, contradicting the assumption that f is one-to-one. Therefore if f is one-to-one, then it is onto. Now suppose that f is an onto function from S to T, sets of the same size. Again let $B_j = \{i|f(i) = j\}$ for each j in T. The size of the union of the sets B_j is, by the sum principle, the sum of their sizes. Since f is onto, each B_j has at least one element. Since the number of sets B_j is the number of elements of S, if one of those sets has more than one element, the the size of their union is more than the size of S, which is a contradiction since they are subsets of S. Therefore each set B_j has exactly one element and therefore f is one-to-one.

Activity 63. There is a **generalized pigeonhole principle** which says that if we partition a set with more than kn elements into n blocks, then at least one block has at least $k + 1$ elements. Prove the generalized pigeonhole principle.

Hint. The statement of the generalized pigeonhole principle involves the number of elements in a block, so a counting principle is likely to help you.

Solution. Suppose we partition a set S of more than kn elements into n blocks. If each block has at most k elements, the by the sum principle the size of S is at most kn. But this is a contradiction, so some block has at least $k + 1$ elements.

Activity 64. All the powers of five end in a five, and all the powers of two are even. Show that for for some integer n, if you take the first n powers of a prime other than two or five, one must have "01" as the last two digits.

Hint. You may choose a specific number for n if you want to. Notice that the last two digits of the powers of a prime other than two cannot represent an even number.

Solution. If we take 40 powers of such a prime, either one will end in "01" or some two, say p^i and p^j with $i > j$ must have the same last two digits by the pigeon hole principle. Then $p^i - p^j = 100k$ for some integer k. Thus $p^j(p^{i-j} - 1)$ must be a multiple of 100, and since neither 2 nor 5 divide p, $p^{i-j} - 1 = 100k'$ for some integer k' then $p^{i-j} = 100k' + 1$, so the last two digits of p^{i-j} must be "01."

Activity 65. Show that in a set of six people, there is a set of at least three people who all know each other, or a set of at least three people none of whom know each other. (We assume that if person 1 knows person 2, then person 2 knows person 1.)

Hint. While this sounds like a pigeonhole principle problem, the ordinary pigeonhole principle doesn't guarantee three of something.

Solution.   By the generalized pigeonhole principle, person 1 either knows at least three people or doesn't know at least three people. Suppose person 1 knows three people. Then either two of these people know each other, giving us, with person 1, three mutual acquaintances, or no two of these people know each other, giving us three mutual strangers. On the other hand if there are three people person 1 does not know, then either two of these people don't know each other, giving us, with person 1, three mutual strangers, or all three of these people know each other, giving us three mutual acquaintances.

Activity 66. Draw five circles labeled Al, Sue, Don, Pam, and Jo. Find a way to draw red and green lines between people so that every pair of people is joined by a line and there is neither a triangle consisting entirely of red lines or a triangle consisting of green lines. What does Problem 65 tell you about the possibility of doing this with six people's names? What does this problem say about the conclusion of Problem 65 holding when there are five people in our set rather than six?

Solution.   In the figure that follows, we use solid lines for red and dashed lines for green. Clearly there is no solid triangle and no dashed triangle.
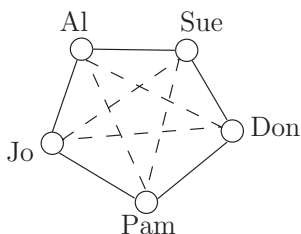
Activity 65 says you can't do this with six people's names. This problem says that the conclusion of Activity 65 does not hold when you have five people.

Ramsey Numbers

Activity 65–66 together show that six is the smallest number R with the property that if we have R people in a room, then there is either a set of (at least) three mutual acquaintances or a set of (at least) three mutual strangers. Another way to say the same thing is to say that six is the smallest number so that no matter how we connect 6 points in the plane (no three on a line) with red and green lines, we can find either a red triangle or a green triangle. There

is a name for this property. The **Ramsey Number** $R(m, n)$ is the smallest number R so that if we have R people in a room, then there is a set of at least m mutual acquaintances or at least n mutual strangers. There is also a geometric description of Ramsey Numbers; it uses the idea of a *complete graph* on R vertices. A *complete graph* on R vertices consists of R points in the plane together with line segments (or curves) connecting each two of the R vertices.[4] The points are called **vertices** and the line segments are called **edges**. In Figure 1.3.2 we show three different ways to draw a complete graph on four vertices. We use K_n to stand for a complete graph on n vertices.

Figure 1.3.2: Three ways to draw a complete graph on four vertices

Our geometric description of $R(3, 3)$ may be translated into the language of graph theory (which is the subject that includes complete graphs) by saying $R(3, 3)$ is the smallest number R so that if we color the edges of a K_R with two colors, then we can find in our picture a K_3 all of whose edges have the same color. The graph theory description of $R(m, n)$ is that $R(m, n)$ is the smallest number R so that if we color the edges of a K_R with red and green, then we can find in our picture either a K_m all of whose edges are red or a K_n all of whose edges are green. Because we could have said our colors in the opposite order, we may conclude that $R(m, n) = R(n, m)$. In particular $R(n, n)$ is the smallest number R such that if we color the edges of a K_R with two colors, then our picture contains a K_n all of whose edges have the same color.

Activity 67. Since $R(3, 3) = 6$, an uneducated guess might be that $R(4, 4) = 8$. Show that this is not the case.

Hint. What usually makes it hard for students to start this problem is the fact that we just defined what $R(4, 4)$ is, and not what it means for a number not to be $R(4, 4)$. So to get started, try to write down what it means to say $R(4, 4)$ is not 8. You will see that there are two things that can keep $R(4, 4)$ from being 8. You need to figure out which one happens and explain why. One such explanation could involve the graph K_8 .

Solution. In the graph of Figure 1.3.3, each vertex has three dashed lines emanating from it, and there are no dashed lines connecting any of the three vertices adjacent to it by dashed lines. Each vertex has four solid lines emanating from it, and no three of the four vertices adjacent to it by solid lines are all adjacent by solid lines. Thus there is no solid line K_4 and there is no dashed line K_4.

---

[4]As you may have guessed, a complete graph is a special case of something called a graph. The word graph will be defined in Subsection .
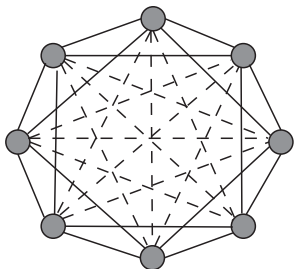
Figure 1.3.3: A graph showing $R(4, 4) > 8$

Activity 68. Show that among ten people, there are either four mutual acquaintances or three mutual strangers. What does this say about $R(4, 3)$?

Hint.   Review Activity 65 and your solution of it.

Solution.   Take a person, say person 1. If person has six acquaintances, then by Activity 65 among them there are either three mutual strangers, in which case we are done, or three mutual acquaintances. These three acquaintances together with person 1 form a set of 4 mutual acquaintances in which case we are again done. Thus we may assume Person 1 has at most 5 acquaintances, and so has four non-acquaintances. Now either all four of these people are acquainted, in which case we are done, or else two of them are not acquainted. Then these two people, together with person 1 make three mutual nonacquaintances. Therefore in every possible case, we have either four mutual acquaintances or three mutual strangers. This means that $R(4, 3) \leq 10$.

Activity 69. Show that among an odd number of people there is at least one person who is an acquaintance of an even number of people and therefore also a stranger to an even number of people.

Hint.   Let a_i be the number of acquaintances of person i. Can you explain why the sum of the numbers a_i is even?

Solution.   Suppose we add, for each person, the number of people with whom he or she is acquainted. Then we get twice the number of acquaintance edges in the graph of acquaintance and non-acquaintance relationships. Thus the sum must be even. But if each person among an odd number of people were acquainted with an odd number of people, then the sum would be odd. Since this is a contradiction, among an odd number of people, there must be at least one who is acquainted with an even number of people. Since the number of people different from this person is even, the number of people with whom this person is not acquainted is also even.

Activity 70. Find a way to color the edges of a K_8 with red and green so that there is no red K_4 and no green K_3.

Hint.   Often when there is a counter-example, there is one with a good deal of symmetry. (Caution: there is a difference between often and always!) One way to help yourself get a symmetric example, if there is one, is to put 8

vertices into a circle. Then, perhaps, you might draw green edges in some sort of regular fashion until it is impossible to draw another green edge between any two of the vertices without creating a green triangle.

Solution. In the graph shown in Figure 1.3.3, there is no K_3 whose edges are dashed, and no K_4 whose edges are solid. By symmetry, to verify this you need only look at vertex 1 and vertices connected to it by either dashed lines or by solid lines.

Activity 71. Find $R(4, 3)$.

Hint. In Activity 68 you showed that $R(4, 3) \leq 10$. In Activity 70 you showed that $R(4, 3) > 8$. Thus $R(4, 3)$ is either 9 or 10. Deciding which is the case is just plain hard. But there is a relevant problem we have done that we haven't used yet.

Solution. $R(4, 3) = 9$. In Activity 70 we showed that $R(4, 3)$ is more than 8. So we must show that if we have nine people, we either have 4 mutual acquaintances or three mutual strangers. By Activity 69 there is at least one person (say person A) who is acquainted with an even number of people. If person A is acquainted with six or more people, then among these six people, there are either three mutual acquaintances or three mutual strangers. If there are three mutual strangers, we are done; if there are three mutual acquaintances, they, together with Person A are four mutual acquaintances. Thus we may assume Person A is acquainted with at most four people. Thus person A is a stranger to at least four people. If two of these people are strangers, then they, together with person A form three mutual strangers and we are done. Otherwise all of these people know each other and we have at least four mutual acquaintances, and so in every possible situation, we have either four mutual acquaintances or three mutual strangers.

As of this writing, relatively few Ramsey Numbers are known. $R(3, n)$ is known for $n < 10$, $R(4, 4) = 18$, and $R(5, 4) = R(4, 5) = 25$.

1.4   Supplementary Chapter Problems

1. (interesting) Remember that we can write n as a sum of n ones. How many plus signs do we use? In how many ways may we write n as a sum of a list of k positive numbers? Such a list is called a *composition* of n into k parts. We use $n - 1$ plus signs. Write down such a sum and choose $k - 1$ of the plus signs. Then each string of ones and plusses between two chosen plus signs, before the first chosen plus sign or after the last chosen one corresponds to a part of a composition of n. Thus the number of compositions of n with k parts is the number of ways to choose the $k - 1$ places, which is $\binom{n-1}{k-1}$.

2. In Problem 1 we defined a composition of n into k parts. What is the total number of compositions of n (into any number of parts). The total

number of compositions is the number of ways to choose a subset of the plus signs which is 2^{n-1}.

3. (essential for this or the next section) Write down a list of all 16 0-1 sequences of length four starting with 0000 in such a way that each entry differs from the precious one by changing just one digit. This is called a Grey Code. That is, a *Grey Code* for 0-1 sequences of length n is a list of the sequences so that each entry differs from the previous one in exactly one place. Can you describe how to get a Grey Code for 0-1 sequences of length five from the one you found for sequences of length 4? Can you describe how to prove that there is a Grey code for sequences of length n? (One of many) 0000, 0001, 0011, 0010, 0110, 0111, 0101, 0100, 1100, 1101, 1111, 1110, 1010, 1011, 1001, 1000. To get a code for sequences of length 5, put a zero at the end of each of the sequences we have. Follow that revised sequence by 10001, and write the remainder of the sequence in reverse order with a 1 at the end of each term. (Don't reverse the individual length four sequences, just the sequence of sequences!) We just, in essence, described the inductive step of an inductive proof that Grey Codes exist for sequences of any length.

4. (interesting) Use the idea of a Grey Code from Problem 3 to prove bijectively that the number of even-sized subsets of an n-element set equals the number of odd-sized subsets of an n-element set. Each sequence in the Grey Code is the characteristic function of a set, and the number of elements of the set is the number of ones in the sequence. Since each sequence differs in just one place from the preceding one, the sequences alternate between having an even number of ones and an odd number of ones. Since the first sequence is all zeros and there are 2^n sequences, the last one has an odd number of zeros. Thus the map that takes each sequence except the last to the next one, and takes the last to the first is a bijection between the characteristic functions of sets with an even number of elements and sets with an odd number of elements.

5. (interesting) A list of parentheses is said to be balanced if there are the same number of left parentheses as right, and as we count from left to right we always find at least as many left parentheses as right parentheses. For example, $(((()()))())$ is balanced and $((())$ and $(()()))(()$ are not. How many balanced lists of n left and n right parentheses are there? The number is the Catalan number: we get a bijection between balanced lists of parentheses and Catalan paths by sending each left paren to an upstep and each right paren to a downstep. The condition that there are always as many left parentheses as right ensures we never go below the x axis.

6. (difficult) Suppose we plan to put six distinct computers in a network as shown in Figure 1.4.1. The lines show which computers can communicate directly with which others. Consider two ways of assigning computers to the nodes of the network different if there are two computers

that communicate directly in one assignment and that don't communicate directly in the other. In how many different ways can we assign computers to the network?
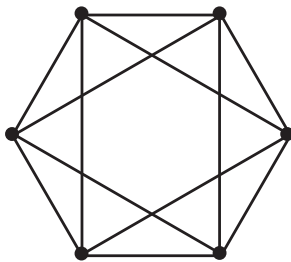
Figure 1.4.1: A computer network.

We consider two assignments of computers to be equivalent if in both assignments, each computer communicates directly with exactly the same computers. This partitions the set of all 6! computer assignments into blocks of 48 computers each. Thus we have $720/48 = 15$ ways to assign the computers to the network.

7. (interesting) In a circular ice cream dish we are going to put four distinct scoops of ice cream chosen from among twelve flavors. Assuming we place four scoops of the same size as if they were at the corners of a square, and recognizing that moving the dish doesn't change the way in which we have put the ice cream into the dish, in how many ways may we choose the ice cream and put it into the dish? Each ice cream arrangement is equivalent to three others, the ones we get by rotating the dish. This divides the arrangements of four flavors of ice cream into blocks of size 4. Thus we may arrange the ice cream we have chosen in the dish in $4!/4 = 6$ ways. We may choose the ice cream in $\binom{12}{4} = 495$ ways, and so we may choose it and put it into the dish in 2970 ways.

8. (interesting) In as many ways as you can, show that $\binom{n}{k}\binom{n-k}{m} = \binom{n}{m}\binom{n-m}{k}$. You can prove this by plugging in the formula for $\binom{n}{k}$ on both sides and cancelling stuff until you get the same thing on both sides. However a much more interesting proof is that the right hand side counts the number of ways to choose a k-element set form an n-element set and then choose an m-element set from what remains. The left hand side counts the number of ways to first chose a k-element subset from the n-element set and then choose an m-element subset from what remains. Thus in both cases you are counting the number of ways to choose an ordered pair consisting of an m-element subset and a disjoint k-element subset from an n-element set. You can also base a proof on the observation that $(x + y + x)^n = \sum_{k=0}^{n} \binom{n}{k}(x + y)^k z^{n-k}$ and $(x + y + z)^n = \sum_{m=0}^{n} \binom{n}{m} x^m (y + z)^{n-m}$ and asking for the coefficient

of $x^m y^{n-m-k} z^k$. You do have to use the binomial theorem with an eye to the result you are looking for, however.

9. (interesting) A tennis club has $4n$ members. To specify a doubles match, we choose two teams of two people. In how many ways may we arrange the members into doubles matches so that each player is in one doubles match? In how many ways may we do it if we specify in addition who serves first on each team? We now have many methods for solving this problem. Perhaps the easiest is to list all $(4n)$ people and take them in groups of four for doubles matches, with the first two in a group of four as one team and the second two as another team. We note that interchanging the n blocks of 4 does not change the matches, nor does interchanging the two people on a team nor interchanging the two teams. Thus we have $(4n)!/n!2^{3n}$ ways to arrange the matches. If we are to say who serves first on each team, we might as well say it is the first of the two listed, so now we have $(4n)!/n!2^n$ ways to arrange the matches.

10. A town has n streetlights running along the north side of main street. The poles on which they are mounted need to be painted so that they do not rust. In how many ways may they be painted with red, white, blue, and green if an even number of them are to be painted green? We can think of first choosing the set of even size of poles to be painted green, and the painting the remaining poles red, white, and blue. We may do this in $\sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} 3^{n-2k}$ ways.

11. We have n identical ping-pong balls. In how many ways may we paint them red, white, blue, and green? We can line up the identical ping-pong balls and break them into four groups, those of each color, by inserting dividers. If we want to paint at least one in each color, we can choose three of the spaces between the balls in which to insert dividers, so we can paint them in $\binom{n-1}{k}$. But the problem didn't require us to use each color, so we can put two dividers adjacent to each other. Thus there are $n+1$ places where we can put the first divider (putting it before all the balls means we use no red, and putting it after all of them means we use no green. Now there are $n+2$ places where we can put the second divider, including before or after the first, and $n+3$ places where we can put the third divider. However if we interchange two dividers we still paint the balls before the first divider red, those between then next two white, and so on. Thus $3! = 6$ of these arrangements of balls and dividers correspond to the same paint job, so the number of ways to paint the balls is $\frac{(n+1)(n+2)(n+3)}{6} = \binom{n+3}{3}$. This suggests that another way to think of the problem is to consider $n+3$ slots in a row, and fill n of them with balls and 3 of them with dividers; since the balls are identical and the dividers might as well be identical, the number of ways to do this is the number of ways to choose the slots that get dividers.

12. We have n identical ping-pong balls. In how many ways may we paint them red, white, blue, and green if we use green paint on an even number

of them? We first decide how many balls to paint green, then paint the remainder with the other three colors as in Problem 11 This gives us

$$\sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n - 2k + 2}{2}$$

ways to paint the balls.

Chapter 2

Applications of Induction and Recursion in Combinatorics and Graph Theory

2.1 Some Examples of Mathematical Induction

In Chapter 1 ($\langle\langle$Unresolved xref, reference "SubsetsByInduction"; check spelling or use "provisional" attribute$\rangle\rangle$), we used the principle of mathematical induction to prove that a set of size n has 2^n subsets. If you were unable to do that problem and you haven't yet read Appendix B (a portion of which is repeated here), you should do so now.

Mathematical induction

The *principle of mathematical induction* states that

> In order to prove a statement about an integer n, if we can
>
> 1. Prove the statement when $n = b$, for some fixed integer b
> 2. Show that the truth of the statement for $n = k - 1$ implies the truth of the statement for $n = k$ whenever $k > b$,
>
> then we can conclude the statement is true for all integers $n \geq b$.

As an example, let us return to $\langle\langle$Unresolved xref, reference "SubsetsByInduction"; check spelling or use "provisional" attribute$\rangle\rangle$. The statement we wish to prove is the statement that "A set of size n has 2^n subsets."

> Our statement is true when $n = 0$, because a set of size 0 is the empty set and the empty set has $1 = 2^0$ subsets. (This step of our

proof is called a *base step*.) Now suppose that $k > 0$ and every set with $k - 1$ elements has 2^{k-1} subsets. Suppose $S = \{a_1, a_2, \ldots a_k\}$ is a set with k elements. We partition the subsets of S into two blocks. Block B_1 consists of the subsets that do not contain a_n and block B_2 consists of the subsets that do contain a_n. Each set in B_1 is a subset of $\{a_1, a_2, \ldots a_{k-1}\}$, and each subset of $\{a_1, a_2, \ldots a_{k-1}\}$ is in B_1. Thus B_1 is the set of all subsets of $\{a_1, a_2, \ldots a_{k-1}\}$. Therefore by our assumption in the first sentence of this paragraph, the size of B_1 is 2^{k-1}. Consider the function from B_2 to B_1 which takes a subset of S including a_k and removes a_k from it. This function is defined on B_2, because every set in B_2 contains a_k. This function is onto, because if T is a set in B_1, then $T \cup \{a_k\}$ is a set in B_2 which the function sends to T. This function is one-to-one because if V and W are two different sets in B_2, then removing a_k from them gives two different sets in B_1. Thus we have a bijection between B_1 and B_2, so B_1 and B_2 have the same size. Therefore by the sum principle the size of $B_1 \cup B_2$ is $2^{k-1} + 2^{k-1} = 2^k$. Therefore S has 2^k subsets. This shows that if a set of size $k - 1$ has 2^{k-1} subsets, then a set of size k has 2^k subsets. Therefore by the principle of mathematical induction, a set of size n has 2^n subsets for every nonnegative integer n.

The first sentence of the last paragraph is called the *inductive hypothesis*. In an inductive proof we always make an inductive hypothesis as part of proving that the truth of our statement when $n = k - 1$ implies the truth of our statement when $n = k$. The last paragraph itself is called the *inductive step* of our proof. In an inductive step we derive the statement for $n = k$ from the statement for $n = k - 1$, thus proving that the truth of our statement when $n = k - 1$ implies the truth of our statement when $n = k$. The last sentence in the last paragraph is called the *inductive conclusion*. All inductive proofs should have a base step, an inductive hypothesis, an inductive step, and an inductive conclusion.

There are a couple details worth noticing. First, in this problem, our base step was the case $n = 0$, or in other words, we had $b = 0$. However, in other proofs, b could be any integer, positive, negative, or 0. Second, our proof that the truth of our statement for $n = k - 1$ implies the truth of our statement for $n = k$ required that k be at least 1, so that there would be an element a_k we could take away in order to describe our bijection. However, condition (2) of the principle of mathematical induction only requires that we be able to prove the implication for $k > 0$, so we were allowed to assume $k > 0$.

Strong Mathematical Induction

One way of looking at the principle of mathematical induction is that it tells us that if we know the "first" case of a theorem and we can derive each other case of the theorem from a smaller case, then the theorem is true in all cases. However the particular way in which we stated the theorem is rather restrictive

in that it requires us to derive each case from the immediately preceding case. This restriction is not necessary, and removing it leads us to a more general statement of the principal of mathematical induction which people often call the *strong principle of mathematical induction.* It states:

> In order to prove a statement about an integer n if we can
>
> 1. prove our statement when $n = b$ and
> 2. prove that the statements we get with $n = b$, $n = b + 1$, … $n = k - 1$ imply the statement with $n = k$,
>
> then our statement is true for all integers $n \geq b$.

You will find some explicit examples of the use of the strong principle of mathematical induction in Appendix B and will find some uses for it in this chapter.

Binomial Coefficients and the Binomial Theorem

Activity 72. When we studied the Pascal Equation and subsets in Chapter 1, it may have appeared that there is no connection between the Pascal relation $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$ and the formula $\binom{n}{k} = \frac{n!}{k!(n-k)!}$. Of course you probably realize you can prove the Pascal relation by substituting the values the formula gives you into the right-hand side of the equation and simplifying to give you the left hand side. In fact, from the Pascal Relation and the facts that $\binom{n}{0} = 1$ and $\binom{n}{n} = 1$, you can actually prove the formula for $\binom{n}{k}$ by induction on n. Do so.

Solution. We wish to prove that $\binom{n}{i} = \frac{n!}{i!(n-i)!}$. We note that since $\binom{n}{0} = 1$ and $\binom{n}{n} = 1$ are in agreement with this formula, we only have to consider the cases in which $0 < i < n$, which by the way, requires that $n \geq 2$. We will prove that our formula holds by induction on n for $n \geq 2$. If $n = 2$, the only i we need to consider is $i = 1$, and we know that $\binom{2}{1} = 2$, the number of one-element subsets of a two-element set. But $\frac{2!}{1!(2-1)!}$ is 2 also, so our formula holds when $n = 2$. Now suppose our formula holds when $n = k - 1$, so that for every i with $0 < i < n - 1$, $\binom{k-1}{i} = \frac{(k-1)!}{i!(k-1-i)!}$. Then by the Pascal Equation

$$\binom{k}{i} = \binom{k-1}{i-1} + \binom{k-1}{i}$$
$$= \frac{(k-1)!}{(i-1)!(k-1-i+1)!} + \frac{(k-1)!}{i!(k-1-i)!}$$
$$= \frac{(k-1)!i + (k-1)!(k-i)}{i!(k-i)!} = \frac{k!}{i!(k-i)!}.$$

Thus the truth of our formula for $n = k - 1$ implies its truth for $n = k$. Therefore by the principle of mathematical induction, our formula is true for all integers $n \geq 2$. We have already seen is is true when $i = 0$ or $i = 1$, so it is true for all nonnegative n and all numbers i with $0 \leq i \leq n$.

Activity 73.

(a) Use the fact that $(x+y)^n = (x+y)(x+y)^{n-1}$ to give an inductive proof of the binomial theorem.

Solution. We prove the binomial theorem by induction on n. When $n = 0$, $(x + y)^n = (x + y)^0 = 1 = \sum_{i=0}^{0} \binom{n}{i} x^{0-i} y^i$ since that last summation consists of the one term $\binom{0}{0} x^0 y^0$.

Now suppose that when $n = k - 1$, $(x + y)^n = \sum_{i=0}^{n} \binom{n}{i} x^{n-i} y^i$. This gives us

$$
\begin{aligned}
(x + y)^k &= (x + y)(x + y)^{k-1} = (x + y) \sum_{i=0}^{k-1} \binom{k-1}{i} x^{k-1-i} y^i \\
&= \sum_{i=0}^{k-1} \binom{k-1}{i} x^{k-i} y^i + \sum_{i=0}^{k-1} \binom{k-1}{i} x^{k-1-i} y^{i+1} \\
&= \sum_{i=0}^{k-1} \binom{k-1}{i} x^{k-i} y^i + \sum_{i=1}^{k} \binom{k-1}{i-1} x^{k-i} y^i \\
&= x^k + \left(\sum_{i=1}^{k-1} \binom{k-1}{i} x^{k-i} y^i + \binom{k-1}{i-1} x^{k-i} y^i \right) + y^k \\
&= x^k + \left(\sum_{i=1}^{k-1} \binom{k}{i} x^{k-i} y^i \right) + y^k = \sum_{i=0}^{k} \binom{k}{i} x^{k-i} y^i.
\end{aligned}
$$

Thus the truth of the binomial theorem for $n = k-1$ implies its truth for $n = k$. Then by the principle of mathematical induction, the binomial theorem must be true for all integers $n \geq 0$.

(b) Suppose that f is a function defined on the nonnegative integers such that $f(0) = 3$ and $f(n) = 2f(n-1)$. Find a formula for $f(n)$ and prove your formula is correct.

Solution. $f(n) = 3 \cdot 2^n$. We prove our formula is correct by induction. When $n = 0$ our formula gives $f(0) = 3$, which is what we were given. Now suppose that when $n = k - 1$, $f(n) = 3 \cdot 2^n$. Then $f(k) = 2 \cdot f(k - 1) = 2 \cdot 3 \cdot 2^{k-1} = 3 \cdot 2^k$. Therefore the truth of our formula when $n = k - 1$ implies its truth when $n = k$ and so by the principle of mathematical induction, $f(n) = 3 \cdot 2^n$ for all nonnegative integers n.

Inductive definition

You may have seen $n!$ described by the two equations $0! = 1$ and $n! = n(n-1)!$ for $n > 0$. By the principle of mathematical induction we know that this pair of equations defines $n!$ for all nonnegative numbers n. For this reason we call

such a definition an *inductive definition*. An inductive definition is sometimes called a *recursive definition*. Often we can get very easy proofs of useful facts by using inductive definitions.

Activity 74. An inductive definition of a^n for nonnegative n is given by $a^0 = 1$ and $a^n = aa^{n-1}$. (Notice the similarity to the inductive definition of $n!$.) We remarked above that inductive definitions often give us easy proofs of useful facts. Here we apply this inductive definition to prove two useful facts about exponents that you have been using almost since you learned the meaning of exponents.

(a) Use this definition to prove the rule of exponents $a^{m+n} = a^m a^n$ for nonnegative m and n.

    Solution. We use induction on n to prove this. When $n = 0$, the formula gives us $a^{m+0} = a^m a^0 = a^m \cdot 1 = a^m$, so the rule of exponents holds when $n = 0$. Now assume it holds when $n = k - 1$ so that $a^{m+k-1} = a^m a^{k-1}$. Then, starting and ending with our inductive definition, we may write

$$a^{m+n} = aa^{m+n-1} = aa^m a^{k-1} = a^m \cdot a \cdot a^{k-1} = a^m a^k.$$

    Thus the truth of our law for $n = k - 1$ implies its truth for $n = k$. Therefore, by the principle of mathematical induction, $a^{m+n} = a^m a^n$ for all nonnegative integers n.

(b) Use this definition to prove the rule of exponents $a^{mn} = (a^m)^n$.

    Solution. We will use induction on n and part (a) of this problem to prove that $a^{mn} = (a^m)^n$. First, when $n = 0$ the left and right hand sides of the equation are both 1, so $a^{mn} = (a^m)^n$ holds when $n = 0$. Now assume that $a^{m(k-1)} = (a^m)^{k-1}$. This may be rewritten as $a^{mk-m} = (a^m)^{k-1}$ Multiply both sides by a^m and apply part (a) of the problem and then the inductive definition (with a^m replacing a) to get

$$a^{mk-m}a^m = (a^m)^{k-1}a^m$$
$$a^{mk} = (a^m)^{k-1}a^m$$
$$a^{mk} = (a^m)^k.$$

    Thus the truth of our formula when $n = k-1$ implies its truth when $n = k$. Therefore by the principle of mathematical induction, the formula is true for all nonnegative integers n.

Activity 75. Suppose that f is a function on the nonnegative integers such that $f(0) = 0$ and $f(n) = n + f(n-1)$. Prove that $f(n) = n(n+1)/2$. Notice that this gives a third proof that $1 + 2 + \cdots + n = n(n+1)/2$, because this

sum satisfies the two conditions for f. (The sum has no terms and is thus 0 when $n = 0$.)

Solution. We prove the formula for f by induction on n. If $n = 0$, then $n(n+1)/2 = 0$ which is what we were given. Now assume that $f(k-1) = (k-1)k/2$. Then $f(k) = k + f(k-1) = k + (k-1)k/2 = (k^2 + 2k - k)/2 = k(k+1)/2$. Therefore the truth of the formula for $n = k - 1$ implies its truth for $n = k$, and thus by the principle of mathematical induction, the formula for f holds for all nonnegative integers n.

Activity 76. Give an inductive definition of the summation notation $\sum_{i=1}^{n} a_i$. Use it and the distributive law $b(a + c) = ba + bc$ to prove the distributive law

$$b \sum_{i=1}^{n} a_i = \sum_{i=1}^{n} ba_i.$$

Solution. We define $\sum_{i=1}^{1} a_i = a_1$ and for $n > 1$, $\sum_{i=1}^{n} a_i = \left(\sum_{i=1}^{n-1} a_i \right) + a_n$. When $n = 1$, $b \sum_{i=1}^{1} a_i = ba_1$ by the base step of our inductive definition. Assume that $k > 1$ and $b \sum_{i=1}^{k-1} a_i = \sum_{i=1}^{k-1} ba_i$. Now we can write

$$b \sum_{i=1}^{k} a_i = b \left[\left(\sum_{i=1}^{k-1} a_i \right) + a_k \right] = \left(b \sum_{i=1}^{k-1} a_i \right) + ba_k = \left(\sum_{i=1}^{k-1} ba_i \right) + ba_k = \sum_{i=1}^{k} ba_i,$$

where the last step is justified by the inductive step of our inductive definition with a_i replaced by ba_i. Thus the truth of our statement for $k - 1$ implies its truth for $i = k$, and therefore by the principle of mathematical induction, for all positive integers n, $b \sum_{i=1}^{n} a_i = \sum_{i=1}^{n} ba_i$.

Proving the general product principle (Optional)

We stated the sum principle as

> If we have a partition of a set S, then the size of S is the sum of the sizes of the blocks of the partition.

In fact, the simplest form of the sum principle says that the size of the sum of two disjoint (finite) sets is the sum of their sizes.

Activity 77. Prove the sum principle we stated for partitions of a set from the simplest form of the sum principle.

Solution. We prove by induction on n that the size of the union of n disjoint sets is the sum of their sizes. We assume that the size of the union of two disjoint sets is the sum of their sizes. Now assume $k > 2$ and the size of the union of $k - 1$ disjoint sets is the sum of their sizes. Then we may write

$$|\cup_{i=1}^{k} S_i| = |\left(\cup_{i=1}^{k-1} S_i \right) \cup S_k| = \left(\sum_{i=1}^{k-1} |S_i| \right) + |S_k| = \sum_{i=1}^{k} |S_i|.$$

Thus whenever the size of the union of $k-1$ disjoint sets is the sum of their sizes, then the size of a union of k disjoint sets is the sum of their sizes. Thus by the principle of mathematical induction, the size of the union of n disjoint sets is the sum of their sizes for all $n > 1$. The statement holds trivially when $n = 1$ as well.

We stated the simplest form of the product principle as

> If we have a partition of a set S into m blocks, each of size n, then S has size mn.

In Problem 14 we gave a more general form of the product principle which can be stated as

> Let S be a set of functions f from $[n]$ to some set X. Suppose that
>
> - there are k_1 choices for $f(1)$, and
> - suppose that for each choice of $f(1)$, $f(2)$, … $f(i-1)$, there are k_i choices for $f(i)$.
>
> Then the number of functions in the set S is $k_1 k_2 \cdots k_n$.

Activity 78. Prove the general form of the product principle from the simplest form of the product principle.

Solution. We prove by induction that if S is a set of functions defined on $[m]$ such that

-+ There are k_1 choices for $f(1)$ and

-+ when $2 \le i \le m$, for each choice of $f(1)$, $f(2)$, … $f(i-1)$, there are k_i choices for $f(i)$,

then there are $\prod_{i=1}^{m} k_i$ functions in S. When $m = 1$, the product is k_1 and there are k_1 functions in S. Now assume inductively that when S' is a set of functions defined on $[m-1]$ such that

-+ There are k_1 choices for $f(1)$ and

-+ when $2 \le i \le m-1$, for each choice of $f(1)$, $f(2)$, … $f(i-1)$, there are k_i choices for $f(i)$,

then there are $\prod_{i=1}^{m-1} k_i$ functions in S'. Now partition S into k_1 sets S_j, where S_j is the set of functions f in S with $f(n) = y_j$ for each of the k_n values y_j that are possible for $f(1)$. Thus S is a union of k_n sets S_j each of size $\prod_{i=1}^{m-1} k_i$ (by the inductive hypothesis), and so by the product principle for unions of sets, S has size $\prod_{i=1}^{m} k_i$. Therefore, by the principle of mathematical induction, we have proved the general product principle.

Double Induction and Ramsey Numbers

In Section we gave two different descriptions of the Ramsey number $R(m, n)$. However if you look carefully, you will see that we never showed that Ramsey numbers actually exist; we merely described what they were and showed that

$R(3,3)$ and $R(3,4)$ exist by computing them directly. As long as we can show that there is some number R such that when there are R people together, there are either m mutual acquaintances or n mutual strangers, this shows that the Ramsey Number $R(m, n)$ exists, because it is the smallest such R. Mathematical induction allows us to show that one such R is $\binom{m+n-2}{m-1}$. The question is, what should we induct on, m or n? In other words, do we use the fact that with $\binom{m+n-3}{m-2}$ people in a room there are at least $m-1$ mutual acquaintances or n mutual strangers, or do we use the fact that with at least $\binom{m+n-3}{n-2}$ people in a room there are at least m mutual acquaintances or at least $n-1$ mutual strangers? It turns out that we use both. Thus we want to be able to simultaneously induct on m and n. One way to do that is to use yet another variation on the principle of mathematical induction, the *Principle of Double Mathematical Induction*. This principle (which can be derived from one of our earlier ones) states that

> In order to prove a statement about integers m and n, if we can
>
> 1. Prove the statement when $m = a$ and $n = b$, for fixed integers a and b
>
> 2. Prove the statement when $m = a$ and $n > b$ and when $m > a$ and $n = b$ (for the same fixed integers a and b),
>
> 3. Show that the truth of the statement for $m = j$ and $n = k-1$ (with $j \geq a$ and $k > j$) and the truth of the statement for $m = j - 1$ and $n = k$ (with $j > a$ and $k \geq b$) imply the truth of the statement for $m = j$ and $n = k$,
>
> then we can conclude the statement is true for all pairs of integers $m \geq a$ and $n \geq b$.

Activity 79. Prove that $R(m, n)$ exists by proving that if there are $\binom{m+n-2}{m-1}$ people in a room, then there are either at least m mutual acquaintances or at least n mutual strangers.

Solution. We use double induction on m and n to prove this for $m, n \geq 2$. Note first that $R(m, 2) = m = \binom{m+2-2}{m-1}$ and $R(2, n) = n = \binom{2+n-2}{1}$. (In words, if there are m people in a room, then either all m people know each other or there are two mutual nonacquaintances, and if there are n people in a room, then either there are two people who know each other or they are all mutual strangers.)[1] Now assume that whenever there are $\binom{m+n-3}{m-1}$ people in a room there are either at least m mutual acquaintances or $n - 1$ mutual strangers, and that whenever there are at least $\binom{m+n-3}{m-2}$ people in a room there are either at least $m - 1$ mutual acquaintances or n mutual strangers. Suppose that we have $\binom{m+n-2}{m-1}$ people in a room. Choose a person, say P. Then since $\binom{m+n-2}{m-1} = \binom{m-n-3}{m-1} + \binom{m+n-3}{m-2}$, person P is, by the generalized pigeonhole principle, either acquainted with $\binom{m+n-3}{m-2}$ people or unacquainted with $\binom{m+n-3}{m-1}$ people. In the first case, among the people with

---

[1]Note that we have covered both steps 1 and 2 of a double induction proof now.

whom P is acquainted, either $m-1$ are mutual acquaintances or n are mutual strangers. If m are mutual strangers, we are done, while if $m-1$ are mutual acquaintances, these $m-1$ people, together with person P, are m mutual acquaintances, in which case we are done as well. In the second case, among the $\binom{m+n-3}{m-1}$ people with whom person P is unacquainted, there are either m mutual acquaintances, in which case we are done, or there are $n-1$ mutual strangers. In this event, these $n-1$ mutual strangers, along with person P make up n mutual strangers. Thus in every case, if we know that with $\binom{m+n-3}{m-2}$ people in a room there are either $m-1$ mutual acquaintances or n mutual strangers, and we know that with $\binom{m+n-3}{m-1}$ people in a room there are either m mutual acquaintances or $n-1$ mutual strangers, we can conclude that with $\binom{m+n-2}{m-1}$ people in a room there are either m mutual acquaintances or n mutual strangers. Therefore by the principle of double mathematical induction we know that for all m and n greater than or equal to 2, if there are $\binom{m+n-2}{m-1}$ people in a room, then there are either m mutual acquaintances of n mutual strangers. This shows that $R(m,n)$ exists and is no more than $\binom{m+n-2}{m-1}$.

Activity 80. Prove that $R(m,n) \leq R(m-1,n) + R(m,n-1)$.

Solution. If there are $R(m-1,n) + R(m,n-1)$ people in a room, choose one person, say person P. By the generalized pigeonhole principle, there are either $R(m-1,n)$ people with whom P is acquainted or $R(m,n-1)$ people with whom person P is unacquainted. In the first case, among the people with whom person P is acquainted, there are either n mutual strangers, in which case we are done, or there are $m-1$ people with whom person P is acquainted. These $m-1$ people and person P form m people who are mutually acquainted, and so we have m mutual acquaintances. On the other hand, if P is unacquainted with $R(m,n-1)$ people, then among these people, there are either m mutually acquainted people, in which case we are done, or among these people there are $m-1$ mutually unacquainted people, and these $m-1$ people together with P make m mutual strangers. Thus in every case, if there are $R(m-1,n) + R(m,n-1)$ people in a room, there are either at least m mutual acquaintances or at least n mutual strangers. Therefore $R(m,n) \leq R(m-1,n) + R(m,n-1)$.

Activity 81.

  (a) What does the equation in Problem 80 tell us about $R(4,4)$?

    Solution. $R(4,4) \leq R(3,4) + R(4,3) = 9 + 9 = 18$.

  (b) Consider 17 people arranged in a circle such that each person is acquainted with the first, second, fourth, and eighth person to the right and the first, second, fourth, and eighth person to the left. can you find a set of four mutual acquaintances? Can you find a set of four mutual strangers?

Solution. You cannot find either. If there were a set of four mutual acquaintances, you could assume by symmetry that it includes person 1, and two people from among those one, two, four, and eight places to the right. Thus you can assume your set of four acquaintances contains person 1 and two from among persons 2, 3, 5, and 9. However persons 2 and 5, 2 and 9 and 3 and 9 are not acquainted. Thus three of the mutually acquainted people are either persons 1, 2, and 3, persons 1, 5 and 9 or persons 1, 3, and 5. However person 5 is not acquainted with the person one, two, or eight places to the left of person 1, so if person 5 is in the set of mutual acquaintances, then person 14 must be as well. However person 3 and person 9 are not acquainted with person 14. Thus our set must contain persons 1, 2, and 3. However person 3 is not acquainted with the person one, two, four, or eight persons to the left of person 1, so there is no set of four mutual acquaintances. A similar argument holds for nonacquaintances.

(c) What is $R(4,4)$?

Solution.

(d) (Optional) Prove the inequality of Problem 79 by induction on $m + n$.

Solution. We want to prove that if $m \geq 2$ and $n \geq 2$, then when there are $\binom{m+n-2}{m-1}$ people in a room, there are either m mutual acquaintances or n mutual strangers. If $m+n = 4$, then $m = 2$ and $n = 2$, and if there are $\binom{2+2-2}{1} = 2$ people in a room, there are either two who know each other or two who don't.

Now assume that when $m + n = k - 1$, it is the case that if there are $\binom{m+n-2}{m-1}$ people in a room, then there are either m mutual acquaintances or n mutual strangers. Suppose that $m' + n' = k$ and there are $\binom{m'+n'-2}{m'-1}$ people in a room. If $n' = 2$, then we know that with $\binom{m'+2-2}{m'-1} = m'$ people in a room, there are either m' mutual acquaintances or two mutual strangers, and similarly if $m' = 2$ there are either two mutual acquaintances or n mutual strangers among $\binom{m'+n'-2}{m-1}$ people. Thus we may assume that both m' and n' are greater than two. Since $\binom{m'+n'-2}{m-1} = \binom{(m'-1)+n-2}{m'-2} + \binom{m'+(n-1)-2}{m-1}$, if there are $\binom{m'+n'-2}{m'-1}$ people in a room, then a given person, say person P, is either acquainted with $\binom{(m'-1)+n'-2}{m'-1}$ of them (call this case 1) or is a stranger with $\binom{m'+(n-1)-2}{m-1}$ of them (call this case 2). Notice that $m' - 1 + n' = k - 1$ and $m' + n' - 1 = k - 1$. Thus in case 1, our inductive hypothesis tells us that either $m' - 1$ of person P's acquaintances are mutually acquainted, in which case they and person P form m' mutual acquaintances, or n' of P's acquaintances are mutual strangers, in which case we have n' mutual strangers. Similarly in case 2 we have either m' mutual acquaintances or n' mutual strangers. Thus by the principle of mathematical induction,

for all values of $m + n$ greater than or equal to 4, if we have $\binom{m+n-2}{m-1}$ people in a room, then we have either m mutual acquaintances or n mutual strangers, so that $R(m, n)$ exists and is no more than $\binom{m+n-2}{m-1}$.

(e) Use Stirling's approximation (Problem 46) to convert the upper bound for $R(n, n)$ that you get from Problem 79 to a multiple of a power of an integer.

Solution. $R(n, n) \leq \binom{n+n-2}{m-1} = \frac{(2n-2)!}{(n-1)!^2}$. For n sufficiently large, this is approximately

$$\frac{\sqrt{2\pi(2n-2)}(2n-2)^{2n-2}/e^{2n-2}}{\sqrt{2\pi(n-1)}(n-1)^{n-1}\sqrt{2\pi(n-1)}(n-1)^{n-1}/e^{n-1}e^{n-1}}$$

$$= \frac{2^{2n-2}(n-1)^{2n-2}}{\sqrt{\pi(n-1)}(n-1)^{2n-2}}$$

$$= \frac{1}{\sqrt{\pi(n-1)}} 2^{2n-2}$$

A bit of asymptotic combinatorics

Problem 81 gives us an upper bound on $R(n, n)$. A very clever technique due to Paul Erdös, called the "probabilistic method," will give a lower bound. Since both bounds are exponential in n, they show that $R(n, n)$ grows exponentially as n gets large. An analysis of what happens to a function of n as n gets large is usually called an *asymptotic analysis.* The *probabilistic method,* at least in its simpler forms, can be expressed in terms of averages, so one does not need to know the language of probability in order to understand it. We will apply it to Ramsey numbers in the next problem. Combined with the result of Problem 81, this problem will give us that $\sqrt{2}^n < R(n, n) < 2^{2n-2}$, so that we know that the Ramsey number $R(n, n)$ grows exponentially with n.

Activity 82. Suppose we have two numbers n and m. We consider all possible ways to color the edges of the complete graph K_m with two colors, say red and blue. For each coloring, we look at each n-element subset N of the vertex set M of K_m. Then N together with the edges of of K_m connecting vertices in N forms a complete graph on n vertices. This graph, which we denote by K_N, has its edges colored by the original coloring of the edges of K_m.

(a) Why is it that if there is no subset $N \subseteq M$ so that all the edges of K_N are colored the same color, then $R(n, n) > m$?

Solution. Another way to say there is no such subset is to say that it is not possible to find a K_n all of whose edges are red or a K_n all of whose edges are blue. This means that $R(n, n) > n$.

(b) To apply the probabilistic method, we are going to compute the average, over all colorings of K_m, of the number of sets $N \subseteq M$ with $|N| = n$ such that K_N *does* have all its edges the same color. Explain why it is that if the average is less than 1, then for some coloring there is no set N such that K_N has all its edges colored the same color. Why does this mean that $R(n,n) > m$?

Solution. If the average of n nonnegative integers is less than one, they cannot all be one or more, so one has to be zero. Thus in this context there must be some coloring that has no set N so that K_N has all its edges colored the same color.

(c) We call a K_N *monochromatic* for a coloring c of K_m if the color $c(e)$ assigned to edge e is the same for every edge e of K_N. Let us define $mono(c, N)$ to be 1 if N is monochromatic for c and to be 0 otherwise. Find a formula for the average number of monochromatic K_Ns over all colorings of K_m that involves a double sum first over all edge colorings c of K_m and then over all n-element subsets $N \subseteq M$ of $mono(c, N)$.

Solution.

$$\frac{1}{2^{\binom{m}{2}}} \sum_{c:c \text{ is a coloring of } K_m} \sum_{N:N \subseteq M, \, |N|=n} mono(c, N).$$

(d) Show that your formula for the average reduces to $2\binom{m}{n} \cdot 2^{-\binom{n}{2}}$

Solution.

$$\frac{1}{2^{\binom{m}{2}}} \sum_{c:c \text{ is a coloring of } K_m} \sum_{N:N \subseteq M, \, |N|=n} mono(c, N)$$

$$=\frac{1}{2^{\binom{m}{2}}} \sum_{N:N \subseteq M, \, |N|=n} \sum_{c:c \text{ is a coloring of } K_m} mono(c, N)$$

$$=2^{-\binom{m}{2}} \sum_{N:N \subseteq M, \, |N|=n} 2 \cdot 2^{\binom{m}{2}-\binom{n}{2}}$$

$$=2\binom{m}{n} 2^{-\binom{n}{2}}$$

(e) Explain why $R(n,n) > m$ if $\binom{m}{n} \leq 2^{\binom{n}{2}-1}$.

Solution. $R(n,n) > m$ if the average above is less than 1. Thus $R(n,n) > m$ if $2\binom{m}{n}2^{-\binom{n}{2}} < 1$, which is equivalent to $\binom{m}{n} < 2^{\binom{n}{2}-1}$.

(f) Explain why $R(n,n) > \sqrt[n]{n!2^{\binom{n}{2}-1}}$.

Solution. $\binom{m}{n} < 2^{\binom{n}{2}-1}$ is the same as $\frac{m^{\underline{n}}}{n!} < 2^{\binom{n}{2}-1}$. And since $m^{\underline{n}} < m^n$, the inequality $\frac{m^{\underline{n}}}{n!} < 2^{\binom{n}{2}-1}$ holds if the inequality $\frac{m^n}{n!} \leq 2^{\binom{n}{2}-1}$ holds. And this last inequality holds if $m \leq \sqrt[n]{n!2^{\binom{n}{2}-1}}$ holds. Thus $R(n,n) > m$ for any m such that $m \leq \sqrt[n]{n!2^{\binom{n}{2}-1}}$, which implies that $R(n,n) > \sqrt[n]{n!2^{\binom{n}{2}-1}}$.

(g) By using Stirling's formula, show that if n is large enough, then $R(n,n) > \sqrt{2^n} = \sqrt{2}^{\,n}$

Solution. Using Stirling's approximation to $n!$ we get

$$R(n,n) > \sqrt[n]{\frac{n^n}{e^n}\sqrt{2\pi n}2^{\frac{n^2-n-2}{2}}} = \frac{n}{e}2^{\frac{n^2-n-2}{2n}}\sqrt[2n]{2\pi n} > 2^n/2 = \sqrt{2}^{\,n}.$$

2.2 Recurrence Relations

We have seen in ⟨⟨Unresolved xref, reference "SubsetsByInduction"; check spelling or use "provisional" attribute⟩⟩ (or Problem 332 in the Appendix on Induction) that the number of subsets of an n-element set is twice the number of subsets of an $(n-1)$-element set.

Activity 83. Explain why it is that the number of bijections from an n-element set to an n-element set is equal to n times the number of bijections from an $(n-1)$-element subset to an $(n-1)$-element set. What does this have to do with Problem 27?

Solution. To specify a bijection f from an n-element set $\{a_1, a_2, \ldots a_n\}$ to an n-element set, we have n choices for $f(a_1)$, and then b_{n-1} choices for how to define f from the elements $\{a_2, a_3, \ldots, a_n\}$ to the remaining $n-1$ elements of our range. By the product principle this gives us $b_n = nb_{n-1}$ for the number b_n of bijections from an n-element set to an n-element set.

We can summarize these observations as follows. If s_n stands for the number of subsets of an n-element set, then

$$s_n = 2s_{n-1}, \tag{2.1}$$

and if b_n stands for the number of bijections from an n-element set to an n-element set, then

$$b_n = nb_{n-1}. \tag{2.2}$$

Equations (2.1) and Equation (2.2) are examples of *recurrence equations* or *recurrence relations*. A *recurrence relation* or simply a *recurrence* is an equation that expresses the nth term of a sequence a_n in terms of values of a_i for $i < n$. Thus Equations (2.1) and Equation (2.2) are examples of recurrences.

Examples of recurrence relations

Other examples of recurrences are

$$a_n = a_{n-1} + 7, \tag{2.3}$$

$$a_n = 3a_{n-1} + 2^n, \tag{2.4}$$

$$a_n = a_{n-1} + 3a_{n-2}, \text{ and} \tag{2.5}$$

$$a_n = a_1 a_{n-1} + a_2 a_{n-2} + \cdots + a_{n-1} a_1. \tag{2.6}$$

A *solution* to a recurrence relation is a sequence that satisfies the recurrence relation. Thus a solution to Recurrence (2.1) is $s_n = 2^n$. Note that $s_n = 17 \cdot 2^n$ and $s_n = -13 \cdot 2^n$ are also solutions to Recurrence (2.1). What this shows is that a recurrence can have infinitely many solutions. In a given problem, there is generally one solution that is of interest to us. For example, if we are interested in the number of subsets of a set, then the solution to Recurrence (2.1) that we care about is $s_n = 2^n$. Notice this is the only solution we have mentioned that satisfies $s_0 = 1$.

Activity 84. Show that there is only one solution to Recurrence (2.1) that satisfies $s_0 = 1$.

Solution. We prove by induction on n that there is one and only one value s_n that satisfies both $s_n = 2s_{n-1}$ for $n > 0$ and $s_0 = 1$. First, there is clearly one and only one value s_0 that satisfies $s_0 = 1$. Now assume that $k > 0$ and there is one and only one value s_{k-1} that satisfies the two equations. Then $s_k = 2s_{k-1}$ is the one and only one value that satisfies the two equations. Therefore by the principle of mathematical induction, for all nonnegative integers n there is one and only one value s_n that satisfies the equations $s_0 = 1$ and $s_k = 2s_{k-1}$ for all $k > 0$. (Note that since we were making a statement about s_n for all nonnegative integers n it was not appropriate to use n as the dummy variable in the recursive equation $s_k = 2s_{k-1}$.)

Activity 85. A first-order recurrence relation is one which expresses a_n in terms of a_{n-1} and other functions of n, but which does not include any of the terms a_i for $i < n - 1$ in the equation.

(a) Which of the recurrences (2.1) through Equation (2.6) are first order recurrences?

   Solution. The recurrences (2.1), Equation (2.2), Equation (2.3), and Equation (2.4) are all examples of first order recurrences. The recurrences (2.5) and Equation (2.6) are not.

(b) Show that there is one and only one sequence a_n that is defined for every nonnegative integer n, satisfies a given first order recurrence, and satisfies $a_0 = a$ for some fixed constant a.

Solution. A first order recurrence will give a_n in terms of a_{n-1}, that is, there will be a function f such that $a_n = f(a_{n-1})$ for all $n > 0$. We prove by induction that there is one and only one solution to a first order recurrence that satisfies $a_0 = a$ for some fixed constant a. First, there is one and only one value for a_0. Now suppose that when $n = k - 1$, there is one and only one value possible for a_{n-1}. Then a_k is uniquely determined by $a_k = f(a_{k-1}$. Thus the truth of the statement that a_{k-1} is uniquely determined by the equations $a_0 = a$ and $a_n = f(a_{n-1})$ implies the truth of the statement that a_k is determined uniquely by the equations $a_0 = a$ and $a_n = f(a_{n-1})$. Therefore by the principle of mathematical induction, a_k is uniquely determined by the equations $a_0 = a$ and $a_n = f(a_{n-1})$ for all nonnegative integers k.
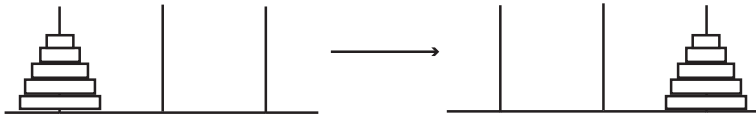
Figure 2.2.1: The Towers of Hanoi Puzzle

Activity 86. The "Towers of Hanoi" puzzle has three rods rising from a rectangular base with n rings of different sizes stacked in decreasing order of size on one rod. A legal move consists of moving a ring from one rod to another so that it does not land on top of a smaller ring. If m_n is the number of moves required to move all the rings from the initial rod to another rod that you choose, give a recurrence for m_n. (Hint: suppose you already knew the number of moves needed to solve the puzzle with $n - 1$ rings.)

Solution. We can solve the puzzle in one step if there is one ring, so $m_1 = 1$. If $n > 0$ and we want to move all the rings from the initial rod to rod 3, then first we solve the problem of moving all but the bottom ring to rod 2; this takes m_{n-1} steps, then we move the bottom ring to rod 3, then we solve the problem of moving all the remaining rings from rod 2 to rod 3. Thus we have $m_n = 2m_{n-1} + 1$.

Activity 87. We draw n mutually intersecting circles in the plane so that each one crosses each other one exactly twice and no three intersect in the same point. (As examples, think of Venn diagrams with two or three mutually intersecting sets.) Find a recurrence for the number r_n of regions into which the plane is divided by n circles. (One circle divides the plane into two regions, the inside and the outside.) Find the number of regions with n circles. For what values of n can you draw a Venn diagram showing all the possible intersections of n sets using circles to represent each of the sets?

Solution. One circle defines two regions, the inside and outside. When we draw a second circle that intersects the first, we can start at one of the intersection points and go inside the first circle, cutting its region into two pieces, and then when we leave it we cut the outside region into two pieces.

This suggests the general pattern. If we have drawn $n-1$ circles and start a new one, each time we enter a circle, we start dividing a region into two pieces. Each time we leave a circle, we also start dividing a region into two pieces. Thus if we have r_n regions with n circles, to get the number of regions, we note that in going from $n-1$ circles to n circles, we start with r_{n-1} regions, and divide $2(n-1)$ of them in half, so we get $2n-2$ new regions. This gives us $r_n = r_{n-1} + 2(n-1)$. Notice that by substitution of the formula $r_{n-1} = r_{n-2} + 2(n-2)$, we get $r_n = r_{n-2} + 2(n-2) + 2(n-1)$, and would guess that $r_n = r_{n-3} + 2(n-3) + 2(n-2) + 2(n-1)$. This leads to the conjecture

$$r_n = r_1 + 2 \cdot 1 + 2 \cdot 2 + \cdots + 2 \cdot (n-1) = r_1 + 2 \sum_{i=1}^{n-1} i = 2 + n(n-1).$$

We can prove this formula by induction. When $n = 1$ we have $2 + 1 \cdot 0$ regions. Assuming that $n-1$ circles give us $2 + (n-1)(n-2)$ regions, for n circles we have $2 + (n-1)(n-2) + 2(n-1) = 2 + n(n-1)$ regions. Thus the correctness of our formula for $n-1$ circles implies its correctness when we have n circles, so for all positive integers n, we get $2 + n(n-1)$ regions determined by n mutually intersecting circles. Two circles cannot touch more than twice, and if we let some of our n circles touch just once, or not at all, that would reduce the number of regions we would get. Similarly, allowing a circle to go through the intersection point of two other circles could only reduce the number of regions. So with n circles we could never have more than $2 + n(n-1)$ regions. In particular with 4 circles we get just 14 regions, rather than the 16 that would be required in a Venn diagram for four sets. We could prove, again by induction, that $2 + n(n-1) < 2^n$ for all $n > 3$, so it is not possible to draw a Venn diagram using circles to illustrate the intersections of four or more sets.

Arithmetic Series (optional)

Activity 88.

(a) A child puts away two dollars from her allowance each week. If she starts with twenty dollars, give a recurrence for the amount a_n of money she has after n weeks and find out how much money she has at the end of n weeks.

Solution. $a_n = a_{n-1} + 2$. Then by substitution $a_n = a_{n-2} + 2 + 2$, and so we conjecture that $a_n = 20 + 2n$. Since she adds two dollars to her savings each week for n weeks, she has adds 2n dollars to her original 20, which proves the formula. We could have used induction to prove it as well.

(b) A sequence that satisfies a recurrence of the form $a_n = a_{n-1} + c$ is called an *arithmetic progression*. Find a formula in terms of the initial

value a_0 and the common difference c for the term a_n in an arithmetic progression and prove you are right.

Solution. $a_n = a_o + cn$. The formula is valid with $n = 0$, and if $a_{n-1} = a_0 + c(n-1)$, then $a_n = a_0 + c(n-1) + c = a_0 + cn$. Therefore the fact that $a_{n-1} = a_0 + ca_{n-1}$ implies the fact that $a_n = a_0 + cn$. Therefore by the principle of mathematical induction, $a_n = a_0 + cn$ for all nonnegative integers n.

(c) A person who is earning \$50,000 per year gets a raise of \$3000 a year for n years in a row. Find a recurrence for the amount a_n of money the person earns over $n + 1$ years. What is the total amount of money that the person earns over a period of $n + 1$ years? (In $n + 1$ years, there are n raises.)

Solution. By $\langle\langle$Unresolved xref, reference "arithmeticprogression"; check spelling or use "provisional" attribute$\rangle\rangle$ we saw that if b_n is the salary in year n, then $b_n = 50,000 + 3000n$. If a_n is the total amount earned over the period of from year 0 through the end of year n, a period of $n + 1$ years, then $a_n = a_{n-1} + b_n = a_{n-1} + 50,000 + 3000n$. Further, $a_n = \sum_{i=0}^{n} b_i = \sum_{i=0}^{n} 50,000 + 3000i = 50,000(n + 1) + 3000 \sum_{i=0}^{n} i = 50,000(n + 1) + 1500(n(n + 1))$.

(d) An *arithmetic series* is a sequence s_n equal to the sum of the terms a_0 through a_n of of an arithmetic progression. Find a recurrence for the sum s_n of an arithmetic progression with initial value a_0 and common difference c (using the language of $\langle\langle$Unresolved xref, reference "arithmeticprogression"; check spelling or use "provisional" attribute$\rangle\rangle$). Find a formula for general term s_n of an arithmetic series.

Solution. $s_n = \sum_{i=0}^{n} a_0 + ci = (n + 1)a_0 + c \sum_{i=0}^{n} i = (n + 1)a_0 + cn(n + 1)/2$.

First order linear recurrences

Recurrences such as those in Equations (2.1) through Equation (2.5) are called *linear recurrences*, as are the recurrences of Problems 86 and Activity 87. A *linear recurrence* is one in which a_n is expressed as a sum of functions of n times values of (some of the terms) a_i for $i < n$ plus (perhaps) another function (called the *driving function*) of n. A linear equation is called *homogeneous* if the driving function is zero (or, in other words, there is no driving function). It is called a *constant coefficient linear recurrence* if the functions that are multiplied by the a_i terms are all constants (but the driving function need not be constant).

Activity 89. Classify the recurrences in Equations (2.1) through Equation (2.5) and Problems 86 and Activity 87 according to whether or not they are constant coefficient, and whether or not they are homogeneous.

Solution. Recurrence (2.1) is first order, linear, constant coefficient, and homogeneous. Recurrence (2.2) is first order, linear, and homogeneous, but not constant coefficient. Recurrence (2.3) is first order, linear, constant coefficient but not homogeneous. Recurrence (2.4) is first order, linear, constant coefficient but not homogeneous. Recurrence (2.5) is not first order (it is second order), is linear, constant coefficient and homogeneous. The recurrence of Problem 86 is first order, linear, and constant coefficient, and that of Problem 87 is first order, linear, and constant coefficient.

Activity 90. As you can see from Problem 89 some interesting sequences satisfy first order linear recurrences, including many that have constant coefficients, have constant driving term, or are homogeneous. Find a formula in terms of b, d, a_0 and n for the general term a_n of a sequence that satisfies a constant coefficient first order linear recurrence $a_n = ba_{n-1} + d$ and prove you are correct. If your formula involves a summation, try to replace the summation by a more compact expression.

Solution. Note that by the formula, $a_{n-1} = ba_{n-2} + d$. Substituting this into the original equation for a_n gives $a_n = b^2 a_{n-2} + bd + d$. Repeating this kind of substitution gives us $a_n = b^3 a_{n-3} + b^2 d + bd + d$. This suggests that $a_n = a_0 b^n + \sum_{i=0}^{n-1} db^i$. We would guess the same formula by writing out the first few values of a_i, namely, a_0, $a_0 b + d$, $a_0 b^2 + db + d$, $a_0 b^3 + b^2 d + bd + d$, and so on. We prove our general formula by induction on n. It is clearly true when $n = 0$ as there are no terms in the sum and $b^0 = 1$. If we assume the formula is true when $n = k - 1$, we may write

$$a_k = ba_{k-1} + d$$

$$= b\left(a_0 b^{k-1} + b\left(\sum_{i=0}^{n-2} db^i\right)\right) + d$$

$$= a_0 b^n + \sum_{i=0}^{n-1} db^i$$

Thus the truth of our formula for $n = k - 1$ implies its truth for $n = k$ and therefore by the principle of mathematical induction, it is true for all nonnegative integers n.

We can give a more compact expression for the sum $\sum_{i=0}^{n-1} db^i = d \sum_{i=0}^{n-1} b^i$. Recall from algebra that $(1 + x)(1 - x) = 1 - x^2$, $(1 + x + x^2)(1 - x) = 1 - x^3$, and in general $(1 + x + x^2 + \cdots x^{n-1}))(1 - x) = 1 - x^n$. If you do not recall this formula, you can prove it by induction, or observe that

$$(1 + x + x^2 + \cdots x^{n-1}))(1 - x)$$
$$= (1 + x + x^2 + \cdots x^{n-1}) \cdot 1 - (1 + x + x^2 + \cdots x^{n-1}) \cdot x$$

$$=1 + x + x^2 + \cdots x^{n-1} - (x + x^2 + \cdots + x^n)\rangle = \rangle 1 - x^n.$$

Dividing the first and last terms by $1 - x$ gives us

$$\sum_{i=1}^{n-1} x^i = \frac{1 - x^n}{1 - x}.$$

Using this in our formula for a_n gives us $a_n = a_o b^n + d\frac{1-b^n}{1-b}$. This is valid except in the case $b = 1$ (in our computation with x above, we would be dividing by 0.) If $b = 1$ we get $a_n = a_0 + nd$ for the sum.

Geometric Series

A sequence that satisfies a recurrence of the form $a_n = ba_{n-1}$ is called a *geometric progression*. Thus the sequence satisfying Equation (2.1), the recurrence for the number of subsets of an n-element set, is an example of a geometric progression. From your solution to Problem 90, a geometric progression has the form $a_n = a_0 b^n$. In your solution to Problem 90 you may have had to deal with the sum of a geometric progression in just slightly different notation, namely $\sum_{i=0}^{n-1} db^i$. A sum of this form is called a *(finite) geometric series*.

Activity 91. Do this problem only if your final answer (so far) to Problem 90 contained the sum $\sum_{i=0}^{n-1} db^i$.

(a) Expand $(1 - x)(1 + x)$. Expand $(1 - x)(1 + x + x^2)$. Expand $(1 - x)(1 + x + x^2 + x^3)$.

Solution. $(1 - x)(1 + x) = 1 - x^2$. $(1 - x)(1 + x + x^2) = 1 - x^3$. $(1 - x)(1 + x + x^2 + x^3) = 1 - x^4$.

(b) What do you expect $(1-b)\sum_{i=0}^{n-1} db^i$ to be? What formula for $\sum_{i=0}^{n-1} db^i$ does this give you? Prove that you are correct.

Solution. We expect $(1 - b)\sum_{i=0}^{n-1} db^i$ to be $d(1 - b^n)$. If $b \neq 1$, this gives us $\sum_{i=0}^{n-1} db^i = d\frac{1-b^n}{1-b}$. We can prove this by induction on n. If $n = 0$ we get 0 for $\frac{1-b^n}{1-b}$, and also for the sum $\sum_{i=0}^{-1} db^i$, since that sum has no terms. Assuming the formula holds when $n = k - 1$, we may write

$$\sum_{i=0}^{k-1} db^i$$

$$=\sum_{i=0}^{k-2} db^i + db^{k-1}$$

$$=d\frac{1 - b^{k-1}}{1 - b} + db^{k-1}$$

$$= \frac{d - db^{k-1} + db^{k-1} - db^k}{1 - b} \rangle =\rangle d\frac{1 - b^k}{1 - b}.$$

Since the truth of the formula for $n = k - 1$ implies its truth for $n = k$, by the principle of mathematical induction the formula is true for all nonnegative integers n. If $b = 1$ we get the formula $\sum_{i=0}^{n-1} db^i = dn$.

In Problem 90 and perhaps 91 you proved an important theorem.

Theorem 2.2.2. *If $b \neq 1$ and $a_n = ba_{n-1} + d$, then $a_n = a_0 b^n + d\frac{1 - b^n}{1 - b}$. If $b = 1$, then, $a_n = a_0 + nd$*

Corollary 2.2.3. *If $b \neq 1$, then $\sum_{i=0}^{n-1} b^i = \frac{1 - b^n}{1 - b}$. If $b = 1$, $\sum_{i=0}^{n-1} b^i = n$.*

2.3 Graphs and Trees

Undirected graphs

In Section we introduced the idea of a directed graph. Graphs consist of vertices and edges. We describe vertices and edges in much the same way as we describe points and lines in geometry: we don't really say what vertices and edges are, but we say what they do. We just don't have a complicated axiom system the way we do in geometry. A *graph* consists of a set V called a vertex set and a set E called an edge set. Each member of V is called a *vertex* and each member of E is called an *edge*. Associated with each edge are two (not necessarily different) vertices called its endpoints. We draw pictures of graphs by drawing points to represent the vertices and line segments (curved if we choose) whose endpoints are at vertices to represent the edges. In Figure 2.3.1 we show three pictures of graphs.
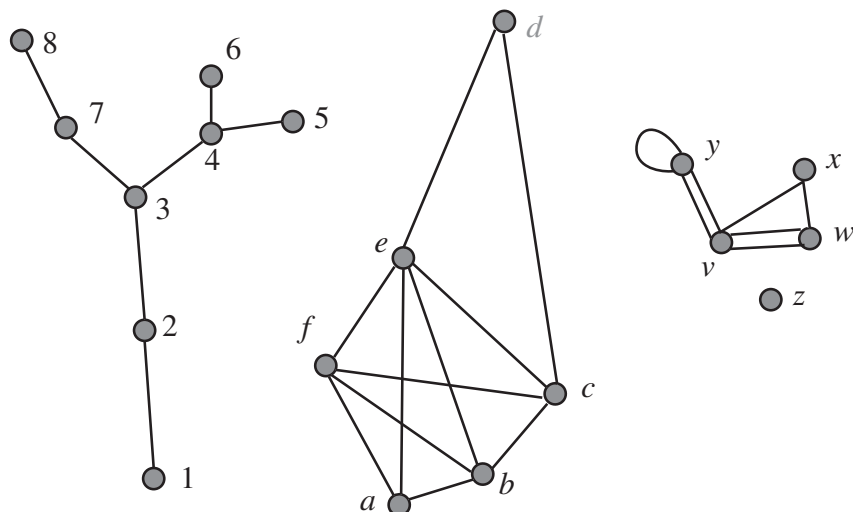
Figure 2.3.1: Three different graphs

Each grey circle in the figure represents a vertex; each line segment represents an edge. You will note that we labelled the vertices; these labels are names we chose to give the vertices. We can choose names or not as we please. The third graph also shows that it is possible to have an edge that connects a vertex (like the one labelled y) to itself or it is possible to have two or more edges (like those between vertices v and y) between two vertices. The *degree* of a vertex is the number of times it appears as the endpoint of edges; thus the degree of y in the third graph in the figure is four.

Activity 92. In the graph on the left in Figure 2.3.1, what is the degree of each vertex?

Solution. The degree of vertex 1 is one, of vertex 2 is two, of vertex 3 is three, of vertex 4 is three, of vertex 5 is one, of vertex 6 is one, of vertex 7 is two, of vertex 8 is one.

Activity 93. For each graph in Figure 2.3.1 is the number of vertices of odd degree even or odd?

Solution. In all three cases it is even.

Activity 94. The sum of the degrees of the vertices of a (finite) graph is related in a natural way to the number of edges.

(a) What is the relationship?

Solution. The sum of the degrees of the vertices is twice the number of edges.

(b) Find a proof that what you say is correct that uses induction on the number of edges. Hint: To make your inductive step, think about what happens to a graph if you delete an edge.

Solution. If a graph has no edges, then the sum of the degrees of the vertices is 0, which is twice the number of edges. Now suppose that whenever a graph has $n-1$ edges, the sum of the degrees of the vertices is twice the number of edges. Let G be a graph with n edges, and delete an edge from G to get G'. The the sum of the degrees of G' is $2(n-1)$, and adding the edge back into G' to get G either increases the degrees of exactly two vertices by one each or increases the degree of one vertex by 2. Thus the sum of the degrees of the vertices of G is $2n$, which is twice the number of edges. Thus by the principle of mathematical induction, for all nonnegative integers n, if a graph has n edges, then the sum of the degrees of the vertices is twice the number of edges.

(c) Find a proof that what you say is correct that uses induction on the number of vertices.

Solution. If a graph has no vertices, then the sum of the degrees of the vertices is 0, which is twice the number of edges. Now suppose that whenever a graph has $n-1$ vertices, the sum of the degrees is twice the number of edges. Let G be a graph with n vertices. Delete one vertex x and all edges it is incident with. By our inductive hypothesis the sum of the degrees of the vertices of the resulting graph is twice the number of edges. Now replace the vertex x and add in the edges connecting it to other vertices in the graph. Each time you add in an edge, you increase the degree of x by one and of one other vertex by one. Finally add in any edges connecting x to itself. Each of these increases the degree of x by two. Thus for each edge we added, the sum of the degrees of the vertices increased by two, so the sum of the degrees of the vertices of G is still twice the number of edges. Therefore by the principle of mathematical induction, for all nonnegative integers n, if a graph has n vertices, then the sum of the degrees of its vertices is even.

(d) Find a proof that what you say is correct that does not use induction.

Solution. The sum of the degrees of the vertices is the sum over all edges of the number of times that edge touches a vertex, which is twice the number of edges.

Activity 95. What can you say about the number of vertices of odd degree in a graph?

Solution. The number of vertices of odd degree must be even, because otherwise the sum of the degrees of the vertices would be odd.

Walks and paths in graphs

A **walk** in a graph is an alternating sequence $v_0 e_1 v_1 \ldots e_i v_i$ of vertices and edges such that edge e_i connects vertices v_{i-1} and v_i. A graph is called connected if, for any pair of vertices, there is a walk starting at one and ending at the other.

Activity 96. Which of the graphs in Figure 2.3.1 is connected?

Solution. They first two connected; the third is not.

Activity 97. A **path** in a graph is a walk with no repeated vertices. Find the longest path you can in the third graph of Figure 2.3.1.

Solution. The path from y to v to x to w is a typical longest path. There are quite a few others. Notice you have two choices for the edge to use to get from y to v.

Activity 98. A **cycle** in a graph is a walk whose first and last vertex are equal but which has no other repeated vertices. Which graphs in Figure 2.3.1 have cycles? What is the largest number of edges in a cycle in the second graph in Figure 2.3.1? What is the smallest number of edges in a cycle in the third graph in Figure 2.3.1?

Solution. The second and third graphs have cycles. The largest number of edges in a cycle in the second graph is six; the smallest number of edges in a cycle in the third graph is one.

Activity 99. A connected graph with no cycles is called a **tree**. Which graphs, if any, in Figure 2.3.1 are trees?

Solution. The first graph is a tree.

Counting vertices, edges, and paths in trees

Activity 100. Draw some trees and on the basis of your examples, make a conjecture about the relationship between the number of vertices and edges in a tree. Prove your conjecture. (Hint: what happens if you choose an edge and delete it, but not its endpoints?)

Solution. The number of edges of a tree is one less than the number of vertices. We prove this by strong induction on the number of edges. First, if a tree has no edges, it can have only one vertex (otherwise it is not connected). Thus the number of edges is one less than the number of vertices. Now suppose that if a tree has fewer than n edges, it the number of edges is one less than the number of vertices. Choose an edge e with endpoints x and y in the tree and remove it. The resulting graph is not connected, for if it were, the path remaining between the endpoints of e, together with e, would form a cycle. Every vertex in the tree is connected to x by a path; those connected by a path not using e still remain connected to x. Similarly those connected to y by a path in the tree that does not use the edge e remain connected to y after

e is deleted. Every vertex is connected to x by a path in the tree and every vertex is connected to y by a path in the tree. If a vertex v were connected to x by a path using the edge e and to y by a path using the edge e, that would mean that we have a path from v to x and then through e to y and also a path from v to y and then through e to x. Thus taking e out of these two paths would give us a path from v to x and a path from v to y, neither of which used e. Thus when we removed the edge e, there would still be a path from x to y, and adding e to this path would give a cycle. Therefore each vertex is connected to either x or y, but not both, by a path that does not use e. Therefore when we remove e, the graph that remains consists of two trees, the tree of all vertices connected to x and the tree of all vertices connected to y. Each of these trees has fewer edges than the original tree, so if they have m_1 and m_2 vertices respectively, they have, by the inductive hypothesis, $m_1 - 1$ and $m_2 - 1$ edges respectively. But together they have all the vertices of the original tree, so the original tree has $m_1 + m_2$ vertices, and has $m_1 - 1 + m_2 - 1 + 1 = m)_1 + m_2 - 1$ edges, the edges of each of the two smaller trees as well as the edge e. Therefore the number of edges of the original tree is one less than the number of vertices. Therefore by the strong principle of mathematical induction, the number of edges of a tree is one less than the number of vertices.

Activity 101. What is the minimum number of vertices of degree one in a finite tree? What is it if the number of vertices is bigger than one? Prove that you are correct.

Solution. The minimum is zero, which happens with a tree with one vertex. If the tree has more than one vertex, the minimum number of vertices of degree one is two. To prove this, we prove that every tree with two or more vertices has at least two vertices of degree two. Note that a tree with two vertices has exactly two vertices of degree 2. Now take a tree with more than two vertices. Remove an edge e without removing its endpoints. As in the solution to Problem 100 this gives two trees. We may assume inductively that each has at least two vertices of degree 1, or else is a single vertex. When we put e back in, it connects one vertex in one tree to one in the other. If both these vertices have degree 1 in their trees, there will be at least one vertex of degree 1 remaining in each tree, so there will be at least two vertices of degree 1 in the tree we get. If exactly one of these vertices is a tree with one vertex after the removal of e, when we connect it to the other tree, we will remove at most one increase the degree of at most one vertex of degree 1 and will create a new vertex of degree 1, so the tree that results still has at least two vertices of degree 1. Therefore by the strong principle of mathematical induction, every tree with more than two vertices has at least two vertices of degree 1. Since a two-vertex tree has two vertices of degree 1, the minimum number of vertices of degree 1 in a tree with two or more vertices is two. (In fact a path with n vertices is a tree and it has exactly two vertices of degree one also.)

Activity 102. In a tree, given two vertices, how many paths can you find between them? Prove that you are correct.

Solution. Exactly one. Suppose there were two distinct paths P_1 and P_2 from x to y. As they leave x, they might leave on the same edge or on different edges. However, since they are different, there must be some first vertex x' on both paths so that when leave x' (as we go from x to y), they leave on different edges. Then since they must both enter y, there must be some first vertex y', following x' on both paths as we go from x to y, such that the two paths enter y' on two different edges. Then the portion of path 1 from x' to y' followed by the portion of path 2 from y' to x' will be a cycle. This is impossible in a tree, so the supposition that there were two distinct paths is impossible.

Activity 103. How many trees are there on the vertex set $\{1, 2\}$? On the vertex set $\{1, 2, 3\}$? When we label the vertices of our tree, we consider the tree which has edges between vertices 1 and 2 and between vertices 2 and 3 different from the tree that has edges between vertices 1 and 3 and between 2 and 3. See Figure 2.3.2.
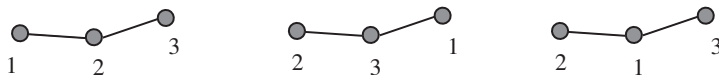
Figure 2.3.2: The three labelled trees on three vertices

How many (labelled) trees are there on four vertices? You don't have a lot of data to guess from, but try to guess a formula for the number of labelled trees with vertex set $\{1, 2, \cdots, n\}$. (If you organize carefully, you can figure out how many labelled trees there are with vertex set $\{1, 2, 3, 4, 5\}$ to help you make your conjecture.) Given a tree with two or more vertices, labelled with positive integers, we define a sequence $b_1, b_2, \ldots$ of integers inductively as follows: If the tree has two vertices, the sequence consists of one entry, namely the label of the vertex with the larger label. Otherwise, let a_1 be the lowest numbered vertex of degree 1 in the tree. Let b_1 be the label of the unique vertex in the tree adjacent to a_1 and write down b_1. Given a_{i-1} and b_{i-1}, let a_i be the lowest numbered vertex of degree 1 in the tree you get by deleting a_1 through a_{i-1} and let b_i be the unique vertex in this new tree adjacent to a_i. We use b to stand for the sequence of b_is we get in this way. For example in the tree (the first graph) in Figure 2.3.1, a_1 is 1 and the sequence b is 234478. (If you are unfamiliar with inductive (recursive) definition, you might want to write down some other labelled trees on eight vertices and construct the sequence of b_is.) How long will the sequence b be if it is applied to a tree with n vertices (labelled with 1 through n)? What can you say about the last member of the sequence of b_is? Can you tell from the sequence of b_is what a_1 is? Find a bijection between labelled trees and something you can "count" that will tell you how many labelled trees there are on n labelled vertices.

Solution. There is one labelled tree on two vertices. We know there are three labelled trees on three vertices, and they all are paths. The difference among them is which vertex is the central vertex on the path. On four vertices

a tree either has a vertex of degree 3 (there are four such trees) or it is a path, in which case there are six choices for the two vertices of degree 2, and for each choice of these two vertices, there are two different ways to attach the remaining vertices to them as vertices of degree 1. Thus there are $12 + 4 = 16$ trees on four vertices. Assuming this is not enough data for a good guess, note that on five vertices, we either have a vertex of degree 4 (there are five such trees), or we have a vertex of degree three which must be adjacent to a vertex of degree two in order to have five vertices. There are $5 \cdot 4 = 20$ ways to choose these two vertices; then three more choices we can make for the degree one vertex attached to the degree 2 vertex. Thus we have 60 trees with a vertex of degree three. If we have neither a vertex of degree four nor a vertex of degree three, then the tree is a path. We have $\binom{5}{2} = 10$ ways to choose the two vertices of degree one, and then there are $3! = 6$ ways to arrange the remaining vertices along the path, so we have 60 paths. Thus we have 125 trees on five vertices. These computations suggest there are n^{n-2} labelled trees on n vertices, which is what we shall prove with the sequences of as and bs.

On a tree with n vertices, the sequence b will have length $n - 1$. The last member of the sequence b will be n. Vertex n can not be in the sequence a, because the tree that remains after we delete an a_i will have at least two vertices of degree 1, so the one of smaller degree will be a_{i+1}. Thus we never delete the vertex n from the tree. Therefore when we choose the last b, we have vertex n and one other vertex, so the other vertex is our a-vertex and n is the vertex adjacent to it. a_1 will be the smallest number that is not in the sequence of b's. This tells us one edge of the tree, namely the edge between a_1 and b_1. The vertex a_2 will be the smallest number different from a_1 not in the sequence b_2 through b_{n-1}. In general, a_i will be the smallest vertex different from a_1 through a_{i-1} not in the sequence b_i through b_{n-1}, which gives us all $n - 1$ edges of the tree (edge i goes from a_i to b_i). Thus there is a bijection between trees and the sequences b_1 through b_{n-1}. But since $b_{n-1} = n$, there is also a bijection between trees and the sequences b_1 through b_{n-2}. But given a sequence of numbers $c_1, c_2, \ldots, c_{n-2}, c_{n-1}$, all between 1 and n and with $c_{n-1} = n$, there is always a smallest number a_1 not in the sequence, and given $a_1, a_2, \ldots a_{i-1}$, there is always a smallest number not in the sequence c_i through c_{n-1} and different from the a_is already chosen, so we can construct the edges from a_i to c_i. Further, if we start with the edge from a_{n-1} to c_{n-1} and work backwards, we will always have a connected graph and will always be adding a vertex of degree 1 to it, so we will have no cycles. Therefore we will get a tree. Thus we have a bijection between labelled trees on n vertices and sequences of length $n - 2$ consisting of members of $[n]$. There are n^{n-2} such sequences, and thus n^{n-2} labelled trees on n vertices.

The sequence $b_1, b_2, \ldots, b_{n-2}$ in Problem 103 is called a *Prüfer coding* or *Prüfer code* for the tree. There is a good bit of interesting information encoded into the Prüfer code for a tree.

Activity 104. What can you say about the Prüfer code for a tree with exactly two vertices of degree 1? Does this characterize such trees?

Solution. It consists of $n - 2$ distinct numbers between 1 and n. Any such Prüfer code is the code of a tree with exactly two vertices of degree one.

Activity 105. What can you determine about the degree of the vertex labelled i from the Prufer code of the tree?

Solution. The degree of a vertex in a tree is one more than the number of times the vertex appears in the Prüfer code of the tree.

Activity 106. What is the number of (labelled) trees on n vertices with three vertices of degree 1? (Assume they are labelled with the integers 1 through n.)

Solution. There are $\binom{n}{3}$ ways to choose the three vertices of degree one. Each of the other $n - 3$ vertices must appear in the Prüfer Code, so one must appear twice. We have $n - 3$ ways to choose that one vertex and $\binom{n-2}{2}\binom{n-4}{1}\binom{n-5}{1}\cdots\binom{1}{1} = \frac{(n-2)!}{2!}$ ways to choose which of the $n - 2$ places to use for which vertices in the Prüfer code. Thus there are $\binom{n}{3}(n - 3)\frac{(n-2)!}{2} = \frac{n!(n-2)(n-3)}{12}$ labelled trees with three vertices of degree one.

Spanning trees

Many of the applications of trees arise from trying to find an efficient way to connect all the vertices of a graph. For example, in a telephone network, at any given time we have a certain number of wires (or microwave channels, or cellular channels) available for use. These wires or channels go from a specific place to a specific place. Thus the wires or channels may be thought of as edges of a graph and the places the wires connect may be thought of as vertices of that graph. A tree whose edges are some of the edges of a graph G and whose vertices are all of the vertices of the graph G is called a *spanning tree* of G. A spanning tree for a telephone network will give us a way to route calls between any two vertices in the network. In Figure 2.3.3 we show a graph and all its spanning trees.
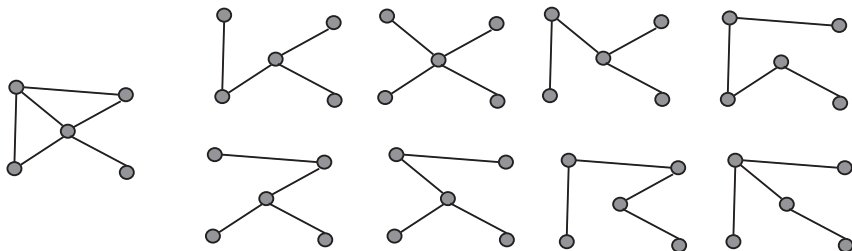

Figure 2.3.3: A graph and all its spanning trees.

Activity 107. Show that every connected graph has a spanning tree. It is possible to find a proof that starts with the graph and works "down" towards the spanning tree and to find a proof that starts with just the vertices and works "up" towards the spanning tree. Can you find both kinds of proof?

Solution.   Here are three proofs:

Start with a connected graph, and if you can find a cycle, remove one edge of that cycle. Repeat this process until you get a tree. You will get a tree, because removing an edge of a cycle reduces the number of cycles but leaves the graph connected. This tree will be a spanning tree of the original graph.

Start with the original vertex set and no edges for a graph H. Go through the edges of the original graph G one at a time, and if you can add an edge of G to the graph H without creating a cycle, do so. Otherwise discard the edge and go on to the next one. This will give you a graph H with no cycles, and if it were not connected, there would be an edge in G between two vertices not yet connected in H. (If there weren't, the graph you just constructed would be connected.) Thus you get a spanning tree.

Start with vertex 1 and no edges as a graph H. Take an edge from it to another vertex. Add that edge and vertex to H. Now take an edge from one of the vertices you currently have to yet another vertex of the original graph. Add that vertex and edge to H. Repeat this process until you can find no such edge. You will get a tree because each edge you add connects a vertex of degree 1 to a tree you have already constructed. Since the original graph is connected, there must always be an edge from the current set of vertices you are considering to something not in that set.

Minimum cost spanning trees

Our motivation for talking about spanning trees was the idea of finding a minimum number of edges we need to connect all the edges of a communication network together. In many cases edges of a communication network come with costs associated with them. For example, one cell-phone operator charges another one when a customer of the first uses an antenna of the other. Suppose a company has offices in a number of cities and wants to put together a communication network connecting its various locations with high-speed computer communications, but to do so at minimum cost. Then it wants to take a graph whose vertices are the cities in which it has offices and whose edges represent possible communications lines between the cities. Of course there will not necessarily be lines between each pair of cities, and the company will not want to pay for a line connecting city i and city j if it can already connect them indirectly by using other lines it has chosen. Thus it will want to choose a spanning tree of minimum cost among all spanning trees of the communications graph. For reasons of this application, if we have a graph with numbers assigned to its edges, the sum of the numbers on the edges of a spanning tree of G will be called the *cost* of the spanning tree.

Activity 108. Describe a method (or better, two methods different in at least one aspect) for finding a spanning tree of minimum cost in a graph whose edges are labelled with costs, the cost on an edge being the cost for including that edge in a spanning tree. Prove that your method(s) work.

Solution.   Start with vertex, number it vertex 1, and choose the least costly edge leaving it. Number the new vertex 2. Given your current set of vertices

and edges, choose the least costly edge leaving that set, and if your set has $i - 1$ vertices, label the new vertex i. You will always have a tree as you go along since you are adding a vertex of degree 1 to a tree you already have. You will get a spanning tree because the graph you start with is connected. If your tree did not have the lowest cost among all spanning trees, there would be some smallest i such that there is an edge from vertex i in a least cost spanning tree that is not in your tree. Since we chose the least i, the edge we just chose from vertex i goes to a higher-numbered vertex. Thus you could have chosen that edge as you were constructing your tree, so there cannot be a tree of lower total cost than the one you chose.

Alternatively, start with the vertex set of the graph and no edges. Choose an edge of least cost. Repeat the following until it cannot be repeated. Given the set of edges you have so far, choose an edge of least cost among all edges that do not form a cycle with edges you already have. The graph you get will have no cycles, and it will have to be connected, because otherwise there would be an edge in the original graph that connects two vertices in the graph you just constructed. Thus you will have a tree. Suppose there is another tree with lower total cost. Choose such a tree with as many edges as common with your tree as possible. Then there is some edge e of this tree of lowest cost among all edges connecting two vertices that are not connected by an edge in your tree. Suppose the cost of this edge is c. Since these vertices are connected by some path in your tree, when you were considering edges of cost c, these two vertices were already connected by a path in your tree. There must be some edge f on that path not in the least cost tree. The edge f was already in your tree while you were considering edges of cost c, so its cost is no more than c. Adding f to the tree of least cost gives a cycle. All edges on that cycle that are not in your tree have cost at least c by our choice of e. But since f was from your tree, there must be some edge g of the cycle that is in the least cost tree that but not in your tree. Removing g from the least cost tree and adding f cannot increase the cost of the tree, but it gives a tree that has one more edge in common with your tree. This contradicts the choice of the least cost tree, so there must have been no tree of lower total cost than the one you constructed.

The method you used in Problem 108 is called a *greedy method*, because each time you made a choice of an edge, you chose the least costly edge available to you.

The deletion/contraction recurrence for spanning trees

There are two operations on graphs that we can apply to get a recurrence (though a more general kind than those we have studied for sequences) which will let us compute the number of spanning trees of a graph. The operations each apply to an edge e of a graph G. The first is called *deletion*; we *delete* the edge e from the graph by removing it from the edge set. Figure 2.3.4 shows how we can delete edges from a graph to get a spanning tree.
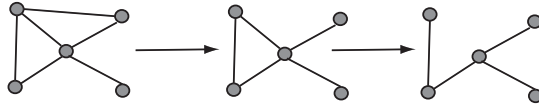
Figure 2.3.4: Deleting two appropriate edges from this graph gives a spanning tree.

The second operation is called *contraction*.
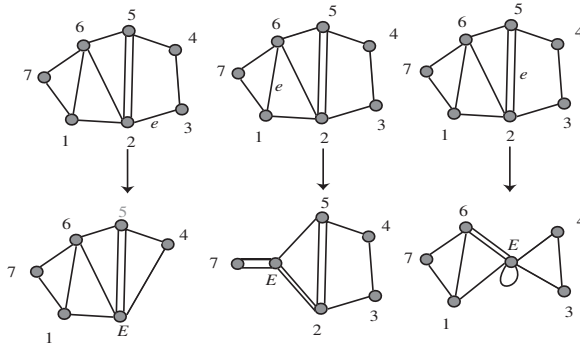
Figure 2.3.5: The results of contracting three different edges in a graph.

Contractions of three different edges in the same graph are shown in Figure 2.3.5. Intuitively, we contract an edge by shrinking it in length until its endpoints coincide; we let the rest of the graph "go along for the ride." To be more precise, we *contract* the edge e with endpoints v and w as follows:

1. remove all edges having either v or w or both as an endpoint from the edge set,

2. remove v and w from the vertex set,

3. add a new vertex E to the vertex set,

4. add an edge from E to each remaining vertex that used to be an endpoint of an edge whose other endpoint was v or w, and add an edge from E to E for any edge other than e whose endpoints were in the set $\{v, w\}$.

We use $G - e$ (read as G minus e) to stand for the result of deleting e from G, and we use G/e (read as G contract e) to stand for the result of contracting e from G.

Activity 109. How do the number of spanning trees of G not containing the edge e and the number of spanning trees of G containing e relate to the number of spanning trees of $G - e$ and G/e? Use $\#(G)$ to stand for the number of spanning trees of G (so that, for example, $\#(G/e)$ stands for the number of spanning trees of G/e). Find an expression for $\#(G)$ in terms of $\#(G/e)$

and $\#(G - e)$. This expression is called the *deletion-contraction recurrence*. Use it to compute the number of spanning trees of the graph in .
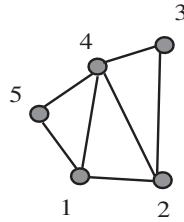
Figure 2.3.6: A graph.

Solution. The number of spanning trees of G not containing e is the number of spanning trees of $G - e$. The number of spanning trees of G containing the edge e is the number of spanning trees of G/e. Therefore $\#(G) = \#(G - e) + \#(G/e)$. Applying the formula twice to G gives

$$\#(G) = \#(G - \{1,2\} - \{2,3\}) + \#((G - \{1,2\})/\{2,3\})$$
$$+ \qquad\qquad\qquad\qquad \#((G/\{1,2\}) - \{2,3\}) + \#($$

We now show the four graphs on the right hand side of the equation.
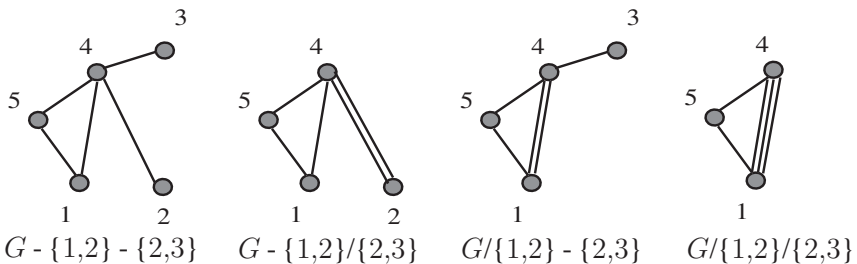
Figure 2.3.7: Various deletions and contractions.

We could now convert each of these graphs to trees with multiple edges by deleting and contracting one more edge, say edge $\{1,5\}$, which would make the analysis easier but the picture twice as big. Since we can easily count spanning trees of a triangle, we can also stop here, noting that the first graph has three spanning trees, the second has six, the third has five, and the fourth has seven, so the total number of spanning trees is $3 + 6 + 5 + 7 = 21$.

Shortest paths in graphs

Suppose that a company has a main office in one city and regional offices in other cities. Most of the communication in the company is between the main office and the regional offices, so the company wants to find a spanning tree that minimizes not the total cost of all the edges, but rather the cost of

communication between the main office and each of the regional offices. It is not clear that such a spanning tree even exists. This problem is a special case of the following. We have a connected graph with nonnegative numbers assigned to its edges. (In this situation these numbers are often called weights.) The *(weighted) length* of a path in the graph is the sum of the weights of its edges. The *distance* between two vertices is the least (weighted) length of any path between the two vertices. Given a vertex v, we would like to know the distance between v and each other vertex, and we would like to know if there is a spanning tree in G such that the length of the path in the spanning tree from v to each vertex x is the distance from v to x in G.

Activity 110. Show that the following algorithm (known as Dijkstra's algorithm) applied to a weighted graph whose vertices are labelled 1 to n gives, for each i, the distance from vertex 1 to i as $d(i)$.

1. Let $d(1) = 0$. Let $d(i) = \infty$ for all other i. Let $v(1)=1$. Let $v(j) = 0$ for all other j. For each i and j, let $w(i,j)$ be the minimum weight of an edge between i and j, or ∞ if there are no such edges. Let $k = 1$. Let $t = 1$.

2. For each i, if $d(i) > d(k) + w(k,i)$ let $d(i) = d(k) + w(k,i)$.

3. Among those i with $v(i) = 0$, choose one with $d(i)$ a minimum, and let $k = i$. Increase t by 1. Let $v(i) = 1$.

4. Repeat the previous two steps until $t = n$

Solution. We prove that the distance from vertex 1 of a vertex i with $v(i) = 1$ is $d(i)$. We use induction on the number t of vertices with $v(i) = 1$. If $t = 1$, then $d(1) = 0$ is the distance from vertex 1 to vertex 1. Now when $t = s - 1$, we have vertices $i_1, i_2, \ldots i_{s-1}$ such that $v(i_p) = 1$. Suppose inductively that $d(i_p)$ is the distance from vertex i_p to vertex 1 for $p = 1, 2, \ldots, s - 1$. When $t = s$, we choose a vertex i with $d(i)$ a minimum. Suppose $u_1, u_2, \ldots, u_r$ is the sequence of vertices of a shortest (least total weight) path from vertex $u_1 = 1$ to vertex $u_r = i$ and that the length (total weight) of this path is less than $d(i)$. Suppose that some vertex u_p has $v(u_p) = 0$, and choose the smallest p such that this is so. Then $d(u_{p-1})$ is the distance from vertex 1 to vertex u_{p-1}, and $d(u_{p-1}) + w(u_{p-1}, u_p)$ is less than $d(i)$ because the length of the path from u_1 to u_r is less than $d(i)$. Then we would not have chosen the vertex i after all, but rather the vertex u_p, which is a contradiction, so all the vertices u_p with $p < r$ must have $v(u_p) = 1$, and, by our inductive hypothesis, $d(u_p)$ must be the distance from vertex 1 to vertex u_p for $p < r$. Thus when we were computing $d(i)$, we would have found that $d(i) \leq d(u_{r-1}) + w(u_{r-1}, i)$. Thus $d(i)$ must be the distance from vertex 1 to vertex i after all. Therefore the fact that Dijkstra's algorithm works when $t = s - 1$ implies that it works when $t = s$, so that, by the principle of mathematical induction, it works for all nonnegative integers t.

Activity 111. Is there a spanning tree such that the distance from vertex 1 to vertex i given by the algorithm in Problem 110 is the distance for vertex 1 to vertex i in the tree (using the same weights on the edges, of course)?

Solution. Yes, when we choose the i with $d(i)$ a minimum, before we change k to i, we add an edge from vertex k to vertex i to the edge set of a graph on the vertex set $[n]$. We get a tree each time we do this step, because we are adding a vertex of degree 1 to a smaller tree. We essentially proved in the solution to Problem 110 that the path from vertex 1 to vertex i in this tree has length (total weight) $d(i)$. Using the same approach we could prove it directly by induction on the number of vertices in our tree.

2.4 Supplementary Problems

1. Use the inductive definition of a^n to prove that $(ab)^n = a^n b^n$ for all nonnegative integers n. If $n = 0$ we get $(ab)^0 = 1$ and $a^0 b^0 = 1$. Assume inductively that $(ab)^{n-1} = a^{n-1} b^{n-1}$. Then by the inductive definition, inductive hypothesis, and commutative law,

$$(ab)^n = (ab)^{n-1} ab = a^{n-1} b^{n-1} ab = a^{n-1} ab^{n-1} b = a^n b^n.$$

Thus the fact that $(ab)^{n-1} = a^{n-1} b^{n-1}$ implies the fact that $(ab)^n = a^n b^n$. Therefore by the principle of mathematical induction, $(ab)^n = a^n b^n$ for all nonnegative integers n.

2. Give an inductive definition of $\bigcup_{i=1}^n S_i$ and use it and the two set distributive law to prove the distributive law $A \cap \bigcup_{i=1}^n S_i = \bigcup_{i=1}^n A \cap S_i$. We define $\bigcup_{i=1}^1 S_i = S_i$ and $\bigcup_{i=1}^n S_i = \bigcup_{i=1}^{n-1} S_i \cup S_n$. Then

$$A \cap \bigcup_{i=1}^1 S_i = A \cap S_1 = \bigcup_{i=1}^1 A \cap S_i.$$

Assume that $n > 1$ and $A \cap \bigcup_{i=1}^{n-1} = \bigcup_{i=1}^{n-1} A \cap S_i$. Now

$$A \cap \bigcup_{i=1}^n S_i = A \cap \left(\bigcup_{i=1}^{n-1} S_i \cup S_n \right) = \left(A \cap \bigcup_{i=1}^{n-1} S_i \right) \cup (A \cap S_n)$$

$$= \left(\bigcup_{i=1}^{n-1} A \cap S_i \right) \cup (A \cap S_n) = \bigcup_{i=1}^n A \cap S_i.$$

Thus the truth of the distributive law for distributing an intersection over a union of $n-1$ sets implies its truth for distributing an intersection over a union of n sets. Therefore by the principle of mathematical induction, the distributive law $A \cap \bigcup_{i=1}^{n} S_i = \bigcup_{i=1}^{n} A \cap S_i$ holds for all positive integers n.

3. (interesting) A hydrocarbon molecule is a molecule whose only atoms are either carbon atoms or hydrogen atoms. In a simple molecular model of a hydrocarbon, a carbon atom will bond to exactly four other atoms and hydrogen atom will bond to exactly one other atom. We represent a hydrocarbon compound with a graph whose vertices are labelled with C's and H's so that each C vertex has degree four and each H vertex has degree one. A hydrocarbon is called an "alkane" (common examples are methane (natural gas), propane, hexane (ordinary gasoline), octane (to make gasoline burn more slowly), etc.) if the graph is a tree. How many vertices are labelled H in the graph of an alkane with exactly n vertices labelled C? We have n vertices of degree four, and so if we have m vertices of degree 1, we get $4n + m = 2(m + n - 1)$ from the fact that the sum of the degrees of the vertices must be twice the number of edges. Thus we have $m = 2n + 2$ hydrogen atoms.

4. (a) Give a recurrence for the number of ways to divide $2n$ people into sets of two for tennis games. (Don't worry about who serves first.)
   $t_{2n} = (2n - 1)t_{2n-2}$

   (b) (interesting) Give a recurrence for the number of ways to divide $4n$ people into sets of four for games of bridge. (Don't worry about how they sit around the bridge table or who is the first dealer.)
   $b_{4n} = \binom{4n-1}{3} b_{4n-4}$.

5. Use induction to prove your result in Supplementary Problem 2 at the end of Chapter 1. A composition of n is an ordered list of positive numbers that adds to n. We wish to show that there are 2^{n-1} compositions of n. There is one composition of the number 1, and $2^{1-1} = 1$. Now assume inductively that there are 2^{n-2} compositions of the number $n - 1$. From a composition of $n - 1$, we can get a composition of n either by making a new last part of size 1, or by adding one to the last part. Clearly these two operations give different partitions of n; what is not so clear is that they give all partitions of n, but they do: Either the last part of a partition of n is 1, in which case it comes from the first kind of operation, or it is larger than one, in which case it comes from the second operation. Thus the number of compositions of n is twice the number of compositions of $n - 1$, and so is $2 \cdot 2^{n-2} = 2^{n-1}$. Therefore the statement that there are 2^{n-2} compositions of $n - 1$ implies the statement that there are 2^{n-1} compositions of n. Thus by the principle of mathematical induction, there are 2^{n-1} compositions of n for every positive integer n.

6. Give an inductive definition of the product notation $\prod_{i=1}^{n} a_i$. $\prod_{i=1}^{1} a_i = a_1$,

   and $\prod_{i=1}^{n} a_i = \left(\prod_{i=1}^{n-1} a_i \right) \cdot a_n$.

7. Using the fact that $(ab)^k = a^k b^k$, use your inductive definition of product notation in Problem 6 to prove that $\left(\prod_{i=1}^{n} a_i \right)^k = \prod_{i=1}^{n} a_i^k$. When $n = 1$ we get $\left(\prod_{i=1}^{1} a_i \right)^k = a_i^k = \prod_{i=1}^{1} a_i^k$. Now assume inductively that $\left(\prod_{i=1}^{m-1} a_i \right)^k = \prod_{i=1}^{m-1} a_i^k$. Then we may write

$$\left(\prod_{i=1}^{m} a_i \right)^k = \left(\left(\prod_{i=1}^{m-1} a_i \right) \cdot a_m \right)^k = \left(\prod_{i=1}^{m-1} a_i^k \right) \cdot a_m^k = \prod_{i=1}^{m} a_i^k.$$

   Thus the correctness of the formula for $n = m - 1$ implies its correctness for $n = m$. Therefore by the principle of mathematical induction, the formula holds for all positive integers n.

8. (interesting) How many labelled trees on n vertices have exactly four vertices of degree 1? The vertices of degree 1 are the vertices that do not appear in the Grey code for the tree. So we first choose four vertices out of n in $\binom{n}{4}$ ways to be our vertices of degree 1, and then we use the remaining $n - 4$ vertices to fill in our list of $n - 2$ vertices, using each of the $n - 4$ at least once. Thus we either use one of them 3 times and the rest once, or two of them twice and the rest once. There are $n - 4$ ways to choose the one we use three times and $\binom{n-2}{3}\binom{n-5}{1}\binom{n-6}{1} \cdots \binom{1}{1} = \frac{(n-2)!}{3!}$ ways to label the $n-2$ places with the chosen vertices. There are $\binom{n-4}{2}$ ways to choose the vertices we would use twice, and $\binom{n-2}{2}\binom{n-4}{2}\binom{n-6}{1}\binom{n-7}{1} \cdots \binom{1}{1} = \frac{(n-2)!}{2!2!}$ ways to assign the chosen vertices to the $n - 2$ places in the Prüfer Code. Thus we have

$$\binom{n}{4}\left((n-4)\frac{(n-2)!}{3!} + \frac{(n-4)(n-3)}{2}\frac{(n-2)!}{4} \right)$$
$$= \frac{n!}{48}(n-2)^{\underline{3}}\left(\frac{1}{3} + \frac{n-3}{4} \right)$$

   possible Prüfer codes and therefore the same number of labelled trees.

9. (interesting and difficult) The degree sequence of a tree is a list of the degrees of the vertices in nonincreasing order. For example the degree

sequence of the first graph in [Figure 2.3.3](#) is 43221. How many labelled trees are there on n vertices with degree sequence $d_1, d_2, \ldots d_n$, or, equivalently, with the degree sequence in which the degree d appears i_d times? We first solve the simpler problem in which we assume d_i is the degree of vertex i. The number of times i appears in the Prüfer code of a tree is one less than the degree of i, so vertex i appears $d_i - 1$ times. Thus the sum of the $d_i - 1$ should be $2n - 2 - n = n - 2$. Of the $n - 2$ places in the Prüfer code, we want to label $d_1 - 1$ of them with 1, $d_2 - 1$ of them with 2 and in general $d_i - 1$ of them with i. There are

$$\binom{n-2}{d_1-1}\binom{n-2-(d_1-1)}{d_2-1}\binom{n-2-(d_1-1+d_2-1)}{d_3-1}\cdots\binom{d_n-1}{d_n-1}$$

ways to do this, so the number of trees in which vertex i has degree d_i is $\frac{(n-2)!}{(d_1-1)!(d_2-1)!\cdots(d_n-1)!}$.

Now we modify the solution by observing that to count all graphs with a given degree sequence, the actual vertices which have the given degrees is irrelevant, so we must multiply the result of the easier problem by the number of ways to assign the degrees to the vertices. To assign the degrees, we can list the vertices in $n!$ ways, choose the first i_1 of these vertices to have degree 1, then next i_2 to have degree 2, and so on. But the order in which we list the vertices of a given degree is irrelevant. Thus the number of ways to assign the degrees is $\frac{n!}{i_1!i_2!\cdots i_n!}$. Once the degrees are assigned, there are $\frac{(n-2)!}{\prod_{d=1}^{n}(d-1)!^{i_d}}$, by translating our easier result. Thus the total number of trees with the degree sequence in which there are i_d vertices of degree d is

$$\frac{n!(n-2)!}{\prod_{j=1}^{n} i_j!(j-1)!^{i_j}}.$$

Chapter 3

Distribution Problems

3.1   The idea of a distribution

Many of the problems we solved in Chapter 1 may be thought of as problems of distributing objects (such as pieces of fruit or ping-pong balls) to recipients (such as children). Some of the ways of viewing counting problems as distribution problems are somewhat indirect. For example, in Problem 37 you probably noticed that the number of ways to pass out k ping-pong balls to n children so that no child gets more than one is the number of ways that we may choose a k-element subset of an n-element set. We think of the children as recipients and objects we are distributing as the identical ping-pong balls, distributed so that each recipient gets at most one ball. Those children who receive an object are in our set. It is helpful to have more than one way to think of solutions to problems. In the case of distribution problems, another popular model for distributions is to think of putting balls in boxes rather than distributing objects to recipients. Passing out identical objects is modeled by putting identical balls into boxes. Passing out distinct objects is modeled by putting distinct balls into boxes.

The twenty-fold way

When we are passing out objects to recipients, we may think of the objects as being either identical or distinct. We may also think of the recipients as being either identical (as in the case of putting fruit into plastic bags in the grocery store) or distinct (as in the case of passing fruit out to children). We may restrict the distributions to those that give at least one object to each recipient, or those that give exactly one object to each recipient, or those that give at most one object to each recipient, or we may have no such restrictions. If the objects are distinct, it may be that the order in which the objects are received is relevant (think about putting books onto the shelves in a bookcase) or that the order in which the objects are received is irrelevant (think about dropping a handful of candy into a child's trick or treat bag). If we ignore

the possibility that the order in which objects are received matters, we have created $2 \cdot 2 \cdot 4 = 16$ distribution problems. In the cases where a recipient can receive more than one distinct object, we also have four more problems when the order objects are received matters. Thus we have 20 possible distribution problems.

The Twentyfold Way: A Table of Distribution Problems

k objects and conditions on how they are received	n recipients and mathem... Distinct
1. Distinct no conditions	n^k functions
2. Distinct Each gets at most one	$n^{\underline{k}}$ k-element permutations
3. Distinct Each gets at least one	? onto functions
4. Distinct Each gets exactly one	$k! = n!$ bijections
5. Distinct, order matters	? ?
6. Distinct, order matters Each gets at least one	? ?
7. Identical no conditions	? ?
8. Identical Each gets at most one	$\binom{n}{k}$ subsets
9. Identical Each gets at least one	? ?
10. Identical Each gets exactly one	1 if $k = n$; 0 otherwise

Table 3.1.1: An incomplete table of the number of ways to distribute k objects to n recipients, with restrictions on how the objects are received

We describe these problems in . Since there are twenty possible distribution problems, we call the table the "Twentyfold Way," adapting terminology suggested by Joel Spencer for a more restricted class of distribu-

tion problems. In the first column of the table we state whether the objects are distinct (like people) or identical (like ping-pong balls) and then give any conditions on how the objects may be received. The conditions we consider are whether each recipient gets at most one object, whether each recipient gets at least one object, whether each recipient gets exactly one object, and whether the order in which the objects are received matters. In the second column we give the solution to the problem and the name of the mathematical model for this kind of distribution problem when the recipients are distinct, and in the third column we give the same information when the recipients are identical. We use question marks as the answers to problems we have not yet solved and models we have not yet studied. We give explicit answers to problems we solved in ⟨⟨Unresolved xref, reference "whatis"; check spelling or use "provisional" attribute⟩⟩ and problems whose answers are immediate. The goal of this chapter is to develop methods that will allow us to fill in the table with formulas or at least quantities we know how to compute, and we will give a completed table at the end of the chapter. We will now justify the answers that are not question marks and replace some question marks with answers as we cover relevant material.

If we pass out k distinct objects (say pieces of fruit) to n distinct recipients (say children), we are saying for each object which recipient it goes to. Thus we are defining a function from the set of objects to the recipients. We saw the following theorem in Problem .

Theorem 3.1.2. *There are n^k functions from a k-element set to an n-element set.*

We proved it in ⟨⟨Unresolved xref, reference "provenumberoffunctionscon-jecture"; check spelling or use "provisional" attribute⟩⟩. If we pass out k distinct objects (say pieces of fruit) to n indistinguishable recipients (say identical paper bags) then we are dividing the objects up into disjoint sets; that is we are forming a partition of the objects into some number, certainly no more than the number k of objects, of parts. Later in this chapter (and again in the next chapter) we shall discuss how to compute the number of partitions of a k-element set into n parts. This explains the entries in row one of our table.

If we pass out k distinct objects to n recipients so that each gets at most one, we still determine a function, but the function must be one-to-one. The number of one-to-one functions from a k-element set to an n element set is the same as the number of one-to-one functions from the set $[k] = \{1, 2, \ldots, k\}$ to an n-element set. In [Problem 20](#) we proved the following theorem.

Theorem 3.1.3. *If $0 \le k \le n$, then the number of k-element permutations of an n-element set is*

$$n^{\underline{k}} = n(n-1)\cdots(n-k+1) = n!/(n-k)!.$$

If $k > n$ there are no one-to-one functions from a k element set to an n element, so we define $n^{\underline{k}}$ to be zero in this case. Notice that this is what

the indicated product in the middle term of our formula gives us. If we are supposed to distribute k distinct objects to n identical recipients so that each gets at most one, we cannot do so if $k > n$, so there are 0 ways to do so. On the other hand, if $k \leq n$, then it doesn't matter which recipient gets which object, so there is only one way to do so. This explains the entries in row two of our table.

If we distribute k distinct objects to n distinct recipients so that each recipient gets at least one, then we are counting functions again, but this time functions from a k-element set *onto* an n-element set. At present we do not know how to compute the number of such functions, but we will discuss how to do so later in this chapter and in the next chapter. If we distribute k identical objects to n recipients, we are again simply partitioning the objects, but the condition that each recipient gets at least one means that we are partitioning the objects into exactly n blocks. Again, we will discuss how compute the number of ways of partitioning a set of k objects into n blocks later in this chapter. This explains the entries in row three of our table.

If we pass out k distinct objects to n recipients so that each gets exactly one, then $k = n$ and the function that our distribution gives us is a bijection. The number of bijections from an n-element set to an n-element set is $n!$ by Theorem 3.1.3. If we pass out k distinct objects of n identical recipients so that each gets exactly 1, then in this case it doesn't matter which recipient gets which object, so the number of ways to do so is 1 if $k = n$. If $k \neq n$, then the number of such distributions is zero. This explains the entries in row four of our table.

We now jump to row eight of our table. We saw in Problem 37 that the number of ways to pass out k identical ping-pong balls to n children is simply the number of k-element subsets of an n-element set. In Problem 39 we proved the following theorem.

Theorem 3.1.4. *If $0 \leq k \leq n$, the number of k-element subsets of an n-element set is given by*

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

We define $\binom{n}{k}$ to be 0 if $k > n$, because then there are no k-element subsets of an n-element set. Notice that this is what the middle term of the formula in the theorem gives us. This explains the entries of row 8 of our table. For now we jump over row 9.

In row 10 of our table, if we are passing out k identical objects to n recipients so that each gets exactly one, it doesn't matter whether the recipients are identical or not; there is only one way to pass out the objects if $k = n$ and otherwise it is impossible to make the distribution, so there are no ways of distributing the objects. This explains the entries of row 10 of our table. Several other rows of our table can be computed using the methods of Chapter 1.

Ordered functions

Activity 112. Suppose we wish to place k distinct books onto the shelves of a bookcase with n shelves. For simplicity, assume for now that all of the books would fit on any of the shelves. Also, let's imagine pushing the books on a shelf as far to the left as we can, so that we are only thinking about how the books sit relative to each other, not about the exact places where we put the books. Since the books are distinct, we can think of a the first book, the second book and so on.

(a) How many places are there where we can place the first book?

   Solution.   There are n places where we can place the first book.

(b) When we place the second book, if we decide to place it on the shelf that already has a book, does it matter if we place it to the left or right of the book that is already there?

   Solution.   Yes.

(c) How many places are there where we can place the second book?

   Hint.   If you decide to put it on a shelf that already has a book, you have two choices of where to put it on that shelf.

   Solution.   Once we have placed it, there are $n + 1$ places where we can place the second book, because on the shelf that has one book, we could put the second book to the left or to the right of the book already there.

(d) Once we have $i - 1$ books placed, if we want to place book i on a shelf that already has some books, is sliding it in to the left of all the books already there different from placing it to the right of all the books already or between two books already there?

   Solution.   All of these are different.

(e) In how many ways may we place the ith book into the bookcase?

   Hint.   Among all the places you could put books, on all the shelves, how many are to the immediate left of some book? How many other places are there?

   Solution.   Once we have $i - 1$ books on the shelves the ith book could go on any shelf to the left of all books there, if any, giving us n places,

or it could go to the immediate right of any book already there, giving us another $i - 1$ places. Thus there are $n + i - 1$ places where we could place book i.

(f) In how many ways may we place all the books?

Solution. From this, we can see that the number of ways to place all the books is

$$\prod_{i=1}^{k} (n + i - 1).$$

Activity 113. Suppose we wish to place the books in Problem 112 (satisfying the assumptions we made there) so that each shelf gets at least one book. Now in how many ways may we place the books? (Hint: how can you make sure that each shelf gets at least one book before you start the process described in Problem 112?)

Hint. How can you make sure that each shelf gets at least one book before you start the process described in Problem 112?

Solution. Choose n books from the k books in $\binom{k}{n}$ ways, and assign them to the n places shelves in $n!$ ways, giving us $k!/(k - n)!$ ways to put a book on each shelf. Now leaving these books at the far left of each shelf, place the remaining books in

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

ways. Thus we have

$$\frac{k!(k - 1)!}{(k - n)!(n - 1)!} = k! \binom{k - 1}{n - 1}$$

ways to place the books. Of course the right hand side of that equation cries out for a combinatorial explanation. Here it is. Imagine lining up the k books in a row. Then there are $k - 1$ places in between them. Choose $n - 1$ of these places, and slide a piece of paper in there as a divider. Now put the books before the first divider on shelf one, and the books after divider i on shelf $i + 1$. This gives an arrangement of the books on the shelves so that every shelf has a book!

The assignment of which books go to which shelves of a bookcase is simply a function from the books to the shelves. But a function doesn't determine which book sits to the left of which others on the shelf, and this information is part of how the books are arranged on the shelves. In other words, the order in which the shelves receive their books matters. Our function must thus assign an ordered list of books to each shelf. We will call such a function

an ordered function. More precisely, an **ordered function** from a set S to a set T is a function that assigns an (ordered) list of elements of S to some, but not necessarily all, elements of T in such a way that each element of S appears on one and only one of the lists.[1] (Notice that although it is not the usual definition of a function from S to T, a function can be described as an assignment of subsets of S to some, but not necessarily all, elements of T so that each element of S is in one and only one of these subsets.) Thus the number of ways to place the books into the bookcase is the entry in the middle column of row 5 of our table. If in addition we require each shelf to get at least one book, we are discussing the entry in the middle column of row 6 of our table. An **ordered onto function** is one which assigns a list to each element of T. In Problem 112 you showed that the number of ordered functions from a k-element set to an n-element set is $\prod_{i=1}^{n}(n+i-1)$.

This product occurs frequently enough that it has a name; it is called the kth **rising factorial power** of n and is denoted by $n^{\overline{k}}$. It is read as "n to the k rising." (This notation is due to Don Knuth, who also suggested the notation for falling factorial powers.) We can summarize with a theorem that adds two more formulas for the number of ordered functions.

Theorem 3.1.5. *The number of ordered functions from a k-element set to an n-element set is*

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

Multisets

In the middle column of row 7 of our table, we are asking for the number of ways to distribute k identical objects (say ping-pong balls) to n distinct recipients (say children).

Activity 114. In how many ways may we distribute k identical books on the shelves of a bookcase with n shelves, assuming that any shelf can hold all the books?

Hint. What is the relationship between the number of ways to distribute identical books and the number of ways to distribute distinct books?

Solution. We saw that we could arrange k distinct books on n shelves in $\prod_{i=1}^{k}(n+i-1)$ ways. We partition these arrangements into blocks by putting two arrangements in the same block if we can get one from the other by permuting the books among themselves. Then the number of blocks is the number of ways to place identical books on the shelves. However, there are $k!$ arrangements per block, so there are

$$\frac{\prod_{i=1}^{k}(n+i-1)}{k!} = (n+k-1)^{\underline{k}} = \binom{n+k-1}{k}$$

---

[1]The phrase ordered function is not a standard one, because there is as yet no standard name for the result of an ordered distribution problem.

ways to arrange identical books.

Activity 115. A **multiset** chosen from a set S may be thought of as a subset with repeated elements allowed. For example the multiset of letters of the word Mississippi is $\{i, i, i, i, m, p, p, s, s, s, s\}$. To determine a multiset we must say how many times (including, perhaps, zero) each member of S appears in the multiset. The number of times an element appears is called its **multiplicity**. The size of a multiset chosen from S is the total number of times any member of S appears. For example, the size of the multiset of letters of Mississippi is 11. What is the number of multisets of size k that can be chosen from an n-element set?

Hint. Look for a relationship between a multiset of shelves and a way of distributing identical books to shelves

Solution. There is a bijection between arrangements of identical books on n shelves and multisets chosen from an n-element set: the multiplicity of element i is the number of books on shelf i. Thus we have $\binom{n+k-1}{k}$ ways to choose a k-element multiset from an n-element set by Problem 114.

Activity 116. Your answer in the previous problem should be expressible as a binomial coefficient. Since a binomial coefficient counts subsets, find a bijection between subsets of something and multisets chosen from a set S.

Hint. Note that $\binom{n+k-1}{k} = \binom{n+k-1}{n-1}$. So we have to figure out how choosing either k elements or $n-1$ elements out of $n+k-1$ elements constitutes the choice of a multiset. We really have no idea what set of $n+k-1$ objects to use, so why not use $[n+k-1]$? If we choose $n-1$ of these objects, there are k left over, the same number as the number of elements of our multiset. Since our multiset is supposed to be chosen from an n-element set, perhaps we should let the n-element set be $[n]$. From our choice of $n-1$ numbers, we have to decide on the multiplicity of 1 through n. For example with $n = 4$ and $k = 6$, we have $n + k - 1 = 9$. Here, shown with underlines, is a selection of $3 = n - 1$ elements from $[9]$: $1, 2, \underline{3}, 4\underline{5}, 6, 7, \underline{8}, 9$. How do the underlined elements give us a multiset of size 6 chosen from an $[4]$-element set? In this case, 1 has multiplicity 2, 2 has multiplicity 1, 3 has multiplicity 2, and 4 has multiplicity 1.

Solution. Note that $\binom{n+k-1}{k} = \binom{n+k-1}{n-1}$. We will show a bijection between ways of choosing $n - 1$ things our of $n + k - 1$ things and multisets. Namely, take $n + k - 1$ objects and line them up in a row. Choose $n - 1$ of them. Now let the multiplicity of element 1 of our multiset be the number of objects before the first thing we chose. if $1 < i < n$, let the multiplicity of element i of our multiset be the number of things between the $i - 1$th thing we chose and the ith thing we choose. Let the multiplicity of the nth element of our multiset be the number of objects after the last one we chose. Another way to say the essentially same thing is to make a list of $n + k - 1$ blank spaces. We choose k of them in which we put ones and $n - 1$ of them in which we put plus signs. Then the multiplicity of element 1 is the number of ones before

the first plus sign, the multiplicity of element n is the number of ones after the last plus sign and if $1 < i < n$, the multiplicity of element i is the number of ones between the $(i-1)$th plus sign and the ith plus sign.

Activity 117. How many solutions are there in nonnegative integers to the equation $x_1 + x_2 + \cdots + x_m = r$, where m and r are constants?

Hint.   A solution to the equations assigns a nonnegative number to each of $1, 2, \ldots, m$ so that the nonnegative numbers add to r. Does such an assignment have a combinatorial meaning?

Solution.   We can think of x_i as the multiplicity of element i of a multiset chosen from among m things. The total number of elements of the multiset will be r. Thus we have $\binom{m+r-1}{r}$ solutions.

Activity 118. In how many ways can we distribute k identical objects to n distinct recipients so that each recipient gets at least m?

Hint.   Can you think of some way of guaranteeing that each recipient gets m objects (assuming $k \geq mn$) right at the beginning of the process of passing the objects out?

Solution.   First give each recipient m objects. This leaves $k - mn$ identical objects to be distributed among n recipients, so we may do this in the number of ways to choose a $(k - mn)$-element multiset from n things. This is $\binom{n+k-mn+1}{k-mn}$, or $\binom{k-(m-1)n+1}{n+1}$.

Multisets explain the entry in the middle column of row 7 of our table of distribution problems.

Compositions of integers

Activity 119. In how many ways may we put k identical books onto n shelves if each shelf must get at least one book?

Hint.   We already know how to place k distinct books onto n distinct shelves so that each shelf gets at least one. Suppose we replace the distinct books with identical ones. If we permute the distinct books before replacement, does that affect the final outcome? There are other ways to solve this problem.

Solution.   In ⟨⟨Unresolved xref, reference "bookcaseeveryshelf"; check spelling or use "provisional" attribute⟩⟩ we showed that with k distinct books we could place the books in $k!\binom{k-1}{n-1}$ ways. We can partition these arangements of distinct books into blocks, where each block consists of all arrangements that we get just by permuting the books among themselves. Thus each block has $k!$ arrangements in it, and each arrangement corresponds to an arrangement of identical books. Thus there are $\binom{k-1}{n-1}$ ways to arrange identical books.

Activity 120. A **composition** of the integer k into n parts is a list of n positive integers that add to k. How many compositions are there of an integer k into n parts?

Hint. Do you see a relationship between compositions and something else we have counted already?

Solution. There is a bijection between compositions of k into n parts and arrangements of k identical books on n shelves so that each shelf gets a book. Namely, the number of books on shelf i is the ith element of the list. Thus the number of compositions of k into n parts is $\binom{k-1}{n-1}$.

Activity 121. Your answer in Problem 120 can be expressed as a binomial coefficient. This means it should be possible to interpret a composition as a subset of some set. Find a bijection between compositions of k into n parts and certain subsets of some set. Explain explicitly how to get the composition from the subset and the subset from the composition.

Hint. If we line up k identical books, how many adjacencies are there in between books?

Solution. If we line up k identical books, there are $k-1$ places in between two books. If we choose $n-1$ of these places and slip dividers into those places, then we have a first clump of books, a second clump of books, and so on. The ith element of our list is the number of books in the ith clump. Clearly using books is irrelevant; we could line up any k identical objects and make the same argument. Our bijection is between compositions and $(n-1)$-element subsets of the set of $k-1$ spaces between our objects.

Activity 122. Explain the connection between compositions of k into n parts and the problem of distributing k identical objects to n recipients so that each recipient gets at least one.

Solution. Since the recipients are distinct, we can think of them as a first recipient, a second, and so on. Given a composition of k into n parts, let the ith element of the list be the number of objects given to recipient number i.

The sequence of problems you just completed should explain the entry in the middle column of row 9 of our table of distribution problems.

Broken permutations and Lah numbers

Activity 123. In how many ways may we stack k distinct books into n identical boxes so that there is a stack in every box? The hints may suggest that you can do this problem in more than one way!

Hint. Imagine taking a stack of k books, and breaking it up into stacks to put into the boxes in the same order they were originally stacked. If you are going to use n boxes, in how many places will you have to break the stack up into smaller stacks, and how many ways can you do this?

Hint. How many different bookcase arrangements correspond to the same way of stacking k books into n boxes so that each box has at least one book?

Solution. We can make a list of the k distinct books in $k!$ ways. Then we have to choose $n-1$ of the $k-1$ places between the lists as the places where we

will break the list. However the order in which we list the boxes is irrelevant, so we have equivalence classes of $n!$ arrangements for each way of putting the books into boxes. Thus we can put the books in boxes in $k!\binom{k-1}{n-1}/n!$ ways.

Alternately, we can take the number of ways to put k books onto n bookshelves so that each shelf gets at least one, and then divide by the number of shelves factorial. That gives us $k!\binom{k-1}{n-1}/n!$ ways to arrange the books.

We can think of stacking books into identical boxes as partitioning the books and then ordering the blocks of the partition. This turns out not to be a useful computational way of visualizing the problem because the number of ways to order the books in the various stacks depends on the sizes of the stacks and not just the number of stacks. However this way of thinking actually led to the first hint in Problem 123. Instead of dividing a set up into nonoverlapping parts, we may think of dividing a *permutation* (thought of as a list) of our k objects up into n ordered blocks. We will say that a set of ordered lists of elements of a set S is a **broken permutation** of S if each element of S is in one and only one of these lists.[2] The number of broken permutations of a k-element set with n blocks is denoted by $L(k, n)$. The number $L(k, n)$ is called a **Lah Number** and, from our solution to Problem 123, is equal to $k!\binom{k-1}{n-1}/n!$.

The Lah numbers are the solution to the question "In how many ways may we distribute k distinct objects to n identical recipients if order matters and each recipient must get at least one?" Thus they give the entry in row 6 and column 6 of our table. The entry in row 5 and column 6 of our table will be the number of broken permutations with less than or equal to n parts. Thus it is a sum of Lah numbers.

We have seen that ordered functions and broken permutations explain the entries in rows 5 and 6 of our table.

In the next two sections we will give ways of computing the remaining entries.

3.2   Partitions and Stirling Numbers

We have seen how the number of partitions of a set of k objects into n blocks corresponds to the distribution of k distinct objects to n identical recipients. While there is a formula that we shall eventually learn for this number, it requires more machinery than we now have available. However there is a good method for computing this number that is similar to Pascal's equation. Now that we have studied recurrences in one variable, we will point out that Pascal's equation is in fact a *recurrence in two variables*; that is it lets us compute $\binom{n}{k}$ in terms of values of $\binom{m}{i}$ in which either $m < n$ or $i < k$ or both. It was the fact that we had such a recurrence and knew $\binom{n}{0}$ and $\binom{n}{n}$ that let us create Pascal's triangle.

---

[2]The phrase broken permutation is not standard, because there is no standard name for the solution to this kind of distribution problem.

Stirling Numbers of the second kind

We use the notation $S(k, n)$ to stand for the number of partitions of a k element set with n blocks. For historical reasons, $S(k, n)$ is called a *Stirling number of the second kind.*

Activity 124.

(a) In a partition of the set $[k]$, the number k is either in a block by itself, or it is not. How does the number of partitions of $[k]$ with n parts in which k is in a block with other elements of $[k]$ compare to the number of partitions of $[k-1]$ into n blocks? Find a two variable recurrence for $S(n, k)$, valid for n and k larger than one.

Solution. The number of partitions of $[k]$ into n parts in which k is in a block with other elements of $[k]$ is equal n times the number of partitions of $[k-1]$ into n blocks, because k could be in any of the n parts, and since it is in a block with other elements of $[k-1]$, removing it leaves a partition of $[k-1]$ into n blocks. The number of partitions of $[k]$ into n blocks in which k is in a block by itself is the number of partitions of $[k]$ into $n-1$ blocks, because you can get any such partition in by deleting the block containing k from a partition of $[k]$ in which k is in a block by itself. Thus $S(k, n)) = S(k-1, n-1) + nS(k-1, n)$.

(b) What is $S(k, 1)$? What is $S(k, k)$? Create a table of values of $S(k, n)$ for k between 1 and 5 and n between 1 and k. This table is sometimes called **Stirling's Triangle (of the second kind)** How would you define $S(k, n)$ for the nonnegative values of k and n that are not both positive? Now for what values of k and n is your two variable recurrence valid?

Solution. $S(k, 1) = 1$ and $S(k, k) = 1$.

$k\backslash n$	0	1	2	3	4	5	6
0	1	0	0	0	0	0	0
1	0	1	0	0	0	0	0
2	0	1	1	0	0	0	0
3	0	1	3	1	0	0	0
4	0	1	7	6	1	0	0
5	0	1	15	25	10	1	0

Table 3.2.1: Stirling's Triangle of the second kind

As you see in the table, we define $S(0, 0) = 1$, and $S(0, n)$ or $S(k, 0)$ to be 0 otherwise. This makes sense because for $n > 0$ there is no partition

of an empty set into n parts, and for $k > 0$ there is no partition of a k-element set into no parts, but saying there is one partition of the empty set into no parts allows us to use our recurrence to compute $S(1, 1)$. This makes our recurrence valid for all nonnegative values of k and n.

(c) Extend Stirling's triangle enough to allow you to answer the following question and answer it. (Don't fill in the rows all the way; the work becomes quite tedious if you do. Only fill in what you need to answer this question.) A caterer is preparing three bag lunches for hikers. The caterer has nine different sandwiches. In how many ways can these nine sandwiches be distributed into three identical lunch bags so that each bag gets at least one?

Solution.   We need $S(9, 3)$. Thus we need to extend our table for four more rows, but only out to the column labelled 3. These rows are 6,0,1,31,90, 7,0,1,63,301, 7,0,1,127,966, 8,0,1,255,3025, 9,0,1,511,9330. Thus there are 9330 ways to distribute the sandwiches into the lunch bags.

(d) The question in ⟨⟨Unresolved xref, reference "sandwiches"; check spelling or use "provisional" attribute⟩⟩ naturally suggests a more realistic question; in how many ways may the caterer distribute the nine sandwich's into three identical bags so that each bag gets exactly three? Answer this question. (Hint, what if the question asked about six sandwiches and two distinct bags? How does having identical bags change the answer?)

Solution. $\binom{9}{3}\binom{6}{3}\binom{3}{3}/3!$. First we choose three sandwiches for bag 1, then three for bag 2, and put the remainder in bag 3. However, it doesn't matter which bags the sandwiches are in so we have counted each partition 3! times.

Activity 125.

(a) In how many ways can we partition k items into n blocks so that we have k_i blocks of size i for each i? (Notice that $\sum_{i=1}^{k} k_i = n$ and $\sum_{i=1}^{k} i k_i = k$.) The sequence $k_1, k_2, \ldots, k_n$ is called the **type vector** of the partition.

Solution. $\frac{n!}{\prod_{i=1}^{n}(i!)^{k_i} k_i!}$. We can make a list in $n!$ ways, and then break it into first k_1 blocks of size 1, then k_2 blocks of size 2, k_3 blocks of size 3 up to k_n blocks of size n. But then we realize that we get the same partition if we permute the $i!$ elements of a block of size i and we get the same partition if we permute the k_i blocks of size i so we apply the quotient principle.

(b) Describe how to compute $S(n, k)$ in terms of quantities given by the formula you found in Problem 125.

Solution. We can find $S(n, k)$ by summing $\frac{n!}{\prod_{i=1}^{n} (i!)^{k_i} k_i!}$ over all type vectors $(k_1, k_2, \ldots, k_n)$ such that $k_1 + k_2 + \cdots + k_n = k$.

Activity 126. Find a recurrence similar to the one in Problem 124 for the Lah numbers $L(k, n)$.

Solution. $L(k, n)$ is the number of broken permutations of a k-element set into n parts. Either k is in an ordered block by itself or it is not. If it is, it can go after any of the $k - 1$ other elements, or it can go at the beginning of any of the n blocks. If it is not, deleting it gives a broken permutation of a $k - 1$-element set into $n-1$ blocks. Thus $L(k, n) = L(k-1, n-1)+(n+k-1)L(k, n)$.

Activity 127. (Relevant in Chapter B.1.) The total number of partitions of a k-element set is denoted by $B(k)$ and is called the k-th *Bell number*. Thus $B(1) = 1$ and $B(2) = 2$.

(a) Show, by explicitly exhibiting the partitions, that $B(3) = 5$.

Solution. The five partitions of $[3]$ are the sets $\{\{1\}, \{2\}, \{3\}\}$, $\{\{1, 2\}, \{3\}\}$, $\{\{1, 3\}, \{2\}\}$, $\{\{1\}, \{2, 3\}\}$, and $\{\{1, 2, 3\}\}$.

(b) Find a recurrence that expresses $B(k)$ in terms of $B(n)$ for $n < k$ and prove your formula correct in as many ways as you can.

Solution. If we delete the block containing k, we get a partition of a subset of $[k - 1]$. Thus $B(k)$ is the sum over all subsets of $[k - 1]$ of the number of partitions of that subset. This gives us $B(k) = \sum_{n=0}^{k-1} \binom{k-1}{n} B(n)$.

Alternatively, we can show by the same sort of argument that $S(k, n) = \sum_{i=0}^{k-1} \binom{k-1}{i} S(i, n - 1)$ and then use the fact that $B(k) = \sum_{n=0}^{k} S(k, n)$ to get the recurrence for $B(k)$.

(c) Find $B(k)$ for $k = 4, 5, 6$.

Solution.

$$B(4) = \binom{3}{0} B_0 + \binom{3}{1} B_1 + \binom{3}{2} B_2 + \binom{3}{3} B_3 = 1 + 3 + 3 \cdot 2 + 5 = 15$$

$$B(5) = \sum_{n=0}^{4} \binom{4}{n} B_n = 1 + 4 + 6 \cdot 2 + 4 \cdot 5 + 15 = 52$$

$$B(6) = \sum_{n=0}^{5} \binom{5}{n} B_n = 1 + 5 + 10 \cdot 2 + 10 \cdot 5 + 5 \cdot 15 + 52 = 203$$

Stirling Numbers and onto functions

Activity 128. Given a function f from a k-element set K to an n-element set, we can define a partition of K by putting x and y in the same block of the partition if and only if $f(x) = f(y)$. How many blocks does the partition have if f is onto? How is the number of functions from a k-element set onto an n-element set related to a Stirling number? Be as precise in your answer as you can.

Solution.   If f is onto, the number of blocks of the partition is n. The number of onto functions from a k-element set onto an n-element set is $S(k,n)n!$, because we have a one-to-one function from the blocks to the n-element set.

Activity 129. Each function from a k-element set K to an n-element set N is a function from K onto *some* subset of N. If J is a subset of N of size j, you know how to compute the number of functions that map onto J in terms of Stirling numbers. Suppose you add the number of functions mapping onto J over all possible subsets J of N. What simple value should this sum equal? Write the equation this gives you.

Solution.   The sum should equal the number of functions, n^k. Thus we get $\sum_{j=0}^{n} \binom{n}{j} S(k,j) j! = n^k$. By using the fact that $\binom{n}{j} = n^{\underline{j}}/j!$, this may be rewritten as $\sum_{j=0}^{n} n^{\underline{j}} S(k,j) = n^k$.

Activity 130. In how many ways can the sandwiches of $\langle\langle$Unresolved xref, reference "sandwiches"; check spelling or use "provisional" attribute$\rangle\rangle$ be placed into three distinct bags so that each bag gets at least one?

Solution. $S(9,3) \cdot 3! = 55,980$.

Activity 131. In how many ways can the sandwiches of $\langle\langle$Unresolved xref, reference "caterer2"; check spelling or use "provisional" attribute$\rangle\rangle$ be placed into distinct bags so that each bag gets exactly three?

Solution.   Choose three sandwiches for bag one in $\binom{9}{3}$ ways, three for bag two in $\binom{6}{3}$ ways and put the remainder in bag 3. This gives us $\binom{9}{3}\binom{6}{3} = \frac{9!}{3!3!3!} = 1680$ ways.

The $\frac{9!}{3!3!3!}$ suggests another solution. We can line up the sandwiches in 9! ways. We take the first three for bag one, the second three for bag two and the last three for bag 3. The order of the sandwiches in the bag does not matter though, so each there are 3!3!3! listings corresponding to each way of putting sandwiches in bags, giving us $\frac{9!}{3!3!3!}$ ways to put the sandwiches in bags.

Activity 132.

(a) How many functions are there from a k-element set K to a set $N = \{y_1, y_2, \ldots y_n\}$ so that y_i is the image of j_i elements of K for each i from 1 to n? This number is called a *multinomial coefficient* and denoted by

$$\binom{k}{j_1, j_2, \ldots, j_n}.$$

Solution. $\binom{k}{j_1, j_2, \ldots, j_n} = \frac{k!}{j_1! j_2! \cdots j_n!}$. We get this either as the product of binomial coefficients

$$\binom{k}{j_1} \binom{k - j_1}{j_2} \binom{k - j_1 - j_2}{j_3} \cdots \binom{j_n}{j_n},$$

or more elegantly, by lining up the elements of the domain in $k!$ ways, taking the first j_1 elements to y_1, the next j_2 elements to y_2 and so on. However the order of the j_i elements that go to y_i is irrelevant, so $j_1! j_2! \cdots j_n!$ lists all correspond to the same function, giving us $\frac{k!}{j_1! j_2!, \cdots j_n!}$ functions.

(b) Explain how to compute the number of functions from a k-element set K onto an n-element set N by using multinomial coefficients.

Solution. Add the multinomial coefficients $\binom{k}{j_1, j_2, \ldots, j_n}$ in which each j_i is different from zero. To see why, let $N = \{y_1, y_2, \ldots, y_n\}$ and note that we are counting functions that send at least one element of K to each element y_i.

Activity 133. What do multinomial coefficients have to do with expanding the kth power of a multinomial $x_1 + x_2 + \cdots + x_n$? This result is called the *multinomial theorem.*

Solution. When we use the distributive law to multiply out $(x_1 + x_2 + \cdots + x_n)^k$, we will get a sum of a bunch of terms of the form $x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}$ where $i_1 + i_2 + \cdots + i_n = k$. The terms with a given sequence $i_1, i_2, \ldots, i_n$ of exponents will arise from choosing, as we apply the distributive law over and over again, x_1 from i_1 of the factors, x_2 from i_2 of the factors, and so on. Thus the number of terms $x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}$ will be the number of ways to label i_1 of the factors with a 1, i_2 of the factors with a 2, ..., and i_n of the factors with an n. The number of ways to do this is a multinomial coefficient, as we now explain. This labeling gives us a function from $[k]$ to $[n]$ as follows. If factor i is labelled j we let $f(i) = j$. Further each function f from $[k]$ to $[n]$ gives us that maps i_j elements of $[k]$ to j will give us such a labelling. Thus the coefficient of $x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}$ will be the multinomial coefficient $\binom{k}{i_1, i_2, \ldots, i_n}$.

Stirling Numbers and bases for polynomials

Activity 134. Find a way to express n^k in terms of $k^{\underline{j}}$ for appropriate values j. You may use Stirling numbers if they help you. Notice that $x^{\underline{j}}$ makes sense for a numerical variable x (that could range over the rational numbers, the real numbers, or even the complex numbers instead of only the nonnegative integers, as we are implicitly assuming n does), just as x^j does. Find a way to express the power x^k in terms of the polynomials $x^{\underline{j}}$ for appropriate values of j and explain why your formula is correct.

Solution. In Problem 129, we saw that $\sum_{j=0}^{n} \binom{n}{j} S(k,j) j! = n^k$. Using the relationship between binomial coefficients and falling factorials, we may rewrite this as $\sum_{j=0}^{n} n^{\underline{j}} S(k,j) = n^k$. This expresses n^k in terms of $k^{\underline{j}}$. To be precise, we define $x^{\underline{j}}$ to be $x(x-1)\cdots(x-j+1)$. At first glance it looks like we could express $x^{\underline{j}}$ in terms of powers of x by simply substituting x for n in the equation $\sum_{j=0}^{n} n^{\underline{j}} S(k,j) = n^k$. However this gives us $\sum_{j=0}^{x} x^{\underline{j}} S(k,j) = x^k$, and we have never defined what we mean by a sum whose upper limit is a variable x. Thus we need to examine the equation $\sum_{j=0}^{n} n^{\underline{j}} S(k,j) = n^k$ to see if we can replace the n that is the upper limit of the sum with something else. Notice that $S(k,j) = 0$ when $j > k$. This means that if $k \leq j$, then $\sum_{j=0}^{n} n^{\underline{j}} S(k,j) = \sum_{j=0}^{k} n^{\underline{j}} S(k,j)$. Notice also that $n^{\underline{j}} = n(n-1)\cdots(n-j+1)$ is zero when $j > n$ because one of its factors is zero then. This implies that if $k > j$, then $\sum_{j=0}^{n} n^{\underline{j}} S(k,j) = \sum_{j=0}^{k} n^{\underline{j}} S(k,j)$. Therefore, regardless of the relative size of k and n, we have that $\sum_{j=0}^{n} n^{\underline{j}} S(k,j) = \sum_{j=0}^{k} n^{\underline{j}} S(k,j)$. Therefore

$$\sum_{j=0}^{k} n^{\underline{j}} S(k,j) = n^k. \tag{3.1}$$

It makes sense to write the polynomial $\sum_{j=0}^{k} x^{\underline{j}} S(k,j)$; this is simply a polynomial of degree k in the variable x. The expression $\sum_{j=0}^{k} x^{\underline{j}} S(k,j) - x^k$ is also a polynomial in x, but it might not be of degree k since we are subtracting a degree k term from a degree k polynomial. In fact for every positive integer value n of x, this polynomial is zero. That is, $\sum_{j=0}^{k} n^{\underline{j}} S(k,j) - n^k = 0$, which is just a restatement of Equation (3.1). But it is a fact of algebra that the number of solutions of a nontrivial polynomial equation is no more than the degree of the polynomial. Since the polynomial equation $\sum_{j=0}^{k} x^{\underline{j}} S(k,j) - x^k$ has infinitely many different solutions, it must be a trivial equation; that is $\sum_{j=0}^{k} x^{\underline{j}} S(k,j) - x^k$ must be zero for every real (and even every complex) number x. Thus $\sum_{j=0}^{k} x^{\underline{j}} S(k,j) = x^k$, and we have expressed x^k in terms of $x^{\underline{j}}$ for $j \leq k$.

You showed in Problem 134 how to get each power of x in terms of the falling factorial powers $x^{\underline{j}}$. Therefore every polynomial in x is expressible in terms of a sum of numerical multiples of falling factorial powers. Using the language of linear algebra, we say that the ordinary powers of x and the falling factorial powers of x each form a basis for the "space" of polynomials, and that the numbers $S(k,n)$ are "change of basis coefficients." If you are not familiar with linear algebra, a **basis** for the **space of polynomials**[3] is a set of polynomials such that each polynomial, whether in that set or not, can be expressed in one and only one way as a sum of numerical multiples of polynomials in the set.

Activity 135. Show that every power of $x + 1$ is expressible as a sum of numerical multiples of powers of x. Now show that every power of x (and

---

[3]The space of polynomials is just another name for the set of all polynomials.

thus every polynomial in x) is a sum of numerical multiples (some of which could be negative) of powers of $x + 1$. This means that the powers of $x + 1$ are a basis for the space of polynomials as well. Describe the change of basis coefficients that we use to express the binomial powers $(x+1)^n$ in terms of the ordinary x^j explicitly. Find the change of basis coefficients we use to express the ordinary powers x^n in terms of the binomial powers $(x + 1)^k$.

Solution. We know that

$$(x + 1)^n = \sum_{i=0}^{n} \binom{n}{i} x^i \tag{3.2}$$

from the binomial theorem. (In the way we stated the binomial theorem, instead of $\binom{n}{i}x^i$ we would have gotten $\binom{n}{i}x^{n-i}$. There are two ways to fix this. One is to observe that the coefficient of x^i in that expansion is $\binom{n}{n-i}$, which equals $\binom{n}{i}$. The other is to observe that when we expand $(1 + x)^n$ according to the binomial theorem we get exactly what we wrote on the right hand side in Equation (3.2).) Therefore every power of $x+1$ is expressible in terms of powers of x.

How do we express powers of x in terms of powers of $x + 1$? Some experimentation would help us guess how to do so; however there is a really nice trick that also isn't hard to see. Namely, we can write

$$x^n = (x + 1 - 1)^n = [(x + 1) - 1]^n = \sum_{i=0}^{n} \binom{n}{i} (x + 1)^{n-i}(-1)^i$$

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

This means that every power of x is expressible in terms of powers of $x+1$ and the change of basis coefficients to express powers of x in terms of powers of $x + 1$ are $(-1)^{n-i}\binom{n}{i}$ while the change of basis coefficients used to express powers of $x + 1$ in terms of powers of x are $\binom{n}{i}$.

Activity 136.

(a) By multiplication, we can see that every falling factorial polynomial can be expressed as a sum of numerical multiples of powers of x. In symbols, this means that there are numbers $s(k, n)$ (notice that this s is lower case, not upper case) such that we may write $x^{\underline{k}} = \sum_{n=0}^{k} s(k, n)x^n$. These numbers $s(k, n)$ are called Stirling Numbers of the first kind. By thinking algebraically about what the formula

$$x^{\underline{k}} = x^{\underline{k-1}}(x - k + 1) \tag{3.3}$$

means, we can find a recurrence for Stirling numbers of the first kind that gives us another triangular array of numbers called Stirling's triangle of the first kind. Explain why Equation (3.3) is true and use it to derive a recurrence for $s(k, n)$ in terms of $s(k - 1, n - 1)$ and $s(k - 1, n)$.

Solution.   Equation (3.3) is effectively the inductive step of an inductive definition of $x^{\underline{k}}$. With this equation we can write

$$\sum_{n=0}^{k} s(k,n)x^n = x^{\underline{k}} = x^{\underline{k-1}}(x-k+1)$$

$$= \left(\sum_{n=0}^{k-1} s(k-1,n)x^n\right)(x-k+1)$$

$$= \sum_{n=0}^{k-1} s(k-1,n)x^{n+1} - \sum_{n=0}^{k-1}(k-1)s(k-1,n)x^n$$

$$= \sum_{n=1}^{k} s(k-1,n-1)x^n - \sum_{n=0}^{k-1}(k-1)s(k-1,n)x^n.$$

Equating the coefficients of x^n in the first and last line of this equation, we get $s(k,n) = s(k-1,n-1) - (k-1)s(k-1,n)$, for n between 1 and $k-1$..

(b) Write down the rows of Stirling's triangle of the first kind for $k=0$ to 6.

	$k\backslash n$	0	1	2	3	4	5	6
	0	1	0	0	0	0	0	0
	1	0	1	0	0	0	0	0
Solution.	2	0	-1	1	0	0	0	0
	3	0	2	-3	1	0	0	0
	4	0	-6	11	-6	1	0	0
	5	0	24	-50	35	-10	1	0
	6	0	-120	274	-225	85	-15	1

By definition, the Stirling numbers of the first kind are also change of basis coefficients. The Stirling numbers of the first and second kind are change of basis coefficients from the falling factorial powers of x to the ordinary factorial powers, and vice versa.

Activity 137. Explain why every rising factorial polynomial $x^{\overline{k}}$ can be expressed in terms of the falling factorial polynomials $x^{\underline{n}}$. Let $b(k,n)$ stand for the change of basis coefficients that allow us to express $x^{\overline{k}}$ in terms of the falling factorial polynomials $x^{\underline{n}}$; that is, define $b(k,n)$ by the equations

$$x^{\overline{k}} = \sum_{n=0}^{k} b(k,n)x^{\underline{n}}.$$

(a) Find a recurrence for $b(k, n)$.

Solution.

$$\sum_{n=0}^{k} b(k, n)x^{\underline{n}} = x^{\overline{k}} = x^{\overline{k-1}}(x + k - 1)$$

$$= \left(\sum_{n=0}^{k-1} b(k - 1, n)x^{\underline{n}}\right)(x + k - 1)$$

$$= \sum_{n=0}^{k-1} b(k - 1, n)x^{\underline{n}}(x + k - 1)$$

$$= \sum_{n=0}^{k-1} b(k - 1, n)x^{\underline{n}}(x - n + n + k - 1)$$

$$= \sum_{n=0}^{k-1} b(k - 1, n)x^{\underline{n+1}} + (n + k - 1)b(k - 1, n)x^{\underline{n}}$$

$$= \sum_{n=1}^{k} b(k - 1, n - 1)x^{\underline{n}} + \sum_{n=0}^{k-1}(n + k - 1)b(k - 1, n)x^{\underline{n}}$$

Thus if n is not 0 or k, we equate the coefficient of $x^{\underline{n}}$ in the first line and last line to get

$$b(k, n) = b(k - 1, n - 1) + (n + k - 1)b(k - 1, n).$$

The trick of subtracting n and adding n in the middle of the computation was the result of wanting to mimic the way in which we increased the power on x in the solution to .

(b) Find a formula for $b(k, n)$ and prove the correctness of what you say in as many ways as you can.

Solution. We will answer the next part of the problem here! The recurrence for $b(k, n)$ is exactly the same as the recurrence for $L(k, n)$. Further, $b(0, 0) = 1 = L(0, 0)$, $b(0, n) = 0 = L(0, n)$ for $n > 0$, and $b(k, k) = L(k, k) = 1$. Thus $b(k, n)$ and $L(k, n)$ are identical. This and the formula from gives one proof that $b(k, n) = k!\binom{k-1}{n-1}/n!$.

A second proof that the change of basis coefficients are Lah numbers goes as follows. $n^{\overline{k}}$ counts the number of ordered functions from a k-element set to an n-element set. One way to determine such an ordered function is to take a broken permutation of the k-element set into n or fewer parts, and then take a one-to-one function from the parts to the n-element set. More informally we assign the parts of the broken permutation to distinct elements of the n-element set. If the broken

permutation has i parts, the number of ways to do this assignment is the number of i-element permutations of an n-element set, $n^{\underline{i}}$. Thus $n^{\overline{k}} = \sum_{i=0}^{n} L(k,i) n^{\underline{i}}$. However we can change the upper limit of the sum to k because $L(k,i)$ is zero when $i > k$ and $n^{\underline{i}}$ is zero when $i > n$. Now we change n to x because we have a polynomial equality which is valid for infinitely many of the values of the variable. This gives us $x^{\overline{k}} = \sum_{i=0}^{k} L(k,i) x^{\underline{i}}$. Thus $b(k,i) = L(k,i)$.

(c) Is $b(k,n)$ the same as any of the other families of numbers (binomial coefficients, Bell numbers, Stirling numbers, Lah numbers, etc.) we have studied?

Solution. As we said in our solution to the previous part, $b(k,n)$ is the Lah number $L(k,n)$.

(d) Say as much as you can (but say it precisely) about the change of basis coefficients for expressing $x^{\underline{k}}$ in terms of $x^{\overline{n}}$.

Solution. There are several ways of finding this relationship, but the most concise way is to observe that $(-x)^{\underline{k}} = (-1)^k x^{\overline{k}}$ and $(-x)^{\overline{k}} = (-1)^k x^{\underline{k}}$. This lets us write

$$(-x)^{\overline{k}} = \sum_{n=0}^{k} b(k,n)(-x)^{\underline{n}}$$

$$(-1)^k x^{\underline{k}} = \sum_{n=0}^{k} (-1)^n b(k,n) x^{\overline{n}}$$

$$x^{\underline{k}} = \sum_{n=0}^{k} (-1)^{n-k} b(k,n) x^{\overline{n}}.$$

Therefore the change of basis coefficients are $(-1)^{n-k} b(k,n)$.

3.3   Partitions of Integers

We have now completed all our distribution problems except for those in which both the objects and the recipients are identical. For example, we might be putting identical apples into identical paper bags. In this case all that matters is how many bags get one apple (how many recipients get one object), how many get two, how many get three, and so on. Thus for each bag we have a number, and the multiset of numbers of apples in the various bags is what determines our distribution of apples into identical bags. A multiset of positive integers that add to n is called a **partition** of n. Thus the partitions of 3 are 1+1+1, 1+2 (which is the same as 2+1) and 3. The number of partitions

of k is denoted by $P(k)$; in computing the partitions of 3 we showed that $P(3) = 3$. It is traditional to use Greek letters like λ (the Greek letter λ is pronounced LAMB duh) to stand for partitons; we might write $\lambda = 1, 1, 1$, $\gamma = 2, 1$ and $\tau = 3$ to stand for the three partitions we just described. We also write $\lambda = 1^3$ as a shorthand for $\lambda = 1, 1, 1$, and we write $\lambda \dashv 3$ as a shorthand for "λ is a partition of three."

Activity 138. Find all partitions of 4 and find all partitions of 5, thereby computing $P(4)$ and $P(5)$.

Solution. $4 = 1 + 1 + 1 + 1$, $4 = 2 + 1 + 1$, $4 = 2 + 1$, $4 = 3 + 1$, $4 = 4$, so that $P(4) = 5$. $5 = 1 + 1 + 1 + 1 + 1$, $5 = 2 + 1 + 1 + 1$, $5 = 2 + 2 + 1$, $5 = 3 + 1 + 1$, $5 = 3 + 2$, $5 = 4 + 1$, $5 = 5$, so that $P(5) = 7$.

The number of partitions of k into n parts

A **partition of the integer k into n parts** is a multiset of n positive integers that add to k. We use $P(k, n)$ to denote the number of partitions of k into n parts. Thus $P(k, n)$ is the number of ways to distribute k identical objects to n identical recipients so that each gets at least one.

Activity 139. Find $P(6, 3)$ by finding all partitions of 6 into 3 parts. What does this say about the number of ways to put six identical apples into three identical bags so that each bag has at least one apple?

Solution. $6 = 4 + 1 + 1$, $6 = 3 + 2 + 1$, $6 = 2 + 2 + 2$, so $P(6, 3) = 3$. This says there are three ways to put six identical apples into three identical bags so that each bag gets at least one apple.

Representations of partitions

Activity 140.

(a) How many solutions are there in the positive integers to the equation $x_1 + x_2 + x_3 = 7$ with $x_1 \geq x_2 \geq x_3$?

   Solution. This problem is asking for $P(7, 3)$ and suggests an organized way to go about finding it: list the partitions starting with the largest part and work down. $7 = 5 + 1 + 1$, $7 = 4 + 2 + 1$, $7 = 3 + 3 + 1$, $7 = 3 + 2 + 2$, and if we have three numbers that add to seven, one must be larger than two, so there are four such solutions.

(b) Explain the relationship between partitions of k into n parts and lists $x_1, x_2, ..., x_n$ of positive integers with $x_1 \geq x_2 \geq ... \geq x_n$. Such a representation of a partition is called a **decreasing list** representation of the partition.

   Solution. There is a bijection between partitions of k into n parts and lists, in nonincreasing order, of n positive integers that add to k, because

each multiset of numbers that adds to k can be listed in nonincreasing order in exactly one way.

Activity 141.

(a) Describe the relationship between partitions of k and lists or vectors $(x_1, x_2, \ldots, x_n)$ such that $x_1 + 2x_2 + \ldots kx_k = k$. Such a representation of a partition is called a **type vector** representation of a partition, and it is typical to leave the trailing zeros out of such a representation; for example $(2, 1)$ stands for the same partition as $(2, 1, 0, 0)$. What is the decreasing list representation for this partition, and what number does it partition?

Solution. The type vector of a partition of k is a way of representing the multiplicity function of the multiset of integers that adds to k. Thus there is a bijection between type vectors and partitions.

(b) How does the number of partitions of k relate to the number of partitions of $k + 1$ whose smallest part is one?

Solution. They are equal, because if we take two different partitions of $k - 1$ and increase the multiplicity of 1 in each (by one), they are still different; also if we take two different partitions of k that have parts of size one, and decrease the multiplicity of 1 in each (by one), they are still different.

When we write a partition as $\lambda = \lambda_1, \lambda_2, \ldots, \lambda_n$, it is customary to write the list of λ_is as a decreasing list. When we have a type vector $(t_1, t_2, \ldots, t_m)$ for a partition, we write either $\lambda = 1^{t_1} 2^{t_2} \cdots m^{t_m}$ or $\lambda = m^{t_m} (m-1)^{t_{m-1}} \cdots 2^{t_2} 1^{t_1}$. Henceforth we will use the second of these. When we write $\lambda = \lambda_1^{i_1} \lambda_2^{i_2} \cdots \lambda_n^{i_n}$, we will assume that $\lambda_i > \lambda_i + 1$.

Ferrers and Young Diagrams and the conjugate of a partition

The decreasing list representation of partitions leads us to a handy way to visualize partitions. Given a decreasing list $(\lambda_1, \lambda_2, \ldots \lambda_n)$, we draw a figure made up of rows of dots that has λ_1 equally spaced dots in the first row, λ_2 equally spaced dots in the second row, starting out right below the beginning of the first row and so on. Equivalently, instead of dots, we may use identical squares, drawn so that a square touches each one to its immediate right or immediately below it along an edge. See Figure 3.3.1 for examples. The figure we draw with dots is called the **Ferrers diagram** of the partition; sometimes the figure with squares is also called a Ferrers diagram; sometimes it is called a **Young diagram**. At this stage it is irrelevant which name we choose and which kind of figure we draw; in more advanced work the squares are handy

because we can put things like numbers or variables into them. From now on we will use squares and call the diagrams Young diagrams.

Figure 3.3.1: The Ferrers and Young diagrams of the partition (5,3,3,2)

Activity 142. Draw the Young diagram of the partition (4,4,3,1,1). Describe the geometric relationship between the Young diagram of (5,3,3,2) and the Young diagram of (4,4,3,1,1).
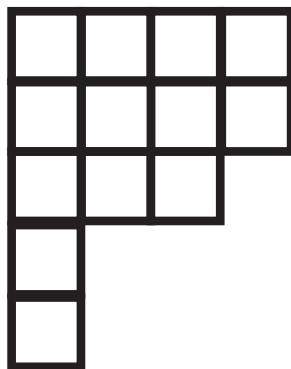
Solution.

Figure 3.3.2: Forming the conjugate of a Young diagram

We get the Young diagram of $(5, 3, 3, 2)$ by flipping the Young diagram of $(4, 4, 3, 1, 1)$ around a line that includes the diagonal of the upper left box; if we think of the top left corner of the diagram as being at the origin, we flip around the line $y = -x$.

Activity 143. The partition $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ is called the **conjugate** of the partition $(\gamma_1, \gamma_2, \ldots, \gamma_m)$ if we obtain the Young diagram of one from the Young diagram of the other by flipping one around the line with slope -1 that extends the diagonal of the top left square. See Figure 3.3.3 for an example.
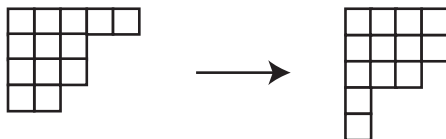

Figure 3.3.3: The Ferrers diagram the partition (5,3,3,2) and its conjugate.

What is the conjugate of (4,4,3,1,1)? How is the largest part of a partition related to the number of parts of its conjugate? What does this tell you about the number of partitions of a positive integer k with largest part m?

Solution. $(5, 3, 3, 2)$. The largest part of a partition equals the number of parts of its conjugate. The number of partitions of k with largest part m equals the number of partitions of k with m parts.

Activity 144.

(a) A partition is called *self-conjugate* if it is equal to its conjugate. Find a relationship between the number of self-conjugate partitions of k and the number of partitions of k into distinct odd parts.

   Solution.   The number of self-conjugate partitons of k equals the number of partitions of k with distinct odd parts. Here is a geometric description of a bijection from self conjugate partitions of k to partitions into distinct odd parts.

   Figure 3.3.4: Transforming a self-conjugate partition

   Take the top row and left column of squares of the Young diagram, and make them into one row in a new diagram. (Only include the square that is in both the row and column once.) Now take the remaining squares in the next row and column and make a new row of the Young diagram of the second partition with them. Continue this process with succeeding rows and columns, not using any squares you have already used. Because the first partition is self conjugate, the diagram has the same number of rows as columns and row i and column i have the same length. Because row i and column i share one square, and we only use that square once when we create a new row, each row we create has odd length. Thus we get a partition with the same number of squares, so it is a partition of k and each part is odd. The parts are distinct because when we take off the squares of a row and column, we reduce the number of squares in each row and column that remains. Given a partition of k into distinct odd parts, we use the fact that each row has a unique middle element, and each is shorter than the one above (by at least two squares) to reverse the process. Thus we have a bijection.

(b) Explain the relationship between the number of partitions of k into even parts and the number of partitions of k into parts of even multiplicity, i.e. parts which are each used an even number of times as in (3,3,3,3,2,2,1,1).

   Solution.   The number of partitions of k into even parts equals the number of partitions of parts of even multiplicity, because if we take the

Young diagram of a partition of k into even parts and conjugate it, the resulting diagram has columns of even length. Thus the difference in heights of two successive columns is an even number, but this difference is the multiplicity of one of the parts of the conjugate. Further the height of the last column of a partition is the multiplicity of the first part. Since the multiplicity of any part of a partition is either the difference in height of two successive columns of the Young diagram or the height of the last column, then each part of the conjugate has even multiplicity. This bijection can be reversed, because if all the differences in height of the columns are even and the height of the last column is even, then when we conjugate this partition, the last row will be an even length, and all differences in length of the rows will be even, so all the parts of the resulting partition will be even.

Activity 145.

(a) Show that the number of partitions of k into 4 parts equals the number of partitions of $3k$ (or $3k + 4$ or $3k - 4$) into 4 parts.

Solution. Think about putting the Young diagram of the partition into the upper left corner of a rectangle that is k units wide and four units high. Subdivide the rectangle into $4k$ squares of unit area. The Young diagram covers k of these squares. The uncovered squares are in rows of length $r_1 \leq r_2 \leq r_3 \leq r_4$. Thus if we list these lengths in the opposite order, we have a decreasing list representation of a partition of $3k$. Even r_1 will have to be positive, because the first part of the original partition will be at most $k - 3$. To get partitions of $3k + 4$, use a rectangle of width $k + 1$, and to get partitions of $3k - 4$, use a rectangle of width $k - 1$. Since the first row of the Young diagram has at most $k - 3$ squares, we will still have four nonzero parts in the partition that results.

(b) The idea of conjugation of a partition could be defined without the geometric interpretation of a Young diagram, but it would seem far less natural without the geometric interpretation. Another idea that seems much more natural in a geometric context is this. Suppose we have a partition of k into n parts with largest part m. Then the Young diagram of the partition can fit into a rectangle that is m or more units wide (horizontally) and n or more units deep. Suppose we place the Young diagram of our partition in the top left-hand corner of an m' unit wide and n' unit deep rectangle with $m' \geq m$ and $n' \geq n$, as in .

Figure 3.3.5: To complement the partition $(5, 3, 3, 2)$ in a 6 by 5 rectangle: enclose it in the rectangle, rotate, and cut out the original Young diagram.

Why can we interpret the part of the rectangle not occupied by our Young diagram, rotated in the plane, as the Young diagram of another partition? This is called the **complement** of our partition in the rectangle. What integer is being partitioned by the complement? What conditions on m' and n' guarantee that the complement has the same number of parts as the original one? What conditions on m' and n' guarantee that the complement has the same largest part as the original one? Is it possible for the complement to have both the same number of parts and the same largest part as the original one? If we complement a partition in an m' by n' box and then complement that partition in an m' by n' box again, do we get the same partition that we started with?

Solution.   If we fill the rectangle with unit squares, those not in the Young diagram of the original partition λ will fall into rows. The lengths of the rows are nonnegative, and are nondecreasing as we move down. Therefore, after we rotate through 180 degrees, these same rows will be listed in the opposite order, lined up along the left sides, and will have nonincreasing length. Thus they will be the Young diagram of a partition. The integer being partitioned will be $m'n' - k$. If $m' > m$ and $n' = n$. the two partitions will have the same number of parts, because we will have a nonzero number of empty squares at the end of each row of the Young diagram of λ. If $m' = m$ and $n' - n$ is the multiplicity of the largest part of λ, they will have the same number of parts. Otherwise, their numbers of parts will differ. If $n' > n$ and $m = m'$, then the two partitions will have the same largest part. If $n' = n$ and $m' - m$ is the smallest part of λ, then they will have the same largest part. Otherwise, their largest parts will differ. Thus for the two partitions to have the same number of parts, either $m' = m$ or $n' = n$. If $m' = m$ and they have the same largest part, then $n' > n$. But this is consistent with $n' - n$ being the multiplicity of the largest part of λ. Thus they can have the same number of parts and the same largest part if $m' = m$ and $n' - n$ is the multiplicity of the largest part of λ, or similarly if $n = n'$ and $m' - m$ is the smallest part of λ.

Activity 146.

(a) Suppose we take a partition of k into n parts with largest part m, complement it in the smallest rectangle it will fit into, complement the result in the smallest rectangle it will fit into, and continue the process until we get the partition 1 of one into one part. What can you say about the partition with which we started?

Solution.   Let us call the process of enclosing λ in the smallest rectangle possible and then forming the complement in that rectangle **encomplementation** (This is short for *en*closure and *complementation* and is not a standard term—there is no standard term for this operation.) and call the result of it the **encomplement** of λ. The result

of two encomplementations on the Young diagram of a partition is to remove all rows of maximum length and all columns of maximum length from the Young diagram. Thus the description of the result of an even number $2j$ of encomplementations is straightforward; we remove all the rows of the j largest distinct lengths and all columns of the j largest distinct lengths. So if an even number of encomplementations brings us to a partition with one block of size one, we should be able to describe the original partition fairly easily. To deal with the result of an odd number of encomplementations, we ask what happens if we encomplement just once. If the complement of λ in the smallest rectangle in which if fits has one square, then $\lambda = \lambda_1^{n_1}\lambda_1 - 1$. Thus we are asking for the partitions which, after an even number of encomplementations, give us either the partition with one block or a partition of the form $\lambda_1^{n_1}(\lambda_1 - 1)$. First we ask what kind of partition results in the second one after two encomplementations. If we get $\lambda_1^{n_1}(\lambda_1 - 1)$ from two encomplementations, the partition we started with had the form

$$\lambda_0^{n_0}(\lambda_1 + \lambda_2)^{n_1}(\lambda_1 + \lambda_2 - 1)\lambda_2^{n_2}.$$

If we get $\lambda_1^{n_1}(\lambda_1 - 1)$ from four encomplementations, then we started with a partition of the form

$$\lambda_{-1}^{n-1}(\lambda_0 + \lambda_3)^{n_0}(\lambda_1 + \lambda_2 + \lambda_3)^{n_1}(\lambda_1 + \lambda_2 + \lambda_3 - 1)(\lambda_2 + \lambda_3)^{n_3}\lambda_3^{n_3}.$$

From this pattern we see that a partition that results in $\lambda_1^{n_1}(\lambda_1 - 1)$ after $2j$ encomplementations has the form

$$\lambda_{1-j}^{n_{1-j}}\lambda_{2-j}^{n_{2-j}}\cdots\lambda_0^{n_0}\lambda_1'^{\,n_1}(\lambda_1' - 1)\lambda_2^{n_2}\cdots\lambda_{j+1}^{n_{j+1}}, \qquad (3.4)$$

where $\lambda_i > \lambda_{i+1}$ and $\lambda_0 > \lambda_1' > \lambda_2 + 1$.

On the other hand, a partition λ that results in 1 after two encomplementations has the form $\lambda_0^{n_0}(\lambda_1 + 1)\lambda_1^{n_1}$, and so a partition that results in 1 after j encomplementations is of the form

$$\lambda_{1-j}^{n_{1-j}}\lambda_{2-j}^{n_{2-j}}\cdots\lambda_0^{n_0}(\lambda_1 + 1)\lambda_1^{n_1}\lambda_2^{n_2}\cdots\lambda_j^{n_j}, \qquad (3.5)$$

where $\lambda_i > \lambda_{i+1}$ and $\lambda_0 > \lambda_1 + 1$. Thus a partition results in a single part of size 1 after some number of encomplementations if and only if it has the form of Equation (3.4) or Equation (3.5).

(b) Show that $P(k, n)$ is at least $\frac{1}{n!}\binom{k-1}{n-1}$.

Solution. The number of compositions of k into n parts is $\binom{k-1}{n-1}$. We can divide the compositions into blocks, where two compositions are in the same block if and only if one is a rearrangement of the other. Then the blocks correspond bijectively to partitions of k into n parts.

However we cannot compute the number of blocks by dividing by the number of compositions per block since the number of compositions per block ranges from 1 to $n!$. But then if we divide the number of compositions by $n!$ we will get a number less than the number of blocks because $n!$ times the number of blocks would be, by the sum principle, greater than the number of partitions.

With the binomial coefficients, with Stirling numbers of the second kind, and with the Lah numbers, we were able to find a recurrence by asking what happens to our subset, partition, or broken permutation of a set S of numbers if we remove the largest element of S. Thus it is natural to look for a recurrence to count the number of partitions of k into n parts by doing something similar. Unfortunately, since we are counting distributions in which all the objects are identical, there is no way for us to identify a largest element. However if we think geometrically, we can ask what we could remove from a Young diagram to get a Young diagram. Two natural ways to get a partition of a smaller integer from a partition of n would be to remove the top row of the Young diagram of the partition and to remove the left column of the Young diagram of the partition. These two operations correspond to removing the largest part from the partition and to subtracting 1 from each part of the partition respectively. Even though they are symmetric with respect to conjugation, they aren't symmetric with respect to the number of parts. Thus one might be much more useful than the other for finding a recurrence for the number of partitions of k into n parts.

Activity 147. In this problem we will study the two operations and see which one seems more useful for getting a recurrence for $P(k, n)$.

(a) How many parts does the remaining partition have when we remove the largest part (more precisely, we reduce its multiplicity by one) from a partition of k into n parts? What can you say about the number of parts of the remaining partition if we remove one from each part?

Solution. Reducing the multiplicity of the largest part by one reduces the number of parts by one. Removing 1 from each part reduces the number of parts by the multiplicity of the smallest part, so it strictly reduces the number of parts, perhaps even to one.

(b) If the largest part of a partition is j and we remove it, what integer is being partitioned by the remaining parts of the partition? If we remove one from each part of a partition of k into n parts, what integer is being partitioned by the remaining parts?

Solution. If we remove the largest part, the integer being partitioned is k minus the largest part. Thus it is a number less than k and at least $n - 1$. If we remove one from each part of the partition, the integer being partitioned is $k - n$.

(c) The last two questions are designed to get you thinking about how we can get a bijection between the set of partitions of k into n parts and some other set of partitions that are partitions of a smaller number. These questions describe two different strategies for getting that set of partitions of a smaller number or of smaller numbers. Each strategy leads to a bijection between partitions of k into n parts and a set of partitions of a smaller number or numbers. For each strategy, use the answers to the last two questions to find and describe this set of partitions into a smaller number and a bijection between partitions of k into n parts and partitions of the smaller integer or integers into appropriate numbers of parts.

Solution. Removing the largest part of a partition of k into n parts gives us a bijection between partitions of k into n parts and and partitions of numbers k' between $n - 1$ and $k - 1$ into $n - 1$ parts of size at most $k - k'$. (Removing the largest part gives us such a partition, and adjoining a part of size $k - k'$ to such a partition gives us a partition of k with n parts.)

Removing one from each part of a partition of k into n parts gives us a bijection between partitions of k into n parts and and partitions $k - n$ into n or fewer parts. (Removing one from each part of a partition of k into n parts gives us such a partition, and, given such a partition, we get a partition of k into n parts by adding one to each part and then creating enough parts of size 1 to have n parts.)

(d) Find a recurrence (which need not have just two terms on the right hand side) that describes how to compute $P(k, n)$ in terms of the number of partitions of smaller integers into a smaller number of parts.

Hint. One of the two sets of partitions of smaller numbers from the previous part is more amenable to finding a recurrence than the other.

Solution. The second bijection is to the set of partitions of $k - 1$ into n or fewer parts, and this makes the second bijection sound easier to work with. We get $P(k, n) = \sum_{i=1}^{n} P(k - n, i)$. The proof is the bijection we already described; in particular a partition of $k - n$ into i parts corresponds to the partition of k we get by adding one to each of the i parts and then creating $n - i$ parts of size one.

(e) What is $P(k, 1)$ for a positive integer k?

Solution. $P(k, 1) = 1$.

(f) What is $P(k, k)$ for a positive integer k?

Solution. $P(k, k) = 1$.

(g) Use your recurrence to compute a table with the values of $P(k, n)$ for values of k between 1 and 7.

$k\backslash n$	1	2	3	4	5	6	7
1	1	0	0	0	0	0	0
2	1	1	0	0	0	0	0
3	1	1	1	0	0	0	0
4	1	2	1	1	0	0	0
5	1	2	2	1	1	0	0
6	1	3	3	2	1	1	0
7	1	3	4	3	2	1	1

Solution. (label to the left of the table, at row 3)

(h) What would you want to fill into row 0 and column 0 of your table in order to make it consistent with your recurrence. What does this say $P(0, 0)$ should be? We usually define a sum with no terms in it to be zero. Is that consistent with the way the recurrence says we should define $P(0, 0)$?

Solution.   We would want to have $P(0, 0) = 1$ and $P(k, 0) = P(0, n) = 0$ for positive integer k or n. Since the sum of the empty multiset of positive integers is zero, this gives us one partition of the number zero, namely the empty multiset of positive integers.

It is remarkable that there is no known formula for $P(k, n)$, nor is there one for $P(k)$. This section was are devoted to developing methods for computing values of $P(n, k)$ and finding properties of $P(n, k)$ that we can prove even without knowing a formula. Some future sections will attempt to develop other methods.

We have seen that the number of partitions of k into n parts is equal to the number of ways to distribute k identical objects to n recipients so that each receives at least one. If we relax the condition that each recipient receives at least one, then we see that the number of distributions of k identical objects to n recipients is $\sum_{i=1}^{n} P(k, i)$ because if some recipients receive nothing, it does not matter which recipients these are. This completes rows 7 and 8 of our table of distribution problems. The completed table is shown in Table 3.3.6. There are quite a few theorems that you have proved which are summarized by Table 3.3.6. It would be worthwhile to try to write them all down!

The Twentyfold Way: A Table of Distribution Problems

k objects and conditions on how they are received	n recipients and mathem... Distinct
1. Distinct no conditions	n^k functions
2. Distinct Each gets at most one	$n^{\underline{k}}$ k-element permutations
3. Distinct Each gets at least one	$S(k,n)n!$ onto functions
4. Distinct Each gets exactly one	$k! = n!$ permutations
5. Distinct, order matters	$(k+n-1)^{\underline{k}}$ ordered functions
6. Distinct, order matters Each gets at least one	$(k)^{\underline{n}}(k-1)^{\underline{k-n}}$ ordered onto functions
7. Identical no conditions	$\binom{n+k-1}{k}$ multisets
8. Identical Each gets at most one	$\binom{n}{k}$ subsets
9. Identical Each gets at least one	$\binom{k-1}{n-1}$ compositions (n parts)
10. Identical Each gets exactly one	1 if $k = n$; 0 otherwise

Table 3.3.6: The number of ways to distribute k objects to n recipients, with restrictions on how the objects are received

Partitions into distinct parts

Often $Q(k, n)$ is used to denote the number of partitions of k into distinct parts, that is, parts that are different from each other.

Activity 148. Show that

$$Q(k, n) \leq \frac{1}{n!} \binom{k-1}{n-1}.$$

Solution. The number of compositions of k into n parts is $\binom{k-1}{n-1}$. Thus the number of compositions of k into n distinct parts is less than $\binom{k-1}{n-1}$. Divide the compositions of k into n distinct parts into blocks with two compositions in the same block if one is a rearrangement of the other. Because the parts are distinct, each block has $n!$ members. Further, there is a bijection between the blocks of this partition and the partitions of k into n distinct parts. Since the number of compositions of k into n distinct parts is less than $\binom{k-1}{n-1}$, the number of partitions of k into n distinct parts is less than $\frac{1}{n!} \binom{k-1}{n-1}$.

Activity 149. Show that the number of partitions of 7 into 3 parts equals the number of partitions of 10 into three distinct parts.

Solution. Given a partition λ of 7 in decreasing list form $\lambda_1, \lambda_2, \lambda_3$, if we add 0 to λ_3, 1 to λ_2 and 2 to λ_1 the resulting partition of 10 has distinct parts. If we take a partition λ' of 10 with distinct parts, then $\lambda'_1 \geq \lambda'_2 + 1$, $\lambda'_1 \geq \lambda'_2 + 2$, and $\lambda'_2 \geq \lambda'_3 + 1$. Therefore if we subtract 2 from λ'_1 to get λ_1, subtract 1 from λ'_2 to get λ_2 and let $\lambda_3 = \lambda'_3$, then $\lambda_1, \lambda_2, \lambda_3$ is the decreasing list representation of a partition of $10 - 3 = 7$. Thus there is a bijection between partitions of 7 into three parts and partitions of 10 into three distinct parts.

Activity 150. There is a relationship between $P(k, n)$ and $Q(m, n)$ for some other number m. Find the number m that gives you the nicest possible relationship.

Solution. The number of partitions of k into n parts is equal to the number of partitions of $k + \binom{n}{2}$ into n distinct parts. The bijection from partitions of k with n parts to partitions of $k + \binom{n}{2}$ with n distinct parts that proves this is the one that takes a partition $\lambda_n \lambda_{n-1} \cdots \lambda_1$ of k with $\lambda_i > \lambda_{i+1}$ and adds $i - 1$ to λ_i to get λ'_i. Then λ' is a partition into distinct parts, and the number it partitions is $k + 1 + 2 + \cdots + n - 1 = k + \binom{n}{2}$. The proof that it is a bijection is the fact that subtracting $n - i$ from the ith part of a partition of k into distinct parts yields a partition of k, because part $i + j$ is at least j smaller than part i.

Activity 151. Find a recurrence that expresses $Q(k, n)$ as a sum of $Q(k - n, m)$ for appropriate values of m.

Solution. Suppose λ is a partition of k into n distinct parts. Either 1 is one of those parts or not. Thus if we subtract 1 from each part, we either get

a partition of $k - n$ into $n - 1$ parts or a partition of $k - n$ into n parts. If λ and λ' are different partitions of k into n distinct parts, they go to different partitions. Each partition of $k - n$ into $n - 1$ parts or n parts can be gotten in this way from a corresponding partition of k into n parts. Thus we have a bijective correspondence and $Q(k, n) = Q(k - n, n - 1) + Q(k - n, n)$.

Activity 152. Show that the number of partitions of k into distinct parts equals the number of partitions of k into odd parts.

Solution. We start by giving a function from the set of partitions of k to the set of partitions of k with (only) odd parts. Clearly such a function cannot be one to one. Then we show that when restricted to the partitions with distinct parts it is one-to-one and onto by constructing an inverse. Given a partition $\lambda_1^{i_1} \lambda_2^{i_2} \cdots \lambda_n^{i_n}$, write $\lambda_i = \gamma_i 2^{k_i}$, where γ_i is odd. (Thus 2^{k_i} is the highest power of 2 that is a factor of λ_i, so it is 1 if λ_i is odd.). It is possible that $\gamma_i = \gamma_j$, for example if $\lambda_i = 36$ and $\lambda_j = 18$, then $\gamma_i = \gamma_j = 9$. We construct a new partition π whose parts are the numbers γ_j as follows: Given an odd number p, let the multiplicity $m(p)$ of p in π be $\sum_{j:\gamma_j=m} 2^{k_j}$. Thus $\sum_{p:m(p)\neq 0} m(p)p = k$. Therefore, π is a partition of k whose parts are all odd.

Now consider a partition π of k whose parts are all odd. Let $\pi = \pi_1^{r_1} \pi_2^{r_2} \cdots \pi_t^{r_t}$, with $\pi_i > \pi_{i+1}$. (In terms of the multiplicity function m, $m(\pi_i) = r_i$, and $\sum_{i=1}^t r_i \pi_i = k$.) We are going to write the binary expansion of each r_i as $r_i = \sum_{j=0}^{\lfloor \log_2 r_i \rfloor} 2^j a_{ij}$, where a_{ij} is 1 if 2^j appears in the binary expansion of r_i, and 0 otherwise. All of the numbers $\pi_i 2^{j a_{ij}}$ are distinct, because a power of two times one odd number cannot equal a power of two times another odd number. The numbers $\pi_i 2^{j a_{ij}}$ add to k, so they are the parts of a partition π' of k into distinct parts. When we apply the function constructed in the first part of the solution to π', we get π, so the correspondence between π and π' is a bijection.

Activity 153. Euler showed that if $k \neq \frac{3j^2+j}{2}$, then the number of partitions of k into an even number of distinct parts is the same as the number of partitions of k into an odd number of distinct parts. Prove this, and in the exceptional case find out how the two numbers relate to each other.

Solution. This solution is taken largely from the book *Introduction to Combinatorics* by Ioan Tomescu (published in London by Collet's in 1975). Tomescu calls a collection of rows in a Young diagram a "trapezoid" if each row contains one less cell than the row above and the number of cells in the rows above and below the trapezoid differ by two or more from the number of cells in rows of the trapezoid. Thus in (8,6,5,4,2,1) we have 3 trapezoids, the first row, the next three rows, and the last two. Since we are dealing with partitions with distinct parts, we don't have to worry about how two equal rows affect the definition of a trapezoid. We will describe a way to transform a partition with an even number of distinct parts into a partition with an odd number of distinct parts and vice versa.

First we describe a transformation on Young diagrams. Here is the first part of the description. Suppose the smallest part m of λ is less than or equal

to the number j of rows in the top trapezoid. Suppose further that if we have only one trapezoid, then $j > m$. Then we construct a partition with one less part by adding 1 to each of the m largest parts and discarding the part m. We still have a diagram for a partition of the same integer, but now the parity of the number of parts has changed, and we *may* have increased the number of trapezoids by 1. The smallest part will now be larger than the number (now m) of rows in the top trapezoid. (Notice that the construction would not work if we had only one trapezoid and $j = m$ because we would first remove one row of the trapezoid and thus have no row to which to attach one of our squares.)

Here is the second part of the description of the transformation. Suppose now that m is larger than the number j of rows of the top trapezoid in the Young diagram. Suppose also that the Young diagram has at least two trapezoids or it has one trapezoid and $j \geq m - 2$. Take one square from each of the j rows of the top trapezoid (which is the whole diagram if there is only one trapezoid) and also add a row of j squares at the bottom of the diagram. (Since $m > j$, this gives us a Young diagram of a partition of the same integer into distinct parts.) The parity of the number of rows has changed, and now the number of rows of the top trapezoid is at least as large as the smallest part of the partition. (Note, two previously distinct trapezoids may have joined together to form one on top.) (Notice that if we have one trapezoid and $j = m + 1$, then the construction yields a partition with two equal parts, which is why we made the special assumption above.) Now let T be the transformation described by the two constructions above. Its domain is all Young diagrams except those with one trapezoid and $m \leq j \leq m + 1$. T^2 is the identity, and so T is a bijection. When restricted to partitions with an odd number of parts, T gives partitions with an even number of parts, so on its domain it gives a bijection between partitions with an even number of parts and partitions with an odd number of parts.

If $m = j$ and the diagram has just one trapezoid, then the diagram has $\frac{3j^2 - j}{2}$ squares, and if $m = j + 1$ and the diagram has just one trapezoid, then the diagram has $\frac{3j^2 + j}{2}$ squares. Thus if $k \neq \frac{3j^2 \pm j}{2}$, the number of partitions of k into distinct even parts equals the number of partitions of k into distinct odd parts.

If $k = \frac{3j^2 \pm j}{2}$ and j is even, then there is one diagram of a partition of k that is not in the domain of the bijection and has an even number of rows, so in this case there will be one more partition with an even number of parts than with an odd number. If $k = \frac{3j^2 \pm j}{2}$ and j is odd, there is one diagram with an odd number of rows not in the domain and so in this case there is one more partition with an odd number of parts than with an even number. This completes the exceptional cases of the problem.

3.4   Supplementary Problems

1. Answer each of the following questions with n^k, k^n, $n!$, $k!$, $\binom{n}{k}$, $\binom{k}{n}$, $n^{\underline{k}}$,

$k^{\underline{n}}$, $n^{\overline{k}}$, $k^{\overline{n}}$, $\binom{n+k-1}{k}$, $\binom{n+k-1}{n}$, $\binom{n-1}{k-1}$, $\binom{k-1}{n-1}$, or "none of the above".

(a) In how many ways may we pass out k identical pieces of candy to n children? $\binom{n+k-1}{k}$

(b) In how many ways may we pass out k distinct pieces of candy to n children? n^k

(c) In how many ways may we pass out k identical pieces of candy to n children so that each gets at most one? (Assume $k \le n$.) $\binom{n}{k}$.

(d) In how many ways may we pass out k distinct pieces of candy to n children so that each gets at most one? (Assume $k \le n$.) $n^{\underline{k}}$

(e) In how many ways may we pass out k distinct pieces of candy to n children so that each gets at least one? (Assume $k \ge n$.) None of the above.

(f) In how many ways may we pass out k identical pieces of candy to n children so that each gets at least one? (Assume $k \ge n$.) $\binom{k-1}{n-1}$

2. The neighborhood betterment committee has been given r trees to distribute to s families living along one side of a street.

(a) In how many ways can they distribute all of them if the trees are distinct, there are more families than trees, and each family can get at most one? $s^{\underline{r}}$

(b) In how many ways can they distribute all of them if the trees are distinct, any family can get any number, and a family may plant its trees where it chooses? s^r

(c) In how many ways can they distribute all the trees if the trees are identical, there are no more trees than families, and any family receives at most one? $\binom{s}{r}$

(d) In how many ways can they distribute them if the trees are distinct, there are more trees than families, and each family receives at most one (so there could be some leftover trees)? $\sum_{k=0}^{s} \binom{s}{k} r^{\underline{k}}$ or $\sum_{k=0}^{s} s^{\underline{k}} \binom{r}{k}$

(e) In how many ways can they distribute all the trees if they are identical and anyone may receive any number of trees? $\binom{r+s-1}{r}$

(f) In how many ways can all the trees be distributed and planted if the trees are distinct, any family can get any number, and a family must plant its trees in an evenly spaced row along the road? $s^{\overline{r}} = (r+s-1)^{\underline{r}}$

(g) Answer the question in Part 2.f assuming that every family must get a tree. $r!\binom{r-1}{s-1}$

(h) Answer the question in Part 2.e assuming that each family must get at least one tree. $\binom{r-1}{s-1}$

3. In how many ways can n identical chemistry books, r identical mathematics books, s identical physics books, and t identical astronomy books be arranged on three bookshelves? (Assume there is no limit on the number of books per shelf.) $\frac{(n+r+s+t+2)!}{n!r!s!t!2!}$

4. (interesting) One formula for the Lah numbers is

$$L(k, n) = \binom{k}{n}(k-1)^{\underline{k-n}}$$

Find a proof that explains this product. First choose the n elements which will be the first member of the part they lie in. (This, in effect, labels the n parts.) Then assign the remaining $k - n$ elements to their parts by making an ordered function of $n - k$ objects to n recipients in $(n + (k - n) - 1)^{\underline{k-n}} = (k-1)^{\underline{k-n}}$ ways.

5. What is the number of partitions of n into two parts? $n/2$ if n is even and $(n-1)/2$ if n is odd, equivalently, $\lfloor n/2 \rfloor$

6. Show that the number of partitions of k into n parts of size at most m equals the number of partitions of $mn - k$ into no more than n parts of size at most $m - 1$. If we take the complement of the Young diagram of a partition of k into n parts of size at most m in an rectangle with n rows and m columns, the number we partition will be $mn - k$, and we will have no more than n parts, each of size at most $m - 1$. And if we take the complement of a partition of this second kind in the same rectangle, we will get a partition of the first kind.

7. Show that the number of partitions of k into parts of size at most m is equal to the number of partitions of of $k + m$ into m parts. Given the first kind of partition, take the conjugate (giving a partition of k into at most m parts), add one to each part, and then add enough parts of size 1 to get a total of m parts. It is straightforward that this process can be reversed.

8. You can say something pretty specific about self-conjugate partitions of k into distinct parts. Figure out what it is and prove it. With that, you should be able to find a relationship between these partitions and partitions whose parts are consecutive integers, starting with 1. What is that relationship? In a self-conjugate partition, the number of parts is the size of the largest part. If these parts are distinct, this means that each number between 1 and the largest part appears once as a part. That is, the parts are a list of consecutive integers, starting with 1.

9. What is $s(k, 1)$? Since $s(k, 1)$ is the coefficient of x^1 in

$$x^{\underline{k}} = x(x-1)(x-2) \cdot (x - (k-1)),$$

it is $(-1)^{k-1}(k-1)!$.

10. Show that the Stirling numbers of the second kind satisfy the recurrence

$$S(k,n) = \sum_{i=1}^{k} S(k-i, n-1)\binom{n-1}{i-1}.$$

A partition of $[k]$ into n blocks has a block containing k. If this block has size i, when you remove it, you get a partition of a set of size $k-i$ into $n-1$ blocks. The number of possible sets of size i containing k is $\binom{k-1}{i-1}$, and i can be any number between 1 and k. Each partition of k into n blocks may be constructed exactly once by first choosing the block containing k and then partitioning the remaining elements into $n-1$ blocks. This proves the formula.

11. (interesting) Let $c(k,n)$ be the number of ways for k children to hold hands to form n circles, where one child clasping his or her hands together and holding them out to form a circle is considered a circle. Find a recurrence for $c(k,n)$. Is the family of numbers $c(k,n)$ related to any of the other families of numbers we have studied? If so, how? The kth child is either in a circle by him/her self, and there are $c(k-1, n-1)$ ways for this to happen, or is in a circle with some other children. In the second case child i can be to the immediate right of any of the other $k-1$ children, so there are $(k-1)c(k-1, n)$ ways for this to happen. Thus $c(k,n) = c(k-1, n-1) + (k-1)c(k-1, n)$. This recurrence is almost the same as the recurrence for $s(k,n)$, except it has a plus sign where the recurrence for the Stirling numbers of the first kind has a minus sign. Further $c(k,1) = (k-1)!$ and $c(k,k) = 1$, which agrees, except for sign, with the Stirling numbers of the first kind. If we experiment with applying the recurrence, we see that whenever we use it to compute $c(k,n)$, we get that $c(k,n) = |s(k,n)|$. It is now straightforward to prove by induction that $c(k,n) = |s(k,n)|$.

# Chapter 4

Generating Functions

4.1 The Idea of Generating Functions

Suppose you are going to choose three pieces of fruit from among apples, pears and bananas for a snack. We can symbolically represent all your choices as

$$🍎🍎🍎 + 🍐🍐🍐 + 🍌🍌🍌 + 🍎🍎🍐 + 🍎🍎🍌 + 🍎🍐🍐 + 🍐🍐🍌 + 🍎🍌🍌 + 🍐🍌🍌 + 🍎🍐🍌.$$

Here we are using a picture of a piece of fruit to stand for taking a piece of that fruit. Thus 🍎 stands for taking an apple, 🍎🍐 for taking an apple and a pear, and 🍎🍎 for taking two apples. You can think of the plus sign as standing for the "exclusive or," that is, 🍎 + 🍌 would stand for "I take an apple or a banana but not both." To say "I take both an apple and a banana," we would write 🍎🍌. We can extend the analogy to mathematical notation by condensing our statement that we take three pieces of fruit to

$$🍎^3 + 🍐^3 + 🍌^3 + 🍎^2🍐 + 🍎^2🍌 + 🍎🍐^2 + 🍐^2🍌 + 🍎🍌^2 + 🍐🍌^2 + 🍎🍐🍌.$$

In this notation 🍎^{3} stands for taking a multiset of three apples, while 🍎^{2}🍌 stands for taking a multiset of two apples and a banana, and so on. What our notation is really doing is giving us a convenient way to list all three element multisets chosen from the set $\{🍎, 🍐, 🍌\}$.[1] Suppose now that we plan to choose between one and three apples, between one and two pears, and between one and two bananas. In a somewhat clumsy way we could describe our fruit selections as

$$🍎🍐🍌 + 🍎^2🍐🍌 + \cdots + 🍎^2🍐^2🍌 + \cdots + 🍎^2🍐^2🍌^2 + 🍎^3🍐🍌 + \cdots + 🍎^3🍐^2🍌 + \cdots + 🍎^3🍐^2🍌^2. \tag{4.1}$$

Activity 154. Using an A in place of the picture of an apple, a P in place of the picture of a pear, and a B in place of the picture of a banana, write

---

[1]This approach was inspired by George Pólya's paper "Picture Writing," in the December, 1956 issue of the *American Mathematical Monthly*, page 689. While we are taking a somewhat more formal approach than Pólya, it is still completely in the spirit of his work.

out the formula similar to Formula (4.1) without any dots for left out terms. (You may use pictures instead of letters if you prefer, but it gets tedious quite quickly!) Now expand the product $(A + A^2 + A^3)(P + P^2)(B + B^2)$ and compare the result with your formula.

Solution. $APB + APB^2 + AP^2B + AP^2B^2 + A^2PB + A^2PB^2 + A^2P^2B + A^2P^2B^2 + A^3PB + A^3PB^2 + A^3P^2B + A^3P^2B^2$

$$(A + A^2 + A^3)(P + P^2)(B + B^2)$$
$$= APB + APB^2 + AP^2B + AP^2B^2 + A^2PB + A^2PB^2 + A^2P^2B$$
$$+ \qquad\qquad\qquad\qquad\qquad\qquad\qquad A^2P^2B^2 + A^3PB + A^3$$

We get the same expression in both cases.

Activity 155. Substitute x for all of A, P and B (or for the corresponding pictures) in the formula you got in Problem 154 and expand the result in powers of x. Give an interpretation of the coefficient of x^n.

Solution. $x^3 + 3x^4 + 4x^5 + 3x^6 + x^7$. There is one way to choose three pieces of fruit, there are three ways to choose four pieces, four ways to chose 5 pieces, three ways to choose 6 pieces , and there is one way to choose 7 pieces of fruit. The coefficient of x^n is the number of ways to choose n pieces of fruit.

If we were to expand the formula

$$(🍎 + 🍎^2 + 🍎^3)(🍐 + 🍐^2)(🍌 + 🍌^2). \qquad (4.2)$$

we would get Formula (4.1). Thus formula (4.1) and formula (4.2) each describe the number of multisets we can choose from the set $\{🍎, 🍐, 🍌\}$ in which 🍎 appears between 1 and three times and 🍐, and 🍌 each appear once or twice. We interpret Formula (4.1) as describing each individual multiset we can choose, and we interpret Formula (4.2) as saying that we first decide how many apples to take, and then decide how many pears to take, and then decide how many bananas to take. At this stage it might seem a bit magical that doing ordinary algebra with the second formula yields the first, but in fact we could define addition and multiplication with these pictures more formally so we could explain in detail why things work out. However since the pictures are for motivation, and are actually difficult to write out on paper, it doesn't make much sense to work out these details. We will see an explanation in another context later on.

Picture functions

As you've seen, in our descriptions of ways of choosing fruits, we've treated the pictures of the fruit as if they are variables. You've also likely noticed that it is much easier to do algebraic manipulations with letters rather than pictures, simply because it is time consuming to draw the same picture over and over again, while we are used to writing letters quickly. In the theory of generating functions, we associate variables or polynomials or even power

series with members of a set. There is no standard language describing how we associate variables with members of a set, so we shall invent some. By a **picture** of a member of a set we will mean a variable, or perhaps a product of powers of variables (or even a sum of products of powers of variables). A function that assigns a picture $P(s)$ to each member s of a set S will be called a **picture function** . The **picture enumerator** for a picture function P defined on a set S will be

$$E_P(S) = \sum_{s:s\in S} P(s).$$

We choose this language because the picture enumerator lists, or enumerates, all the elements of S according to their pictures. Thus Formula (4.1) is the picture enumerator the set of all multisets of fruit with between one and three apples, one and two pears, and one and two bananas.

Activity 156. How would you write down a polynomial in the variable A that says you should take between zero and three apples?

Solution. $A^0 + A^1 + A^2 + A^3$.

Activity 157. How would you write down a picture enumerator that says we take between zero and three apples, between zero and three pears, and between zero and three bananas?

Solution. $(A^0 + A^1 + A^2 + A^3)(P^0 + P^1 + P^2 + P^3)(B^0 + B^1 + B^2 + B^3)$.

Activity 158. (Used in Chapter 6.) Notice that when we used A^2 to stand for taking two apples, and P^3 to stand for taking three pears, then we used the product A^2P^3 to stand for taking two apples and three pears. Thus we have chosen the picture of the ordered pair (2 apples, 3 pears) to be the product of the pictures of a multiset of two apples and a multiset of three pears. Show that if S_1 and S_2 are sets with picture functions P_1 and P_2 defined on them, and if we define the picture of an ordered pair $(x_1, x_2) \in S_1 \times S_2$ to be $P((x_1, x_2)) = P_1(x_1)P_2(x_2)$, then the picture enumerator of P on the set $S_1 \times S_2$ is $E_{P_1}(S_1)E_{P_2}(S_2)$. We call this the *product principle for picture enumerators*.

Solution.

$$E_P(S_1 \times S_2) = \sum_{(x_1,x_2)\in S_1\times S_2} P(x_1)P(x_2)$$

$$= \sum_{x_1:x_1\in S_1} \sum_{x_2:x_2\in S_2} P(x_1)P(x_2)$$

$$= \sum_{x_1\in S_1} P(x_1) \sum_{x_2\in S_2} P(x_2)$$

$$= \sum_{x_1\in S_1} P(x_1)E_{P_2}(S_2)$$

$$= \left(\sum_{x_1 \in S_1} P(x_1) \right) E_{P_2}(S_2)$$

$$= E_{P_1}(S_1) E_{P_2}(S_2)$$

Generating functions

Activity 159. Suppose you are going to choose a snack of between zero and three apples, between zero and three pears, and between zero and three bananas. Write down a polynomial in one variable x such that the coefficient of x^n is the number of ways to choose a snack with n pieces of fruit.

Hint. Substitute something for A, P and B in your formula from Activity 157.

Solution. $(1 + x + x^2 + x^3)^3$

Activity 160. Suppose an apple costs 20 cents, a banana costs 25 cents, and a pear costs 30 cents. What should you substitute for A, P, and B in Problem 157 in order to get a polynomial in which the coefficient of x^n is the number of ways to choose a selection of fruit that costs n cents?

Solution.   Substitute x^{20} for A, x^{25} for B and x^{30} for P.

Activity 161. Suppose an apple has 40 calories, a pear has 60 calories, and a banana has 80 calories. What should you substitute for A, P, and B in Problem 157 in order to get a polynomial in which the coefficient of x^n is the number of ways to choose a selection of fruit with a total of n calories?

Solution.   Substitute x^{40} for A, x^{60} for P, and x^{80} for B.

Activity 162. We are going to choose a subset of the set $[n] = \{1, 2, \dots, n\}$. Suppose we use x_1 to be the picture of choosing 1 to be in our subset. What is the picture enumerator for either choosing 1 or not choosing 1? Suppose that for each i between 1 and n, we use x_i to be the picture of choosing i to be in our subset. What is the picture enumerator for either choosing i or not choosing i to be in our subset? What is the picture enumerator for all possible choices of subsets of $[n]$? What should we substitute for x_i in order to get a polynomial in x such that the coefficient of x^k is the number of ways to choose a k-element subset of n? What theorem have we just reproved (a special case of)?

Solution.   The picture enumerator for choosing 1 or not choosing 1 is $x_1 + 1$. The picture enumerator for choosing or not choosing i is $x_i + 1$. The picture enumerator for choosing all possible subsets of $[n]$ is $(x_1+1)(x_2+1)\cdots(x_n+1)$. We should substitute x for x_i, thus getting $(1+x)^n$. Since the number of ways to choose an n-element subset is $\binom{n}{k}$, we have just proved the version of the binomial theorem that says

$$(x + 1)^n = \sum_{i=0}^{n} \binom{n}{i} x^i.$$

In Problem 162 we see that we can think of the process of expanding the polynomial $(1+x)^n$ as a way of "generating" the binomial coefficients $\binom{n}{k}$ as the coefficients of x^k in the expansion of $(1+x)^n$. For this reason, we say that $(1+x)^n$ is the "generating function" for the binomial coefficients $\binom{n}{k}$. More generally, the **generating function** for a sequence a_i, defined for i with $0 \leq i \leq n$ is the expression $\sum_{i=0}^{n} a_i x^i$, and the **generating function** for the sequence a_i with $i \geq 0$ is the expression $\sum_{i=0}^{\infty} a_i x^i$. This last expression is an example of a power series. In calculus it is important to think about whether a power series converges in order to determine whether or not it represents a function. In a nice twist of language, even though we use the phrase generating function as the name of a power series in combinatorics, we don't require the power series to actually represent a function in the usual sense, and so we don't have to worry about convergence.[2] Instead we think of a power series as a convenient way of representing the terms of a sequence of numbers of interest to us. The only justification for saying that such a representation is convenient is because of the way algebraic properties of power series capture some of the important properties of some sequences that are of combinatorial importance. The remainder of this chapter is devoted to giving examples of how the algebra of power series reflects combinatorial ideas.

Because we choose to think of power series as strings of symbols that we manipulate by using the ordinary rules of algebra and we choose to ignore issues of convergence, we have to avoid manipulating power series in a way that would require us to add infinitely many real numbers. For example, we cannot make the substitution of $y + 1$ for x in the power series $\sum_{i=0}^{\infty} x^i$, because in order to interpret $\sum_{i=0}^{\infty} (y+1)^i$ as a power series we would have to apply the binomial theorem to each of the $(y+1)^i$ terms, and then collect like terms, giving us infinitely many ones added together as the coefficient of y^0, and in fact infinitely many numbers added together for the coefficient of any y^i. (On the other hand, it would be fine to substitute $y + y^2$ for x. Can you see why?)

Power series

For now, most of our uses of power series will involve just simple algebra. Since we use power series in a different way in combinatorics than we do in calculus, we should review a bit of the algebra of power series.

Activity 163. In the polynomial $(a_0 + a_1 x + a_2 x^2)(b_0 + b_1 x + b_2 x^2 + b_3 x^3)$, what is the coefficient of x^2? What is the coefficient of x^4?

Solution. $a_0 b_2 + a_1 b_1 + a_2 b_0$ is the coefficient of x^2. $a_1 b_3 + a_2 b_2$ is the coefficient of x^4.

---

[2]In the evolution of our current mathematical terminology, the word function evolved through several meanings, starting with very imprecise meanings and ending with our current rather precise meaning. The terminology "generating function" may be thought of as an example of one of the earlier usages of the term function.

Activity 164. In Problem 163 why is there a b_0 and a b_1 in your expression for the coefficient of x^2 but there is not a b_0 or a b_1 in your expression for the coefficient of x^4? What is the coefficient of x^4 in

$$(a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4)(b_0 + b_1x + b_2x^2 + b_3x^3 + b_4x^4)?$$

Express this coefficient in the form

$$\sum_{i=0}^{4} \text{something},$$

where the something is an expression you need to figure out. Now suppose that $a_3 = 0$, $a_4 = 0$ and $b_4 = 0$. To what is your expression equal after you substitute these values? In particular, what does this have to do with Problem 163?

Solution.   There is a b_0 because it can be paired with an a_2 to give the term $a_2b_0x^4$. Similarly there is a b_1 because it can be paired with a_1 for the same purpose. However there is no a_i that we can pair with b_0 to get a coefficient of x^4 and no a_i that we can pair with b_3 to get a coefficient of x^4.

The coefficient of x^4 in

$$(a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4)(b_0 + b_1x + b_2x^2 + b_3x^3 + b_4x^4)$$

is $\sum_{i=0}^{4} a_ib_{4-i}$. If we substitute $a_3 = 0$, $a_4 = 0$, and $b_4 = 0$, we get the coefficient of x^4 in $(a_0 + a_1x + a_2x^2)(b_0 + b_1x + b_2x^2 + b_3x^3)$. This exemplifies the idea that we can get a uniform formula for the coefficient of x^i (namely, sum all a_jb_{i-j} from $j = 0$ to i) in a product of two polynomials if we are willing to say that the coefficient of a power of x that does not appear in a polynomial is 0.

Activity 165. The point of the Problems 163 and Activity 164 is that so long as we are willing to assume $a_i = 0$ for $i > n$ and $b_j = 0$ for $j > m$, then there is a very nice formula for the coefficient of x^k in the product

$$\left(\sum_{i=0}^{n} a_ix^i\right)\left(\sum_{j=0}^{m} b_jx^j\right).$$

Write down this formula explicitly.

Solution. $\sum_{i=0}^{k} a_ib_{k-i}$.

Activity 166. Assuming that the rules you use to do arithmetic with polynomials apply to power series, write down a formula for the coefficient of x^k in the product

$$\left(\sum_{i=0}^{\infty} a_ix^i\right)\left(\sum_{j=0}^{\infty} b_jx^j\right).$$

Solution. $\sum_{i=0}^{k} a_ib_{k-i}$.

We use the expression you obtained in Problem 166 to *define* the product of power series. That is, we define the product

$$\left(\sum_{i=0}^{\infty} a_i x^i\right) \left(\sum_{j=0}^{\infty} b_j x^j\right)$$

to be the power series $\sum_{k=0}^{\infty} c_k x^k$, where c_k is the expression you found in Problem 166. Since you derived this expression by using the usual rules of algebra for polynomials, it should not be surprising that the product of power series satisfies these rules.[3]

Product principle for generating functions

Each time that we converted a picture function to a generating function by substituting x or some power of x for each picture, the coefficient of x had a meaning that was significant to us. For example, with the picture enumerator for selecting between zero and three each of apples, pears, and bananas, when we substituted x for each of our pictures, the exponent i in the power x^i is the number of pieces of fruit in the fruit selection that led us to x^i. After we simplify our product by collecting together all like powers of x, the coefficient of x^i is the number of fruit selections that use i pieces of fruit. In the same way, if we substitute x^c for a picture, where c is the number of calories in that particular kind of fruit, then the i in an x^i term in our generating function stands for the number of calories in a fruit selection that gave rise to x^i, and the coefficient of x^i in our generating function is the number of fruit selections with i calories. The product principle of picture enumerators translates directly into a product principle for generating functions.

Activity 167. Suppose that we have two sets S_1 and S_2. Let v_1 (v stands for value) be a function from S_1 to the nonnegative integers and let v_2 be a function from S_2 to the nonnegative integers. Define a new function v on the set $S_1 \times S_2$ by $v(x_1, x_2) = v_1(x_1) + v_2(x_2)$. Suppose further that $\sum_{i=0}^{\infty} a_i x^i$ is the generating function for the number of elements x_1 of S_1 of value i, that is with $v_1(x_1) = i$. Suppose also that $\sum_{j=0}^{\infty} b_j x^j$ is the generating function for the number of elements of x_2 of S_2 of value j, that is with $v_2(x_2) = j$. Prove that the coefficient of x^k in

$$\left(\sum_{i=0}^{\infty} a_i x^i\right) \left(\sum_{j=0}^{\infty} b_j x^j\right)$$

is the number of ordered pairs (x_1, x_2) in $S_1 \times S_2$ with total value k, that is with $v_1(x_1) + v_2(x_2) = k$. This is called the **product principle for generating functions**.

---

[3]Technically we should explicitly state these rules and prove that they are all valid for power series multiplication, but it seems like overkill at this point to do so!

Solution.   The generating function for ordered pairs of total value k will have the number of ordered pairs of total value k as the coefficient of x^k. But we get a total value k by taking something of value i in S_1 and something of value $k - i$ in j. And since values cannot be negative, the only is available to us are the ones between 0 and k. By the product principle for pairs, the number of ordered pairs (x, y) with $v_1(x) = i$ and $v_2(y) = k - i$ is $a_i b_{k-i}$. To get the number of pairs of total value k, we have to sum over all possible pairs $(i, k - i)$ of values, that is, we have to take the sum $\sum_{i=0}^{k} a_i b_{k-i}$. And this is the coefficient of x^k in the product

$$\left(\sum_{i=0}^{\infty} a_i x^i \right) \left(\sum_{j=0}^{\infty} b_j x^j \right).$$

This proves the product principle for generating functions.

Activity 168. Let i denote an integer between 1 and n.

(a) What is the generating function for the number of subsets of $\{i\}$ of each possible size? (Notice that the only subsets of $\{i\}$ are $\emptyset$ and $\{i\}$.)

   Solution. $1 + x$.

(b) Use the product principle for generating functions to prove the binomial theorem.

   Solution.   The generating function for number of n-tuples of 0s and 1s with k ones is, by the product principle for generating functions (extended to n factors) $(1 + x)^n$. However the number of n-tuples of 0s and 1s with k ones is the number of k-element subsets of an n-element set. Therefore $(1 + x)^n = \sum_{i=0}^{n} \binom{n}{i} x^i$. To get the $(x + y)^n$ form of the binomial theorem, note that $x + y = y(1 + \frac{x}{y})$, so that

$$(x + y)^n = y^n \left(\frac{x}{y} + 1 \right)^n$$

$$= y^n \sum_{i=0}^{n} \binom{n}{i} \frac{x^i}{y^i}$$

$$= \sum_{i=0}^{n} \binom{n}{i} x^i y^{n-i}.$$

The extended binomial theorem and multisets

Activity 169. Suppose once again that i is an integer between 1 and n.

(a) What is the generating function in which the coefficient of x^k is the number of multisets of size k chosen from $\{i\}$? This series is an example of what is called an **infinite geometric series**.

Solution. $1 + x + x^2 + \cdots + x^i + \cdots = \sum_{i=0}^{\infty} x^i$.

(b) Express the generating function in which the coefficient of x^k is the number of multisets chosen from $[n]$ as a power of a power series. What does Problem 115 (in which your answer could be expressed as a binomial coefficient) tell you about what this generating function equals?

Solution. The generating function is $\left(\sum_{i=0}^{\infty} x^i\right)^n$. Problem 115 tells us that this equals $\sum_{i=0}^{\infty} \binom{n+i-1}{i} x^i$.

Activity 170. What is the product $(1-x)\sum_{k=0}^{n} x^k$? What is the product

$$(1-x)\sum_{k=0}^{\infty} x^k?$$

Solution.

$$(1-x)\sum_{k=0}^{n} x^k = 1 - x + x - x^2 + \cdots + x^{n-1} - x^n + x^n - x^{n+1} = 1 - x^{n+1}.$$

$$(1-x)\sum_{k=0}^{\infty} x^k = \sum_{i=0}^{\infty} x^i - \sum_{i=0}^{\infty} x^{i+1} = \sum_{i=0}^{\infty} x^i - \sum_{i=1}^{\infty} x^i = 1.$$

Activity 171. Express the generating function for the number of multisets of size k chosen from $[n]$ (where n is fixed but k can be any nonnegative integer) as a 1 over something relatively simple.

Solution. Since $(1-x)\sum_{k=0}^{\infty} x^k = 1$, we have that

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}.$$

Therefore $\left(\sum_{k=0}^{\infty} x^k\right)^n = \frac{1}{(1-x)^n}$ is the generating function for multisets of size k chosen from an n element set.

Activity 172. Find a formula for $(1+x)^{-n}$ as a power series whose coefficients involve binomial coefficients. What does this formula tell you about how we should define $\binom{-n}{k}$ when n is positive?

Solution.

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

We want the coefficient of x^k in $(1+x)^{-n}$ to be $\binom{-n}{k}$, so we want $\binom{-n}{k} = (-1)^k \binom{n+k-1}{k}$.

Activity 173. If you define $\binom{-n}{k}$ in the way you described in Problem 172, you can write down a version of the binomial theorem for $(x + y)^n$ that is valid for both nonnegative and negative values of n. Do so. This is called the **extended binomial theorem**.

Solution. $(x + y)^n = \sum_{i=0}^{\infty} \binom{n}{i} x^i$. The proof consists of writing $(x + y) = y(\frac{x}{y}+1)$ and applying the Problem 172 when n is negative. When n is positive, we recall that $\binom{n}{k}$ is zero when $k > n$, so replacing the upper limit of n in the standard version of the binomial theorem by ∞ doesn't change the value of the sum.

Activity 174. Write down the generating function for the number of ways to distribute identical pieces of candy to three children so that no child gets more than 4 pieces. Write this generating function as a quotient of polynomials. Using both the extended binomial theorem and the original binomial theorem, find out in how many ways we can pass out exactly ten pieces. Use one of our earlier counting techniques to verify your answer.

Solution. $(1 + x + x^2 + x^3 + x^4)^3$. We can write

$$(1 + x + x^2 + x^3 + x^4)^3\rangle =\rangle \left(\frac{1 - x^5}{1 - x}\right)^3$$

$$=(1 - x^5)^3(1 - x)^{-3}$$

$$=(1 - 3x^5 + 3x^{10} - x^{15}) \sum_{i=0}^{\infty} \binom{3 + i - 1}{i} x^i$$

$$=(1 - 3x^5 + 3x^{10} - x^{15}) \sum_{i=0}^{\infty} \binom{2 + i}{i} x^i$$

The coefficient of x^{10} is the number of ways to pass out ten pieces of candy, and is $\binom{12}{10} - 3\binom{7}{5} + 3\binom{2}{0}$. Thus the number of ways to pass out ten pieces of candy is $66 - 3 \cdot 21 + 3 = 6$.)

A second solution technique that is appropriate is inclusion-exclusion counting. We let property i be that child i gets more than four pieces of fruit. The number of ways to pass out n pieces of fruit with a set J of size j properties is $\binom{n-5j+3-1}{n-5j} = \binom{n-5j+2}{2}$. By Equation (5.2) we have $N_e(\emptyset) = \sum_{j=0}^{3}(-1)^j\binom{3}{j}\binom{n-5j+2}{2}$. Using $n = 10$ gives us $\binom{12}{2} - 3\binom{7}{2} + 3\binom{2}{2} = 6$. (You can either note that $\binom{-3}{2}$ is 0 or note that with 10 pieces of candy, there is no way to pass out five pieces to each of three children. In this second case we would revise our statement about the number of ways to pass out n pieces of fruit to three children with the properties in J to note that it is zero if $5j > n$, and otherwise is $\binom{n-5j+2}{2}$.)

Activity 175. What is the generating function for the number of multisets chosen from an n-element set so that each element appears at least j times and less than m times. Write this generating function as a quotient of polynomials, then as a product of a polynomial and a power series.

Solution.

$$
\begin{aligned}
(x^j + x^{j+1} + \cdots + x^{m-1}) &= x^j \left(\sum_{i=0}^{m-j-1} x^i \right)^n \\
&= \left(x^j \frac{1 - x^{m-j}}{1 - x} \right)^n \\
&= \left(\frac{x^j - x^m}{1 - x} \right)^n \\
&= (x^j - x^m)^n \sum_{i=0}^{\infty} \binom{n+i-1}{i} x^i.
\end{aligned}
$$

Generating functions for integer partitions

Activity 176. If we have five identical pennies, five identical nickels, five identical dimes, and five identical quarters, give the picture enumerator for the combinations of coins we can form and convert it to a generating function for the number of ways to make k cents with the coins we have. Do the same thing assuming we have an unlimited supply of pennies, nickels, dimes, and quarters.

Solution. $(1 + P + P^2 + P^3 + P^4 + P^5)(1 + N + N^2 + N^3 + N^4 + N^5)(1 + D + D^2 + D^3 + D^4 + D^5)(1 + Q + Q^2 + Q^3 + Q^4 + Q^5)$. Substituting x for P, x^5 for N, x^{10} for D and x^{25} for Q gives us

$$
\sum_{i=0}^{5} x^i \sum_{i=0}^{5} x^{5i} \sum_{i=0}^{5} x^{10i} \sum_{i=0}^{5} x^{25i} = \frac{1 - x^6}{1 - x} \cdot \frac{1 - x^{30}}{1 - x^5} \cdot \frac{1 - x^{60}}{1 - x^{10}} \cdot \frac{1 - x^{150}}{1 - x^{25}}.
$$

Although we could write this as a polynomial times a product of four power series, doing so would not significantly increase our understanding, though it would let us make some painful computations of the number of ways to make a certain number of cents. If we actually wanted such numbers we would be better off asking a computer algebra package to expand the product of the polynomials on the left. With unlimited supplies the generating function becomes

$$
\frac{1}{1 - x} \cdot \frac{1}{1 - x^5} \cdot \frac{1}{1 - x^{10}} \cdot \frac{1}{1 - x^{25}}.
$$

Again, we could write this as a product of power series, and that would allow us to painfully compute the number of ways to create a certain number of cents. If we actually wanted to know the number of ways to make up 200 cents, say, it would be more sensible to ask a computer algebra package to extract the coefficient of x^{200} in the product of the four quotients.

Activity 177. Recall that a partition of an integer k is a multiset of numbers that adds to k. In Problem 176 we found the generating function for the number of partitions of an integer into parts of size 1, 5, 10, and 25. Give the generating function for the number partitions of an integer into parts of size

one through ten. Give the generating function for the number of partitions of an integer k into parts of size at most m. (Where m is fixed but k may vary.) Notice this is the generating function for partitions whose Young diagram fits into the space between the line $x = 0$ and the line $x = m$ in a coordinate plane. (We assume the boxes in the Young diagram are one unit by one unit.) When working with generating functions for partitions, it is becoming standard to use q rather than x as the variable in the generating function. Write your answers in this notation.[4]

Solution. $\prod_{i=1}^{10} \frac{1}{1-q^i}, \prod_{i=1}^{m} \frac{1}{1-q^i}$.

Activity 178. In Problem 177 you gave the generating function for the number of partitions of an integer into parts of size at most m. Explain why this is also the generating function for partitions of an integer into at most m parts. Notice that this is the generating function for the number of partitions whose Young diagram fits into the space between the line $y = 0$ and the line $y = m$.

Solution. Conjugation is a bijection between partitions with largest part at most m and partitions with at most m parts. Thus the coefficient of q^i (the number of partitions of i into parts of size at most m) in the generating function for the number of partitions of integers into parts of size at most m will be the coefficient of q^i (the number of partitions of i with at most m parts) in the generating function for the number of partitions of integers into parts of size at most m. Thus the two generating functions are the same.

Activity 179. Give the generating function for the number of partitions of an integer into parts of any size. Don't forget to use q rather than x as your variable. This generating function involves an infinite product. Describe the kinds of terms you actually multiply and add together to get the last generating function. Rewrite any power series that appear in your product as quotients of polynomials or as integers divided by polynomials.

Solution. We start with

$$(1 + q + q^2 + \cdots)(1 + q^2 + q^4 + \cdots) \cdots (1 + q^i + q^{2i}) \cdots,$$

which we can write more precisely as

$$\prod_{i=1}^{\infty} \sum_{j=0}^{\infty} q^{ij}.$$

To get the coefficient of q^k in this product, we look at all ways of choosing one summand from each of the infinite series and multiplying them together to get q^k, and add all these products up. That is, the coefficient of q^k is the number of ways of making these choices of one summand from each series so

---

[4]The reason for this change in the notation relates to the subject of finite fields in abstract algebra, where q is the standard notation for the size of a finite field. While we will make no use of this connection, it will be easier for you to read more advanced work if you get used to the different notation.

that the product of our choices is q^k. We can rewrite the infinite product as

$$\prod_{i=0}^{\infty} \frac{1}{1-q^i}.$$

Activity 180. In Problem 179, we multiplied together infinitely many power series. Here are two notations for infinite products that look rather similar:

$$\prod_{i=1}^{\infty} 1 + x + x^2 + \cdots + x^i \text{ and } \prod_{i=1}^{\infty} 1 + x^i + x^{2i} + \cdots + x^{i^2}.$$

However, one makes sense and one doesn't. Figure out which one makes sense and explain why it makes sense and the other one doesn't. If we want a product of the form

$$\prod_{i=1}^{\infty} 1 + p_i(x),$$

where each $p_i(x)$ is a nonzero polynomial in x to make sense, describe a relatively simple assumption about the polynomials $p_i(x)$ that will make the product make sense. If we assumed the terms $p_i(x)$ were nonzero power series, is there a relatively simple assumption we could make about them in order to make the product make sense? (Describe such a condition or explain why you think there couldn't be one.)

Solution. $\prod_{i=1}^{\infty} 1 + x^i + x^{2i} + \cdots + x^{i^2}$ makes sense because when we look for ways of choosing one summand from each factor so that the summands multiply together to give us x^k, we will find only finitely many ways of making those choices, so the coefficient of x^k can be taken to be the number of such choices. On the other hand, in the expression $\prod_{i=1}^{\infty} 1 + x + x^2 + \cdots + x^i$, there are infinitely many ways to choose x from one term and 1 from all the rest of the terms so that the product of these summands is x. Thus we can't even specify what the coefficient of x is in the product. On the basis of this analysis, we see that for $\prod_{i=1}^{\infty} 1 + p_i(x)$ to make sense, we need to assume that for each possible positive integer n, there are only a finite number of polynomials p_i whose lowest degree term has degree less than or equal to n. In that way, for each positive integer, there will be only finitely many ways to chose a summand from each factor so that the product of the summands is a multiple of x^k. The same assumption works when the p_i are power series, for the same reason.

Activity 181. What is the generating function (using q for the variable) for the number of partitions of an integer in which each part is even?

Solution. $(1 + q^2 + q^4 + \cdots)(1 + q^4 + q^8 + \cdots)(1 + q^6 + q^{12} + \cdots) \cdots$, which can be written more precisely as $\displaystyle\prod_{i=1}^{\infty} \sum_{j=0}^{\infty} q^{2ij}$.

Activity 182. What is the generating function (using q as the variable) for the number of partitions of an integer into distinct parts, that is, in which each part is used at most once?

Solution. $(1+q)(1+q^2)(1+q^3)\cdots = \prod_{i=1}^{\infty}(1+q^i)$.

Activity 183. Use generating functions to explain why the number of partitions of an integer in which each part is used an even number of times equals the generating function for the number of partitions of an integer in which each part is even.

Solution.   In the generating function $\prod_{i=1}^{\infty}\sum_{j=0}^{\infty}q^{2ij}$, we may interpret the $2ij$ in q^{2ji} the value of using $2i$ as a part j times or as the value of using i as a part $2j$ times. Therefore this is the generating function both for the number of partitions of integers into parts that are even and the number of partitions into parts that are used an even number of times. Therefore the number of partitions of n in which each part is even equals the number of partitions of n in which each part is used an even number of times.

Activity 184.

(a) Use the fact that
$$\frac{1-q^{2i}}{1-q^i} = 1 + q^i$$
and the generating function for the number of partitions of an integer into distinct parts to show how the number of partitions of an integer k into distinct parts is related to the number of partitions of an integer k into odd parts.

Solution. $\prod_{i=1}^{\infty}1 + q^i = \prod_{i=1}^{n}\frac{1-q^{2i}}{1-q^i} = \frac{\prod_{i=1}^{\infty}1-q^{2i}}{\prod_{j=1}^{\infty}1-q^j} = \prod_{i=j}^{\infty}\frac{1}{1-q^{2j-1}},$
because all the terms in the numerator cancel with alternate terms in the denominator leaving only terms with odd powers of q. But
$$\prod_{j=1}^{\infty}\frac{1}{1-q^{2j-1}}\rangle =\rangle \prod_{j=1}^{\infty}\sum_{i=0}^{\infty}(q^{2j-1})^i$$
$$=(1 + q + q^{2\cdot 1} + \cdots)(1 + q^3 + q^{2\cdot 3} + \cdots)(1 + q^5 + q^{2\cdot 5} + \cdots)\cdots,$$

which is the generating function for partitions of integers into parts that are odd numbers.

(b) Write down the generating function for the number of ways to partition an integer into parts of size no more than m, each used an odd number of times. Write down the generating function for the number of partitions of an integer into parts of size no more than m, each used an even number of times. Use these two generating functions to get a relationship between the two sequences for which you wrote down the generating functions.

Solution.

$$\prod_{i=1}^{m}(q^i+q^{3i}+q^{5i}+\cdots) = q^{1+2+\cdots m}\prod_{i=1}^{m}(1+q^{2i}+q^{4i}+\cdots) = q^{\binom{m+1}{2}}\prod_{i=1}^{m}\frac{1}{1-q^{2i}}$$

is the generating function for the number of ways to partition an integer into parts of size at most m, each used an odd number of times.

$$\prod_{i=1}^{m}(1 + q^{2i} + q^{4i} + q^{6i} + \cdots) = \prod_{i=1}^{m}\frac{1}{1 - q^{2i}}$$

is the generating function for the number of partitions of an integer into parts of size no more than m, each used an even number of times. Therefore the number of partitions of k into parts of size no more than m, each used an even number of times is the number of partitions of $k + \binom{m+1}{2}$, into parts of size no more than m, each used an odd number of times.

Activity 185. In Problem 177 and Problem 178 you gave the generating functions for, respectively, the number of partitions of k into parts the largest of which is at most m and for the number of partitions of k into at most m parts. In this problem we will give the generating function for the number of partitions of k into at most n parts, the largest of which is at most m. That is we will analyze $\sum_{i=0}^{\infty} a_k q^k$ where a_k is the number of partitions of k into at most n parts, the largest of which is at most m. Geometrically, it is the generating function for partitions whose Young diagram fits into an m by n rectangle, as in Problem 145. This generating function has significant analogs to the binomial coefficient $\binom{m+n}{n}$, and so it is denoted by $\left[\begin{smallmatrix} m+n \\ n \end{smallmatrix}\right]_q$. It is called a *q*-**binomial coefficient**.

(a) Compute $\left[\begin{smallmatrix} 4 \\ 2 \end{smallmatrix}\right]_q = \left[\begin{smallmatrix} 2+2 \\ 2 \end{smallmatrix}\right]_q$.

   Solution. A partition with up to two parts of size up to two can have no parts, one part of size 1, which makes it a partition of 1, one part of size 2 which makes it a partition of 2, two parts of size 1, which makes it a partition of 2, a part of size 2 and a part of size 1, which makes it a partition of 3, or two parts of size 2, which makes it a partition of 4. Thus $\left[\begin{smallmatrix} 4 \\ 2 \end{smallmatrix}\right]_q = 1 + q + 2q^2 + q^3 + q^4$.

(b) Find explicit formulas for $\left[\begin{smallmatrix} n \\ 1 \end{smallmatrix}\right]_q$ and $\left[\begin{smallmatrix} n \\ n-1 \end{smallmatrix}\right]_q$.

   Solution. Both are $1 + q + q^2 + \cdots + q^{n-1} = \frac{1-q^{n-1}}{1-q}$, because they are the generating function for the number of partitions whose Young diagram fits into a rectangle $n - 1$ units wide and 1 unit deep or into a rectangle 1 unit wide and $n - 1$ units deep respectively.

(c) How are $\left[\begin{smallmatrix} m+n \\ n \end{smallmatrix}\right]_q$ and $\left[\begin{smallmatrix} m+n \\ n \end{smallmatrix}\right]_q$ related? Prove it. (Note this is the same as asking how $\left[\begin{smallmatrix} r \\ s \end{smallmatrix}\right]_q$ and $\left[\begin{smallmatrix} r \\ r-s \end{smallmatrix}\right]_q$ are related.)

Solution.   By conjugation, $\left[\begin{smallmatrix} m+n \\ n \end{smallmatrix}\right]_q = \left[\begin{smallmatrix} m+n \\ m \end{smallmatrix}\right]_q$.

(d) So far the analogy to $\binom{m+n}{n}$ is rather thin! If we had a recurrence like the Pascal recurrence, that would demonstrate a real analogy. Is $\left[\begin{smallmatrix} m+n \\ n \end{smallmatrix}\right]_q = \left[\begin{smallmatrix} m+n-1 \\ n-1 \end{smallmatrix}\right]_q + \left[\begin{smallmatrix} m+n-1 \\ n \end{smallmatrix}\right]_q$?

Solution.   No. $\left[\begin{smallmatrix} 4 \\ 2 \end{smallmatrix}\right]_q = 1+q+2q^2+q^3+q^4$, but $\left[\begin{smallmatrix} 3 \\ 1 \end{smallmatrix}\right]_q = 1+q+q^2$ and $\left[\begin{smallmatrix} 3 \\ 2 \end{smallmatrix}\right]_q = 1+q+q^2$.

(e) Recall the two operations we studied in Problem 147. The largest part of a partition counted by $\left[\begin{smallmatrix} m+n \\ n \end{smallmatrix}\right]_q$ is either m or is less than or equal to $m-1$. In the second case, the partition fits into a rectangle that is at most $m-1$ units wide and at most n units deep. What is the generating function for partitions of this type? In the first case, what kind of rectangle does the partition we get by removing the largest part sit in? What is the generating function for partitions that sit in this kind of rectangle? What is the generating function for partitions that sit in this kind of rectangle after we remove a largest part of size m? What recurrence relation does this give you?

Solution.   The generating function for partitions that arise in the second case is $\left[\begin{smallmatrix} m+n-1 \\ m-1 \end{smallmatrix}\right]_q = \left[\begin{smallmatrix} m+n-1 \\ n \end{smallmatrix}\right]_q$. In the first case, after we delete the largest part, the Young diagram sits in a rectangle of width m and depth $n-1$. The generating function for partitions that arise in the first case *after* we have deleted a part of size m is $\left[\begin{smallmatrix} m+n-1 \\ n-1 \end{smallmatrix}\right]_q$. Thus the generating function for partitions that arise in the first case (*before* we delete the part of size m) is $q^m \left[\begin{smallmatrix} m+n-1 \\ n-1 \end{smallmatrix}\right]_q$. Thus by the sum principle the generating function for all partitions that fit into a rectangle of width m and depth n is given by

$$\left[\begin{matrix} m+n \\ n \end{matrix}\right]_q = q^m \left[\begin{matrix} m+n-1 \\ n-1 \end{matrix}\right]_q + \left[\begin{matrix} m+n-1 \\ n \end{matrix}\right]_q.$$

If you don't use the symmetry $\left[\begin{smallmatrix} m+n-1 \\ m-1 \end{smallmatrix}\right]_q = \left[\begin{smallmatrix} m+n-1 \\ n \end{smallmatrix}\right]_q$, you get

$$\left[\begin{matrix} m+n \\ n \end{matrix}\right]_q = q^m \left[\begin{matrix} m+n-1 \\ n-1 \end{matrix}\right]_q + \left[\begin{matrix} m+n-1 \\ m-1 \end{matrix}\right]_q.$$

While this doesn't quite look like the Pascal recurrence, it is still a correct answer.

(f) What recurrence do you get from the other operation we studied in Problem 147?

Solution. We either have exactly k parts or fewer than k parts. In the first case, removing one from each part gives us a partition whose Young diagram fits into a $m-1$ by n box, while in the second case doing nothing gives us a partition that fits into a m by $n-1$ box. In the first case the partition we get partitions $k-n$, and in the second case it still partitions k. Thus we get

$$\begin{bmatrix} m+n \\ n \end{bmatrix}_q = q^n \begin{bmatrix} m+n-1 \\ n \end{bmatrix}_q + \begin{bmatrix} m+n-1 \\ n-1 \end{bmatrix}_q.$$

(g) It is quite likely that the two recurrences you got are different. One would expect that they might give different values for $\begin{bmatrix} m+n \\ n \end{bmatrix}_q$. Can you resolve this potential conflict?

Solution. Yes, by conjugation.

(h) Define $[n]_q$ to be $1+q+\cdots+q^{n-1}$ for $n > 0$ and $[0]_q = 1$. We read this simply as n-sub-q. Define $[n]!_q$ to be $[n]_q[n-1]_q\cdots[3]_q[2]_q[1]_q$. We read this as n cue-torial, and refer to it as a *q*-**ary factorial**. Show that

$$\begin{bmatrix} m+n \\ n \end{bmatrix}_q = \frac{[m+n]!_q}{[m]!_q[n]!_q}.$$

Solution. Note that $\begin{bmatrix} n \\ 0 \end{bmatrix}_q = \begin{bmatrix} n \\ n \end{bmatrix}_q = 1$, because only the partition with no parts sits in a rectangle of width or depth 0. But $\frac{[m+0]!_q}{[m]!_q[0]!_q} = 1$ and $\frac{[0+n]!_q}{[0]!_q[n]!_q} = 1$. Thus the formula holds when $m = 0$ or $n = 0$. But

$$q^m \frac{[m+n-1]!_q}{[n-1]!_q[m]!_q} + \frac{[m+n-1]!_q}{[n]!_q[m-1]!_q}$$

$$= [m+n-1]!_q \left(\frac{q^m}{[n-1]!_q[m]!_q} + \frac{1}{[n]!_q[m-1]!_q} \right)$$

$$[m+n-1]!_q \left(\frac{[n]_q q^m}{[n]!_q[m]_q} + \frac{[m]_q}{[n]!_q[m]!_q} \right)$$

$$= \frac{[m+n-1]!_q}{[n]!_q[m]!_q} \left((1+q+\cdots+q^{n-1})q^m + 1+q+\cdots+q^{m-1} \right)$$

$$= \frac{[m+n-1]!_q[m+n]_q}{[n]!_q[m]!_q}$$

$$= \begin{bmatrix} m+n \\ n \end{bmatrix}_q.$$

Thus $\frac{[m+n]!_q}{[m]!_q[n]!_q}$ satisfies our recurrence and so by the principle of mathematical induction, $\begin{bmatrix} m+n \\ n \end{bmatrix}_q = \frac{[m+n]!_q}{[m]!_q[n]!_q}$.

(i) Now think of q as a variable that we will let approach 1. Find an explicit formula for

   (a) $\lim\limits_{q \to 1} [n]_q$.

   (b) $\lim\limits_{q \to 1} [n]!_q$.

   (c) $\lim\limits_{q \to 1} \begin{bmatrix} m+n \\ n \end{bmatrix}_q$.

   Why is the limit in Part 3 equal to the number of partitions (of any number) with at most n parts all of size most m? Can you explain bijectively why this quantity equals the formula you got?

   Solution.

   (a) n

   (b) $n!$

   (c) $\binom{m+n}{n}$

   Since the generating function is a finite sum (we are talking about partitions whose Young diagram into a finite rectangle), the limit is obtained by setting $q = 1$, and this sums the number of partitions of each possible number k that have at most n parts all of size at most m. We want a bijection between such partitions and the n element subsets of an $m+n$ element set. Recall that there is a bijection between subsets of an n element set and lattice paths from $(0,0)$ to m, n in a coordinate plane. If we draw our rectangle of width m and depth n with its lower left corner at $(0,0)$, then each Young diagram gives us such a lattice path and each such lattice path gives us a Young diagram.

(j) What happens to $\begin{bmatrix} m+n \\ n \end{bmatrix}_q$ if we let q approach -1?

   Solution. $\begin{bmatrix} m+n \\ n \end{bmatrix}_q$ is the generating function for the number of partitions whose Young diagram fits into an m by n rectangle. That is,

$$\begin{bmatrix} m+n \\ n \end{bmatrix}_q = \sum_{k=0}^{\infty} a_k q^k,$$

where a_k is the number of partitions of k whose Young diagram fits into an m by n rectangle. In particular $a_k = 0$ if $k > mn$, because the Young diagram of a partition of a number larger than mn certainly cannot fit into an m by n rectangle. Thus

$$\begin{bmatrix} m+n \\ n \end{bmatrix}_q = \sum_{k=0}^{mn} a_k q^k.$$

Furthermore, $a_k = a_{mn-k}$, because the complement in an m by n rectangle of a partition of k is a partition of $mn - k$, and complementation in an m by n rectangle is a bijection between partitions of k that fit into the rectangle and partitions of $mn - k$ that fit into the rectangle. Thus $\begin{bmatrix} m+n \\ n \end{bmatrix}_q$ is a polynomial of degree mn in which the coefficient of q_k equals the coefficient of q^{mn-k}. For this reason we can compute the limit as q approaches -1 simply by substituting -1 for q in the polynomial; the only trouble being that the formula we know for $\begin{bmatrix} m+n \\ n \end{bmatrix}_q$ is quotient of polynomials that has zero in both the numerator and denominator when we substitute in $q = -1$. Nonetheless, when we substitute -1 for q we get the alternating sum $\sum_{i=0}^{mn}(-1)^i a_i$. Thus if a_i and a_{mn-i} have opposite sign, they will cancel out. If mn is even, i is even exactly when $mn - i$ is even, and so a_i and a_{mn-i} have the same sign. However when mn is odd, a_i and a_{mn-i} have opposite signs in the sum $\sum_{i=0}^{mn}(-1)^i a_i$, and so the sum is zero. Thus the polynomial $\begin{bmatrix} m+n \\ n \end{bmatrix}_q$ is zero at $q = -1$ whenever mn is odd.

Our experience with binomial coefficients might lead us to hope that the alternating sum of the coefficients a_k will always be zero. However, we computed that $\begin{bmatrix} 4 \\ 2 \end{bmatrix}_q = 1 + q + 2q^2 + q^3 + q^4$, so

$$\lim_{q \to -1} \begin{bmatrix} 4 \\ 2 \end{bmatrix}_q = 1 - 1 + 2 - 1 + 1 = 1.$$

We could compute some more values of the limit by going back to the definition in this way, but it seems unlikely that we could get enough data to make a good conjecture. However we have the recurrence, and setting $q = -1$ in the recurrence gives us Table **??**. From the table, it is clear that we get binomial coefficients interspersed with 0s in the even numbered rows of our table and repeated binomial coefficients in the odd numbered rows of our table. In particular when $m + n$ and n are both even, $\begin{bmatrix} m+n \\ n \end{bmatrix}_{-1} = \binom{(m+n)/2}{n/2}$, and if $m+n$ is even and n is odd, $\begin{bmatrix} m+n \\ n \end{bmatrix}_{-1} = 0$. In our table, at least, if $m + n$ is odd, it appears that we get $\begin{bmatrix} m+n \\ n \end{bmatrix}_{-1} = \binom{\lfloor (m+n)/2 \rfloor}{\lfloor n/2 \rfloor}$. In fact, using the recurrence just as we used it to construct the table, we can prove by induction that

$$\begin{bmatrix} m+n \\ n \end{bmatrix}_{-1} = \begin{cases} 0 & \text{if } mn \text{ is odd} \\ \binom{\lfloor (m+n)/2 \rfloor}{\lfloor n/2 \rfloor} & \text{otherwise.} \end{cases}$$

$m+n\backslash n$	0	1	2	3	4	5	6	7	8	9
0	1	0	0	0	0	0	0	0	0	0
1	1	1	0	0	0	0	0	0	0	0
2	1	0	1	0	0	0	0	0	0	0
3	1	1	1	1	0	0	0	0	0	0
4	1	0	2	0	1	0	0	0	0	0
5	1	1	2	2	1	1	0	0	0	0
6	1	0	3	0	3	0	1	0	0	0
7	1	1	3	3	3	3	1	1	0	0
8	1	0	4	0	6	0	4	0	1	0
9	1	1	4	4	6	6	4	4	1	1

(Note that mn is odd if and only if $m + n$ is even and n is odd.) It would be fascinating to know what we are counting here!

4.2   Generating Functions and Recurrence Relations

Recall that a recurrence relation for a sequence a_n expresses a_n in terms of values a_i for $i < n$. For example, the equation $a_i = 3a_{i-1} + 2^i$ is a first order linear constant coefficient recurrence.

How generating functions are relevant

Algebraic manipulations with generating functions can sometimes reveal the solutions to a recurrence relation.

Activity 186. Suppose that $a_i = 3a_{i-1} + 3^i$.

(a) Multiply both sides by x^i and sum both the left hand side and right hand side from $i = 1$ to infinity. In the left-hand side use the fact that

$$\sum_{i=1}^{\infty} a_i x^i = \left(\sum_{i=0}^{\infty} x^i\right) - a_0$$

and in the right hand side, use the fact that

$$\sum_{i=1}^{\infty} a_{i-1} x^i = x \sum_{i=1}^{\infty} a_i x^{i-1} = x \sum_{j=0}^{\infty} a_j x^j = x \sum_{i=0}^{\infty} a_i x^i$$

(where we substituted j for $i - 1$ to see explicitly how to change the limits of summation, a surprisingly useful trick) to rewrite the equation in terms of the power series $\sum_{i=0}^{\infty} a_i x^i$. Solve the resulting equation for the power series $\sum_{i=0}^{\infty} a_i x^i$.

Solution.

$$\sum_{i=1}^{\infty} a_i x^i = 3 \sum_{i=1}^{\infty} a_{i-1} x^i + \sum_{i=1}^{i} 3^i x^i$$

$$\sum_{i=1}^{\infty} a_i x^i = 3x \sum_{i=1}^{\infty} a_{i-1} x^{i-1} + \sum_{i=0}^{\infty} 3^i x^i - 3^0 x^0$$

$$\sum_{i=0}^{\infty} a_i x^i - a_0 = 3x \sum_{i=0}^{\infty} a_i x^i + \frac{1}{1-3x} - 1$$

$$(1-3x) \sum_{i=0}^{\infty} a_i x^i = a_0 + \frac{1}{1-3x} - 1$$

$$\sum_{i=0}^{\infty} a_i x^i = \frac{a_0 - 1}{1-3x} + \frac{1}{(1-3x)^2}$$

(b) Use the previous part to get a formula for a_i in terms of a_0.

Solution.

$$\sum_{i=0}^{\infty} a_i x^i = \frac{a_0 - 1}{1-3x} + \frac{1}{(1-3x)^2}$$

$$= (a_0 - 1) \sum_{i=0}^{\infty} 3^i x^i + \sum_{i=0}^{\infty} \binom{i+1}{i} 3^i x^i,$$

which gives us $a_n = (a_0 - 1)3^i + (i+1)3^i = (a_0 + i)3^i$.

(c) Now suppose that $a_i = 3a_{i-1} + 2^i$. Repeat the previous two steps for this recurrence relation. (There is a way to do this part using what you already know. Later on we shall introduce yet another way to deal with the kind of generating function that arises here.)

Solution.

$$\sum_{i=1}^{\infty} a_i x^i = 3 \sum_{i=1}^{\infty} a_{i-1} x^i + \sum_{i=1}^{\infty} 2^i x^i$$

$$\sum_{i=0}^{\infty} a_i x^i - a_0 = 3x \sum_{i=1}^{\infty} a_{i-1} x^{i-1} + \sum_{i=0}^{\infty} (2x)^i - 1$$

$$\sum_{i=0}^{\infty} a_i x^i - a_0 = 3x \sum_{i=0}^{\infty} a_i x^i + \frac{1}{1-2x} - 1$$

$$(1-3x) \sum_{i=0}^{\infty} a_i x^i = a_0 + \frac{1}{1-2x} - 1$$

$$\sum_{i=0}^{\infty} a_i x^i = \frac{a_0 - 1}{1 - 3x} + \frac{1}{(1 - 2x)(1 - 3x)}$$

$$\sum_{i=0}^{\infty} a_i x^i = (a_0 - 1)\sum_{i=0}^{\infty} 3^i x^i + \sum_{i=0}^{\infty} 2^i x^i \sum_{j=0}^{\infty} 3^j x^j$$

But

$$\sum_{i=0}^{\infty} 2^i x^i \sum_{j=0}^{\infty} 3^j x^j = \sum_{k=0}^{\infty}\sum_{i=0}^{k} 2^i 3^{k-i} x^k = \sum_{k=0}^{\infty} 3^k x^k \sum_{i=0}^{k} \frac{2^i}{3^i} = \sum_{k=0}^{\infty} \frac{1 - \left(\frac{2}{3}\right)^{k+1}}{1 - \frac{2}{3}} 3^k x^k = \sum_{k=0}^{\infty}(3$$

Substituting this into the equation for $\sum_{i=0}^{\infty} a_i x^i$ gives us $a_i = (a_0 + 2)3^i - 2^{i+1}$.

Activity 187. Suppose we deposit \$5000 in a savings certificate that pays ten percent interest and also participate in a program to add \$1000 to the certificate at the end of each year (from the end of the first year on) that follows (also subject to interest.) Assuming we make the \$5000 deposit at the end of year 0, and letting a_i be the amount of money in the account at the end of year i, write a recurrence for the amount of money the certificate is worth at the end of year n. Solve this recurrence. How much money do we have in the account (after our year-end deposit) at the end of ten years? At the end of 20 years?

Solution. $a_n = 1.1a_{n-1} + 1000$, and $a_0 = 5000$.

$$\sum_{i=1}^{\infty} a_i x^i = 1.1x \sum_{i=1}^{\infty} a_{i-1} x^{i-1} + 1000 \sum_{i=1}^{\infty} x^i$$

$$\sum_{i=0}^{\infty} a_i x^i - a_0 = 1.1x \sum_{i=0}^{\infty} a_i x^i + 1000(\sum_{i=0}^{\infty} x^i - 1)$$

$$(1 - 1.1x) \sum_{i=0}^{\infty} a_i x^i = a_0 + 1000 \sum_{i=10}^{\infty} x^i - 1000$$

$$\sum_{i=0}^{\infty} a_i x^i = a_0 \sum_{i=0}^{\infty}(1.1)^i x^i + 1000 \sum_{i=0}^{\infty}(1.1)^i x^i \sum_{i=0}^{\infty} x^i - 1000$$

$$\sum_{i=0}^{\infty} a_i x^i = a_0 \sum_{i=0}^{\infty}(1.1)^i x^i + 1000 \sum_{k=0}^{\infty}\left(\sum_{i=0}^{k}(1.1)^i\right)x^k - 1000\sum_{i=0}^{\infty}(1.1)^i x^i$$

This gives us that $a_n = 4000(1.1)^n + 1000\frac{1-(1.1)^{n+1}}{-.1}$. This simplifies to $a_n = 4000(1.1)^n + 10000(1.1)^{n+1} - 10000 = 15,000(1.1)^n - 10000$. Courtesy of Maple, after ten years we have \$28,906.14 and after 20 years we have \$90,912.50.

Fibonacci Numbers

The sequence of problems that follows describes a number of hypotheses we might make about a fictional population of rabbits. We use the example of

a rabbit population for historic reasons; our goal is a classical sequence of numbers called Fibonacci numbers. When Fibonacci[5] introduced them, he did so with a fictional population of rabbits.

Second order linear recurrence relations

Activity 188. Suppose we start (at the end of month 0) with 10 pairs of baby rabbits, and that after baby rabbits mature for one month they begin to reproduce, with each pair producing two new pairs at the end of each month afterwards. Suppose further that over the time we observe the rabbits, none die. Let a_n be the number of rabbits we have at the end of month n. Show that $a_n = a_{n-1} + 2a_{n-2}$. This is an example of a **second order** *linear* recurrence with constant coefficients. Using a method similar to that of Problem 186, show that

$$\sum_{i=0}^{\infty} a_i x^i = \frac{10}{1 - x - 2x^2}.$$

This gives us the generating function for the sequence a_i giving the population in month i; shortly we shall see a method for converting this to a solution to the recurrence.

Solution.

$$\sum_{n=2}^{\infty} a_n x^n = \sum_{n=2}^{\infty} a_{n-1} x^n + 2 \sum_{n=2}^{\infty} a_{n-2} x^n$$

$$\sum_{n=0}^{\infty} a_n x^n - a_0 - a_1 x = x \sum_{n=2}^{\infty} a_{n-1} x^{n-1} + 2x^2 \sum_{n=2}^{\infty} a_{n-2} x^{n-2}$$

$$\sum_{n=0}^{\infty} a_n x^n - a_0 - a_1 x = x \sum_{n=1}^{\infty} a_n x^n + 2x^2 \sum_{n=0}^{\infty} a_n x^n$$

$$\sum_{n=0}^{\infty} a_n x^n - a_0 - a_1 x = x \left(\sum_{n=0}^{\infty} a_n x^n - a_0 \right) + 2x^2 \sum_{n=0}^{\infty} a_n x^n$$

$$(1 - x - 2x^2) \sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x - a_0 x$$

$$\sum_{n=0}^{\infty} a_n x^n = \frac{a_0 + a_1 x - a_0 x}{(1 - x - 2x^2)}$$

In this problem $a_0 = a_1 = 10$ because we start with ten pairs of baby rabbits, so they have to mature for a month before they begin reproducing. Thus

$$\sum_{n=0}^{\infty} a_n x^n = \frac{10}{(1 - x - 2x^2)}$$

Activity 189. In Fibonacci's original problem, each pair of mature rabbits produces one new pair at the end of each month, but otherwise the situation

---

[5]Apparently Leanardo de Pisa was given the name Fibonacci posthumously

is the same as in Problem 188. Assuming that we start with one pair of baby rabbits (at the end of month 0), find the generating function for the number of pairs of rabbits we have at the end on n months.

Solution.   Our recurrence becomes $a_n = a_{n-1} + a_{n-2}$, and following the pattern of Problem 188 we get

$$\sum_{n=2}^{\infty} a_n x^n = \sum_{n=2}^{\infty} a_{n-1} x^n + \sum_{n=2}^{\infty} a_{n-2} x^n$$

$$\sum_{n=0}^{\infty} a_n x^n - a_0 - a_1 x = x \left(\sum_{n=0}^{\infty} a_n x^n - a_0 \right) + x^2 \sum_{n=0}^{\infty} a_n x^n$$

$$(1 - x - x^2) \sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x - a_0 x$$

$$\sum_{n=0}^{\infty} a_n x^n = \frac{a_0 + a_1 x - a_0 x}{(1 - x - x^2)}$$

Since now $a_0 = a_1 = 1$, we have $\displaystyle\sum_{n=0}^{\infty} a_n x^n = \frac{1}{1 - x - x^2}$.

Activity 190. Find the generating function for the solutions to the recurrence

$$a_i = 5a_{i-1} - 6a_{i-2} + 2^i.$$

Solution.

$$\sum_{i=2}^{\infty} a_i x^i = \sum_{i=2}^{\infty} 5a_{i-1} x^i - \sum_{i=2}^{\infty} 6a_{i-2} x^i + \sum_{i=2}^{\infty} 2^i x^i$$

$$\sum_{i=0}^{\infty} a_i x^i - a_0 - a_1 x = 5x \left(\sum_{i=0}^{\infty} a_i x^i - a_0 \right) - 6x^2 \sum_{i=0}^{\infty} a_i x^i + \sum_{n=0}^{\infty} 2^i x^i - 1 - 2x$$

$$(1 - 5x + 6x^2) \sum_{i=0}^{\infty} a_i x^i = a_0 - 1 + (a_1 - 5a_0 - 2)x + \sum_{n=0}^{\infty} 2^i x^i$$

$$\sum_{i=0}^{\infty} a_i x^i = \frac{a_0 - 1 + (a_1 - 5a_0 - 2)x}{1 - 5x + 6x^2} + \frac{1}{1 - 5x + 6x^2} \cdot \frac{1}{1 - 2x}$$

The recurrence relations we have seen in this section are called **second order** because they specify a_i in terms of a_{i-1} and a_{i-2}, they are called **linear** because a_{i-1} and a_{i-2} each appear to the first power, and they are called **constant coefficient recurrences** because the coefficients in front of a_{i-1} and a_{i-2} are constants.

Partial fractions

The generating functions you found in the previous section all can be expressed in terms of the reciprocal of a quadratic polynomial. However without a power

series representation, the generating function doesn't tell us what the sequence is. It turns out that whenever you can factor a polynomial into linear factors (and over the complex numbers such a factorization always exists) you can use that factorization to express the reciprocal in terms of power series.

Activity 191. Express $\frac{1}{x-3} + \frac{2}{x-2}$ as a single fraction.

Solution. $\frac{x-2+2(x-3)}{(x-3)(x-2)} = \frac{3x-8}{x^2-5x+6}$

Activity 192. In Problem 191 you see that when we added numerical multiples of the reciprocals of first degree polynomials we got a fraction in which the denominator is a quadratic polynomial. This will always happen unless the two denominators are multiples of each other, because their least common multiple will simply be their product, a quadratic polynomial. This leads us to ask whether a fraction whose denominator is a quadratic polynomial can always be expressed as a sum of fractions whose denominators are first degree polynomials. Find numbers c and d so that

$$\frac{5x+1}{(x-3)(x+5)} = \frac{c}{x-3} + \frac{d}{x+5}.$$

Solution.

$$\frac{5x+1}{(x-3)(x+5)} = \frac{c}{x-3} + \frac{d}{x+5}$$

$$\frac{5x+1}{(x-3)(x-5)} = cx + 5c + dx - 3d$$

gives us

$$5x = cx + dx$$
$$1 = 5c - 3d$$
$$5 = c + d$$
$$1 = 5c - 3d$$

Which gives us $16 = 8c$ so that $c = 2$ and then $d = 3$.

Activity 193. In Problem 192 you may have simply guessed at values of c and d, or you may have solved a system of equations in the two unknowns c and d. Given constants a, b, r_1, and r_2 (with $r_1 \neq r_2$), write down a system of equations we can solve for c and d to write

$$\frac{ax+b}{(x-r_1)(x-r_2)} = \frac{c}{x-r_1} + \frac{d}{x-r_2}.$$

Solution. To have

$$\frac{ax+b}{(x-r_1)(x-r_2)} = \frac{c}{x-r_1} + \frac{d}{x-r_2}$$

we must have
$$cx - r_2c + dx - r_2d = ax + b.$$

This gives us the equations $cx + dx = ax$ and $-r_2c - r_1d = b$. Since x can be any value, in particular it can be nonzero, so we can divide by it. This gives us the equations $c + d = a$ and $r_2c + r_1d = -b$.

Writing down the equations in Problem 193 and solving them is called the *method of partial fractions*. This method will let you find power series expansions for generating functions of the type you found in Problems 188 to Activity 190. However you have to be able to factor the quadratic polynomials that are in the denominators of your generating functions.

Activity 194. Use the method of partial fractions to convert the generating function of Problem 188 into the form
$$\frac{c}{x - r_1} + \frac{d}{x - r_2}.$$

Use this to find a formula for a_n.

Solution. $\frac{10}{(1-x-2x^2)} = \frac{10}{(1-2x)(1+x)} = \frac{c}{1-2x} + \frac{d}{1+x}$. This gives us the equations $c + d = 10$ and $c - 2d = 0$. Thus $3d = 10$, so $d = \frac{10}{3}$, and $c = 2d$ so $c = \frac{20}{3}$. Thus

$$\sum_{i=0}^{\infty} a_i x^i = \frac{10}{1 - x - 2x^2}$$
$$= \frac{20/3}{1 - 2x} + \frac{10/3}{1 + x}$$
$$= \frac{20}{3} \sum_{i=0}^{\infty} (2x)^i + \frac{10}{3} \sum_{i=0}^{\infty} (-1)^i x^i$$

Thus $a_i = \frac{20}{3} 2^i + \frac{10}{3}(-1)^i$.

Activity 195. Use the quadratic formula to find the solutions to $x^2 + x - 1 = 0$, and use that information to factor $x^2 + x - 1$.

Solution. $r_1 = \frac{-1+\sqrt{5}}{2}$, $r_2 = \frac{-1-\sqrt{5}}{2}$. These roots give us the equation $x^2 + x - 1 = (x - \frac{-1+\sqrt{5}}{2})(x - \frac{-1-\sqrt{5}}{2})$.

Activity 196. Use the factors you found in Problem 195 to write
$$\frac{1}{x^2 + x - 1}$$

in the form
$$\frac{c}{x - r_1} + \frac{d}{x - r_2}.$$

Hint. You can save yourself a tremendous amount of frustrating algebra if you arbitrarily choose one of the solutions and call it r_1 and call the other solution r_2 and solve the problem using these algebraic symbols in place of the actual roots.[6] Not only will you save yourself some work, but you will get

---

[6]We use the words roots and solutions interchangeably.

a formula you could use in other problems. When you are done, substitute in the actual values of the solutions and simplify.

Solution. $\frac{1}{x^2+x-1} = \frac{c}{x-r_1} + \frac{d}{x-r_2}$ gives us $cx - cr_2 + dx - dr_1 = 1$. Thus $c + d = 0$, and $cr_2 + dr_1 = -1$. This gives us $d = -c$ and so $cr_2 - cr_1 = -1$, which yields $c = \frac{1}{r_1-r_2}$, and $d = \frac{1}{r_2-r_1}$. By substitution, $c = 1/\sqrt{5}$ and $d = -1/\sqrt{5}$. This gives us

$$\frac{1}{x^2 + x - 1} = \frac{1/\sqrt{5}}{x - \frac{-1+\sqrt{5}}{2}} + \frac{-1/\sqrt{5}}{x - \frac{-1-\sqrt{5}}{2}}.$$

Activity 197.

(a) Use the partial fractions decomposition you found in Problem 195 to write the generating function you found in Problem 189 in the form

$$\sum_{n=0}^{\infty} a_n x^i$$

and use this to give an explicit formula for a_n.

Hint. Once again it will save a lot of tedious algebra if you use the symbols r_1 and r_2 for the solutions as in Problem 196 and substitute the actual values of the solutions once you have a formula for a_n in terms of r_1 and r_2.

Solution.

$$\sum_{n=0}^{\infty} a_n x^n = \frac{1}{1 - x - x^2}\rangle = \rangle - \frac{1}{x^2 + x - 1}$$

$$= \frac{1}{\sqrt{5}} \cdot \frac{1}{r_1 - x} - \frac{1}{\sqrt{5}} \cdot \frac{1}{r_2 - x}$$

$$= \frac{1}{r_1\sqrt{5}} \cdot \frac{1}{1 - x/r_1} - \frac{1}{r_2\sqrt{5}} \cdot \frac{1}{1 - x/r_2}$$

$$= \frac{1}{r_1\sqrt{5}} \sum_{n=0}^{\infty} \left(\frac{x}{r_1}\right)^n - \frac{1}{r_2\sqrt{5}} \sum_{n=0}^{\infty} \left(\frac{x}{r_2}\right)^n$$

This gives us that

$$a_n = \frac{1}{\sqrt{5} \cdot r_1^{n+1}} + \frac{1}{\sqrt{5} \cdot r_2^{n+1}}$$

$$= \frac{2^{n+1}}{\sqrt{5}(-1 + \sqrt{5})^{n+1}} + \frac{2^{n+1}}{\sqrt{5}(-1 - \sqrt{5})^{n+1}}$$

$$= \frac{2^{n+1}(1 + \sqrt{5})^{n+1}}{\sqrt{5} \cdot 4^{n+1}} - \frac{2^{n+1}(1 - \sqrt{5})^{n+1}}{\sqrt{5} \cdot 4^{n+1}}$$

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

(b) When we have $a_0 = 1$ and $a_1 = 1$, i.e. when we start with one pair of baby rabbits, the numbers a_n are called *Fibonacci Numbers*. Use either the recurrence or your final formula to find a_2 through a_8. Are you amazed that your general formula produces integers, or for that matter produces rational numbers? Why does the recurrence equation tell you that the Fibonacci numbers are all integers?

Solution. Using the recurrence, the Fibonacci numbers from a_0 to a_8 are 1, 1, 2, 3, 5, 8, 13, 21, 34. The recurrence says each term is the sum of the two preceding terms, and since the first two terms are integers, all the sums must be integers.

(c) Find an algebraic explanation (not using the recurrence equation) of what happens to make the square roots of five go away. Explain why there is a real number b such that, for large values of n, the value of the nth Fibonacci number is almost exactly (but not quite) some constant times b^n. (Find b and the constant.)

Solution. From the binomial theorem,

$$\frac{1}{\sqrt{5}} \left(\frac{1 + \sqrt{5}}{2} \right)^{n+1} - \frac{1}{\sqrt{5}} \left(\frac{1 - \sqrt{5}}{2} \right)^{n+1}$$

$$= \frac{1}{2^{n+1}\sqrt{5}} \left[\sum_{i=0}^{n+1} \binom{n+1}{i} \left(\sqrt{5} \right)^i - \sum_{i=0}^{n+1} \binom{n+1}{i} (-1)^i \left(\sqrt{5} \right)^i \right]$$

$$= \frac{1}{2^{n+1}\sqrt{5}} \sum_{i:i\in[n+1],\ i\text{ is odd}} \binom{n+1}{i} \left(\left(\sqrt{5} \right)^i - (-1)^i \left(\sqrt{5} \right)^i \right)$$

$$= \frac{1}{2^{n+1}\sqrt{5}} \sum_{i:i\in[n+1],\ i\text{ is odd}} 2\binom{n+1}{i} \left(\sqrt{5} \right)^i$$

$$= \frac{1}{2^n} \sum_{i:i\in[n+1],\ i\text{ is odd}} \binom{n+1}{i} \left(\sqrt{5} \right)^{i-1}$$

$$= \frac{1}{2^n} \sum_{i:i\in[n],\ i\text{ is even}} \binom{n+1}{i+1} 5^{i/2}$$

$$= \frac{1}{2^n} \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n+1}{2k+1} 5^k,$$

which makes it clear that a_n is at least a rational number. It is not clear from this new formula why the result is always an integer. Since $\left| \frac{1-\sqrt{5}}{2} \right| < 1$, $\frac{1}{\sqrt{5}} \left(\frac{1-\sqrt{5}}{2} \right)^{n-1}$ approaches 0 as n becomes large. Therefore if we take b to be $\left(\frac{1+\sqrt{5}}{2} \right)^n$ and we take c to be $\frac{1+\sqrt{5}}{2\sqrt{5}}$ then then nth

Fibonacci number is almost exactly cb^n when n is large. In particular, it is the nearest integer to cb^n.

(d) As a challenge (which the author has not yet done), see if you can find a way to show algebraically (not using the recurrence relation, but rather the formula you get by removing the square roots of five) that the formula for the Fibonacci numbers yields integers.

Solution. None is yet available.

Activity 198.

Solve the recurrence $a_n = 4a_{n-1} - 4a_{n-2}$.

Solution.

$$\sum_{n=2}^{\infty} a_n x^n = 4\sum_{n=2}^{\infty} a_{n-1}x^n - 4\sum_{n=2}^{\infty} a_{n-2}x^n$$

$$\sum_{n=0}^{\infty} a_n x^n - a_0 - a_1 x = 4x(\sum_{n=0}^{\infty} a_n x^n - a_0) - 4x^2 \sum_{n=0}^{\infty} a_n x^n$$

$$(1 - 4x + 4x^2)\sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x - 4a_0 x$$

$$\sum_{n=0}^{\infty} a_n x^n = \frac{a_0 + a_1 x - 4a_0 x}{(1 - 4x + 4x^2)}$$

$$\sum_{n=0}^{\infty} a_n x^n = \frac{a_0 + a_1 x - 4a_0 x}{(1 - 2x)^2}$$

$$\sum_{n=0}^{\infty} a_n x^n = (a_0 + a_1 x - 4a_0 x)\sum_{n=0}^{\infty} \binom{n+2-1}{n} 2^n x^n$$

$$\sum_{n=0}^{\infty} a_n x^n = (a_0 + a_1 x - 4a_0 x)\sum_{n=0}^{\infty} (n+1)2^n x^n$$

Thus $a_n = a_0(n+1)2^n + (a_1 - 4a_0)n2^{n-1} = a_0 2^n + (a_1 - 2a_0)n2^{n-1}$.

Catalan Numbers

Activity 199. Using either lattice paths or diagonal lattice paths, explain why the Catalan Number c_n satisfies the recurrence

$$C_n = \sum_{i=1}^{n-1} C_{i-1}C_{n-i}.$$

Show that if we use y to stand for the power series $\sum_{i=0}^{\infty} c_n x^n$, then we can find y by solving a quadratic equation. Solve for y. Taylor's theorem from

calculus tells us that the extended binomial theorem

$$(1+x)^r = \sum_{i=0}^{\infty} \binom{r}{i} x^i$$

holds for any number real number r, where $\binom{r}{i}$ is defined to be

$$\frac{r^{\underline{i}}}{i!} = \frac{r(r-1)\cdots(r-i+1)}{i!}.$$

Use this and your solution for y (note that of the two possible values for y that you get from the quadratic formula, only one gives an actual power series) to get a formula for the Catalan numbers.

Solution.   Recall that a Dyck Path is a diagonal lattice path that never goes lower than its starting point and a Catalan Path of length $2n$ is a Dyck path that goes from $(0,0)$ to $(2n,0)$ The Catalan number C_n is the number of Catalan paths of length $2n$. We take $C_0 = 1$. A Catalan path could touch the x-axis several times before it reaches $(2n,0)$. Its first touch can be any point $(2i,0)$ between $(2,0)$ and $(2n,0)$. For the path to touch first at $(0,2i)$, the path must start with an upstep and then proceed as a Dyck path from $(1,1)$ to $(2i-1,1)$. From there it must take a downstep. But the number of Dyck paths from $(1,1)$ to $(2i-1,1)$ is the same as the number of Catalan paths from $(0,0)$ to $(2i-2,0)$. The number of Catalan paths is the number whose first touch of the x-axis is at $x=2$ plus the number whose first touch is at $x=4,...$, through the number whose first touch is at $2n$. After the first touch at $x=2i$, the path then behaves as a Catalan path from $(2i,0)$ to $(2n,0)$. The number of such paths is C_{n-i}. By the product principle, the number of Catalan paths whose first touch is at $x=2i$ is $C_{i-1}C_{n-i}$. Then by the sum principle, the number of Catalan paths of length 1 or more is

$$C_n = \sum_{i=1}^{n} C_{i-1}C_{n-i}.$$

To solve for C_n, write

$$\sum_{n=0}^{\infty} C_n x^n = 1 + \sum_{n=1}^{\infty} \sum_{i=1}^{n} C_{i-1}C_{n-i}x^n$$

$$\sum_{n=0}^{\infty} C_n x^n = 1 + x\sum_{n=1}^{\infty} \sum_{i=1}^{n} C_{i-1}x^{i-1}C_{n-i}x^{n-i}$$

$$\sum_{n=0}^{\infty} C_n x^n = 1 + x\sum_{i=1}^{\infty} C_{i-1}x^{i-1} \sum_{j=0}^{\infty} C_j x^j$$

$$y = 1 + xy^2$$

This gives us $xy^2 - y + 1 = 0$, and solving for y by the quadratic formula gives us $y = \frac{1 \pm \sqrt{1-4x}}{2x}$. By the extended binomial theorem,

$$\sqrt{1-4x} = (1-4x)^{1/2} = \sum_{i=0}^{\infty} \binom{1/2}{i}(-4x)^i = \sum_{i=0}^{\infty} \frac{(1/2)^{\underline{i}}}{i!}(-1)^i 4^i x^i.$$

The first term of this power series is 1, so to get a power series for y, we must take the negative square root so that the x in the denominator will cancel out. Thus $y = -\frac{1}{2}\sum_{i=1}^{\infty}\frac{(1/2)^{\underline{i}}}{i!}(-1)^i 4^i x^i$. But $(1/2)^{\underline{i}} = (\frac{1}{2})(\frac{-1}{2})(\frac{-3}{2})(\frac{-5}{2})\cdots(\frac{-2i+3}{2})$, so

$$
\begin{aligned}
y &= -\frac{1}{2}\sum_{i=1}^{\infty}\frac{1\cdot 3\cdot 5\cdots(2i-3)}{i!}(-1)^{2i-1}2^i x^{i-1}\\
&= \sum_{i=1}^{\infty}\frac{(2i-2)!}{(i-1)!2^i i!}2^i x^{i-1}\\
&= \sum_{i=1}^{\infty}\frac{2i-2!}{i!(i-1)!}x^{i-1}\\
&= \sum_{j=0}^{\infty}\frac{2j!}{(j+1)!j!}x^j
\end{aligned}
$$

giving us $C_j = \frac{1}{j+1}\binom{2j}{j}$, which is our earlier formula for the Catalan Number C_j.

4.3 Supplementary Problems

Activity 200. What is the generating function for the number of ways to pass out k pieces of candy from an unlimited supply of identical candy to n children (where n is fixed) so that each child gets between three and six pieces of candy (inclusive)? Use the fact that

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

to find a formula for the number of ways to pass out the candy.

Solution. $(x^3 + x^4 + x^5 + x^6)^n$;

$$
\begin{aligned}
(x^3 + x^4 + x^5 + x^6)^n &= x^{3n}(1 + x + x^2 + x^3)^n\\
&= x^{3n}\left(\frac{1-x^4}{1-x}\right)^n\\
&= x^{3n}\sum_{j=0}^{n}(-1)^j\binom{n}{j}x^{4j}\sum_{i=0}^{\infty}\binom{n+i-1}{i}x^i
\end{aligned}
$$

The number of ways to pass out k pieces of candy is the coefficient of x^k in this expression. Thus the answer is zero if $k < 3n$ because of the x^{3n} in front. Otherwise the answer is the coefficient of x^{k-3n} in $\sum_{j=0}^{n}(-1)^j\binom{n}{j}x^{4j}\sum_{i=0}^{\infty}\binom{n+i-1}{i}x^i$, which is

$$\sum_{i,j:4j+i=k-3n}(-1)^j\binom{n}{j}\binom{n+i-1}{i} = \sum_{j=0}^{\lfloor(k-3n)/4\rfloor}(-1)^j\binom{n}{j}\binom{k-2n-4j-1}{n-1}.$$

Activity 201.

(a) In paying off a mortgage loan with initial amount A, annual interest rate $p\%$ on a monthly basis with a monthly payment of m, what recurrence describes the amount owed after n months of payments in terms of the amount owed after $n-1$ months? Some technical details: You make the first payment after one month. The amount of interest included in your monthly payment is $.01p/12$. This interest rate is applied to the amount you owed immediately after making your last monthly payment.

Solution. $a_n = (1 + \frac{.01p}{12})a_{n-1} - m$.

(b) Find a formula for the amount owed after n months.

Solution.   From Problem 90 or by applying generating functions we have

$$
\begin{aligned}
a_n &= A(1 + \frac{.01p}{12})^n - m\frac{1 - (1 + \frac{.01p}{12})^n}{1 - (1 + \frac{.01p}{12})} \\
&= \left(A - \frac{1200m}{p}\right)\left(1 + \frac{.01p}{12}\right)^n + \frac{1200m}{p}
\end{aligned}
$$

(c) Find a formula for the number of months needed to bring the amount owed to zero. Another technical point: If you were to make the standard monthly payment m in the last month, you might actually end up owing a negative amount of money. Therefore it is ok if the result of your formula for the number of months needed gives a non-integer number of months. The bank would just round up to the next integer and adjust your payment so your balance comes out to zero.

Solution. $\left(A - \frac{1200m}{p}\right)\left(1 + \frac{.01p}{12}\right)^n + \frac{1200m}{p} = 0$ gives us the equation $\left(1 + \frac{.01p}{12}\right)^n = \frac{1200m}{1200m - Ap}$. Taking logarithms to any base we choose gives us $n\log(1 + \frac{.01p}{12}) = \log 1200m - \log(1200m - Ap)$ and so $n = \frac{\log 1200m - \log(1200m - Ap)}{\log(1 + \frac{.01p}{12})}$.

(d) What should the monthly payment be to pay off the loan over a period of 30 years?

Solution. $360 = \frac{\log 1200m - \log(1200m - Ap)}{\log(1 + \frac{.01p}{12})}$ is the equation we need to solve for m, the monthly payment. We need to chose some base for the logarithm so we can write its inverse function; suppose we use logs to the base 10. Then

$$
360\log(1 + \frac{.01p}{12}) = \log 1200m - \log(1200m - Ap)
$$

$$10^{360 \log\left(1+\frac{.01p}{12}\right)} = 10^{\log 1200m/(1200m - Ap)}$$

$$\left(1 + \frac{.01p}{12}\right)^{360} = 1200m/(1200m - Ap)$$

$$(1200m - Ap)\left(1 + \frac{.01p}{12}\right)^{360} = 1200m$$

$$1200m\left(\left(1 + \frac{.01p}{12}\right)^{360} - 1\right) = Ap\left(1 + \frac{.01p}{12}\right)^{360}$$

$$m = \frac{Ap\left(1 + \frac{.01p}{12}\right)^{360}}{1200\left(\left(1 + \frac{.01p}{12}\right)^{360} - 1\right)}$$

is our monthly payment.

Activity 202. We have said that for nonnegative i and positive n we want to define $\binom{-n}{i}$ to be $\binom{n+i-1}{i}$. If we want the Pascal recurrence to be valid, how should we define $\binom{-n}{-i}$ when n and i are both positive?

Solution. The number $\binom{n}{k}$ is the number in row n and column k of the Pascal (right) triangle. We have said we want to have $\binom{-n}{0} = \binom{n+0-1}{0}$, so we want ones everywhere in that column. Now the Pascal recurrence gives us that $\binom{-n}{0} = \binom{-n-1}{-1} + \binom{-n-1}{0}$, so that $\binom{-n}{-1} = 0$, as does $\binom{0}{-1}$. Applying the Pascal recurrence again gives us $\binom{-n}{-1} = \binom{-n-1}{-2} + \binom{-n-1}{-1}$, so we have $\binom{-n-1}{-2} = 0$ as well. Following this pattern, we can prove by induction that $\binom{-n}{-k}$ is zero whenever k and n are positive.

Activity 203. Find a recurrence relation for the number of ways to divide a convex n-gon into triangles by means of non-intersecting diagonals. How do these numbers relate to the Catalan numbers?

Solution. Let d_n be the number of ways to divide an n-gon into triangles by means of nonintersecting diagonals. Take an n-gon and label its vertices cyclically from 1 to n. Any triangulation must have a triangle containing the edge $1n$ between vertex 1 and vertex n. The third vertex can be any number between 2 and $n-1$. We consider two cases. First, if the third vertex is 2 or $n-1$, then we have divided our polygon up into a triangle and an $(n-1)$-gon, and any triangulation of that $(n-1)$-gon joins with our original triangle to give us a triangulation of the n-gon. Second, if the third vertex of our original triangle is vertex i with $3 \le i \le n-2$ then we have divided our polygon into the polygon with the i edges 12, 23, ..., $(i-1)i$, $i1$, the polygon with $n-i+1$ edges ni, $i(i+1)$, ..., $(n-1)n$, and the original triangle with edges $n1$, $1i$, in. Triangulations of the first two of these polygons join with the original triangle to give us a triangulation of the original polygon.

The number of triangulations of the original polygon that we get from case 1 is $2d_{n-1}$. The number of triangulations we get from the second case is $\sum_{i=3}^{n-2} d_i d_{n-i+1}$. Thus the total number of triangulations is $2d_{n-1} + d_{n-2}d_3 + d_{n-3}d_4 + \cdots + d_3 d_{n-2}$. If we take $d_2 = 1$, then we may write $d_n = d_{n-1}d_2 + d_{n-2}d_3 + \cdots + d_3 d_{n-2} + d_2 d_{n-1} = \sum_{i=2}^{n-1} d_i d_{n-i+1}$. This is very similar to

the recurrence in Problem 199 for the Catalan numbers. We could apply the generating function method we used with the Catalan numbers to find a formula for d_n. We could also experiment with the first few Catalan numbers and the first few "triangulation" numbers to see if they are related. We have $C_0 = 1$, $C_1 = 1$, $C_2 = 2$, $C_3 = C_0C_2 + C_1C_1 + C_2C_0 = 5$, and $C_4 = C_0C_3 + C_1C_2 + C_2C_1 + C_3C_0 = 14$. We have that $d_2 = 1$, $d_3 = 1$, $d_4 = 2$, $d_5 = d_4d_2 + d_3d_3 + d_2d_4 = 5$, and $d_6 = d_5d_2 + d_4d_3 + d_3d_4 + d_5d_2 = 14$. This makes pretty convincing evidence that $d_i = C_{i-2}$. We have already done a base case (and more) for an inductive proof. So assume inductively that $d_i = C_{i-2}$ for $i < n$. Then

$$
\begin{aligned}
d_n &= \sum_{i=2}^{n-1} d_i d_{n-i+1} \\
&= \sum_{i=2}^{n-1} C_{i-2} C_{n-i+1-2} \\
&= \sum_{i=2}^{n-1} C_{i-2} C_{n-i-1} \\
&= \sum_{k=1}^{n-2} C_{k-1} C_{n-(k+1)-1} \\
&= \sum_{k=1}^{n-2} C_{k-1} C_{(n-2)-k} \\
&= C_{n-2}
\end{aligned}
$$

Thus by the principle of mathematical induction, $d_n = C_{n-2}$ for all integers $n \geq 2$.

Activity 204. How does $\sum_{k=0}^{n} \binom{n-k}{k}$ relate to the Fibonacci Numbers?

Solution. We begin by computing a few values of $a_n = \sum_{k=0}^{n} \binom{n-k}{k}$. We have $a_0 = 1$, $a_1 = 1$, $a_2 = 2$, $a_3 = 1+2 = 3$, $a_4 = 1+3+1 = 5$, $a_5 = 1+4+3 = 8$ and $a_6 = 1 + 5 + 6 + 1 = 13$. So far the sequence agrees with the Fibonacci Numbers. Each term of the sequence is the sum of the two preceding terms, so it makes sense to try to prove that $a_n = a_{n-1} + a_{n-2}$. We may write

$$
\begin{aligned}
a_{n-1} + a_{n-2} &= \sum_{k=0}^{n-1} \binom{n-1-k}{k} + \sum_{k=0}^{n-2} \binom{n-2-k}{k} \\
&= \sum_{k=0}^{n-1} \binom{n-1-k}{k} + \sum_{j=1}^{n-1} \binom{n-1-j}{j-1} \\
&= 1 + \sum_{k=1}^{n-1} \binom{n-1-k}{k} + \binom{n-1-k}{k-1}
\end{aligned}
$$

$$= 1 + \sum_{k=1}^{n-1} \binom{n-k}{k} \rangle \Rightarrow \sum_{k=0}^{n-1} \binom{n-k}{k} \rangle \Rightarrow a_n$$

Thus the sequence satisfies the same recurrence as the Fibonacci numbers and its first two values are the same as the Fibonacci numbers. This lets us prove by induction that a_n is the nth Fibonacci number. More generally, given any second order recurrence, if two sequences satisfy that recurrence and have the same first two values, then they are equal.

Activity 205. Let m and n be fixed. Express the generating function for the number of k-element multisets of an n-element set such that no element appears more than m times as a quotient of two polynomials. Use this expression to get a formula for the number of k-element multisets of an n-element set such that no element appears more than m times.

Solution. $(1 + x + x^2 + \cdots + x^m)^n = \frac{(1-x^{m+1})^n}{(1-x)^n}$. Expanding this gives us $\frac{(1-x^{m+1})^n}{(1-x)^n} = \sum_{i=0}^{n}(-1)^i \binom{n}{i} x^{(m+1)i} \sum_{j=0}^{\infty} \binom{n+j-1}{j}$. The coefficient of x^k in this product is the number of k-element multisets chosen from an n-element set in which no element appears more than m times. This coefficient is

$$\sum_{i,j:(m+1)i+j=k} (-1)^i \binom{n}{i} \binom{n+j-1}{j} = \sum_{i=1}^{\lfloor \frac{k}{m+1} \rfloor} (-1)^i \binom{n}{i} \binom{n+k-(m+1)i-1}{n-1}.$$

Activity 206. One natural but oversimplified model for the growth of a tree is that all new wood grows from the previous year's growth and is proportional to it in amount. To be more precise, assume that the (total) length of the new growth in a given year is the constant c times the (total) length of new growth in the previous year. Write down a recurrence for the total length a_n of all the branches of the tree at the end of growing season n. Find the general solution to your recurrence relation. Assume that we begin with a one meter cutting of new wood (from the previous year) which branches out and grows a total of two meters of new wood in the first year. What will the total length of all the branches of the tree be at the end of n years?

Solution. $a_n = a_{n-1} + c(a_{n-1} - a_{n-2}) = (1+c)a_{n-1} - ca_{n-2}$.

$$\sum_{n=2}^{\infty} a_n x^n = \sum_{n=2}^{\infty}(1+c)a_{n-1}x^n - c\sum_{n=2}^{\infty} a_{n-2}x^n$$

$$(1 - (1+c)x + cx^2)\sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x - a_0(1+c)x$$

$$\sum_{n=0}^{\infty} a_n x^n = \frac{a_0 + (a_1 - a_0(1+c))x}{1 - (1+c)x + cx^2}$$

$$= \frac{a_0 + (a_1 - a_0(1+c))x}{(1-x)(1-cx)}$$

Assuming $c \neq 1$ and using the method of partial fractions gives us

$$\frac{a_0 + (a_1 - a_0(1+c))x}{(1-x)(1-cx)}$$

$$=(a_0 + (a_1 - a_0(1+c))x)\left[\frac{1/(1-c)}{(1-x)} - \frac{c/(1-c)}{1-cx}\right]$$

$$=(a_0 + (a_1 - a_0(1+c))x)\left[\frac{1}{1-c}\sum_{i=0}^{\infty}x^i - \frac{c}{1-c}\sum_{i=0}^{\infty}c^ix^i\right]$$

$$=\frac{(a_0 + (a_1 - a_0(1+c))x)}{1-c}\sum_{i=0}^{\infty}(1 - c^{i+1})x^i.$$

From this we get that

$$a_i = \frac{a_0}{1-c}(1 - c^{i+1}) + \frac{a_1 - a_0(1+c)}{1-c}(1 - c^i).$$

Assuming that we begin with one meter of new wood means $a_0 = 1$, and assuming we have a total of two meters of new wood at the end of the first year means $c = 2$ and $a_1 = 3$. Substituting these into our formula for a_i gives us $a_i = 2^{i+1} - 1$

Activity 207. (Relevant to Chapter B.1) We have some chairs which we are going to paint with red, white, blue, green, yellow and purple paint. Suppose that we may paint any number of chairs red or white, that we may paint at most one chair blue, at most three chairs green, only an even number of chairs yellow, and only a multiple of four chairs purple. In how many ways may we paint n chairs?

Solution. The generating function for the number of ways to paint n chairs is

$$(1 + x + x^2 + \cdots)^2(1+x)(1 + x + x^2 + x^3)(1 + x^2 + x^4 + \cdots)(1 + x^4 + x^8 + \cdots)$$

$$=\frac{(1+x)(1 + x + x^2 + x^3)}{(1-x)^2(1-x^2)(1-x^4)}\rangle =\rangle \frac{1}{(1-x)^4}$$

Thus the number of ways to paint n chairs is $\binom{n+4-1}{n} = \binom{n+3}{n}$

Activity 208. What is the generating function for the number of partitions of an integer in which each part is used at most m times? Why is this also the generating function for partitions in which consecutive parts (in a decreasing list representation) differ by at most m?

Solution.

$$(1 + q + \cdots + q^m)(1 + q^2 + \cdots + q^{2m})\cdots = \prod_{i=1}^{\infty}\sum_{j=0}^{m}q^{ij} = \prod_{i=1}^{\infty}\frac{1 - q^{i(m+1)}}{1 - q^i}$$

This is also the generating function for the number of partitions of an integer in which consecutive parts differ by at most m, because when we conjugate a partition in which each part is used at most m times, we get a partition in which each distinct column of the Young diagram occurs at most m times, which means that the difference between two consecutive parts (in the decreasing list representation) is at most m.

Chapter 5

The Principle of Inclusion and Exclusion

5.1 The Principle of Inclusion and Exclusion

The size of a union of sets

One of our very first counting principles was the *sum principle* which says that the size of a union of disjoint sets is the sum of their sizes. Computing the size of overlapping sets requires, quite naturally, information about how they overlap. Taking such information into account will allow us to develop a powerful extension of the sum principle known as the "principle of inclusion and exclusion."

Activity 209. In a biology lab study of the effects of basic fertilizer ingredients on plants, 16 plants are treated with potash, 16 plants are treated with phosphate, and among these plants, eight are treated with both phosphate and potash. No other treatments are used. How many plants receive at least one treatment? If 32 plants are studied, how many receive no treatment?

Solution. The number of plants receiving treatment was $16 + 16 - 8 = 24$. The number of plants receiving no treatment was eight.

Activity 210. Give a formula for the size of the union $A \cup B$ of two sets A in terms of the sizes $|A|$ of A, $|B|$ of B, and $|A \cap B|$ of $A \cap B$. If A and B are subsets of some "universal" set U, express the size of the complement $U - (A \cup B)$ in terms of the sizes $|U|$ of U, $|A|$ of A, $|B|$ of B, and $|A \cap B|$ of $A \cap B$.

Solution. $|A \cup B| = |A| + |B| - |A \cap B|$.
$\quad |U - (A \cup B)| = |U| - |A| - |B| + |A \cap B|$.

Activity 211. In , there were just two fertilizers used to treat the sample plants. Now suppose there are three fertilizer treatments, and

15 plants are treated with nitrates, 16 with potash, 16 with phosphate, 7 with nitrate and potash, 9 with nitrate and phosphate, 8 with potash and phosphate and 4 with all three. Now how many plants have been treated? If 32 plants were studied, how many received no treatment at all?

Solution. $15 + 16 + 16 - 7 - 9 - 8 + 4 = 27$ plants were treated and five received no treatment.

Activity 212. Give a formula for the size of $A_1 \cup A_2 \cup A_3$ in terms of the sizes of A_1, A_2, A_3 and the intersections of these sets.

Solution. $|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|$.

Activity 213. Conjecture a formula for the size of a union of sets

$$A_1 \cup A_2 \cup \cdots \cup A_n = \bigcup_{i=1}^{n} A_i$$

in terms of the sizes of the sets A_i and their intersections.

Solution.

$$\left| \bigcup_{i=1}^{n} A_i \right| = \sum_{i=1}^{n} |A_i| - \sum_{i=1}^{n} \sum_{j=i+1}^{n} |A_i \cap A_j| + \cdots + (-1)^{n-1} |A_1 \cap A_2 \cap \cdots \cap A_n|.$$

The difficulty of generalizing Problem 212 to Problem 213 is not likely to be one of being able to see what the right conjecture is, but of finding a good notation to express your conjecture. In fact, it would be easier for some people to express the conjecture in words than to express it in a notation. Here is some notation that will make your task easier. Let us define

$$\bigcap_{i:i \in I} A_i$$

to mean the intersection over all elements i in the set I of A_i. Thus

$$\bigcap_{i:i \in \{1,3,4,6\}} = A_1 \cap A_3 \cap A_4 \cap A_6. \tag{5.1}$$

This kind of notation, consisting of an operator with a description underneath of the values of a dummy variable of interest to us, can be extended in many ways. For example

$$\sum_{I:I \subseteq \{1,2,3,4\}, \, |I|=2} |\cap_{i \in I} A_i| = |A_1 \cap A_2| + |A_1 \cap A_3| + |A_1 \cap A_4|$$

$$+ \qquad\qquad\qquad |A_2 \cap A_3| + |A_2 \cap A_4| + |A$$

Activity 214. Use notation something like that of Equation (5.1) and ⟨⟨Unresolved xref, reference "notationsolution"; check spelling or use "provisional" attribute⟩⟩ to express the answer to Problem 213. Note there are many different correct

ways to do this problem. Try to write down more than one and choose the nicest one you can. Say why you chose it (because your view of what makes a formula nice may be different from somebody else's). The nicest formula won't necessarily involve all the elements of Equations (5.1) and ⟨⟨Unresolved xref, reference "notationsolution"; check spelling or use "provisional" attribute⟩⟩.

Solution.
$$\left| \bigcup_{i=1}^{n} A_i \right| = \sum_{S:S\subseteq[n],S\neq\emptyset} (-1)^{|S|-1} |\bigcap_{i:i\in S} A_i|$$

I chose this way of writing the formula partly because it is efficient with symbols; for example, it uses only one sum sign. But more importantly I chose it because it captures what I would want to say in words: "You sum, over all ways of choosing an intersection of the sets A_i, the size of the intersection times a sign factor that is -1 if you are intersecting an even number of sets and 1 if you are intersecting an odd number." If I were writing my solution out in words, I would probably assume that nobody would think about the possibility of an intersection of the empty set of the A_is, but I had to put the $S \neq \emptyset$ in my formula because otherwise the formula would have had us consider the possibility that S was empty.

Activity 215. A group of n students goes to a restaurant carrying backpacks. The manager invites everyone to check their backpack at the check desk and everyone does. While they are eating, a child playing in the check room randomly moves around the claim check stubs on the backpacks. What is the probability that, at the end of the meal, at least one student receives his or her own backpack? In other words, in what fraction of the total number of ways to pass the backpacks back does at least one student get his or her own backpack back? (Hint: For each student, how big is the set of backpack distributions in which that student gets the correct backpack? It might be a good idea to first consider cases with $n = 3$, 4, and 5.) What is the probability that no student gets his or her own backpack?

Solution. First we compute the number of ways to pass out the backpacks so that at least one student gets the correct one, then we divide by $n!$ the total number of ways to pass out the backpacks to get the probability that at least one student gets the correct one. If we let A_i be the set of backpack distributions in which student i gets the correct backpack, then the number we want to compute is the size of the union of the sets A_i. For this purpose we need to know, for every subset $S \subseteq [n]$ the size of $\cap_{i\in S}A_i$. That is, we need to know the number of ways to pass out the backpacks so that student i gets the correct one for each i in S. It won't matter whether or not other students get the correct backpacks, so we can just assume that for each $i \in S$, student i gets the correct backpack and then hand out the remaining $n - |S|$ backpacks to the remaining $n - |S|$ students in $(n - |S|)!$ ways. Thus $(n - |S|)!$ is $|\cap_{i:i\in S} A_i|$. Using our formula from Problem 214 we get

$$\left| \bigcup_{i=1}^{n} A_i \right| = \sum_{S:S\subseteq[n],S\neq\emptyset} (-1)^{|S|-1} \left| \bigcap_{i:i\in S} A_i \right|$$

$$= \sum_{S: S \subseteq [n], S \neq \emptyset} (-1)^{|S|-1}(n - |S|)!$$

$$= \sum_{s=1}^{n} \binom{n}{s}(-1)^{s-1}(n - s)!$$

$$= \sum_{s=1}^{n} (-1)^{s-1}\frac{n!}{s!}$$

Dividing this by $n!$, the total number of ways to pass back the backpacks, we get $\sum_{s=1}^{n} \frac{(-1)^{s-1}}{s!}$ for the probability that at least one student gets the correct backpack. Subtracting from 1 to get the probability that no student gets the correct backpack gives us

$$1 - \sum_{s=1}^{n} \frac{(-1)^{s-1}}{s!} = \sum_{s=0}^{n} \frac{(-1)^{s}}{s!}.$$

Activity 216. As the number of students becomes large, what does the probability that no student gets the correct backpack approach?

Solution. From calculus, we know that $e^x = \sum_{j=0}^{\infty} \frac{x^j}{j!}$. Substituting $x = -1$ gives us $e^{-1} = \sum_{j=0}^{\infty} \frac{(-1)^j}{j!}$ which is the limit as n becomes infinite of the probability in the solution to Problem 215. Thus the probability approaches $1/e$.

The formula you have given in Problem 214 is often called *the principle of inclusion and exclusion* for unions of sets. The reason is the pattern in which the formula first adds (includes) all the sizes of the sets, then subtracts (excludes) all the sizes of the intersections of two sets, then adds (includes) all the sizes of the intersections of three sets, and so on. Notice that we haven't yet proved the principle. We will first describe the principle in an apparently more general situation that doesn't require us to translate each application into the language of sets. While this new version of the principle might seem more general than the principle for unions of sets; it is equivalent. However once one understands the notation used to express it, it is more convenient to apply.

Problem 215 is "classically" called the *hatcheck problem*; the name comes from substituting hats for backpacks. If is also sometimes called the *derangement problem*. A *derangement* of an n-element set is a permutation of that set (thought of as a bijection) that maps no element of the set to itself. One can think of a way of handing back the backpacks as a permutation f of the students: $f(i)$ is the owner of the backpack that student i receives. Then a derangement is a way to pass back the backpacks so that no student gets his or her own.

The hatcheck problem restated

The last question in Problem 215 requires that we compute the number of ways to hand back the backpacks so that nobody gets his or her own backpack. We can think of the set of ways to hand back the backpacks so that student i gets the correct one as the set of permutations of the backpacks with the property that student i gets his or her own backpack. Since there are $n-1$ other students and they can receive any of the remaining $n-1$ backpacks in $(n-1)$ ways, the number of permutations with the property that student i gets the correct backpack is $(n-1)!$. How many permutations are there with the properties that student i gets the correct backpack *and* student j gets the correct backpack? (Let's call these properties i and j for short.) Since there are $n-2$ remaining students and $n-2$ remaining backpacks, the number of permutations with properties i and j is $(n-2)!$. Similarly, the number of permutations with properties $i_1, i_2, \ldots, i_k$ is $(n-k)!$. Thus when we compute the size of the union of the sets

$$S_i = \{f : f \text{ is a permutation with property } i\},$$

we are computing the number of ways to pass back the backpacks so that at least one student gets the correct backpack. This answers the first question in Problem 215. The last question in Problem 215 is asking us for the number of ways to pass back the backpacks that have *none* of the properties. To say this in a different way, the question is asking us to compute the number of ways of passing back the backpacks that have exactly the *empty set*, $\emptyset$, of properties.

Basic counting functions: $N_{atleast}$ and $N_{exactly}$

Notice that the quantities that we were able to count easily were the number of ways to pass back the backpacks so that we satisfy a certain subset $K = \{i_1, i_2, \ldots, i_k\}$ of our properties. In fact, among the $(n-k)$ ways to pass back the backpacks with this particular set K of properties is the permutation that gives each student the correct backpack, and has not just the properties in K, but the whole set of properties. Similarly, for any set M of properties with $K \subseteq M$, the permutations that have all the properties in M are among the $(n-k)!$ permutations that have the properties in the set K. Thus we can think of $(n-k)!$ as counting the number of permutations that have *at least* the properties in K. In particular, $n!$ is the number of ways to pass back the backpacks that have at least the empty set of properties. We thus write $N_{\text{at least}}(\emptyset) = n!$, or $N_{\text{a}}(\emptyset) = n!$ for short. For a k-element subset K of the properties, we write $N_{\text{at least}}(K) = (n-k)!$ or $N_{\text{a}}(K) = (n-k)!$ for short.

The question we are trying to answer is "How many of the distributions of backpacks have exactly the empty set of properties?" For this purpose we introduce one more piece of notation. We use $N_{\text{exactly}}(\emptyset)$ or $N_{\text{e}}(\emptyset)$ to stand for the number of backpack distributions with exactly the empty set of properties, and for any set K of properties we use $N_{\text{exactly}}(K)$ or $N_{\text{e}}(K)$

to stand for the number of backpack distributions with exactly the set K of properties. Thus $N_{\mathrm{e}}(K)$ is the number of distributions in which the students represented by the set K of properties get the correct backpacks back and no other students do.

The principle of inclusion and exclusion for properties

For the principle of inclusion and exclusion for properties, suppose we have a set of arrangements (like backpack distributions) and a set P of properties (like student i gets the correct backpack) that the arrangements might or might not have. We suppose that we know (or can easily compute) the numbers $N_{\mathrm{a}}(K)$ for every subset K of P. We are most interested in computing $N_{\mathrm{e}}(\emptyset)$, the number of arrangements with none of the properties, but it will turn out that with no more work we can compute $N_{\mathrm{e}}(K)$ for every subset K of P. Based on our answer to Problem 215 we expect that

$$N_{\mathrm{e}}(\emptyset) = \sum_{S:S\subseteq P} (-1)^{|S|} N_{\mathrm{a}}(S) \qquad (5.2)$$

and it is a natural guess that, for every subset K of S,

$$N_{\mathrm{e}}(K) = \sum_{S:K\subseteq S\subseteq P} (-1)^{|S|-|K|} N_{\mathrm{a}}(S). \qquad (5.3)$$

Equations (5.2) and Equation (5.3) are called *the principle of inclusion and exclusion for properties.*

Activity 217. Verify that the formula for the number of ways to pass back the backpacks in Problem 215 so that nobody gets the correct backpack has the form of Equation (5.2).

Solution. Suppose we take property i to be the property that student i gets the correct backpack. In our solution to Problem 215 we thought of S as a subset of $[n]$. Since we are numbering our properties from 1 to n, each set S determines a corresponding set S' of properties: those properties numbered by the elements of S. We will use P to stand for the set of all properties, that is property 1 through property n. The number of ways to pass back the hats so that nobody gets the correct hat is $n!$ minus the number of ways to pass back the hats so that somebody does get the correct hat back. From our solution to Problem 215, we see that one way to express this is as

$$n! - \sum_{S:S\subseteq[n],S\neq\emptyset} (-1)^{|S|-1}(n-|S|)! = n! + \sum_{S:S\subseteq[n],S\neq\emptyset} (-1)^{|S|}(n-|S|)!$$

$$= \sum_{S:S\subseteq[n]} (-1)^{|S|}(n-|S|)!$$

$$= \sum_{S:S\subseteq[n]} (-1)^{|S|} N_{\mathrm{a}}(S')$$

$$= \sum_{S':S'\subseteq P} (-1)^{|S|} N_{\mathrm{a}}(S')$$

because $N_{\mathrm{a}}(S')$ is the number of ways to pass back the backpacks so that at least the students in S get the correct backpacks, which is the same as the size of $\bigcap_{i:i\in S} A_i$, and which we computed to be $(n-|S|)!$.

Activity 218. Find a way to express $N_{\mathrm{a}}(S)$ in terms of $N_{\mathrm{e}}(J)$ for subsets J of P containing S. In particular, what is the equation that expresses $N_{\mathrm{a}}(\emptyset)$ in terms of $N_{\mathrm{e}}(J)$ for subsets J of P?

Solution. $N_{\mathrm{a}}(S) = \sum_{J:S\subseteq J\subseteq P} N_{\mathrm{e}}(J)$. In particular, we get by substitution that $N_{\mathrm{a}}(\emptyset) = \sum_{J:J\subseteq P} N_{\mathrm{e}}(J)$.

Activity 219.

(a) Substitute the formula for N_{a} from Problem 218 into the right hand sides of the formulas of Equations (5.2) and Equation (5.3) and simplify what you get to show that for Equations (5.2) and Equation (5.3) the right-hand sides are indeed equal to the left-hand sides. This will prove that those equations are true. (Hint: You will get a double sum. If you can figure out how to reverse the order of the two summations, the binomial theorem may help you simplify the formulas you get.)

Solution. Since Equation (5.2) is a special case of Equation (5.3), we will just substitute the formula from Problem 218 into the right-hand side of Equation (5.3). This gives us

$$\sum_{S:K\subseteq S\subseteq P} (-1)^{|S|-|K|} N_{\mathrm{a}}(S) = \sum_{S:K\subseteq S\subseteq P} (-1)^{|S|-|K|} \sum_{J:S\subseteq J\subseteq P} N_{\mathrm{e}}(J)$$

$$= \sum_{S:K\subseteq S\subseteq P} \sum_{J:S\subseteq J\subseteq P} (-1)^{|S|-|K|} N_{\mathrm{e}}(J)$$

$$= \sum_{J:K\subseteq J\subseteq P} \sum_{S:K\subseteq S\subseteq J} (-1)^{|S|-|K|} N_{\mathrm{e}}(J)$$

$$= \sum_{J:K\subseteq J\subseteq P} N_{\mathrm{e}}(J) \sum_{S:K\subseteq S\subseteq J} (-1)^{|S|-|K|}$$

$$= \sum_{J:K\subseteq J\subseteq P} N_{\mathrm{e}}(J) \sum_{s=|K|}^{|J|} \binom{|J|-|K|}{s-|K|} (-1)^{s-|K|}$$

$$= \sum_{J:K\subseteq J\subseteq P} N_{\mathrm{e}}(J) \sum_{i=0}^{|J|-|K|} \binom{|J|-|K|}{i} (-1)^{i}$$

$$= \sum_{J:K\subseteq J\subseteq P} N_{\mathrm{e}}(J)(1-1)^{|J|-|K|}$$

$$= N_{\mathrm{e}}(K).$$

In the fourth-to-last line of our equations, we used the fact that the number of subsets S of J that contain K is the number of ways to choose the elements of the set $S - K$ of elements in S but not K from the elements of $J - K$, the set of elements of J that are not in K. In the second-to-last line of the equations, we recognized that the second sum in the third-to-last line is the kind of sum we would get by applying the binomial theorem to something of the form $(x+y)^{|J|-|K|}$ for a suitable x and y, and then saw that $x = 1$ and $y = -1$ would give us exactly the second sum in the third-to-last line. In going from the second-to-last line to the last line we used the facts that $0^n = o$ if $n > 0$ and $0^0 = 1$. This proves that Equation (5.3) is true for all K, and in particular when $K = \emptyset$, which also proves Equation (5.2). We have just proved the Principle of inclusion and exclusion.

(b) In how many ways may we distribute k identical apples to n children so that no child gets more than three apples?

Solution. Let property i be "Child i gets four or more apples." Then we are asking for the number of ways to pass out the apples that have none of the properties, so we are asking for $N_e(\emptyset)$. From Equation (5.2) we see that to find this number we need to know $N_a(S)$ for every subset S of our set of properties. But if S has size s, then having the properties in S means that all the children in a particular set of size s will get four or more apples. We already know how to pass out the apples so that the children in a particular set $\hat{S}$ of children get at least four apples: we give everyone in $\hat{S}$ four apples, and then pass out the remaining $k - 4s$ apples to the children in $\binom{k-4s+n-1}{k-4s} = \binom{k-4s+n-1}{n-1}$ ways. This counts the number of ways to give at least four apples to every child in $\hat{S}$, and maybe give four apples to some other children as well. Thus $N_a(S) = \binom{k-4|S|+n-1}{n-1}$. Applying Equation (5.2) gives us

$$N_e(\emptyset) = \sum_{S:S\subseteq P} (-1)^{|S|} N_a(S)$$

$$= \sum_{s=0}^{k} \binom{k}{s} (-1)^s \binom{k-4s+n-1}{n-1}$$

$$= \sum_{s=0}^{k} (-1)^s \binom{k}{s} \binom{k-4s+n-1}{n-1}$$

Activity 220. A group of n married couples comes to a group discussion session where they all sit around a round table. In how many ways can they sit so that no person is next to his or her spouse? (Note that two people of the same sex can sit next to each other.)

Solution. Let Property i be that couple i sits together. We are interested in the number of seating arrangements that have none of the properties. Thus

for a set S of properties, we need to compute $N_a(S)$. If we let each couple described by S sit together, we will seat $|S|$ couples and $2n - 2|S|$ individuals around the table. We can do this in $2^{|S|}(|S| + 2n - 2|S| - 1)!$ ways, because once we chose a place for a couple (i.e. two adjacent seats) there are two ways the couple can sit down. Thus $N_a(S) = 2^{|S|}(2n - |S| - 1)!$. Substituting this into Equation (5.2) with P as the set of all properties gives us

$$N_e(\emptyset) = \sum_{S:S \subseteq P} (-1)^{|S|} N_a(S)$$

$$= \sum_{s=0}^{n} (-1)^s \binom{n}{s} 2^s (2n - s - 1)!$$

Activity 221. A group of n married couples comes to a group discussion session where they all sit around a round table. In how many ways can they sit so that no person is next to his or her spouse or a person of the same sex? This problem is called the *menage problem*. (Hint: Reason somewhat as you did in Problem 220, noting that if the set of couples who do sit side-by-side is nonempty, then the sex of the person at each place at the table is determined once we seat one couple in that set.)

Solution. We are going to consider arrangements of the couples alternating sex around the table. Property i of an arrangement is that couple i sits together. We are interested in the number of arrangements that have exactly none of the properties. Thus for each subset S of the properties, we consider the number of arrangements that has at least the set S of properties. We distinguish the case that S is empty from the others. $N_a(\emptyset)$ is just the number of ways to seat $2n$ couples around the table, alternating sex, but with no other restrictions. We can arrange one of the sexes in a circle in $n - 1!$ ways and then assign the members of the opposite sex to the places between them in $n!$ ways, so $N_a(\emptyset) = (n - 1)!n!$. (Another way to get this result is to let one person sit down. This determines the sex of the person at each place of the table, so there are $(n - 1)!$ ways to assign the people of the same sex of the first person, and $n!$ ways to assign the people of the opposite. It appears that there are $2n$ choices for where the first person sits, but we can break the seating charts up into blocks of $2n$ seating charts, each of which gives the same circular arrangement. Thus there are $(n - 1)!n!$ inequivalent ways to seat the people with at least the empty set of properties.)

Now if S is nonempty, and has s members, we seat one of the couples that must sit together (say the first in alphabetical order), and this determines the sex of the person that must sit at each other place. There are $2n$ pairs of adjacent seats where we can seat that couple and two ways they can sit in the pair of adjacent seats that we choose. Then we have $s - 1$ couples, $n - s$ men and $n - s$ women to seat in the remaining places. First we arrange the $s - 1$ couples and $2n - 2s$ identical empty chairs in places at the table in $(2n - 2s + s - 1)!/(2n - 2s)! = (2n - s - 1)!/(2n - 2s)!$ ways. Each couple can sit in only one way in the places they have chosen, because the sex of the person in a given place has been determined by how the first couple sits. The

sex of the person in each of the remaining chairs has been determined, so we assign the men to their seats in $(n - s)!$ ways and we assign the women to their seats in $(n - s)!$ ways. Thus we have $2 \cdot 2n(2n - s - 1)!(n - s)!^2/(2n - 2s)!$ ways to place the people. But we can partition the placements into blocks of $2n$ equivalent placements, because shifting everyone the same number of places to the right or left gives an equivalent placement. Thus the number of inequivalent seating arrangements is

$$\frac{2(2n - s - 1)!(n - s)!^2}{(2n - 2s)!} = \frac{2(2n - s - 1)!(n - s)!^2}{2(n - s)(2n - 2s - 1)!}$$
$$= \frac{(2n - s - 1)!(n - s)!(n - s - 1)!}{(2n - 2s - 1)!}.$$

Notice that if we take $s = 0$, this formula reduces to $(n - 1)!n!$. Thus for all sets S

$$N_{\mathrm{a}}(S) = \frac{(2n - s - 1)!(n - s)!(n - s - 1)!}{(2n - 2s - 1)!}.$$

Then from Equation (5.2)

$$N_{\mathrm{e}} = \sum_{S:S \subseteq P} (-1)^{|S|} \frac{(2n - |S| - 1)!(n - |S|)!(n - |S| - 1)!}{(2n - 2|S| - 1)!}$$
$$= \sum_{s=0}^{n} (-1)^s \binom{n}{s} \frac{(2n - s - 1)!(n - s)!(n - s - 1)!}{(2n - 2s - 1)!}$$
$$= \sum_{s=0}^{n} (-1)^s \frac{n!}{s!(n - s)!} \frac{(2n - s - 1)!(n - s)!(n - s - 1)!}{(2n - 2s - 1)!}$$
$$= \sum_{s=0}^{n} (-1)^s \frac{n!(2n - s - 1)!(n - s - 1)!}{s!(2n - 2s - 1)!}$$
$$= \sum_{s=0}^{n} (-1)^s \binom{2n - s - 1}{s} n!(n - s - 1)!$$

is the number of ways to seat the people, alternating sex, so that no couple sits together.

Counting onto functions

Activity 222. Given a function f from the k-element set K to the n-element set $[n]$, we say f has property i if $f(x) \neq i$ for every x in K. How many of these properties does an onto function have? What is the number of functions from a k-element set onto an n-element set?

Solution.   An onto function has none of these properties. Thus, using this set P of properties, the number of onto functions is $N_{\mathrm{e}}(\emptyset)$. For a set S of these properties, $N_{\mathrm{a}}(S)$ is the number of functions from K to $[n] - S'$, where

S' is the set of all i such that property i is in S. This number of functions is $(n - |S|)^k$. Thus by Equation (5.2)

$$N_e(\emptyset) = \sum_{S:s\subseteq P} (-1)^{|S|}(n - |S|)^k$$

$$= \sum_{s=0}^{n}(-1)^s \binom{n}{s}(n - s)^k$$

is the number of functions from K onto $[n]$.

Activity 223. Find a formula for the Stirling number (of the second kind) $S(k, n)$.

Solution. Since the number of functions from $[k]$ onto $[n]$ is $S(k, n)n!$, we get from the solution to Problem 222

$$S(k, n) = \frac{1}{n!}\sum_{s=0}^{n}(-1)^s \binom{n}{s}(n - s)^k.$$

The chromatic polynomial of a graph

We defined a graph to consist of set V of elements called vertices and a set E of elements called edges such that each edge joins two vertices. A *coloring* of a graph by the elements of a set C (of colors) is an assignment of an element of C to each vertex of the graph; that is, a function from the vertex set V of the graph to C. A coloring is called *proper* if for each edge joining two distinct vertices[1], the two vertices it joins have different colors. You may have heard of the famous four color theorem of graph theory that says if a graph may be drawn in the plane so that no two edges cross (though they may touch at a vertex), then the graph has a proper coloring with four colors. Here we are interested in a different, though related, problem: namely, in how many ways may we properly color a graph (regardless of whether it can be drawn in the plane or not) using k or fewer colors? When we studied trees, we restricted ourselves to connected graphs. (Recall that a graph is connected if, for each pair of vertices, there is a walk between them.) Here, disconnected graphs will also be important to us. Given a graph which might or might not be connected, we partition its vertices into blocks called *connected components* as follows. For each vertex v we put all vertices connected to it by a walk into a block together. Clearly each vertex is in at least one block, because vertex v is connected to vertex v by the trivial walk consisting of the single vertex v and no edges. To have a partition, each vertex must be in one and only one block. To prove that we have defined a partition, suppose that vertex v is in the blocks B_1 and B_2. Then B_1 is the set of all vertices connected by walks to some vertex v_1 and B_2 is the set of all vertices connected by walks to some vertex v_2.

---

[1] If a graph had a loop connecting a vertex to itself, that loop would connect a vertex to a vertex of the same color. It is because of this that we only consider edges with two distinct vertices in our definition.

Activity 224. (Relevant in Appendix B.1 as well as this section.) Show that $B_1 = B_2$.

Solution. Since v is in B_1, there is a walk from v_1 to v. Since there is a walk from every vertex in B_1 to v_1, there is a walk from every vertex in in B_1 to v. But there is a walk from v to v_2 since $v \in B_2$. Thus there is a walk from every vertex in B_1 to v_2. Then by our description of B_2 just before the problem, every vertex in B_1 is also in B_2. A similar argument shows that every vertex in B_2 is also in B_1. Thus $B_1 = B_2$.

Since $B_1 = B_2$, these two sets are the same block, and thus all blocks containing v are identical, so v is in only one block. Thus we have a partition of the vertex set, and the blocks of the partition are the connected components of the graph. Notice that the connected components depend on the edge set of the graph. If we have a graph on the vertex set V with edge set E and another graph on the vertex set V with edge set E', then these two graphs could have different connected components. It is traditional to use the Greek letter γ (gamma)[2] to stand for the number of connected components of a graph; in particular, $\gamma(V, E)$ stands for the number of connected components of the graph with vertex set V and edge set E. We are going to show how the principle of inclusion and exclusion may be used to compute the number of ways to properly color a graph using colors from a set C of c colors.

Activity 225. Suppose we have a graph G with vertex set V and edge set E. Suppose F is a subset of E. Suppose we have a set C of c colors with which to color the vertices.

(a) In terms of $\gamma(V, F)$, in how many ways may we color the vertices of G so that each edge in F connects two vertices of the same color?

   Solution. For each edge in F to connect two vertices of the same color, we must have all the vertices in a connected component of the graph with vertex set V and edge set F colored the same color. Thus the number of such colorings is $c^{\gamma(V,F)}$.

(b) Given a coloring of G, for each edge e in E, let us consider the property that the endpoints of e are colored the same color. Let us call this property "property e." In this way each set of properties can be thought of as a subset of E. What set of properties does a proper coloring have?

   Solution. A proper coloring has none of the properties.

(c) Find a formula (which may involve summing over all subsets F of the edge set of the graph and using the number $\gamma(V, F)$ of connected components of the graph with vertex set V and edge set F) for the number of proper colorings of G using colors in the set C.

---

[2]The greek letter gamma is pronounced gam-uh, where gam rhymes with ham.

Solution. $N_e(\emptyset) = \sum_{F:F\subseteq E}(-1)^{|F|}c^{\gamma(V,F)}$.

The formula you found in Problem is a formula that involves powers of c, and so it is a polynomial function of c. Thus it is called the **chromatic polynomial** of the graph G. Since we like to think about polynomials as having a variable x and we like to think of c as standing for some constant, people often use x as the notation for the number of colors we are using to color G. Frequently people will use $\chi_G(x)$ to stand for the number of way to color G with x colors, and call $\chi_G(x)$ the **chromatic polynomial** of G.

Activity 226.

(a) In Chapter 2 we introduced the deletion-contraction recurrence for count-ing spanning trees of a graph. Figure out how the chromatic polynomial of a graph is related to those resulting from deletion of an edge e and from contraction of that same edge e. Try to find a recurrence like the one for counting spanning trees that expresses the chromatic polynomial of a graph in terms of the chromatic polynomials of $G - e$ and G/e for an arbitrary edge e. Use this recurrence to give another proof that the number of ways to color a graph with x colors is a polynomial function of x.

Solution. The number of colorings of $G - e$ is equal to the number of proper colorings of G plus the number of colorings of G that are proper except for giving both ends of e the same color. But the number of colorings of G that are proper except for giving both ends of e the same color is the number of proper colorings of G/e. Therefore $\chi_{G-e}(x) = \chi_G(x) - \chi_{G/e}(x)$. This gives us $\chi_G(x) = \chi_{G-e}(x) - \chi_{G/e}(x)$. We can use this to prove inductively that $\chi_G(x)$ is a polynomial in x. If G has one vertex, then the number of ways to color G properly with x colors is x. This is a polynomial in x. Now suppose inductively that G has more than one vertex and whenever a graph H has fewer vertices than G, the function $\chi_H(x)$ is a polynomial function in x. Then $\chi_G(x) = \chi_{G-e}(x) - \chi_{G/e}(x)$, is a difference of of two polynomial functions in x, so it is a polynomial function in x. Therefore by the principle of mathematical induction, for all graphs G on a finite vertex set, the number of ways to properly color G in x colors is a polynomial in x.

(b) Use the deletion-contraction recurrence to compute the chromatic poly-nomial of the graph in Figure 5.1.2. (You can simplify your compu-tations by thinking about the effect on the chromatic polynomial of deleting an edge that is a loop, or deleting one of several edges between the same two vertices.)

Solution. If a graph has a loop it has no proper colorings. The graph in Figure 5.1.2 has no loops and no multiple edges between two ver-tices. The only way we could get a loop is by contracting one of sev-eral multiple edges between two vertices, and the resulting graph would

have no contribution to the chromatic polynomial of the original graph. Thus whenever a contraction gives us a graph with multiple edges between two vertices, we can replace the multiple edges by one edge and go on with our computation from there. The graphs we get when we delete and contract the edges $\{1, 2\}$ and $\{2, 3\}$ are $(G - \{1, 2\}) - \{2, 3\}$, $(G - \{1, 2\})/\{2, 3\}$, $(G - \{2, 3\})/\{1, 2\}$, and $(G/\{2, 3\})/\{1, 2\}$. These are shown in the following picture.
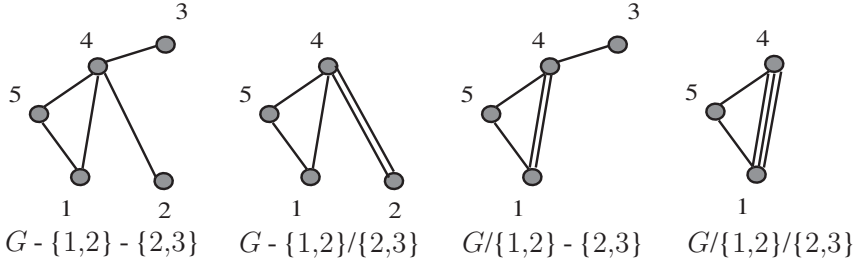

G - $\{1,2\}$ - $\{2,3\}$ G - $\{1,2\}/\{2,3\}$ $G/\{1,2\}$ - $\{2,3\}$ $G/\{1,2\}/\{2,3\}$

Figure 5.1.1: Graphs obtained through deletions and contractions

The chromatic polynomial of a triangle is $x(x - 1)(x - 2)$ because for one vertex we have x colors, for a second we have $x - 1$, and for the third vertex, because it is adjacent to both of the other vertices, we have $x - 2$ choices of colors. For a vertex of degree 1 there are $x - 1$ choices of colors, those colors not used on the one vertex it is adjacent to. As we mentioned, the extra edges do not change the chromatic polynomial, so we have that the chromatic polynomial of $(G - \{1, 2\}) - \{2, 3\}$ is $x(x - 1)^3(x - 2)$, the chromatic polynomial of $(G - \{1, 2\})/\{2, 3\}$ is $x(x - 1)^2(x - 2)$, as is that of $(G - \{2, 3\})/\{1, 2\}$, and the chromatic polynomial of $(G/\{2, 3\})/\{1, 2\}$ is $(x - 1)(x - 2)(x - 3)$. Using the deletion-contraction recurrence, we get that

$$
\begin{aligned}
\chi_G(x) =& \chi_{G-\{1,2\}}(x) - \chi_{G/\{1,2\}}(x) \\
=& \chi_{G-\{1,2\}-\{2,3\}}(x) - \chi_{(G-\{1,2\})/\{2,3\}}(x) - \chi_{(G/\{1,2\})-\{2,3\}}(x) \\
& + \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \chi_{G/\{1,2\}/\{2,3\}}(x) \\
=& x(x - 1)^3(x - 2) - 2x(x - 1)^2(x - 2) + x(x - 1)(x - 2) \\
=& x(x - 1)(x - 2)(x^2 - 2x + 1 + 2x - 2 + 1) \\
=& x^3(x - 1)(x - 2)
\end{aligned}
$$

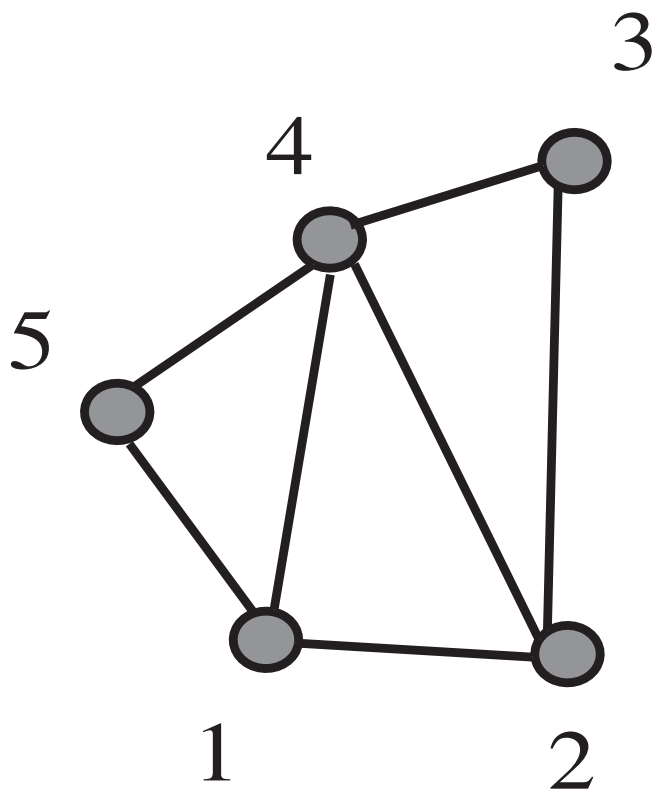for the chromatic polynomial of G.

Figure 5.1.2: A graph.

Activity 227.

(a) In how many ways may you properly color the vertices of a path on n vertices with x colors? Describe any dependence of the chromatic polynomial of a path on the number of vertices. In how many ways may you properly color the vertices of a cycle on n vertices with x colors? Describe any dependence of the chromatic polynomial of a cycle on the number of vertices.

Solution. To color the vertices of a path, start at one end. There are x colors for that vertex, and $x - 1$ colors for each of the next $n - 1$, since each of them must be different from the preceding one. Thus the chromatic polynomial of a path on n vertices is $x(x - 1)^{n-1}$. The dependence on the number of vertices appears in the exponent on $x - 1$. If we use C_n to stand for a path on n vertices and P_n to stand for a path on n vertices, then by the deletion-contraction recurrence, we may write

$$\chi_{C_n}(x) = \chi_{P_n}(x) - \chi_{C_{n-1}}(x)$$
$$= \chi_{P_n}(x) - \chi_{P_{n-1}}(x) + \chi_{C_{n-2}}(x)$$

$$=\chi_{P_n}(x) - \chi_{P_{n-1}}(x) + \chi_{P_{n-3}}(x) - \cdots + (-1)^{n-3}(\chi_{P_3}(x) - \chi_{C_2}(x))$$
$$=x(x-1)^{n-1} - x(x-1)^{n-2} + x(x-1)^{n-3}\cdots$$
$$+ \qquad\qquad\qquad\qquad (-1)^{n-3}[x(x-$$
$$=x(x-1)\sum_{i=0}^{n-2}(x-1)^i(-1)^{n-2-i}$$
$$=x(x-1)(-1)^{n-2}\sum_{i=0}^{n-2}(1-x)^i$$
$$=x(x-1)(-1)^{n-2}\frac{1-(1-x)^{n-1}}{1-(1-x)}$$
$$=(x-1)[(x-1)^{n-1} + (-1)^n].$$

Here the dependence on n is interesting; effectively, we are taking $(x-1)$ times the result of dropping the constant term from $(x-1)^{n-1}$.

(b) In how many ways may you properly color the vertices of a tree on n vertices with x colors?

Solution. Color an arbitrary vertex; you have x choices for the color of that vertex. No two vertices adjacent to it are adjacent (otherwise we'd have a cycle), so for for each of them you have $x - 1$ choices of colors. No two vertices adjacent to colored vertices are adjacent to each other, nor is one of them adjacent to two colored vertices (in either case you'd have a cycle), so for each of them you'd have $x - 1$ colors. You can continue this argument until all vertices are colored, so you have $x(x-1)^{n-1}$ ways to color the vertices.

You can also prove by induction that the chromatic polynomial of a tree is $x(x-1)^{n-1}$. This is clearly true if there is one vertex. Otherwise, choose a vertex of degree 1 in an n-vertex tree and remove it. You may inductively assume that the chromatic polynomial of the remaining tree is $x(x-1)^{n-2}$. Now there are $x - 1$ choices for the color of the vertex you removed since it has degree 1, and so the chromatic polynomial of the tree is $x(x-1)^{n-1}$. There is also an inductive argument in which you delete and contract an arbitrary edge.

Activity 228. What do you observe about the signs of the coefficients of the chromatic polynomial of the graph in Figure 5.1.2? What about the signs of the coefficients of the chromatic polynomial of a path? Of a cycle? Of a tree? Make a conjecture about the signs of the coefficients of a chromatic polynomial and prove it.

Solution. not all powers of x appear, but the signs alternate as the power of x increases; that is, the sign of x^i is opposite that of x^{i+1}. More precisely, if c_i is the coefficient of x^i, then $(-1)^{n-i}c_i \geq 0$. To prove this, note it is trivially

true for a graph with no edges. Choose an edge e of G. Then $\chi_G(x) = \chi_{G-e}(x) - \chi_{G/e}(x)$. In $G - e$, we may assume inductively that $(-1)^{n-i}c_i' \geq 0$ and in $G - /e$ we can assume inductively that $c_i''(-1)^{n-1-i} \geq 0$, where we use c_i' and c_i'' as the coefficient of x^i in $\chi_{G-e}(x)$ and $\chi_{G/e}(x)$, respectively. Then $c_i = c_i' - c_i''$, and

$$c_i(-1)^{n-i} = c_i'(-1)^{n-i} - c_i''(-1)^{n-i} = c_i'(-1)^{n-i} + c_i''(-1)^{n-1-i} \geq 0.$$

Therefore by the principle of mathematical induction, $c_i(-1)^i \geq 0$ for all finite graphs.

5.2 Supplementary Problems

Activity 229. Each person attending a party has been asked to bring a prize. The person planning the party has arranged to give out exactly as many prizes as there are guests, but any person may win any number of prizes. If there are n guests, in how many ways may the prizes be given out so that nobody gets the prize that he or she brought?

Solution. We use inclusion and exclusion. Property P_i is that person i gets the prize he or she brought. We are interested in $N_e(\emptyset)$. If P is the set of all properties, we need to compute $N_a(S)$ for every subset S of P. But $N_a(S)$ is the number of functions from the prizes to the people that give each prize represented by a property in S to the person who brought it. Think in terms of distributing those prizes first. Then there are $n - |S|$ other prizes that we may pass out to the n people as we please, so we may do that in $n^{n-|S|}$ ways. Thus $N_a(S) = n^{n-|S|}$. Applying Equation (5.2), we get

$$N_e(\emptyset) = \sum_{S:s\subseteq P} (-1)^{|S|} n^{n-|S|} = \sum_{s=0}^{n} (-1)^{|S|} \binom{n}{s} n^{n-s}.$$

Activity 230. There are m students attending a seminar in a room with n seats. The seminar is a long one, and in the middle the group takes a break. In how many ways may the students return to the room and sit down so that nobody is in the same seat as before?

Solution. We use inclusion and exclusion. We let property P_i be the property that student i sits in the same seat as before. We are interested in $N_e(\emptyset)$. For this purpose, for each subset S of the set P of all properties, we need to compute $N_a(S)$, the number of ways for the students to return so that every student represented by a property in S sits in his or her previous seat. This leaves us with $n - |S|$ seats to be filled in a one-to-one fashion in $m - |S|$ students. There are $(n - |S|)^{\underline{m-|S|}}$ such seating arrangements, so $N_a(S) = (n - |S|)^{\underline{m-|S|}}$. Thus we have

$$N_e(\emptyset) = \sum_{S:S\subseteq P} (-1)^{|S|} (n - |S|)^{\underline{m-|S|}} = \sum_{s=0}^{m} (-1)^s \binom{m}{s} (n - s)^{\underline{m-s}}$$

ways for the students to return so that nobody sits in his or her previous place.

Activity 231. What is the number of ways to pass out k pieces of candy from an unlimited supply of identical candy to n children (where n is fixed) so that each child gets between three and six pieces of candy (inclusive)? If you have done Activity 200 compare your answer in that problem with your answer in this one.

Solution. As an inclusion-exclusion problem, we would let property i be that child i gets more than six pieces of candy. We would then observe that the number of ways to pass out the candy so that the children determined by a set S of properties all get more than six pieces, and everyone else gets at least 3, is the number of ways to pass out the remaining candy after giving 7 pieces to each child identified by S and 3 pieces to each other child. This number is $\binom{k-7|S|-3(n-|S|)-1}{n-1} = \binom{k-2n-4|S|-1}{n-1}$. From here we would substitute into Equation (5.2), make any simplifications we could, and we would be done.

Activity 232. In how many ways may k distinct books be arranged on n shelves so that no shelf gets more than m books?

Solution. We use inclusion and exclusion. Let property P_i be that shelf i gets more than m books. Then the number of arrangements of books with at least the properties in a subset S of the set P of all properties is $N_a(S) = k^{\underline{(m+1)|S|}} n^{k-(m+1)|S|}$, because in order to have at least the properties in S we may choose $(m+1)|S|$ books and arrange $m+1$ of them on each of the shelves represented by the properties in S, after which we arrange the remainder of the books. Thus

$$N_e(\emptyset) = \sum_{S:S\subseteq P} (-1)^{|S|} k^{\underline{(m+1)|S|}} n^{k-(m+1)|S|} = \sum_{s=0}^{n} (-1)^s \binom{n}{s} k^{\underline{(m+1)s}} n^{k-(m+1)s}$$

is the number of ways to arrange the books so that no shelf gets more than m.

Activity 233. Suppose that n children join hands in a circle for a game at nursery school. The game involves everyone falling down (and letting go). In how many ways may they join hands in a circle again so that nobody is to the right of the same child that was previously to his or her right?

Solution. We use inclusion and exclusion, with property i being the property that child i has the same child to the right the second time they join hands. Given a set S of properties, we can think of arranging units consisting of individual children and strings of children holding hands in a circle. If we have s properties, then we have $n - s$ of these units. Each string of children can be arranged in only one way, because our set specifies who has to have the same child on the right. Thus $N_a(S) = (n - s - 1)!$. This gives us

$$N_e(\emptyset) = \sum_{S:S\subseteq P} (-1)^{|S|}(n - |S| - 1)! = \sum_{s=0}^{n} (-1)^s \binom{n}{s}(n - s - 1)!$$

$$= \sum_{s=0}^{n} (-1)^s \frac{n!}{s!(n-s)}$$

ways for the children to join hands the second time so that none of them has the same child to the right.

Activity 234. Suppose that n people link arms in a folk-dance and dance in a circle. Later on they let go and dance some more, after which they link arms in a circle again. In how many ways can they link arms the second time so that no-one is next to a person with whom he or she linked arms before.

Solution. We use the Principle of Inclusion and Exclusion. Property P_i will be the property that person i links arms with someone previously to his or her right. (Saying it is the person to the right gives us more control over our formulas.) Given a subset S of the set P of all properties, the number of ways for the people in that set to link arms with the people previously on their right is the number of ways to arrange $n - |S|$ strings of people around a circle with strings of length more than 1 having two ways to arrange themselves. (Once we have two or more people linked, another person can be added to this string only at one end or not at all, because this person must have been to the right of one of the people on an end of the string. However a string of length two or more can unlink and then link in the opposite order, and each person will still be linked to exactly the same people.) Thus $N_a(S) = (n - |S| - 1)!2^{m(S)}$, where $m(S)$ is the number of strings of length more than one determined by S. The number $m(S)$ can be any number from 1 to S, so long as S is not too big; namely so long as $|S| \leq \lfloor n/1 \rfloor$. (This is because if $m(S) = |S|$, then each person determined by a property in S must be adjacent to a person not determined by a property in S.) In particular, $N_a(S)$ is not completely determined by the size of S, as in all our other inclusion-exclusion problems. How do we compute $m(S)$? Let us call a subset R of S a run if

1. the people determined by R sit together in a row in both seatings, and

2. no other person in S is in a row with these people in both seatings.

Some runs might determine just one person, but a run could also equal all of S. Each run will have one more person not in S who was originally to the right of the person in the run who was rightmost in the first seating, and so this person will have to sit in a row with the people in R in the second seating as well. Thus the number r of runs in S is the number of strings $m(S)$ that may be seated in two ways, and there are $n - |S| - r$ people who do not have to be seated with runs. Thus $N_a(S) = (n - |S| - 1)!2^r$, because the total number of strings of people (including strings of just one person) we need to seat is $n - |S| - r$, and there are $(k-1)!$ ways to arrange k objects in a circle.. If we try to use the information we have so far to compute $N_e(\emptyset)$, we get

$$N_e(\emptyset) = \sum_{S:S \subseteq P} (-1)^{|S|} (n - |S|_1)! 2^r = \sum_{s=0}^{n} \sum_{r=1}^{|S|} N(s,r)(n - s - 1)! 2^r,$$

in which $N(s,r)$ stands for the number of property sets with size s and r runs.

Picking out runs in a circular arrangement adds a layer of difficulty, so to compute $N(s,r)$, we first compute how many subsets of $[n]$ we have with r runs and then adjust for putting 1 through n around a circle in order. Imagine writing 1 through n in a straight line, each integer occupying one unit of distance along the line. We now place r sticks whose lengths add to s (each stick has positive integer length) along that line. Each stick picks out a set of consecutive integers, as many as its length, so the sticks together pick out s integers. In order to be sure the sticks correspond to runs, we need to make sure they do not touch each other, so we place $n - s$ identical stones along the line too, making sure there is at least one stone between any two sticks. The stones thus pick out the integers not in S. The sticks are not quite identical, though the sticks of a given length are. In other words, which lengths of sticks are in which places is what matters. So the sticks give us a composition of s, a list of distinct positive integers that add to s. We know there are $\binom{s-1}{r-1}$ such compositions. Once we have chosen an ordering for the sticks, we need to distribute the stones among the sticks so that no two sticks are adjacent. Since the stones are identical, we can do this by putting one stone between each pair of sticks in our composition, and then distribute the remaining $n - s - r + 1$ stones in any way we want among the $r - 1$ places between the sticks and the two places to the left and right of all the sticks.. We can do this in

$$\binom{r + 1 + (n - s - r + 1) - 1}{n - s - r + 1} = \binom{n - s + 1}{n - s - r + 1} = \binom{n - s + 1}{r}$$

ways. Thus there are $\binom{s-1}{r-1}\binom{n-s+1}{r}$ ways to choose a subset S of $[n]$ that has r runs.

Now we have to deal with the fact that our n people (who we have replaced with the integers 1 through n in order) were arranged around a circle. That means that a run is now a set of consecutive integers on the circle, where n and 1 are considered consecutive. Recall that the set S is picked out by the sticks. If we arrange 1 through n around a circle in order, the set S that originally had r runs will have $r - 1$ runs if sticks covered both the first and last integer (1 and n), but otherwise it will still have n runs. Thus the number of subsets of $[n]$ that have n runs when 1 through n are arranged in a circle is the number of subsets of $[n]$ with $r + 1$ runs that have both 1 and n in S plus the number of subsets of $[n]$ with r runs that do not have both 1 and n in S. To compute the number of subsets S that *do* contain both 1 and n, we compute the number of arrangements of r sticks and $n - s$ stones that do start and end with a stick; that means that after we choose our composition into r parts to get our arrangement of sticks and place one stone between each pair of previously adjacent sticks, we now place the remaining $n - s - r + 1$ stones in the $r - 1$ places between previously adjacent sticks in

$$\binom{r - 1 + (n - s - r + 1) - 1}{n - s - r + 1} = \binom{n - s - 1}{n - s - r + 1} = \binom{n - s - 1}{r - 2}$$

ways. For the sticks and stones to determine a subset we must assign lengths to the sticks; the number of ways to do this is, as above, $\binom{s-1}{r-1}$, the number of compositions of s with r parts. Thus there are $\binom{s-1}{r-1}\binom{n-s-1}{r-2}$ subsets of $[n]$ that have r runs and include both 1 and n. For our computation we will also want the number of subsets of $[n]$ that have $r+1$ runs and contain both 1 and n this is $\binom{s-1}{r}\binom{n-s-1}{r-1}$.

On the other hand, the number of subsets of $[n]$ that have r runs and do not contain both 1 and n is the total number of subsets with r runs minus the number that do contain both 1 and n; this is

$$\binom{s-1}{r-1}\left(\binom{n-s+1}{r}-\binom{n-s-1}{r-2}\right).$$

This gives us

$$N(s,r) = \binom{s-1}{r}\binom{n-s-1}{r-1} + \binom{s-1}{r-1}\left(\binom{n-s+1}{r}-\binom{n-s-1}{r-2}\right)$$

ways to choose an s-element subset of $[n]$ that has r runs when $[n]$ is arranged around a circle. Thus there are $\sum_{s=0}^{n}\sum_{r=1}^{s}(-1)^s\left[\binom{s-1}{r}\binom{n-s-1}{r-1}+\binom{s-1}{r-1}\left(\binom{n-s+1}{r}-\binom{n-s-}{r-2}\right)\right]$ $s-1)!2^r$ ways for people to arrange themselves in the second circle so that no-one is adjacent to anyone he or she was previously adjacent to.

Activity 235. (A challenge; the author has not tried to solve this one!) Redo Problem 234 in the case that there are n men and n women and when people arrange themselves in a circle they do so alternating gender.

Activity 236. Suppose we take two graphs G_1 and G_2 with disjoint vertex sets, we choose one vertex on each graph, and connect these two graphs by an edge e to get a graph G_{12}. How does the chromatic polynomial of G_{12} relate to those of G_1 and G_2?

Solution.   By the deletion-contraction recurrence,

$$\chi_{G_{12}}(x) = \chi_{G_{12}-e}(x) - \chi_{G_{12}/e}(x).$$

Now $\chi_{G_{12}-e}(x) = \chi_{G_1}(x) \times \chi_{G_2}(x)$ because each ordered pair of proper colorings of G_1 and G_2 is a proper coloring of $G_{12} - e$. G_{12}/e is the graph we get by identifying the endpoint of e in G_1 with the endpoint of e in G_2. Notice that if you fix one vertex of a graph G, fix one color, and ask how many proper colorings with x colors G has in which the fixed vertex is the fixed color, you get $\chi_G(x)/x$. (By the quotient principle.) Thus $\chi_{G_2}(x)/x$ is the number of ways to extend a proper coloring of G_1 to a proper coloring of G_{12}/e. Then, by the product principle, the number of proper colorings of G_{12}/e with x colors is $\chi_{G_1}(x)\chi_{G_2}(x)/x$. Therefore by the deletion-contraction recurrence, $\chi_{G_{12}}(x) = \chi_{G_1}(x)\chi_{G_2}(x)(1 - \frac{1}{x})$.

Chapter 6

Groups acting on sets

6.1 Permutation Groups

Until now we have thought of permutations mostly as ways of listing the elements of a set. In this chapter we will find it very useful to think of permutations as functions. This will help us in using permutations to solve enumeration problems that cannot be solved by the quotient principle because they involve counting the blocks of a partition in which the blocks don't have the same size. We begin by studying the kinds of permutations that arise in situations where we have used the quotient principle in the past.
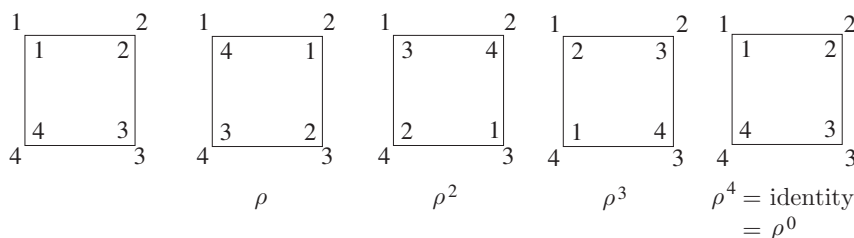
The rotations of a square

Figure 6.1.1: The four possible results of rotating a square and maintaining its position.

In Figure 6.1.1 we show a square with its four vertices labelled 1, 2, 3, and 4. We have also labeled the spot in the plane where each of these vertices falls with the same label. Then we have shown the effect of rotating the square clockwise through 90, 180, 270, and 360 degrees (which is the same as rotating through 0 degrees). Underneath each of the rotated squares we have named the function that carries out the rotation. We use ρ, the Greek letter pronounced "row," to stand for a 90 degree clockwise rotation. We use ρ^2 to

stand for two 90 degree rotations, and so on. We can think of the function ρ as a function on the four element set[1] $\{1, 2, 3, 4\}$. In particular, for any function φ (the Greek letter phi, usually pronounced "fee," but sometimes "fie") from the plane back to itself that may move the square around but otherwise leaves it in the same place, we let $\varphi(i)$ be the label of the place where vertex previously in position i is now. Thus $\rho(1) = 2$, $\rho(2) = 3$, $\rho(3) = 4$ and $\rho(4) = 1$. Notice that ρ is a permutation on the set $\{1, 2, 3, 4\}$.

Activity 237. Find ρ^2 of 1, 2, 3, and 4. Find ρ^3 of 1, 2, 3, and 4. Are ρ^2 and ρ^3 permutations of $\{1, 2, 3, 4\}$?

Solution. $\rho^2(1) = 3$, $\rho^2(2) = 4$, $\rho^2(3) = 1$, $\rho^2(4) = 2$. $\rho^3(1) = 4$, $\rho^3(2) = 1$, $\rho^3(3) = 2$, $\rho^3(4) = 3$. Yes, ρ^2 and ρ^3 are permutations of $\{1, 2, 3, 4\}$.

Activity 238. The composition $f \circ g$ of two functions f and g is defined by $f \circ g(x) = f(g(x))$. Is ρ^3 the composition of ρ and ρ^2? Does the answer depend on the order in which we write ρ and ρ^2? How is ρ^2 related to ρ?

Solution. Yes, ρ^3 is the composition of ρ and ρ^2, and also of ρ^2 and ρ, so it doesn't matter in which order we write them. $\rho^2 = \rho \circ \rho$.

Activity 239. Is the composition of two permutations always a permutation?

Solution. Yes, because the composition of one-to-one functions is one-to-one and the composition of onto functions is onto.

In Problem 238 you see that we can think of $\rho^2 \circ \rho$ as the result of first rotating by 90 degrees and then by another 180 degrees. In other words, the composition of two rotations is the same thing as first doing one and then doing the other. If we first rotate by 90 degrees and then by 270 degrees then we have rotated by 360 degrees, which does nothing visible to the square. Thus we say that ρ^4 is the "identity function." In general the *identity function* on a set S, denoted by ι (the Greek letter iota, pronounced eye-oh-ta) is the function that takes each element of the set to itself. In symbols, $\iota(x) = x$ for every x in S. Of course the identity function on a set is a permutation of that set.

Groups of Permutations

Activity 240. For any function φ from a set S to itself, we define φ^n (for nonnegative integers n) inductively by $\varphi^0 = \iota$ and $\varphi^n = \varphi^{n-1} \circ \varphi$ for every positive integer n. If φ is a permutation, is φ^n a permutation? Based on your experience with previous inductive proofs, what do you expect $\varphi^n \circ \varphi^m$ to be? What do you expect $(\varphi^m)^n$ to be? There is no need to prove these last two answers are correct, for you have, in effect, already done so in Chapter 2.

Solution. It is a permutation because the computation of permutations is a permutation. (You could be more precise and use the inductive definition of φ^n to prove inductively that φ^n is a permutation.) We expect $\varphi^m \circ \varphi^n = \varphi^{m+n}$, and we expect $(\varphi^m)n = \varphi^{mn}$, and we would prove this by induction.

---

[1]What we are doing is restricting the rotation ρ to the set $\{1, 2, 3, 4\}$.

Activity 241. If we perform the composition $\iota \circ \varphi$ for any function φ from S to S, what function do we get? What if we perform the composition $\varphi \circ \iota$?

Solution. $\iota \circ \varphi = \varphi \circ \iota = \varphi$

What you have observed about iota in Problem 241 is called the *identity property* of iota. In the context of permutations, people usually call the function ι "the identity" rather than calling it "iota."

Since rotating first by 90 degrees and then by 270 degrees has the same effect as doing nothing, we can think of the 270 degree rotation as undoing what the 90 degree rotation does. For this reason we say that in the rotations of the square, ρ^3 is the "inverse" of ρ. In general, a function $\varphi : T \to S$ is called an *inverse* of a function $\sigma : S \to T$ (the lower case Greek letter sigma) if $\varphi \circ \sigma = \sigma \circ \varphi = \iota$.

Activity 242. Does every permutation have an inverse? If so, why, and is it unique, i.e. could a permutation have two distinct inverse functions? If not, give an example of a permutation without an inverse or a permutation with two distinct inverses.

Solution. Let φ be a permutation of S. Then for each S in S there is one and only one x in S such that $\varphi(x) = s$ because φ is one-to-one and onto. Thus if we define $\sigma(s)$ to be the unique x such that $\varphi(x) = s$, we have defined a function, and

$$\varphi \circ \sigma(s) = \varphi(\sigma(s)) = \varphi(x) = s.$$

Thus we have that $\varphi \circ \sigma = \iota$. To show that σ is an inverse to φ, we must show that $\sigma \circ \varphi = \iota$ as well. However $\sigma(\varphi(y))$ is the unique x such that $\varphi(x) = \varphi(y)$, that is $\sigma(\varphi(y)) = y$. Thus $\sigma \circ \varphi = \iota$, and σ is an inverse to φ.

Activity 243. Show that if σ is an inverse of the permutation φ of S, then σ is a permutation of S also.

Solution. The function σ must be one-to-one because if $\sigma(x) = \sigma(y)$, then $x = \varphi \circ \sigma(x) = \varphi \circ \sigma(y) = y$. The function σ must be onto because for each y in its range, $\sigma(\varphi(y)) = y$

We use φ^{-1} to denote the inverse of the permutation φ. We've seen that the rotations of the square are functions that return the square to its original position but may move the vertices to different places. In this way we create permutations of the vertices of the square. We've observed three important properties of these permutations.

- (Identity Property) These permutations include the identity permutation.

- (Inverse Property) Whenever these permutations include φ, they also include φ^{-1}.

- (Closure Property) Whenever these permutations include φ and σ, they also include $\varphi \circ \sigma$.

A set of permutations with these three properties is called a *permutation group*[2] or a group of permutations. We will denote the group of permutations corresponding to rotations of the square by R_4 and call it the *rotation group* of the square. There is a similar rotation group with n elements for any regular n-gon.

Activity 244. If a finite set of permutations satisfies the closure property is it a permutation group?

Solution.   Yes, because the permutations σ, σ^2, ..., σ^n, ... must eventually start repeating. But if $\sigma^i = \sigma^j$, and $i < j$, then $\sigma^{j-i} = 1$, so σ^{j-i-1} is the inverse of σ. This means that every member of the set of permutations has an inverse in that set. But then since $\sigma \circ \sigma^{-1} = \iota$, we have that the identity is in the set as well. Thus it is a permutation group.

Activity 245.

(a) How should we define φ^{-n} for an element φ of a permutation group?

   Solution.   We define $\varphi^{-n} = \left(\varphi^{-1}\right)^n$.

(b) Will the two standard rules for exponents

   $$a^m a^n = a^{m+n} \text{ and } (a^m)^n = a^{mn}$$

   still hold if one or more of the exponents may be negative?

   Solution.   Yes.

(c) What would we have to prove to show that the rules still hold?

   Solution.   To prove this we need to show that for nonnegative m and n we have that $\varphi^m \circ \varphi^{-n} = \varphi^{m-n}$, that $\varphi^{-m} \varphi^{-n} = \varphi^{-m-n}$, that $(\varphi^m)^{-n} = \varphi^{-mn}$, that $(\varphi^{-m})^n = \varphi^{-mn}$ and that $(\varphi^{-m})^{-n} = \varphi^{mn}$.

(d) If the rules hold, give enough of the proof to show that you know how to do it; otherwise give a counterexample.

   Solution.   We give the first and last proof. We will induct on n in the first proof. If $n = 0$ the formula automatically holds. Assume inductively that $\varphi^m \circ \varphi^{-(n-1)} = \varphi^{m-(n-1)}$. Compose both sides of this

---

[2]The concept of a permutation group is a special case of the concept of a *group* that one studies in abstract algebra. When we refer to a group in what follows, if you know what groups are in the more abstract sense, you may use the word in this way. If you do not know about groups in this more abstract sense, then you may assume we mean permutation group when we say group.

equation on the right by σ^{-1} and use the associative law for composition of functions to get, in the case that $m - n$ is nonnegative

$$\varphi^m \circ (\varphi^{-(n-1)} \circ \varphi^{-1}) = \varphi^{m-(n-1)} \circ \varphi^{-1}$$
$$\varphi^m \circ (\varphi^{-1})^{n-1}\varphi^{-1} = (\varphi^{m-n} \circ \varphi) \circ \varphi^{-1}$$
$$\varphi^m \circ (\varphi^{-1})^n = \varphi^{m-n}.$$

We do the same in the case that $m - n$ is negative, getting

$$\varphi^m \circ (\varphi^{-(n-1)} \circ \varphi^{-1}) = \varphi^{m-(n-1)} \circ \varphi^{-1}$$
$$\varphi^m \circ (\varphi^{-1})^{n-1}\varphi^{-1} = (\varphi^{-1})^{n-m-1} \circ \varphi^{-1}$$
$$\varphi^m \circ (\varphi^{-1})^n = (\varphi^{-1})^{n-m}$$
$$\varphi^m \circ \varphi^{-n} = \varphi^{m-n}.$$

To prove the last statement we need to prove, we write the following, in which the line ⟨⟨Unresolved xref, reference "crucialline"; check spelling or use "provisional" attribute⟩⟩ follows from the first equation we proved and the line ⟨⟨Unresolved xref, reference "modestlyimportant"; check spelling or use "provisional" attribute⟩⟩ follows from the fact that $(\varphi^{-1})^{-1} = \varphi$.

$$(\varphi^{-m})^{-n} = ([\varphi^{-1})^m]^{-1})^n$$
$$= [(\varphi^{-1})^{-m}]^n$$
$$= ([(\varphi^{-1})^{-1}]^m)^n$$
$$= (\varphi^m)^n$$
$$= \varphi^{mn}$$

Activity 246. There are three dimensional geometric motions of the square that return it to its original position but move some of the vertices to other positions. For example, if we flip the square around a diagonal, most of it moves out of the plane during the flip, but the square ends up in the same place. Draw a figure like Figure 6.1.1 that shows all the possible results of such motions, including the ones shown in Figure 6.1.1. Do the corresponding permutations form a group?

Activity 247. If $f : S \to T$, $g : T \to X$, and $h : X \to Y$, is $h \circ (g \circ f) = (h \circ g) \circ f$? What does this say about the status of the *associative law*

$$\rho \circ (\sigma \circ \varphi) = (\rho \circ \sigma) \circ \varphi$$

in a group of permutations?

Solution. Since $(h \circ (g \circ f))(x) = h((g \circ f)(x)) = h(g(f(x)))$ and $((h \circ g) \circ f)(x) = (h \circ g)(f(x)) = h(g(f(x)))$, we have that

$$h \circ (g \circ f) = (h \circ g) \circ f.$$

This says that the associative law holds for the composition operation in a group of permutations.

Activity 248. If σ and φ are permutations, why must $\sigma \circ \varphi$ have an inverse? Is $(\sigma \circ \varphi)^{-1} = \sigma^{-1}\varphi^{-1}$? (Prove or give a counter-example.) Is $(\sigma \circ \varphi)^{-1} = \varphi^{-1}\sigma^{-1}$? (Prove or give a counter-example.)

Solution. One could either say that the composition of one-to-one and onto functions is one-to-one and onto, or note that $(\sigma \circ \varphi) \circ (\varphi^{-1} \circ \sigma^{-1}) = \iota$ by the associative law. Note that this says that $(\sigma \circ \varphi)^{-1} = \varphi^{-1} \circ \sigma^{-1}$. In the dihedral group D_4, $\rho^{-1} = \rho^3$ and $\varphi_{1/3}^{-1} = \varphi 1/3$. Thus

$$(\rho \circ \varphi_{1/3}) \circ (\rho^{-1} \circ \varphi_{1/3}^{-1}) = \rho \circ \varphi_{1/3}) \circ (\rho^3 \varphi_{1/3})$$

$$= \varphi_{12/34} \circ \varphi_{14/23}$$

$$= \rho^2.$$

This shows that $(\sigma \circ \varphi)^{-1} \neq \sigma^{-1}\varphi^{-1}$.

Activity 249. Explain why the set of all permutations of four elements is a permutation group. How many elements does this group have? This group is called the *symmetric group on four letters* and is denoted by S_4.

Solution. It is a finite set of permutations that satisfies the closure property. It has $4! = 24$ elements.

The symmetric group

In general, the set of all permutations of an n-element set is a group. It is called the *symmetric group on* **n** *letters*. We don't have nice geometric descriptions (like rotations) for all its elements, and it would be inconvenient to have to write down something like "Let $\sigma(1) = 3$, $\sigma(2) = 1$, $\sigma(3) = 4$, and $\sigma(4) = 1$" each time we need to introduce a new permutation. We introduce a new notation for permutations that allows us to denote them *reasonably* compactly and compose them reasonably quickly. If σ is the permutation of $\{1, 2, 3, 4\}$ given by $\sigma(1) = 3$, $\sigma(2) = 1$, $\sigma(3) = 4$ and $\sigma(4) = 2$, we write

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}.$$

We call this notation the *two row notation* for permutations. In the two row notation for a permutation of $\{a_1, a_2, \ldots, a_n\}$, we write the numbers a_1 through a_n in a one row and we write $\sigma(a_1)$ through $\sigma(a_n)$ in a row right below, enclosing both rows in parentheses. Notice that

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

although the second ordering of the columns is rarely used.

If φ is given by

$$\varphi = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{pmatrix},$$

then, by applying the definition of composition of functions, we may compute $\sigma \circ \varphi$ as shown in Figure 6.1.2.
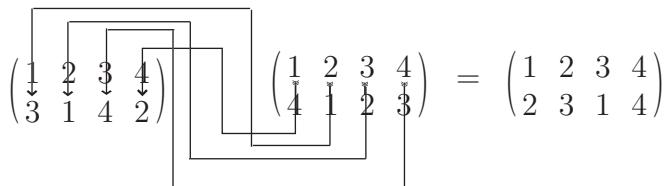
$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix}$$

Figure 6.1.2: How to multiply permutations in two-row notation.

We don't normally put the circle between two permutations in two row notation when we are composing them, and refer to the operation as multiplying the permutations, or as the product of the permutations. To see how Figure 6.1.2 illustrates composition, notice that the arrow starting at 1 in φ goes to 4. Then from the 4 in φ it goes to the 4 in σ and then to 2. This illustrates that $\varphi(1) = 4$ and $\sigma(4) = 2$, so that $\sigma(\varphi(1)) = 2$.

Activity 250. For practice, compute $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 1 & 5 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 3 & 5 & 1 & 2 \end{pmatrix}$.

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

The dihedral group

We found four permutations that correspond to rotations of the square. In Problem 246 you found four permutations that correspond to flips of the square in space. One flip fixes the vertices in positions 1 and 3 and interchanges those in positions 2 and 4. (Notice we did not say it fixes the vertices labelled 1 and 3.)[3] Let us denote it by $\varphi_{1/3}$. One flip fixes the vertices in positions 2 and 4 and interchanges those in positions 1 and 3. Let us denote it by $\varphi_{2/4}$. One flip interchanges the vertices in positions 1 and 2 and also interchanges those in positions 3 and 4. Let us denote it by $\varphi_{12/34}$. The fourth flip interchanges the vertices in positions 1 and 4 and interchanges those in positions 2 and 3. Let us denote it by $\varphi_{14/23}$. Notice that $\varphi_{1/3}$ is a permutation that takes vertex 1 to vertex 1 and vertex 3 to vertex 3, while $\varphi_{12/34}$ is a permutation that takes the edge from 1 to 2 to the edge from 1 to 2 and takes the edge from 3 to 4 to the edge from 3 to 4.

Activity 251.

---

[3]There is room for tremendous confusion here, and professional mathematicians sometimes suffer from it. Just try to remember that when we describe the axis of a flip, we describe it relative to points given in space, not points on the square or other object we are flipping.

(a) Write down the two-row notation for ρ^3, $\varphi_{2/4}$, $\varphi_{12/34}$ and $\varphi_{2/4} \circ \varphi_{12/34}$. Remember that $\sigma(i)$ stands for the position where the vertex that originated in position i is after we apply σ.

Solution. $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}.$

(b) (You may have already done this problem in Problem 246, in which case you need not do it again!) In Problem 246, if a rigid motion of three-dimensional space returns the square to its original position, in how many places can vertex number one land? Once the location of vertex number one is decided, how many possible locations are there for vertex two? Once the locations of vertex one and vertex two are decided, how many locations are there for vertex three? Answer the same question for vertex four. What does this say about the relationship between the four rotations and four flips described above and the permutations you described in Problem 246?

Solution. Vertex number one can go in four places. After the place for vertex 1 is chosen, there are two choices for vertex 2. After vertices 1 and 2 are placed, one side of the square has been placed, and there is only one way the rest of the square can fill in the same place the square used to be, so there is only one choice for where vertex 3 goes and one choice for where vertex 4 goes. Thus the permutations described in Problem 246 are exactly the rotations and flips.

The four rotations and four flips of the square described before Problem 251 form a group called the dihedral group of the square. Sometimes the group is denoted D_8 because it has eight elements, and sometimes the group is denoted by D_4 because it deals with four vertices! Let us agree to use the notation D_4 for the dihedral group of the square. There is a similar dihedral group, denoted by D_n, of all the rigid motions of three-dimensional space that return a regular n-gon to its original position (but might put the vertices in different places.)

Activity 252. How many elements does the group D_n have? Prove that you are correct.

Solution. D_n has $2n$ elements, because once you have chosen one of the n places for vertex one, there are two choices for vertex two, and the remaining vertices can go in only one place each.

Group tables

We can always figure out the composition of two permutations of the same set by using the definition of composition, but if we are going to work with a given permutation group again and again, it is worth making the computations once and recording them in a table. For example the group of rotations of

the square may be represented as in Table 6.1.3. We list the elements of our group, with the identity first, across the top of the table and down the left side of the table, using the same order both times. Then in the row labeled by the group element σ and the column labelled by the group element φ, we write the composition $\sigma \circ \varphi$, expressed in terms of the elements we have listed on the top and on the left side. Since a group of permutations is closed under composition, the result $\sigma \circ \varphi$ will always be expressible as one of these elements.

$\circ$	ι	ρ	ρ^2	ρ^3
ι	ι	ρ	ρ^2	ρ^3
ρ	ρ	ρ^2	ρ^3	ι
ρ^2	ρ^2	ρ^3	ι	ρ
ρ^3	ρ^3	ι	ρ	ρ^2

Table 6.1.3: The group table for the rotations of a square.

Activity 253.

(a) In Table 6.1.3, all the entries in a row (not including the first entry, the one to the left of the line) are different. Will this be true in any group table for a permutation group? Why or why not? Also in Table 6.1.3 all the entries in a column (not including the first entry, the one above the line) are different. Will this be true in any group table for a permutation group? Why or why not?

Solution. It will always be the case that all the entries of a row or column below and to the left of the lines in a group table are different. If two entries of the row labelled by σ were equal, that would mean $\sigma \circ \varphi_1 = \sigma \circ \varphi_2$ for two different elements φ_1 and φ_2. But if we multiply both sides of the equation by σ^{-1} we get $\sigma^{-1}(\sigma \circ \varphi_1) = \sigma^{-1}(\sigma \circ \varphi_2)$, and by using the associative law and the identity property, we get $\varphi_1 = \varphi_2$, a contradiction. The same sort of argument (with σ on the right) works for columns.

(b) In Table 6.1.3, every element of the group appears in every row (even if you don't include the first element, the one before the line). Will this be true in any group table for a permutation group? Why or why not? Also in Table 6.1.3 every element of the group appears in every column (even if you don't include the first entry, the one before the line). Will this be true in any group table for a permutation group? Why or why not?

Solution.   Every element appears in each row and every element appears in each column (after and below the lines). This is because the number of entries in a row or column is the number of elements of the group. By the pigeonhole principle, if not all the elements appear in a row, then two are the same, so $\tau \circ \sigma_1 = \tau \circ \sigma_2$ for some σ_1, σ_2, and τ (τ is the greek letter tau that rhymes with cow) in the group with $\sigma_1 \neq \sigma_2$. Multiplying on the right gives us $\sigma_1 = \sigma_2$, a contradiction. The same kind of argument works on columns, though you now have $\sigma_1 \circ \tau = \sigma_2 \circ \tau$ for some σ_1, σ_2, and τ in the group with $\sigma_1 \neq \sigma_2$.

Activity 254. Write down the group table for the dihedral group D_4. Use the φ notation described above to denote the flips. (Hints: Part of the table has already been written down. Will you need to think hard to write down the last row? Will you need to think hard to write down the last column? When you multiply a product like $\varphi_{1/3} \circ \rho$ remember that we defined $\varphi_{1/3}$ to be the flip that fixes the vertex in position 1 and the vertex in position 3, *not* the one that fixes the vertex on the square labelled 1 and the vertex on the square labelled 3.)

Solution.

$\circ$	ι	ρ	ρ^2	ρ^3	$\varphi_{1/3}$	$\varphi_{2/4}$	$\varphi_{12/34}$	$\varphi_{14/23}$
ι	ι	ρ	ρ^2	ρ^3	$\varphi_{1/3}$	$\varphi_{2/4}$	$\varphi_{12/34}$	$\varphi_{14/23}$
ρ	ρ	ρ^2	ρ^3	ι	$\varphi_{12/34}$	$\varphi_{23/14}$	$\varphi_{2/4}$	$\varphi_{1/3}$
ρ^2	ρ^2	ρ^3	ι	ρ	$\varphi_{2/4}$	$\varphi_{1/3}$	$\varphi_{14/23}$	$\varphi_{12/34}$
ρ^3	ρ^3	ι	ρ	ρ^2	$\varphi_{23/14}$	$\varphi_{12/34}$	$\varphi_{1/3}$	$\varphi_{2/4}$
$\varphi_{1/3}$	$\varphi_{1/3}$	$\varphi_{14/23}$	$\varphi_{2/4}$	$\varphi_{12/34}$	ι	ρ^2	ρ	ρ^3
$\varphi_{2/4}$	$\varphi_{2/4}$	$\varphi_{12/34}$	$\varphi_{1/3}$	$\varphi_{14/23}$	ρ^2	ι	ρ^3	ρ
$\varphi_{12/34}$	$\varphi_{13/24}$	$\varphi_{1/3}$	$\varphi_{14/23}$	$\varphi_{2/4}$	ρ	ρ^3	ι	ρ^2
$\varphi_{14/23}$	$\varphi_{14/23}$	$\varphi_{2/4}$	$\varphi_{12/34}$	$\varphi_{1/3}$	ρ^3	ρ	ρ^2	ι

You may notice that the associative law, the identity property, and the inverse property are three of the most important rules that we use in regrouping parentheses in algebraic expressions when solving equations. There is one property we have not yet mentioned, the *commutative law* which would say that $\sigma \circ \varphi = \varphi \circ \sigma$. It is easy to see from the group table of R_4 that it satisfies the commutative law.

Activity 255. Does the commutative law hold in all permutation groups?

Solution.   No. In the group D_3 a nontrivial rotation does not commute with a flip.

In your group table for the dihedral group D_4, you have a copy of the group of rotations of the square. When one group G of permutations of a set S is a subset of another group G' of permutations of S, we say that G is a *subgroup* of G'. The reason why we introduce the new word subgroup is to

emphasize that the composition operation gives the same result whether it is performed in the larger group or the smaller group.

Activity 256.

(a) Find all subgroups of the group D_4.

   Solution. $\{\rho^2,\iota\}$, $\{\varphi_{1/3},\iota\}$, $\{\varphi_{2/4},\iota\}$, $\{\varphi_{12/34},\iota\}$, $\{\varphi_{14/23},\iota\}$, $\{\iota,\rho,\rho^2,\rho^3\}$, $\{\iota,\varphi_{1/3},\rho^2,\varphi_{2/4}\}$, $\{\iota,\varphi_{12/34},\rho^2,\varphi_{12/34}\}$, $\{\iota\}$, and all of D_4. Notice that once a subgroup has ρ or ρ^3 in it, it has all powers of ρ, and if it has all powers of ρ and just one flip, it has all the flips in it. But the product of two "different kinds" of flips is ρ or ρ^3, and this shows that our list of subgroups is complete.

(b) Can you find subgroups of the symmetric group S_4 with two elements? Three elements? Four elements? Six elements? (For each positive answer, describe a subgroup. For each negative answer, explain why not.)

   Solution. Yes to all of the above. For two elements, $\{\iota, \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix}\}$, for three elements, $\{\iota, \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 2 & 4 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix}\}$ (which is the rotations of a triangle on $\{1,2,3\}$), for four elements, the rotations of a square as above, and for six elements, the rotations and flips of a triangle on $\{1,2,3\}$. (There is no subgroup with five elements, but at this point, that is too hard to verify.)

The cycle structure of a permutation

There is an even more efficient way to write down permutations. Notice that the product in Figure 6.1.2 is $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix}$. We have drawn the directed graph of this permutation in Figure 6.1.4.
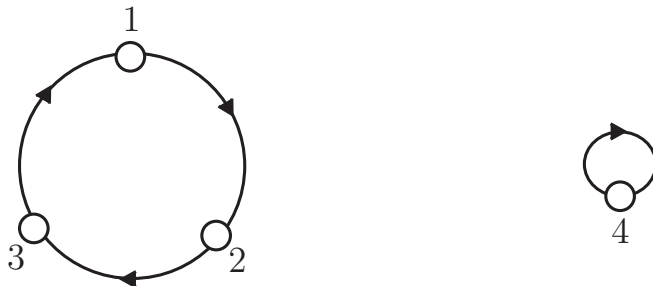

Figure 6.1.4: The directed graph of a permutation .

You see that the graph has two directed cycles, the rather trivial one with vertex 4 pointing to itself, and the nontrivial one with vertex 1 pointing to vertex 2 pointing to vertex 3 which points back to vertex 1. A permutation is called a *cycle* if its digraph consists of exactly one cycle. Thus $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ is a cycle but $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix}$ is not a cycle by our definition. We write $(1\ 2\ 3)$ or $(2\ 3\ 1)$ or $(3\ 1\ 2)$ to stand for the cycle $\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$. We can describe cycles in another way as well. A *cycle* of the permutation σ is a list $(i\ \sigma(i)\ \sigma^2(i)\ \ldots\ \sigma^n(i))$ that does not have repeated elements while the list $(i\ \sigma(i)\ \sigma^2(i)\ \ldots\ \sigma^n(i))\ \sigma^{n+1}(i))$ does have repeated elements. We say that the cycles $(i\ \sigma(i)\ \sigma^2(i)\ \ldots\ \sigma^n(i))$ and $(j\ \sigma(j)\ \sigma^2(j)\ \ldots\ \sigma^n(j))$ are *equivalent* if there is an integer k such that $j = \sigma^k(i)$.

Activity 257.

(a) Find the cycles of the permutations ρ, $\varphi_{1/3}$ and $\varphi_{12/34}$ in the group D_4.

   Solution. ρ has one cycle, $(1\ 2\ 3\ 4)$, $\varphi_{1/3}$ has one cycle $(2\ 4)$, $\varphi_{12/34}$ has two cycles, namely $(1\ 2)$ and $(3\ 4)$.

(b) Find the cycles of the permutation $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 4 & 6 & 2 & 9 & 7 & 1 & 5 & 8 \end{pmatrix}$.

   Solution. $\begin{pmatrix} 1 & 3 & 6 & 7 \end{pmatrix}$, $\begin{pmatrix} 2 & 4 \end{pmatrix}$, and $\begin{pmatrix} 5 & 9 & 8 \end{pmatrix}$.

(c) Show that if $(i\ \sigma(i)\ \sigma^2(i)\ \ldots\ \sigma^n(i))$ does not have repeated elements while the list $(i\ \sigma(i)\ \sigma^2(i)\ \ldots\ \sigma^n(i))\ \sigma^{n+1}(i))$ does have repeated elements, then $\sigma^{n+1}(i) = i$, and so all other $\sigma^s(i)$ are in the cycle.

   Solution.   If $\sigma^{n+1}(i) = \sigma^p(i)$ and $0 \le p \le n$, then $\sigma^{n+1-p}(i) = i$, and so the first repeat was not $\sigma^{n+1}(i)$ after all. Therefore $\sigma^{n+1}(i) = i$, and acting on i with succeeding powers of σ just carries us through the list again and again.

Activity 258. The support set of the cycle $(a_1\ a_2\ \ldots\ a_n)$ is the set $\{a_1, a_2, \ldots, a_n\}$. Show that the support sets of the cycles of a permutation of the set S form a partition of S.

Solution.   Every element is in a cycle by the definition of cycles, so suppose two support sets have an element in common. If the cycles are $(i\ \sigma(i)\ \sigma^2(i)\ \ldots\ \sigma^n(i))$ and $(j\ \sigma(j)\ \sigma^2(j)\ \ldots\ \sigma^m(j))$ and $\sigma^p(i) = \sigma^q(j)$ is a common element, then, $j = \sigma^{p-q}(i)$. Thus if $p \ge q$, j is in the first cycle. Otherwise i is in the second cycle. If j is in the first cycle, it is $\sigma^k(i)$ for some i, and as we take successive powers, we get, by that i is $\sigma^h(j)$, and then successive powers of

σ just give us the rest of the cycle beginning with i. Therefore the two support sets are identical. A similar argument works if i is in the second cycle. Thus the support sets partition S.

We regard a cycle $(i \; \sigma(i) \; \sigma^2(i) \; \ldots \; \sigma^n(i))$ as standing for the permutation

$$\begin{pmatrix} i & \sigma(i) & \cdots & \sigma^{n-1}(i) & \sigma^n(i) \\ \sigma(i) & \sigma^2(i) & \cdots & \sigma^n(i) & i \end{pmatrix}.$$

Since interchanging the columns in the two row notation for a permutation does not change the permutation[4], this means that equivalent cycles represent the same permutation. Thus we consider equivalent cycles to be equal in the same way we consider $\frac{1}{2}$ and $\frac{2}{4}$ to be equal. In particular, this means that $(i_1 \; i_2 \; \ldots i_n) = (i_j \; i_{j+1} \; \ldots \; i_n \; i_1 \; i_2 \; \ldots \; i_{j-1})$.

If γ_1, γ_2, ..., γ_n are the (inequivalent) cycles of a permutation σ of the set S, then we write

$$\sigma = \gamma_1 \gamma_2 \cdots \gamma_n = \gamma_{i_1} \gamma_{i_2} \cdots \gamma_{i_n},$$

that is, we say that σ is the product of its cycles in any order. There are two ways to think of this notation. You can think of it as a brand new kind of multiplication where we multiply a permutation of one set by a permutation of a disjoint set to get a new permutation of the union of the two sets. In this notation $\sigma(x)$ is determined by which cycle γ_k contains x in its support set and $\sigma(x) = \gamma_k(x)$. You can also think of identifying γ with the product of γ with a collection of one-cycles, one one-cycle for each element of S not in the support set of γ. For example we identify the cycle $(2 \; 4 \; 5)$ in S_6 with $(1)(2 \; 4 \; 5)(3)(6)$. Similarly we identify (3) with $(1)(2)(3)(4)(5)(6)$. But when we use (3) as a notation for $(1)(2)(3)(4)(5)(6)$, we still say its support set is $\{3\}$. Then the product above is simply composition of functions. We can summarize this discussion with the following theorem[5]. The proof of the theorem is essentially your solution to Problem 258.

Theorem 6.1.5. *Every permutation can be written, up to order, in one and only one way as a product of disjoint cycles.*

We usually don't bother writing down one-cycles when we write a permutation as a product of cycles; it is a convention to make less work for us. We just have to remember that they are implicitly there. You should recognize that the first interpretation we gave for the multiplication of disjoint cycles applies only when the cycles are disjoint. From now on we will use the phrase "product of permutations" to mean the same thing as "composition of permutations," and will use either the notation $\sigma \circ \varphi$ or $\sigma\varphi$ for the composition of σ and φ, whichever is convenient.

Activity 259.

---

[4]Note that not all ways of interchanging the columns list the bottom row in the order of a cycle of σ, though.

[5]In the statement of the theorem we use the word disjoint; two cycles are said to be *disjoint* if their support sets are disjoint.

(a) Write the permutations ρ, $\varphi_{1/3}$ and $\varphi_{12/34}$ in D_4 as products of disjoint cycles.

Solution. $\rho = (1\ 2\ 3\ 4)$, $\varphi_{1/3} = (2\ 4) = (1)(2\ 4)(3)$ (either side of the equation is a right answer), $\varphi_{12/34} = (1\ 2)(3\ 4)$.

(b) Write the permutation $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 4 & 6 & 2 & 9 & 7 & 1 & 5 & 8 \end{pmatrix}$ as a product of disjoint cycles.

Solution. $(1\ 3\ 6\ 7)(1\ 4)(5\ 9\ 8)$.

(c) Explain why, if γ and γ' are two disjoint cycles, then $\gamma\gamma' = \gamma'\gamma$.

Solution.   When we compute $\gamma\gamma'(i)$, the element i is fixed (i.e. mapped to itself) by either γ or γ'. But if γ' fixes i and $\gamma(i) = j$, with $j \neq i$ then j is in the support set of γ, so it is not in the support set of γ'. Thus $(\gamma\circ\gamma')(i) = j$. However since γ' fixes j as well, then $(\gamma'\circ\gamma)(i) = \gamma'(j) = j$ also. If $\gamma(i) = i$ also, then $\gamma\gamma'(i) = \gamma'\gamma(i) = i$. A similar argument works if γ fixes i, so in all cases, $\gamma\gamma'(i) = \gamma'\gamma(i)$, which is what it means to say that $\gamma\gamma' = \gamma'\gamma$.

Activity 260. Write a recurrence for the number $c(k, n)$ for the number of permutations of $[k]$ that have exactly n cycles, including 1-cycles. Use it to write a table of $c(k, n)$ for k between 1 and 7 inclusive. Can you find a relationship between $c(k, n)$ and any of the other families of special numbers such as binomial coefficients, Stirling numbers, Lah numbers, etc. we have studied? If you find such a relationship, prove you are right.

Solution.   The element k is either in a cycle by itself or it isn't. The number of permutations in which it is in a cycle by itself is $c(k - 1, n - 1)$. If it is in a cycle with something else, it can come after any of the $k - 1$ elements, and each choice of which one it comes after gives a different permutation. (You can always shift a cycle around so that k doesn't come first.) Thus $c(k, n) = c(k - 1, n - 1) + (k - 1)c(k - 1, n)$. The ta-

ble is:

k/n	0	1	2	3	4	5	6	7
0	1	0	0	0	0	0	0	0
1	0	1	0	0	0	0	0	0
2	0	1	1	0	0	0	0	0
3	0	2	3	1	0	0	0	0
4	0	6	11	6	1	0	0	0
5	0	24	50	35	10	1	0	0
6	0	120	274	225	85	15	1	0
7	0	720	1764	1624	735	175	21	1

We see that the

table is identical with our earlier table of the Stirling numbers of the first kind, except that in that table there was an alternating pattern of minus signs, offset by one place in successive rows. Thus it must be the case that $c(k, n) = |s(k, n)|$. To prove this, note that it is the case when $k = 0$, and also when $n = 1$. Now assume inductively it is true when $k = m - 1$ and notice that if $n > 0$, $s(m, n) = s(m - 1, n - 1) - (m - 1)s(m - 1, n)$, that $s(m - 1, n - 1)$ and $s(m - 1, n)$ have opposite signs, so

$$
\begin{aligned}
|s(m - 1) - (m - 1)s(m - 1, n)| &= |s(m - 1| + (m - 1)|s(m - 1, n)| \\
&= c(m - 1, n - 1) + (m - 1)c(m - 1, n) \\
&= c(m, n).
\end{aligned}
$$

Thus by the principle of mathematical induction $c(k, n) = |s(k, n)|$ for all nonnegative numbers k (and all n between 0 and k.)

Activity 261.

(a) (Relevant to Appendix B.1.) A permutation σ is called an *involution* if $\sigma^2 = \iota$. When you write an involution as a product of disjoint cycles, what is special about the cycles?

   Solution. They are all 2-cycles or 1-cycles.

(b) Write the product $(1\ 2)(1\ 3)(1\ 4)$ in two row notation and then write it in terms of of *disjoint* cycles.

   Solution. $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{pmatrix}$, $(1\ 4\ 3\ 2)$.

Signs of permutations and the alternating group

Activity 262.

(a) How many permutations in S_n may be written as a product of not necessarily disjoint two cycles?

   Solution. All of them, because each cycle can be written as a product on non-disjoint cycles.

(b) (May be time-consuming.) Consider the permutations of $\{1, 2, 3, 4\}$ which are ι or have one of the forms $(i\ j\ k)$ or $(i\ j)(h\ k)$, where i, j, k, and h are distinct. Do these permutations form a subgroup of the group S_4? How many of these permutations are there?

   Solution. Note that two cycles in S_4 of the form $(i\ j\ k)$ have two or three elements in common, so they have the same kind of overlap

as (1 2 3) and (2 3 4), (1 2 3) and (3 2 4), (1 2 3) and (1 3 2), or
(1 2 3) and (1 2 3). The products of these four pairs are, respectively,
(2 1)(3 4), (2 4 1), (1)(2)(3) = (1)(2)(3)(4) = ι, and (1 3 2). If we take a
three-cycle times a product of two two-cycles it will have the same form
as (1 2 3)(1 2)(3 4) or (1 2 3)(1 3)(2 4). The first of these products is
(1 3 4) and the second is (2 4 3). If we have a product of two elements
each of which is a product of two two-cycles, it will have the same form
as either (1 2)(3 4)(1 2)(3 4) or (1 2)(3 4)(1 3)(2 4). The first of these is
ι and the second is (1 4)(2 3). Thus our set is closed under composition,
and so it is a subgroup of S_4. The subgroup has size 12.

Activity 263. The *sign* of a 1-cycle, 3-cycle, ..., $(2k - 1)$-cycle is defined to
be 1, while the *sign* of a 2-cycle, 4-cycle, ..., $2k$-cycle is defined to be -1. The
sign of a product of disjoint cycles is the product of the signs of its cycles.

(a) What happens to the sign of an n-cycle if you multiply it by a 2-cycle
(which might or might not be disjoint from it)?

Solution.   If the m-cycle and 2-cycle are disjoint, then by definition,
the sign changes to the opposite value. If they are not disjoint, and the
two cycle is $(i\ j)$, either i or j or both are in the original cycle. Suppose
that only i is in the cycle (and the same argument works if only j is
in the cycle), and the original cycle is $(i\ i_2\ \ldots\ i_n)$. Then the product
$(i\ i_2\ \ldots\ i_n)(i\ j)$ is $(i\ j\ i_2\ \ldots\ i_n)$, and the product $(i\ j)(i\ i_2\ \ldots\ i_n)$ is
$(j\ i\ i_2\ \ldots\ i_n)$. Thus the sign changes to the opposite sign. Suppose that
i and j are both in the original cycle. Then $j = i_k$ for some k. Then
$(i\ i_2\ \ldots\ i_n)(i\ i_k) = (ii_{k+1}\ i_{k+2}\ \ldots\ i_n)(i_k\ i_2\ \ldots\ i_{k-1})$. If the original
sign was 1, then we started with an cycle on an odd number of entries
and so now have an cycle on an even number of entries times an cycle
on an odd number of entries, so the sign has changed. It the original
sign was -1, then we started with an cycle on an even number of entries,
so we now have a product of two cycles on an odd number of entries,
which gives sign 1, or a product of two cycles on an even number of
entries, which gives us sign 1. Thus in every case, the sign changes to
the opposite sign.

(b) What happens to the sign of a product of an n-cycle and a disjoint m-
cycle if you multiply it by a two cycle which is disjoint from neither of
the other two?

Solution.   The product becomes an $m + n$ cycle, and so if both cycles
were on an even number of entries or both were on an odd number, we
now have an cycle on an even number of entries, so its sign is -1, but
previously the product had sign 1. Thus the sign changes. If one of the
original cycles was on an even number of entries and one was on an odd
number, we now have a cycle on an odd number of entries, so it has sign
1, while the original product has sign -1. Thus the sign changes.

(c) What happens to the sign of a permutation if you multiply it by a 2-cycle?

Solution. The cases we just analyzed cover all the possibilities, so the sign changes.

(d) A 2-cycle is often called a *transposition*. What is the sign of a product of k (not necessarily disjoint) transpositions? (Don't forget to prove you are correct.)

Solution. $(-1)^k$. The proof is probably easiest by induction, for the sign of one transposition is -1, and if we take a product of i transpositions and multiply by one more transposition, the sign changes. Thus if the product of i transpositions has sign $(-1)^i$, the product of $i + 1$ transpositions has sign $(-1)^{i+1}$, and so by the principle of mathematical induction, the sign of a product of k transpositions if $(-1)^k$ for all positive integers k. In fact, had we taken ι for our base case, we could have concluded the same result for all nonnegative integers.

(e) Can you compute the sign of a product of two permutations from the signs of the permutations? Why or why not?

Solution. You can compute it; it is the product of the signs. Just write them as products of transpositions, multiply all those transpositions together and use the fact that, for transpositions, the sign of the product is the product of the signs.

Activity 264. Do the elements of S_n with sign -1 form a subgroup of S_n? Do the elements of S_n with sign 1 form a subgroup of S_n?

Solution. The elements of sign -1 are not closed under composition so they do not form a subgroup. The elements of sign 1 are closed under composition so they do form a subgroup.

Activity 265. Describe a subgroup of S_n of size $n!/2$. This subgroup is called the *alternating group A_n*. Describe explicitly a partition of S_n that you can use with the quotient principle to prove bijectively that the subgroup you described has size $n!/2$.

Solution. Partition S_n into the permutations of sign 1 and sign -1. Choose a transposition. Multiplying the elements in one block gives you elements in the other block. Multiplying by the transposition twice gives you the same permutation back, so the function that multiplies by this transposition has an inverse and so is a bijection.

6.2   Groups Acting on Sets

We defined the rotation group R_4 and the dihedral group D_4 as groups of permutations of the vertices of a square. These permutations represent rigid motions of the square in the plane and in three dimensional space respectively. The square has geometric features of interest other than its vertices; for example its diagonals, or its edges. Any geometric motion of the square that returns it to its original position takes each diagonal to a possibly different diagonal, and takes each edge to a possibly different edge.

We could similarly define the group R of rigid motions of a cube as the group of permutations of its eight vertices that correspond to three dimensional motions that return the cube to its original position, but perhaps with vertices in different places than they were originally. Now each of these motions also takes the 12 edges of the cube to (possibly different) edges and takes the six faces of the cube to (possible different) faces. So we can think of our group R as acting not only on the vertices of our geometric figure, but on edges, diagonals, faces, or perhaps other geometric features.

We have seen above that the fact that we have defined a permutation group as the permutations of some specific set doesn't preclude us from thinking of the elements of that group as permuting the elements of some other set as well. In order to keep track of which permutations of which set we are using to define our group and which other set is being permuted as well, we introduce some new language. We are going to say that the group D_4 "acts" on the edges and diagonals of a square and the group R of permutations of the vertices of a cube that arise from rigid motions of the cube "acts" on the edges, faces, diagonals, etc. of the cube. Instead of talking about permutations of these edges, faces, etc., we will use an equivalent word and talk about bijections of these objects. In this way it will be easier for us to keep track of which permutations form our original group and how this group relates to other objects of interest.
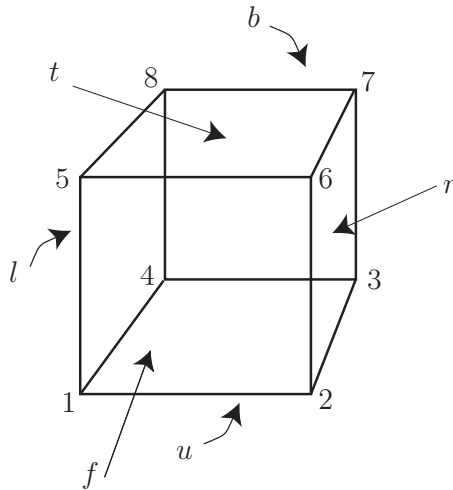
Figure 6.2.1: A cube with the positions of its vertices and faces labelled. The curved arrows point to the positions that are blocked by the cube.

Activity 266. In Figure 6.2.1 we show a cube with the positions of its vertices and faces labelled. *Some of the arrows may be missing in this image. Cross check.* As with motions of the square, we let we let $\varphi(x)$ be the label of the place where vertex previously in position x is now.

(a) Write in two row-notation the permutation ρ of the vertices that corresponds to rotating the cube 90 degrees around a vertical axis through the faces t (for top) and u (for underneath). (Rotate in a right-handed fashion around this axis, meaning that vertex 6 goes to the back and vertex 8 comes to the front.) Write in two-row notation the bijection β_ρ of the faces that corresponds to this member ρ of R.

Solution. $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 3 & 4 & 1 & 6 & 7 & 8 & 5 \end{pmatrix}$ $\begin{pmatrix} t & f & r & b & l & u \\ t & r & b & l & f & u \end{pmatrix}$

(b) Write in two-row notation the permutation φ that rotates the cube 120 degrees around the diagonal from vertex 1 to vertex 7 and carries vertex 8 to vertex 6. Write in two-row notation the bijection β_φ of the faces that corresponds to this member of R.

Solution. $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 4 & 8 & 5 & 2 & 3 & 7 & 6 \end{pmatrix}$ $\begin{pmatrix} t & f & r & b & l & u \\ r & u & b & t & f & l \end{pmatrix}$

(c) Compute the two-row notation for $\rho \circ \varphi$, for $\beta_\rho \circ \beta_\varphi$ (β_ρ was defined in Part), and write in two-row notation the bijection $\beta_{\rho \circ \varphi}$ of the faces that corresponds to the action of the permutation $\rho \circ \varphi$ on the faces of the cube (for this question it helps to think geometrically about what

motion of the cube is carried out by $\rho \circ \varphi$). What do you observe about $\beta_{\rho \circ \varphi}$ and $\beta_\rho \circ \beta_\varphi$?

Solution.

$$
\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 3 & 4 & 1 & 6 & 7 & 8 & 5 \end{pmatrix}
\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 4 & 8 & 5 & 2 & 3 & 7 & 6 \end{pmatrix}
= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 1 & 5 & 6 & 3 & 4 & 8 \end{pmatrix}
$$

$$
\begin{pmatrix} t & f & r & b & l & u \\ t & r & b & l & f & u \end{pmatrix}
\begin{pmatrix} t & f & r & b & l & u \\ r & u & b & t & f & l \end{pmatrix}
= \begin{pmatrix} t & f & r & b & l & u \\ b & u & l & t & r & f \end{pmatrix}
$$

$$
\beta_{\rho \circ \varphi} = \begin{pmatrix} t & f & r & b & l & u \\ b & u & l & t & r & f \end{pmatrix} = \beta_\rho \circ \beta_\varphi
$$

Activity 267. How many permutations are in the group R? R is sometimes called the "rotation group" of the cube. Can you justify this?

Solution.   There are eight places where vertex one can go. Once vertex 1 is placed, there are three ways the three edges "sticking out" of it can be placed, and then the location of all the vertices is determined. Thus there are 24 elements of the group. Of them, we have the identity, three rotations about each of the three axes through the midpoints of the faces, one additional rotation about each of the six axes joining opposite faces, and $2 \cdot 4 = 8$ additional rotations about each of the four axes joining a vertex to the diagonally (in three dimensions) opposite vertex. This gives us a total of 10+6+8=24 rotations, so every element of the group is a rotation.

We say that a permutation group G *acts* on a set S if, for each member σ of G there is a bijection β_σ of S such that

$$
\beta_{\sigma \circ \varphi} = \beta_\sigma \circ \beta_\varphi
$$

for every member σ and φ of G. In Problem you saw one example of this condition. If we think intuitively of ρ and φ as motions in space, then following the action of φ by the action of ρ does give us the action of $\rho \circ \varphi$. We can also reason directly with the permutations in R to show that R acts on the faces of the cube.

Activity 268. Suppose that σ and φ are permutations in the group R of rigid motions of the cube. We have argued already that each rigid motion sends a face to a face. Thus σ and φ both send the vertices on one face to the vertices on another face. Let $\{h, i, j, k\}$ be the set of vertices on a face F.

(a) What are the vertices of the face F' that F is sent to by φ?

Solution. $\varphi(h)$, $\varphi(i)$, $\varphi(j)$, $\varphi(k)$.

(b) What are the vertices of the face F'' that F' is sent to by σ?

Solution. $\sigma(\varphi(h))$, $\sigma(\varphi(i))$, $\sigma(\varphi(j))$, $\sigma(\varphi(k))$.

(c) What are the vertices of the face F''' that F is sent to by $\sigma \circ \varphi$?

Solution. $(\sigma \circ \varphi)(h) = \sigma(\varphi(h))$, $\sigma(\varphi(i))$, $\sigma(\varphi(j))$, $\sigma(\varphi(k))$.

(d) How have you just shown that the group R acts on the faces?

Solution.   We have just shown that for each face F, $\beta_{\sigma \circ \varphi}(F) = \beta_\sigma \circ \beta_\varphi(F)$, so that $\beta_{\sigma \circ \varphi} = \beta_\sigma \circ \beta_\varphi$.

Activity 269. The group D_4 is a group of permutations of $\{1, 2, 3, 4\}$ as in Problem 246. We are going to show in this problem and in Problem 270 how this group acts in a natural way on the two-element subsets of $\{1, 2, 3, 4\}$. In particular, for each two-element subset $\{i, j\}$ of $\{1, 2, 3, 4\}$ and each member σ of D_4 we define $\beta_\sigma(\{i, j\}) = \{\sigma(i), \sigma(j)\}$. Show that with this definition of β, the group D_4 acts on the two-element subsets of $\{1, 2, 3, 4\}$.

Solution.   The action has been defined for us, so all we need to show is that it is indeed an action. We must show that for each permutation σ in D_4, β_σ is a bijection and in addition we must show that $\beta_\sigma \circ \beta_\tau = \beta_{\sigma \circ \tau}$. If $\beta_\sigma(\{i, j\}) = \beta_\sigma(\{h, k\})$, then either $\sigma(i) = \sigma(h)$ and $\sigma(j) = \sigma(k)$ or else $\sigma(i) = \sigma(k)$ and $\sigma(j) = \sigma(h)$. Since σ is a permutation, in the first case we get $i = h$ and $j = k$, so that $\{i, j\} = \{h, k\}$, and in the second case we get that $i = k$ and $j = h$ so that $\{i, j\} = \{k, h\} = \{h, k\}$. Thus in either case $\{i, j\} = \{h, k\}$, so that β_σ a one-to-one function from the finite set of two element subsets of $\{1, 2, 3, 4\}$ to itself, and so it is a bijection.
    To show that $\beta_\sigma \circ \beta_\tau = \beta_{\sigma \circ \tau}$, we note that

$$(\beta_\sigma \circ \beta_\tau)(\{i, j\}) = \beta_\sigma(\beta_\tau(\{i, j\})) = \beta_\sigma(\{\tau(i).\tau(j)\}) =$$
$$\sigma(\tau(i), \tau(j)) = \{(\sigma \circ \tau)(i), (\sigma \circ \tau)(j)\} =$$
$$\beta_{\sigma \circ \tau}(\{i, j\})$$

Activity 270. In Problem 269, what is the set of two element subsets that you get by computing $\beta_\sigma(\{1, 2\})$ for all σ in D_4? What is the set of two-element subsets you get by computing $\beta_\sigma(\{1, 3\})$ for all σ in D_4? Describe these two sets geometrically in terms of the square.

Solution.   From $\{1, 2\}$ we get $\{1, 2\}$, $\{2, 3\}$, $\{3, 4\}$, and $\{3, 4\}$. From $\{1, 3\}$ we get $\{1, 3\}$ and $\{2, 4\}$. Thus we get the set of edges and the set of diagonals of the square.

Activity 271. Using the notation of Problem 254, what is the effect of a 180 degree rotation on the diagonals of a square? What is the effect of the flip $\varphi_{1/3}$ on the diagonals of a square? How many elements of D_4 send each

diagonal to itself? How many elements of D_4 interchange the diagonals of a square? The 180 degree rotation sends the diagonals to themselves, i.e., it fixes the diagonals. $\varphi_{1/3}$ fixes the diagonals. So does $\varphi_{2/4}$ and the identity. However, ρ, ρ^3, $\varphi_{12/34}$ and $\varphi_{14/23}$ all exchange the diagonals. Thus four elements fix each diagonal and four elements interchange them.
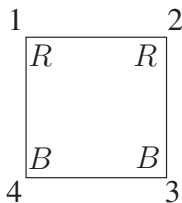

Figure 6.2.2: The colored square corresponding to the function f with $f(1) = R$, $f(2) = R$, $f(3) = B$, $f(4) = B$.

Activity 272. Recall that when you were asked in Problem 45 to find the number of ways to place two red beads and two blue beads at the corners of a square free to move in three dimensional space, you were not able to apply the quotient principle to answer the question. This is a simple prototype of the problems we will solve by using permutation groups. Though we don't yet have the tools to solve it, we can get an interesting example of a group acting on a set from it. An assignment of red and blue beads to the corners of a square can be thought of as a function f from the vertices of the square to the set $\{R, B\}$. For example, in Figure 6.2.2 we would have $f(1) = R$, $f(2) = R$, $f(3) = B$, $f(4) = B$. What is the original position of the vertex that σ moves to position j? In terms of f, σ and j, give an expression that represents the color of the vertex in position j after we apply the permutation σ to the vertices.

Solution. The original position of the vertex that σ moves to position j is $\sigma^{-1}(j)$. Thus the color of the vertex in position j after we apply σ is $f(\sigma^{-1}(j))$, which is also $(f \circ \sigma^{-1})(j)$.

Activity 273. Problem 272 suggests that when we have a group of permutations of a set S and we want to know how it acts on the functions from S to a set C (of "colors"), we should define $\beta_\sigma(f)$ to be $f \circ \sigma^{-1}$. For the coloring function f in Figure 6.2.2, let $g = f \circ \rho^{-1}$ and compute $g(1)$, $g(2)$, $g(3)$, and $g(4)$ so you can see how in this case $f \circ \rho^{-1}$ is the coloring that results from rotating the colored square by the rotation ρ. Show if a group G of permutations acts on a set S, then $\beta_\sigma(f) = f \circ \sigma^{-1}$ satisfies the definition of of G acting on the set of functions f from S to C, while $\beta_\sigma(f) = f \circ \sigma$ does not.

Solution. $g(1) = B$, $g(2) = R$, $g(3) = R$, and $g(4) = B$.
    If $\beta_\sigma(f) = f \circ \sigma^{-1}$, then

$$(\beta_\sigma \circ \beta_\tau)(f) = \beta_\sigma(\beta_\tau(f))$$

$$=\beta_\sigma(f \circ \tau^{-1})$$
$$=(f \circ \tau^{-1}) \circ \sigma^{-1}$$
$$=f \circ (\tau^{-1} \circ \sigma^{-1})$$
$$=f \circ (\sigma \circ \tau)^{-1}$$
$$=\beta_{\sigma \circ \tau}(f).$$

Therefore $\beta_\sigma(f) = f \circ \sigma^{-1}$ defines an action of a group of permutations of a set X on functions defined from the set X to a set Y.

On the other hand, defining $\beta_\sigma(f) = f \circ \sigma$ does not define a group action. For example, if we consider the function $f(i) = i$ on the four numbered corners of the square in Figure 6.2.2, and take the permutations ρ and $\varphi_{1/3}$, we know from the fact that f is the identity function and ρ and $\varphi_{1/3}$ do not commute that $\beta_{\varphi_{1/3} \circ \rho}(f) = f \circ (\varphi_{1/3} \circ \rho) \neq f \circ (\rho \circ \varphi_{1/3}) = (f \circ \rho) \circ \varphi_{1/3} = \beta_\rho(f) \circ \varphi_{1/3} = \beta_{\varphi_{13}}(\beta_\rho(f)) = (\beta_{\varphi_{1/3}} \circ \beta_\rho)(f)$. Thus defining $\beta_\sigma(f) = f \circ \sigma$ does not give an action of D_4 on functions from the vertices of the square to $\{1, 2, 3, 4\}$.

Activity 274. If we are interested in something like the number of ways of painting the faces of a cube, and we know how the rotation group of the cube acts on the faces of the cube, then we are interested in how the rotation group acts on functions from the set of faces of the cube to the set of colors. The difference between this problem and Problem 273 is that in Problem 273 it was the members σ of our permutation group that were permuting the elements of S, while now we have bijections β_σ that are permuting the elements of S, in this case the faces of the cube. We want to define a new action β'_σ of our permutation group on functions from S to a set C. How should we define $\beta'_\sigma(f)$? (Hint: think about the example of the faces of the square. After we act on the faces with β_σ, what is the color of the vertex of the cube that is now in position x?)

Solution.   We should define $\beta'_\sigma(f) = f \circ \beta_\sigma^{-1}$. The color of the vertex now in position x is $f(\beta_\sigma^{-1}(x))$ (which is the same as $(f \circ \beta_\sigma^{-1})(x)$.

Orbits

In Problem 270 you saw that the action of the dihedral group D_4 on two element subsets of $\{1, 2, 3, 4\}$ seems to split them into two sets, one with two elements and one with 4. We call these two sets the "orbits" of D_4 acting on the two elements subsets of $\{1, 2, 3, 4\}$. More generally, the *orbit* of a permutation group G determined by an element x of a set S on which G acts is

$$\{\beta_\sigma(x) | \sigma \in G\},$$

and is denoted by Gx.

When we used the quotient principle to count circular seating arrangements or necklaces, we partitioned up a set of lists of people or beads into blocks of equivalent lists. In the case of seating n people around a round table, what made two lists equivalent was, in retrospect, the action of the rotation
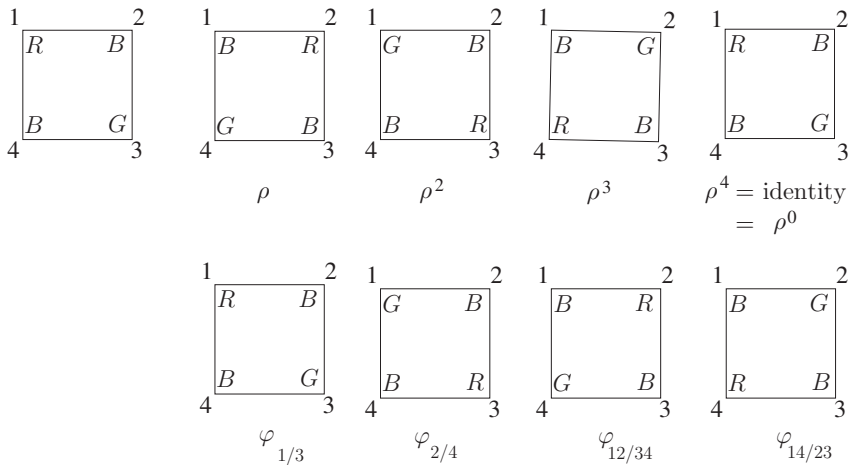
group R_n. In the case of stringing n beads on a string to make a necklace, what made two lists equivalent was the action of the dihedral group. Thus the blocks of our partitions were orbits of the rotation group or the dihedral group, and we were counting the number of orbits of the group action. With this understanding, we will aim to develop tools that allow us to count the number of orbits of a group acting on a set even when the orbits have different sizes. First, though, we have to learn to analyze what the possible sizes of orbits are.

Figure 6.2.3: The four possible results of rotating a colored square and maintaining its position.

In Figure 6.2.3 we show a square which has had its vertices colored with three colors, and we show how the rotation group R_4 acts on this coloring. If we denote our original coloring by f, then $g = f \circ \rho^{-1}$ is the function with $g(1) = B$, $g(2) = R$, $g(3) = B$, and $g(4) = G$. The other functions in the orbit containing f are shown in Figure 6.2.3.

Activity 275. Draw a figure like Figure 6.2.3 that illustrates the action of the dihedral group on the function f above. How many elements are in the orbit of f under the action of dihedral group? How many elements of the dihedral group fix f; that is, for how many $\sigma \in D_4$ is $\beta_\sigma(f) = f$? How many elements of the dihedral group take f to g (where $g = f \circ \rho^{-1}$ as earlier)?

Solution.

Top row of squares:

ρ — square with vertices 1,2,4,3; cells: R B / B G

ρ^2 — square: B R / G B

ρ^3 — square: G B / B R

square: B G / R B

$\rho^4 = $ identity $= \rho^0$ — square: R B / B G

Bottom row of squares:

$\varphi_{1/3}$ — square: R B / B G

$\varphi_{2/4}$ — square: G B / B R

$\varphi_{12/34}$ — square: B R / G B

$\varphi_{14/23}$ — square: B G / R B

Figure 6.2.4: Motions of a colored square

There are four elements in the orbit. Two elements fix f. Two elements take f to g.

Activity 276. If $f_1(1) = R$, $f_1(2) = B$, $f_1(3) = G$ and $f_1(4) = G$, how many elements are in the orbit of f_1 under the action of the dihedral group? How many elements of the dihedral group fix f_1; that is, for how many $\sigma \in D_4$ is $\beta_\sigma(f_1) = f_1$? How many elements of the dihedral group take f_1 to $f_1 \circ \rho^{-1}$?

Solution.

Activity 277. If $f_2(1) = R$, $f_2(2) = B$, $f_2(3) = R$ and $f_2(4) = B$, how many elements are in the orbit of f_2 under the action of the dihedral group? How many elements of the dihedral group fix f_2; that is, for how many $\sigma \in D_4$ is $\beta_\sigma(f_2) = f_2$? How many elements of the dihedral group take f_2 to $f_2 \circ \rho^{-1}$?

Solution.

Activity 278. The rotation group of the square with vertices 1, 2, 3, and 4 in Figure 6.2.1 can be thought of as a subgroup of the rotation group of the cube. Namely it corresponds to the rotations of the cube around the axis that joins the center of the bottom face with the center of the top face. Recall that we have used R to stand for the rotation group of the cube. Does $\beta_{\rho^i}(\sigma) = \rho^i \circ \sigma$ define an action of R_4 on R? Does $\beta_{\rho^i}(\sigma) = \sigma \circ \rho^i$ define a group action of R_4 on R? Does $\beta_{\rho^i}(\sigma) = \sigma \circ \rho^{-i}$ define a group action of R_4 on R? For each group action that you found, describe the orbits as simply as you can. How many orbits are there and what are their sizes? (Hint: to do this problem, you don't need to know anything about the rotation group of the cube except for the fact that it is a group that has R_4 as a subgroup and the sizes of R_4 and R to answer this question.)

Solution. Yes, $\beta_{\rho^i}(\sigma) = \rho^i \circ \sigma$ does define an action of R_4 on R. The orbits are sets of the form $\{\rho^i \circ \sigma | i = 0, 1, 2, 3\}$ for fixed elements σ of R. (For each

fixed element σ we get an orbit, but four different elements σ can give the same orbit.) Yes, $\beta_{\rho^i}(\sigma) = \sigma \circ \rho^i$ does define a group action of R_4 on R. Notice that this gives a group action only because the group R_4 is commutative. The orbits are sets of the form $\{\sigma \circ \rho^i | i = 0, 1, 2, 3\}$ for fixed elements σ of R. Yes, $\beta_{\rho^i}(\sigma) = \sigma \circ \rho^{-i}$ does define a group action of R_4 on R. The orbits consist again of sets of the form $\{\sigma \circ \rho^i | i = 0, 1, 2, 3\}$ for fixed elements σ of R. In each case the orbits have size 4 (because if $\rho^i \circ \sigma = \rho^j \circ \sigma$, then $\rho^i = \rho^j$, and similarly for $\sigma \circ \rho^i$), and there are six of them.

Activity 279. What are the sizes of the orbits, and how many have each size, for the action of D_6 on the two element subsets of $\{1, 2, 3, 4, 5, 6\}$.

Solution.   There are two orbits of size 6 and one orbit of size 3. (Think of the two-element subsets as the edges and diagonals of a hexagon. The edges form one orbit of size six; the diagonals that jump over two edges form another orbit of size six; and the diagonals connecting opposite vertices form an orbit of size three.

Activity 280. Suppose we draw identical circles at the vertices of a regular hexagon. Suppose we color these circles with two colors, red and blue. We may think of a coloring as a function from the set $\{1, 2, 3, 4, 5, 6\}$ to the set $\{R, B\}$ by numbering the vertices from one to six consecutively around the hexagon. To do this problem it will be helpful to use the notation $RBRRBB$ to stand for the function f with $f(1) = R$, $f(2) = B$, $f(3) = R$, $f(4) = R$, $f(5) = B$, $f(6) = B$. How many functions are there from the set $\{1, 2, 3, 4, 5, 6\}$ to the set $\{R, B\}$? These functions are partitioned into orbits by the action of the rotation group on the hexagon. Using our simplified notation, write down all these orbits and observe what the possible sizes of orbits are.

Solution.   There are 64 functions from a six-element set to a two-element set.

   $\{RRRRRR\}$,
   $\{RRRRRB, BRRRRR, RBRRRR, RRBRRR, RRRBRR, RRRRBR\}$,
   $\{RRRRBB, BRRRRB, BBRRRR, RBBRRR, RRBBRR, RRRBBR\}$,
   $\{RRRBRB, BRRRBR, RBRRRB, BRBRRR, RBRBRR, RRBRBR\}$,
   $\{RRRBBB, BRRRBB, BBRRRB, BBBRRR, RBBBRR, RRBBBR\}$,
   $\{RRBRRB, BRRBRR, RBRRBR\}$,
   $\{RRBRBB, BRRBRB, BBRRRB, RBBRRB, BRBBRR, RBRBBR\}$,
   $\{RRBBRR, RRRBBR, RRRRBB, BRRRRB, BBRRRR, RBBRRR\}$,
   $\{RRBBRB, BRRBBR, RBRRBB, BRBRRB, BBRBRR, RBBRBR\}$
   $\{RRBBBB, BRRBBB, BBRRBB, BBBRRB, BBBBRR, RBBBBR\}$
   $\{RBRBRB, BRBRBR\}$
   $\{RBBRBB, BRBBRB, BBRBBR\}$
   $\{RBBBBB, BRBBBB, BBRBBB, BBBRBB, BBBBRB, BBBBBR\}$
   $\{BBBBBB\}$
   We have two orbits of size one, one of size two, two of size three, and nine of size six.

Activity 281. Either show that, when G is a group acting on a set S, the orbits of G partition S or give a counter-example.

Solution. Each element x is in at least one orbit, namely Gx. But if $x \in Gy$, then $x = \sigma y$ for some $\sigma \in G$, so

$$Gx = \{\tau x | \tau \in G\} = \{\tau \sigma y | \tau \in G\} = \{\tau' y | \tau' \in G\} = Gy.$$

Thus each element of S is in one and only one orbit of G, so the orbits of G partition S.

Activity 282. In Problems 275, Activity 276, and Activity 277, what set of elements fixes f and what set of elements takes f to g, what set of elements fixes f_1 and what set of elements takes f_1 to $f_1 \circ \rho^{-1}$, and what set of elements fixes f_2 and what set of elements takes f_2 to $f_2 \circ \rho^{-1}$?

Solution. In Problem 275, the set of elements fixing f is $\{\iota = \rho^0, \varphi_{1/3}\}$. The set of orbits taking f to g is $\{\rho, \varphi_{12/34}\}$. In Problem 276 the set of elements fixing f_1 is $\{\iota\}$, and the set of elements taking f_1 to $f_1 \circ \rho^{-1}$ is $\{\rho\}$. In Problem 277, the set of elements fixing f_2 is $\{\iota, \rho^2, \varphi_{1/3}, \varphi_{2/4}\}$, and the set of elements taking f_2 to $f_2 \circ \rho^{-1}$ is $\{\rho, \rho^3, \varphi_{12/34}, \varphi_{14/23}\}$.

Activity 283.

(a) In Problem 282 the subsets of D_4 that fix a function have a special property. What is it?

Solution. In each case, the subsets that fix a function form a subgroup. To prove this, note that the subset of group elements fixing a function is closed under composition.

(b) Could there be a function h from $\{1, 2, 3, 4\}$ to the set $\{R, B, G\}$ such that the orbit of h under the action of D_4 has size 3 or 5? If so, find one. If not, explain why not.

Solution. This is difficult to do at this stage except by "brute force." Here is an organized brute force solution. There is no orbit of size 3 or 5. If a function is constant, then clearly its orbit has size one. If a function has exactly one R (or B or G), then its orbit under the rotations has size 4. Since every flip is a rotation times $\varphi_{1/3}$, if $\varphi_{1/3}$ fixes the function, then the orbit has size 4, but if $\varphi_{1/3}$ does not take the function to the result of a rotation, then neither does a rotation times $\varphi_{1/3}$, so the orbit has size 8. If the function has exactly three R's, then it has exactly one element of some other color, so the argument is the same. If the function has exactly two R's and exactly two elements of another color, then if the elements of the same color are adjacent, the orbit under the rotations has size 4, and the result of a flip is the same as the result of a rotation so the orbit has size 4. If the two R's are not adjacent, then

the orbit under the rotations has size 2 and the result of each flip is the same as the result of a rotation, so the orbit has size 2. If the function has exactly two R's, one G, and one B, then the size of the orbit is the same as the orbit in Problem 246 or 254. Any other case is equivalent to one we just covered by changing the names of the colors. Thus there are no orbits of size 3 or 5 (or 6 or 7 either). One can also argue as we will in Problem 284.

Activity 284. Make a conjecture about how the size of an orbit of the permutation group G acting on the set S relates to the size of G. Prove your conjecture.

Solution.   The size of an orbit divides the size of the group. To prove the conjecture, the most natural approach is to try to use the quotient principle. This requires that we divide the group up into disjoint subsets of the same size. The elements of the group that fix a given element x form a subgroup H of G. In each case we have seen, the elements that take x to $\sigma(x)$ were of the form $\sigma \circ \tau$, where τ was in the subgroup fixing H. Clearly any such element takes x to $\sigma(x)$. The number of such elements is the size of H. And if $\sigma'(x) = \sigma(x)$, then $\sigma^{-1} \circ \sigma'$ fixes x and so is a member τ of the subgroup. But then $\sigma' = \sigma \circ \tau$, where $\tau = \sigma^{-1}\sigma'$. Thus the number of elements σ that take x to an element $\sigma(x)$ is the size of H. And the set of elements that take x to $\sigma(x)$ has nothing in common with the set of elements that take x to $\sigma_1(x)$, so we have partitioned up the group into subsets each of which has the same size as H, and therefore the size of H time the number of subsets is the size of G. However the number of subsets is the number of orbits, and so the number of orbits divides the size of the group.

Activity 285. Make a conjecture about how the size a subgroup of the permutation group G relates to the size of G.

Solution.   The size of a subgroup of a permutation group G divides the size of G.

Problems 265, Activity 278, Activity 282 and Activity 283 (and less directly Problems 272, Activity 273 and Activity 274) have a common theme.

In Problem 265, a natural way to show that the group of even permutations, those with sign 1, has size $n!/2$ is to observe that for any two-cycle (or any odd permutation) τ, the element $\sigma\tau$ is odd whenever σ is even and is even whenever σ is odd. Thus if we let τA_n stand for the set

$$\tau A_n = \{\tau\sigma | \sigma \in A_n\},$$

then A_n and tA_n are two disjoint sets that partition S_n.

Activity 286. Explain why, for any subgroup H of a permutation group G, H and τH have the same size.

Solution.   There is a bijection between H and τH given by $f(\sigma) = \tau\sigma$. (It is a bijection because τ has an inverse.)

By problem 286 A_n and τA_n have the same size and since their union has $n!$ elements, each has size $n!/2$. The set τA_n is called a *left coset* of the subgroup A_n of S_n.

In Problem 278 you probably saw that $\beta_{\rho^i}(\sigma) = \rho^i \circ \sigma$ and $\beta_{\rho^i}(\sigma) = \sigma \circ \rho^{-i}$ both describe actions of R_4 on R. The orbits for the first action are of the form

$$R_4\sigma = \{\rho^i\sigma | i = 1, 2, 3, 4\},$$

for various elements σ of R. The orbits of the second action are

$$\{\sigma r^{-i} | i = 1, 2, 3, 4\} = \{\sigma r^i | i = 1, 2, 3, 4\} = \sigma R_4,$$

for various elements σ of R. The set $R_4\sigma$ is called a *right coset* of R_4 in R, and as before, the set σR_4 is called a left coset of R_4 in R. In general when H is a subgroup of G and $\sigma \in G$, we call

$$\sigma H = \{\sigma\varphi | \varphi \in H\}$$

a *left coset* of H in G and we call

$$H\sigma = \{\varphi\sigma | \varphi \in H\}$$

a *right coset* of H in G. Notice that H is a coset of itself, since we may take $\sigma = \iota$. In Problem 286 you showed that a subgroup and all its (left) cosets have the same size. Of course the right cosets have that size too.

When a group G acts on a set S, we say that an element $\sigma \in G$ fixes an element $x \in S$ if $\beta_\sigma(x) = x$. We say that an element σ takes x to y if $\beta_\sigma(x) = y$. In Problem 283 the subset of D_4 that fixed f was a subgroup of D_4, the subset of D_4 fixing f_2 was a subgroup of D_4 and the subset of D_4 fixing f_3 was a subgroup of D_4. The set of group elements that sent f to g, f_2 to $f_2 \circ \rho^{-1}$ and f_3 to $f_3 \circ \rho^{-1}$ was a coset of the subset that fixed the function in each case.

Activity 287. Can two distinct cosets of a subgroup H of a group G have an element in common?

Solution. No, because if $\tau \in \sigma H$, then $\tau H = \sigma H$ and if $\tau \in \varphi H$, then $\tau H = \varphi H$, so that $\sigma H = \varphi H$.

Activity 288. Prove your conjecture in Problem 285.

Solution. By Problem 287, the cosets of a subgroup of G partition G. But by problem 286, the cosets all have the same size, namely the size of H. Thus by the quotient principle, the size of H is a factor of the size of G.

Activity 289. Show that whenever a group G acts on a set S, then for each x in S, the set of all elements σ fixing x is a subgroup of G. We will denote this subgroup by $Fix(x)$.

Solution. The set of elements fixing x is closed under composition, and therefore $Fix(x)$ is a subgroup of G.

Activity 290. Show that whenever a group G acts on a set S, then for each x and y the set of all $\tau \in G$ such that $\beta_\tau(x) = y$ is a coset of the subgroup $Fix(x)$ of all elements of G fixing x.

Solution.   First, if $\beta_\tau(x) = y$ and β_σ fixes x, then $\beta_\tau \circ \beta_\sigma = \beta_{\tau \circ \sigma}$ takes x to y. Second, if $\beta_\varphi(x) = y$ as well, then $\beta_{\tau^{-1}} \circ \beta_\varphi = \beta_{\tau^{-1} \circ \varphi}$ fixes x and so is in $Fix(x)$, so that φ is τ times an element (namely $\tau^{-1}\varphi$) of $Fix(x)$. Thus every element of the coset $\tau Fix(x)$ takes x to y, and every element that takes x to y is an element of the coset $\tau Fix(x)$. Thus the set of elements τ such that $\beta_\tau(x) = y$ is a coset of $Fix(x)$.

Activity 291. When a group G acts on a set S, how is the number of elements in the orbit containing the element x of S related to the size of G and the size of $Fix(x)$? Find a bijection that proves that what you say is correct.

Solution.   The number of elements in the orbit containing the element x is the number of cosets of $Fix(x)$, which is also $|G|/|Fix(x)|$. A bijection that proves this is the function f that takes the coset $\varphi Fix(x)$ to the element $\varphi(x)$. This function is well-defined because all the elements in $\varphi Fix(x)$ take x to $\varphi(x)$. It is onto because if $\tau(x) = y$, then the coset $\tau Fix(x)$ is mapped to y. It is one-to-one because if $f(\varphi Fix(x)) = f(\tau Fix(x))$, then $\varphi \in \tau Fix(x)$ and $\tau \in \varphi Fix(x)$, which means the two cosets are identical. The cosets all have the same size, namely $|Fix(x)|$ and therefore by the quotient principle, the number of cosets is $|G|/|Fix(x)|$.

We can summarize our findings in this section with a theorem.

Theorem 6.2.5. *When a group G acts on a set, the set $Fix(x)$ of group elements fixing any given any given element x of S is a subgroup of G. The set of group elements τ such that $\beta_\tau(x) = y$ is a coset of $Fix(x)$, and the size of the orbit of x under the action of G is $|G|/|Fix(x)|$.*

Multiorbits

Activity 292. Draw figures like Figure 6.2.3 to show the eight results of acting with the dihedral group on the squares corresponding to the functions f, f_1, and f_2 of Problems 275, Activity 276, and Activity 277.

Solution.

Row 1:

1 2	1 2	1 2	1 2	1 2
R B	B R	G B	B G	R B
B G	G B	B R	R B	B G
4 3	4 3	4 3	4 3	4 3

$$\rho \qquad \rho^2 \qquad \rho^3 \qquad \rho^4 = \text{identity} = \rho^0$$

Row 2:

1 2	1 2	1 2	1 2
R B	G B	B R	B G
B G	B R	G B	R B
4 3	4 3	4 3	4 3

$$\varphi_{1/3} \qquad \varphi_{2/4} \qquad \varphi_{12/34} \qquad \varphi_{14/23}$$

Row 3:

1 2	1 2	1 2	1 2	1 2
R B	G R	G G	B G	R B
G G	G B	B R	R G	G G
4 3	4 3	4 3	4 3	4 3

$$\rho \qquad \rho^2 \qquad \rho^3 \qquad \rho^4 = \text{identity} = \rho^0$$

Row 4:

1 2	1 2	1 2	1 2
R G	G B	B R	G G
B G	G R	G G	R B
4 3	4 3	4 3	4 3

$$\varphi_{1/3} \qquad \varphi_{2/4} \qquad \varphi_{12/34} \qquad \varphi_{14/23}$$

Row 5:

1 2	1 2	1 2	1 2	1 2
R B	B R	G B	B G	R B
B G	G B	B R	R B	B G
4 3	4 3	4 3	4 3	4 3

$$\rho \qquad \rho^2 \qquad \rho^3 \qquad \rho^4 = \text{identity} = \rho^0$$

Row 6:

1 2	1 2	1 2	1 2
R B	G B	B R	B G
B G	B R	G B	R B
4 3	4 3	4 3	4 3

$$\varphi_{1/3} \qquad \varphi_{2/4} \qquad \varphi_{12/34} \qquad \varphi_{14/23}$$

As Problem 292 shows, in some ways it is more natural to think of an orbit as a multiset rather than a set. For example with the function f_2 from

Problem 277 we get f_2 itself four times when we act with the dihedral group and we get $f_2 \circ \rho^{-1}$ four times when we act on f_2 with the dihedral group. If we think only about the orbit of the dihedral group, then, we lose some information. To avoid this, we define the *multiorbit* of an element x of a set S under the action of a group G to be the multiset

$$Gx_{\text{multi}} = \{\beta_\sigma(x)|\sigma \in G\}_{\text{multi}}.$$

It is immediate that the size of each multiorbit is then exactly the size of G.

Activity 293. In how many different multiorbits of the action of G on S will a given element of S appear?

Solution.   Only one, because the orbits partition S, and there is a bijection between the orbits and the multiorbits given by $f(Gx) = Gx_{\text{multi}}$.

Activity 294. What will the multiplicity of x be in Gx_{multi}?

Solution. $|Fix(x)|$.

It is also immediate from Problem 293 that the number of orbits is the number of multiorbits, because we get the orbits from the multiorbits by deleting repeated elements, and the problem shows that each orbit corresponds to exactly one multiorbit[6]. We define the union of multisets M_1, M_2, ..., M_n to be the multiset in which the multiplicity of an element x is the sum of its multiplicities in the individual multisets M_i. Thus

$$\{a, a, b, b, b\} \cup \{a, b, b, c, c, c\} = \{a, a, a, b, b, b, b, b, c, c, c\}.$$

We will now get a formula for the number of multiorbits by using the fact that they all have the same size and the idea of multiset union.

Activity 295. How does the size of the union of the set of multiorbits of a group G acting on a set S relate to the number of multiorbits and the size of G?

Solution.   It is simply the product of the number of multiorbits and the size of G.

Activity 296. How does the size of the union of the set of multiorbits of a group G acting on a set S relate to the numbers $|Fix(x)|$?

Solution.   Since it is the sum of the multiplicities of the elements of S in the union, it is the sum of $|Fix(x)|$ over all x in S.

---

[6]You might want to argue that we have a multiset of multiorbits, one for each element of S, but that would defeat our purpose since we have bijection between the set of orbits and the *set* of multiorbits, and this bijection will let us count the number of orbits by computing the size of the set of multiorbits.

Activity 297. In Problems 295 and Activity 296 you computed the size of the union of the set of multiorbits of a group G acting on a set S in two different ways, getting two different expressions which must be equal. Write the equation that says they are equal and solve for the number of multiorbits, and therefore the number of orbits.

Solution. Using m for the number of multiorbits, we get

$$m|G| = \sum_{x:x \in S} |Fix(x).$$

Therefore,

$$m = \frac{1}{|G|} \sum_{x:x \in S} |Fix(x)|.$$

The Cauchy-Frobenius-Burnside Theorem

Activity 298. In Problem 297 you stated and proved a theorem that expresses the number of orbits in terms of the number of group elements fixing each element of S. It is often easier to find the number of elements fixed by a given group element than to find the number of group elements fixing an element of S. For this purpose,

(a) Let $\chi(\sigma, x) = 1$ if $\sigma(x) = x$ and let $\chi(\sigma, x) = 0$ otherwise. Use χ to convert the single summation in Problem 297 into a double summation over elements x of S and elements σ of G.

Solution.

$$m = \frac{1}{|G|} \sum_{x:x \in S} \sum_{\sigma:\sigma \in G} \chi(\sigma, x).$$

(b) Reverse the order of the previous summation in order to convert it into a single sum involving the function χ given by[7]

$$\chi(\sigma) = \text{the number of elements of } S \text{ left fixed by } \sigma.$$

Solution.

$$m = \frac{1}{|G|} \sum_{\sigma:\sigma \in G} \chi(\sigma).$$

In Problem 298 you gave a formula for the number of orbits of a group G acting on a set X. This formula was first worked out by Cauchy in the case of the symmetric group, and then for more general groups by Frobenius. In his pioneering book on Group Theory, Burnside used this result as a lemma, and

---

[7]The reason for using the Greek letter χ is that $\chi(\sigma)$ is an example of what is called a group character. Character theory is a major ingredient of the representation theory of groups, an advanced subject that combines abstract beauty with amazing utility.

while he attributed the result to Cauchy and Frobenius in the first edition of his book, in later editions, he did not. Later on, other mathematicians who used his book named the result "Burnside's Lemma," which is the name by which it is still most commonly known. Let us agree to call this result the Cauchy-Frobenius-Burnside Theorem, or CFB Theorem for short in a compromise between historical accuracy and common usage.

Activity 299. In how many ways may we string four (identical) red, six (identical) blue, and seven (identical) green beads on a necklace?

Solution. We are stringing 17 beads on our necklace, so we are asking for the number of orbits of the group D_{17} on lists of four Rs, six Bs, and seven Gs. For a rotation ρ^i to fix a list, it must take an R to a R. The powers of ρ^i form a group. Thus the set of places that contain an R in a list that is fixed by ρ^i must be an orbit of that group. But the size of the orbit must be a divisor of the size of the subgroup, which must be a divisor of seventeen, so the size of the orbit is one or 17. If it is 1, then $i = 0$. Thus no elements are fixed by any nontrivial rotation, and all $\binom{17}{4,6,7}$ lists of Rs, Bs and Gs are fixed by the identity. Each flip will be a flip around a line from a bead to the spot between the two "opposite" beads. This line divides the list into the eight beads on its left, the eight beads on its right and the one bead it goes through. If a flip fixes a list, the eight beads on the left must be identical to the eight beads on the right, meaning that the bead the line goes through must be one of the seven green beads. Thus the number of lists that are fixed by a flip is the number of ways to place three green beads, two red beads, and three blue beads in 8 slots, which is $\binom{8}{3,2,3}$, or $\frac{8!}{3!2!3!}$. There are 17 flips, so there are $\frac{17\cdot8!}{3!2!3!}$ elements fixed by flips. Therefore we have

$$\frac{1}{2\cdot 17}\left(\frac{17!}{4!6!7!} + \frac{17\cdot 8!}{3!2!2!}\right) = \frac{8\cdot 15!}{4!6!7!} + \frac{4\cdot 7!}{3!2!3!} = 120,400$$

necklaces. The numerical answer, which is unimportant here, was obtained from Maple.

Activity 300. If we have an unlimited supply of identical red beads and identical blue beads, in how many ways may we string 17 of them on a necklace?

Solution. We are asking for the number of orbits of D_{17} on the set of functions from $[17]$ to $\{R, B\}$. Every function is fixed by the identity. The only functions fixed by a nontrivial rotation are the constant functions. Each flip is around a line from a bead to the space between the two "opposite" beads. The bead the line goes through can be either color, and then the eight beads to the left of this one must be identical to the eight beads to the right of this one. There are 2^8 ways to assign beads to the positions on the left, so a flip fixes 2^9 functions. Therefore by the CFB theorem, we have

$$\frac{1}{2\cdot 17}\left(2^{17} + 17\cdot 2^9 + 16\cdot 2\right) = \frac{2^{16}}{17} + 2^8 + \frac{2^4}{17} = 16\frac{4097}{17} + 256 = 4112$$

necklaces.

Activity 301. If we have five (identical) red, five (identical) blue, and five (identical) green beads, in how many ways may we string them on a necklace?

Solution. Here we need to consider the action of D_{15} on lists (functions from $[15]$ to $\{R, B, G\}$) with five Rs, five Bs, and five Gs. The identity will fix $\binom{15}{5,5,5}$ lists. A rotation through 3,6,9,or 12 places will fix any function that has the same value in places 1, 4, 7, 10, and 13, the same value in places 2, 5, 8, 11, and 14, and the same value in places 3, 6, 9, 12, and 15. There are 3! such lists. There is no other rotation that fixes any lists. Each flip is around an axis that goes from bead to the space between two opposite beads. If it fixed a list, the seven colors to the left of the axis would have to equal the seven colors to the right of the axis. Thus the number of beads of each color on the left and right sides would have to be equal. So except for the color of the bead the axis goes through, we would have to have an even number of beads of each color for a flip to fix a list of beads. Thus no flip fixes any functions. Therefore by the CFB theorem we have

$$\frac{1}{2 \cdot 15} \left(\frac{15!}{5!5!5!} + 4 \cdot 3! \right) = 25,226$$

necklaces. (The numerical answer, which was obtained from Maple, is not important here.)

Activity 302.

(a) In how many ways may we paint the faces of a cube with six different colors, using all six?

Solution. Here we must consider the action of the rotation group of the cube on lists of six distinct colors. But no nontrivial rotation will fix a cube with all its faces colored differently. The identity rotation will fix all 6! lists, and there are 24 members of the rotation group, so we have $6!/24 = 6 \cdot 5 = 30$ ways to paint the faces of a cube with six distinct colors, using each color

(b) In how many ways may we paint the faces of a cube with two colors of paint? What if both colors must be used?

Solution. We must consider the action of the rotation group of the cube on functions from $[6]$ to $\{R, B\}$ or some other two-element set of colors. There are five kinds of elements in the rotation group of the cube. There is one identity, there are six rotations by 90 degrees or 270 degrees around an axis connecting the centers of two opposite faces, there are three rotations of 180 degrees around such an axis, there are 8 rotations (of 120 degrees and 240 degrees, respectively) around an axis connecting two diagonally opposite vertices, and there are 6 rotations of 180 degrees around an axis connecting the centers of two opposite edges. The identity fixes 2^6 functions. There are eight functions fixed by a 90

degree or 270 degree rotation. There are 16 functions fixed by a 180 degree rotation along an axis through two faces. There are 8 functions fixed by a 180 degree rotation along an axis joining the centers of two opposite sides. There are 4 functions fixed by a 120 degree or 240 degree rotation. Thus by the CFB theorem, we have

$$\frac{1}{24}(64 + 6 \cdot 8 + 3 \cdot 16 + 8 \cdot 4 + 6 \cdot 8) = 10$$

ways to paint the faces of a cube with two colors of paint. Two of these colorings use only one color, so there are eight colorings that use both colors.

Activity 303. In how many ways may we color the edges of a (regular) $(2n + 1)$-gon free to move around in the *plane* (so it cannot be flipped) if we use red n times and blue $n+1$ times? If this is a number you have seen before, identify it.

Solution.   The set of all powers of a rotation is a subgroup of the rotation group of the $(2n + 1)$-gon. If a given rotation fixes a coloring, all powers of that rotation fix the coloring. The set of edges to which a given edge is taken by that rotation is an orbit of the group of powers of the rotation. The size of this orbit is a divisor of the size of the subgroup, which is a divisor of $2n + 1$. If the edge is colored red, the size of the orbit is also a divisor of n, and if the edge is colored blue, the size of the orbit is also a divisor of $n+1$. But neither n nor $n + 1$ have common divisors with $2n + 1$, except for one. Therefore the only rotation that fixes a coloring is the identity rotation, and it fixes all $\binom{2n+1}{n}$ colorings. Thus the number of orbits is

$$\frac{1}{2n + 1}\binom{2n + 1}{n} = \frac{1}{n + 1}\binom{2n}{n},$$

a Catalan number.

Activity 304. In how many ways may we color the edges of a (regular) $(2n + 1)$-gon free to move in *three-dimensional space* so that n edges are colored red and $n + 1$ edges are colored blue. Hint: your answer may depend on whether n is even or odd.

Solution.   The set of all powers of a rotation is a subgroup of the dihedral group of the $(2n + 1)$-gon. If a given rotation fixes a coloring, all powers of that rotation fix the coloring. The set of edges to which a given edge is taken by that rotation is an orbit of the group of powers of the rotation. The size of this orbit is a divisor of the size of the subgroup, which is a divisor of $2n + 1$. If the edge is colored red, the size of the orbit is also a divisor of n, and if the edge is colored blue, the size of the orbit is also a divisor of $n+1$. But neither n nor $n + 1$ have common divisors with $2n + 1$, except for one. Therefore the only rotation that fixes a coloring is the identity rotation, and it fixes all $\binom{2n+1}{n}$ colorings.

A flip, on the other hand can fix some colorings. In particular, if n is even, we color one edge blue, leaving an even number n of edges to be colored blue and an even number n of edges to be colored red. It the flip over the axis perpendicular to the side we picked fixes the coloring, the n edges to the right of the chosen edge must be colored identically with the n edges to the left of the chosen edge. There are $\binom{n}{n/2}$ ways to color the n edges to the right of the chosen edge (choose which edges get red), and so this is the number of colorings fixed by this flip. We have $2n + 1$ flips, and since we have one flip of the type described for each edge, all flips have the form given above. Thus each flip fixes $\binom{n}{n/2}$ colorings.

Thus the number of orbits is

$$\frac{1}{2(2n + 1)} \left(\binom{2n + 1}{n} + (2n + 1) \binom{n}{n/2} \right).$$

If n is odd, we color one edge red, leaving an even number $n - 1$ of edges to be colored red and an even number $n + 1$ of edges to be colored blue. With an argument similar to the previous one we see that there are

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

orbits.

Activity 305. (Not unusually hard for someone who has worked on chromatic polynomials.) How many different proper colorings with four colors are there of the vertices of a graph which is cycle on five vertices? (If we get one coloring by rotating or flipping another one, they aren't really different.)

Solution. We are asking for the number of orbits of D_5 on lists of five colors chosen from the given four with no two adjacent colors equal; we consider the first and last position adjacent as well. The identity fixes all such colorings. For those who are familiar with the chromatic polynomial, as in Problem 226, the number of such colorings is the chromatic polynomial of the cycle on five vertices evaluated at 4. The number of ways to properly color a cycle on five vertices is the number of ways to color a path on five vertices minus the number of ways to color a path on five vertices so that its first and last vertices are identical, which is the number of ways to color a cycle on four vertices. The number of ways to color a cycle on four vertices is the number of ways to color a path on four vertices minus the number of ways to color a path on four vertices so that the first and last vertices are the same color, which is the same as the number of ways to color a cycle on three vertices. The number of ways to properly color a path on five vertices with four colors is $4 \cdot 3^4$. The number of ways to properly color a path on four vertices with four colors is $4 \cdot 3^3$. The number ways to properly color a cycle on three vertices with four colors is $4 \cdot 3 \cdot 2$. Thus the number of proper colorings of a five vertex cycle with four colors is $4 \cdot 3^4 - 4 \cdot 3^3 + 4 \cdot 3 \cdot 2 = 240$. This is the number of colorings fixed by the identity. No proper coloring is fixed by a nontrivial rotation, because if a rotation of the five cycle fixes a coloring, all the vertices

must have the same color. No proper coloring is fixed by a flip, because for a flip to fix a coloring, the coloring must give the same color to two adjacent vertices. Thus the number of really different proper four-colorings of a cycle on 5 vertices is $\frac{240}{10} = 24$.

Activity 306. How many different proper colorings with four colors are there of the graph in Figure 6.2.6? Two graphs are the same if we can redraw one of the graphs, not changing the vertex set or edge set, so that it is identical to the other one. This is equivalent to permuting the vertices in some way so that when we apply the permutation to the endpoints of the edges to get a new edge set, the new edge set is equal to the old one. Such a permutation is called an *automorphism* of the graph. Thus two colorings are different if there is no automorphism of the graph that carries one to the other one.
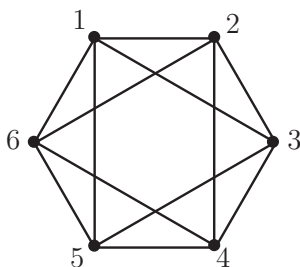
Figure 6.2.6: A graph on six vertices.

Solution.	We want the number of orbits of the set of proper colorings under the action of the group of automorphisms of the graph. An automorphism σ maps vertex 1 to any of six vertices. Vertex 2 can be mapped to any of the four vertices adjacent to the image of vertex 1. Vertex 3 is adjacent to vertices 1 and 2, so it must be mapped to a vertex adjacent to the images of both vertex 1 and vertex 2; by checking cases you can see that there are always exactly two vertices adjacent to the images of vertex 1 and vertex 2, and mapping vertex 3 to either of these vertices preserves all the edges among vertices 1, 2, and 3. However each of the other three vertices is adjacent to exactly two vertices of the set $\{1, 2, 3\}$, and thus it must be mapped to the unique vertex adjacent to the corresponding two of $\sigma(1)$, $\sigma(2)$ and $\sigma(3)$. (It is always the case that each vertex is adjacent to exactly two of these, as you can see by considering the cases with $\sigma(1) = 1$.) Thus there are $6 \cdot 4 \cdot 2 = 48$ elements in the group. Now to apply the CFB theorem we would need to know how many proper colorings are fixed by each group element, so we would need to know what the group elements are. We have observed that a permutation that preserves the edges is determined by where the triangle $\{1, 2, 3\}$ goes. We can see eight triangles in the graph, triangles of the form $\{i, i+1, i+2\}$, where we identify 7 with 1 and 8 with 2, and the triangles $\{1, 3, 5\}$ and $\{2, 4, 6\}$. We can map the set $\{1, 2, 3\}$ to any of these eight sets by six one-to-one maps, so each group element is determined uniquely by one of these mappings. However focusing

on these triangles makes our job here simpler in another way. In a proper coloring, vertices 1, 2, and 3 must be colored differently. We have four choices for the color of vertex 1, three different ones for vertex 2 and two still different ones for vertex 3, so there are 24 ways to color this triangle. Clearly the only difference among these ways is the actual names of the colors. That is, we can assume that vertex 1 is colored red, vertex 2 is colored blue and vertex 3 is colored green, then determine the proper colorings starting with these three colors, and up to changing the names of the colors, we will have determined all the proper colorings. Then we can ask which group elements fix a coloring rather than which colorings are fixed by a group element. This turns out to be easier. Let us write $RBGRBG$ for the coloring that colors vertices 1 and 4 red, two and five blue, and three and six green. An examination of the figure shows that this is a proper coloring. In fact, it is the only proper coloring that starts RBG and uses only three colors. Suppose we were to use a fourth color, Y. Then among vertices 4, 5, and 6, it could be used in just one place, because those three vertices are mutually adjacent. Each of the other two vertices is adjacent to two of the original three vertices colored RBG, and so there is only one color available to use on it. In summary, the colorings that start RBG are

1. $RBGRBG$

2. $RBGYBG$

3. $RBGRYG$

4. $RBGRBY$.

Thus for any of the 24 choices of colorings of the first three vertices, there are four ways to complete it to a proper coloring of the whole graph, so there are 96 proper colorings of the graph. Among the ones that start RGB, let us analyze which group elements fix them. Note that the 2-cycles (1 4), (2 5) and (3 6) are all permutations that fix coloring one. Further, interchanging vertices 1 and 4 does not change the endpoints of any edges, nor does interchanging 2 and 5 nor 3 and 6. So all these two cycles are automorphisms of the graph. A composition of automorphisms must be an automorphism(this follows directly from the definition of automorphism) and so the eight permutations ι, (1 4), (2 5), (3 6), (1 4)(2 5), (1 4)(3 6), (2 5)(3 6), and (1 4)(2 5)(3 6) all are automorphisms of the graph, and they are all in the subgroup of the automorphism group that fixes the coloring in Item 1. Any permutation not in the list will take some vertex to a vertex of another color, and so the eight permutations we listed are the subgroup fixing the coloring in Item 1. The subgroup fixing the coloring in Item 2 is ι, (1 3), (2 4), and (1 3)(2 4). The subgroups fixing the colorings in Item 3 and Item 4 also have size 4. Thus there are $8 + 12 = 20$ pairs of a coloring with R, B, and G, in that order, on vertices 1, 2, and 3 and an automorphism fixing that coloring. Since there are $4 \cdot 3 \cdot 2 = 24$ ways to color vertices 1, 2, and 3 properly, and each gives rise to 20 pairs of a proper coloring and an automorphism fixing that coloring, there

are $20 \cdot 24 = 480$ such pairs. Since the automorphism group has size 48, this means that there are 10 proper colorings of this graph, up to automorphisms. Note that we did not really use the CFB theorem, though we did use the fact that its formula is proved by dividing the number of ordered pairs of a set member and a group element fixing that member by the size of the group.

6.3   Pólya-Redfield Enumeration Theory

George Pólya and Robert Redfield independently developed a theory of generating functions that describe the action of a group G on functions from a set S to a set T when we know the action of G on S. Pólya's work on the subject is very accessible in its exposition, and so the subject has become popularly known as Pólya theory, though Pólya-Redfield theory would be a better name. In this section we develop the elements of this theory.

Our language will be more intuitive if we think of T as a set of "colors." To illustrate that using this language is not restrictive, the set S might be the positions in a hydrocarbon molecule which are occupied by hydrogen, and the group could be the group of spatial symmetries of the molecule (that is, the group of permutations of the atoms of the molecule that move the molecule around so that in its final position the molecule cannot be distinguished from the original molecule). The colors could then be radicals (including hydrogen itself) that we could substitute for each hydrogen position in the molecule. Then the number of orbits of colorings is the number of chemically different compounds we could create by using these substitutions.[8]

So think intuitively about some "figure" that has places to be colored. (Think of the faces of a cube, the beads on a necklace, circles at the vertices of an n-gon, etc.) How can we picture the coloring? If we number the places to be colored, say 1 to n, then a function from $[n]$ to the colors is exactly our coloring; if our colors are blue, green and red, then $BBGRRGBG$ describes a typical coloring of 8 such places. Unless the places are somehow "naturally" numbered, this idea of a coloring imposes structure that is not really there. Even if the structure is there, visualizing our colorings in this way doesn't "pull together" any common features of different colorings; we are simply visualizing all possible functions. We have a group (think of it as symmetries of the figure you are imagining) that acts on the places. That group then acts in a natural way on the colorings of the places and we are interested in orbits of the colorings. Thus we want a picture that pulls together the common features of the colorings in an orbit. One way to pull together similarities of colorings

---

[8]There is a fascinating subtle issue of what makes two molecules different. For example, suppose we have a molecule in the form of a cube, with one atom at each vertex. If we interchange the top and bottom faces of the cube, each atom is still connected to exactly the same atoms as before. However we cannot achieve this permutation of the vertices by a member of the rotation group of the cube. It could well be that the two versions of the molecule interact with other molecules in different ways, in which case we would consider them chemically different. On the other hand if the two versions interact with other molecules in the same way, we would have no reason to consider them chemically different. This kind of symmetry is an example of what is called *chirality* in chemistry.

would be to let the letters we are using as pictures of colors commute as we did with our pictures in Chapter 4; then our picture $BBGRRGBG$ becomes $B^3G^3R^2$, so our picture now records simply how many times we use each color. If you think about how we defined the action of a group on a set of functions, you will see that a group element won't change how many times each color is used; it simply moves colors to different places. Thus the picture we now have of a given coloring is an equally appropriate picture for each coloring in an orbit. One natural question for us to ask is "How many orbits have a given picture?" We can think of a multivariable generating function in which the letters we use to picture individual colors are the variables, and the coefficient of a picture is the number of orbits with that picture. Such a generating function is an answer to our natural question, and so it is this sort of generating function we will seek. Since the CFB theorem was our primary tool for saying how many orbits we have, it makes sense to think about whether the CFB theorem has an analog in terms of pictures of orbits.

The Orbit-Fixed Point Theorem

Activity 307. Suppose that P_1 and P_2 are picture functions on sets S_1 and S_2 in the sense of [Section](). Define P on $S_1 \times S_2$ by $P(x_1, x_2) = P_1(x_1)P_2(x_2)$. How are E_{P_1}, E_{P_1}, and E_P related? (Hint: you may have already done this problem in another context!)

Solution.

$$E_P(S_1 \times S_2) = \sum_{x_1 \in S_1, \ x_2 \in S_2} P(x_1)P(x_2) = \sum_{x_1 \in S_1} \sum_{x_2 \in S_2} P(x_1)P(x_2)$$

$$= \sum_{x_1 \in S_1} P_1(x_1) \sum_{x_2 \in S_2} P_2(x_2) = E_{P_1}(S_1)E_{P_2}(S_2).$$

Activity 308. Suppose P is a picture function on a set T. Suppose that we define the picture of a function from some other set S to the set T to be the product of the pictures of the values of f, i.e.

$$\hat{P}(f) = \prod_{x:x \in S} P(f(x)).$$

How does the picture enumerator $E_{\hat{P}}$ of the set T^S of all functions from S to T relate to the picture enumerator of P on the set T? (You may assume that both S and T are finite.)

Solution. $E_{\hat{P}}(T^S) = E_P(T)^{|S|}$. To prove this, note that we can think of a function f from S to T as an $|S|$-tuple of values,

$$(f(x_1), f(x_2), \ldots, f(x_{|S|})),$$

then think of an $|S|$-tuple an the ordered pair

$$(f(x_1), (f(x_2), f(x_3), \ldots f(x_{|S|}))),$$

and apply induction and [Problem 307]().

Activity 309. Suppose now we have a group G acting on a set and we have a picture function on that set with the additional feature that for each orbit of the group, all its elements have the same picture. In this circumstance we define the picture of an orbit or multiorbit to be the picture of any one of its members. The *orbit enumerator* $Orb(G, S)$ is the sum of all the pictures of all the orbits. The *fixed point enumerator* $Fix(G, S)$ is the sum of all the pictures of all the fixed points of all the elements of G. We are going to construct a generating function analog of the CFB theorem. The main idea of the proof of the CFB theorem was to try to compute in two different ways the number of elements (i.e. the sum of all the multiplicities of the elements) in the union of all the multiorbits of a group acting on a set. Suppose instead we try to compute the sum of all the pictures of all the elements in the union of the multiorbits of a group acting on a set. By thinking about how this sum relates to $Orb(G, S)$ and $Fix(G, S)$, find an analog of the CFB theorem that relates these two enumerators. State and prove this theorem.

Solution.   Let E, for enumerator, be the sum of all the pictures of all the elements in the union of the multiorbits of G acting on a set S. Recall that for any multiorbit M the picture $P(M)$ is the picture $P(x)$ of any element x of M, and the number of elements of a multiorbit M is always the size of G. This lets us write

$$E = \sum_{M: M \text{ is a multiorbit of } G} \ \sum_{x: x \in M} P(x)$$

$$= \sum_{M: M \text{ is a multiorbit of } G} |G| P(M)$$

$$= |G| \sum_{M: M \text{ is a multiorbit of } G} P(M)$$

$$= |G| Orb(G, S)$$

Recall also that the multiplicity of an element x in its multiorbit, and thus in the union of the multiorbits is $|Fix(x)|$. This lets us write

$$E = \sum_{x: x \in S} |Fix(x)| P(x)$$

$$= \sum_{x: x \in S} \sum_{\sigma: \sigma \in G} \chi(\sigma, x) P(x)$$

$$= \sum_{\sigma: \sigma \in G} \sum_{x: \sigma x = x} P(x)$$

$$= Fix(G, S).$$

Setting these two values of E equal and solving for $Orb(G, S)$ gives us

$$Orb(G, S) = \frac{1}{|G|} Fix(G, S).$$

Activity 310.

(a) We will call the theorem of Problem 309 the *Orbit-Fixed Point Theorem*. Use it to determine the Orbit Enumerator for the colorings, with two colors (red and blue), of six circles placed at the vertices of a hexagon which is free to move in the plane. Compare the coefficients of the resulting polynomial with the various orbits you found in Problem 280.

Solution. Let us take the pictures of red and blue to be R and B. Since the hexagon is free to move in the plane, our group G is the group R_6 of rotations of a regular hexagon. There are four kinds of elements of G: the identity, the rotation ρ through 60 degrees and the rotation ρ^5 through 300 degrees (when written as a product of disjoint cycles, the corresponding permutations are six-cycles), the rotations ρ^2 and ρ^4 through 120 and 240 degrees respectively (when written as a product of disjoint cycles, the corresponding permutations are products of two three cycles), and the rotation ρ^3 through 180 degrees (when written as a product of disjoint cycles this is a product of three two-cycles). All 2^6 functions are fixed by the identity. By Problem 308 the enumerator for these functions is $(R+B)^6$. Only the two constant functions (that color every circle the same color) are fixed by ρ or ρ^5. The picture enumerator of the two constant functions is $R^6 + B^6$. If the vertices are numbered one through six clockwise, then ρ^2 is $(1\,3\,5)(2\,4\,6)$. Thus for a function f to be fixed by ρ^2, $f(1) = f(3) = f(5)$ and $f(2) = f(4) = f(6)$. Thus the color of vertex 1 is repeated on vertex 3 and 5, and the color of vertex 2 is repeated on vertices 4 and 6. Therefore the picture enumerator for functions fixed by ρ^2 and ρ^4 is $R^3R^3 + R^3B^3 + B^3R^3 + B^3B^3 = (R^3 + B^3)^2$. We can also think of this as the picture enumerator for functions defined on the cycles of the permutation. There are two possible pictures of a colored cycle, R^3 and B^3, so the picture enumerator for one cycle is $R^3 + B^3$. That is, the picture of assigning red to a cycle is R^3 and the picture of assigning blue to a cycle is B^3. By Problem 308, the picture enumerator for functions defined on the two cycles is then $(R^3 + B^3)^2$. If a function f is fixed by $\rho^3 = (1\,4)(2\,5)(3\,6)$, then $f(1) = f(4)$, $f(2) = f(5)$, and $f(3) = f(6)$, so f is determined by the three values $f(1)$, $f(2)$, and $f(3)$. The picture enumerator for these functions is, by Problem 308 applied to colorings of the cycles, $(R^2 + B^2)^3$. Therefore the fixed point enumerator for the action of G on the colorings is

$$Fix(G, S) = (R + B)^6 + (R^2 + B^2)^3 + 2(R^3 + B^3) + 2(R^6 + B^6).$$

Then the orbit enumerator for the action of G on the colorings is

$$Orb(G, S) = \frac{1}{6}\left((R + B)^6 + (R^2 + B^2)^3 + 2(R^3 + B^3)^2 + 2(R^6 + B^6)\right).$$

Expanding this gives us

$$R^6 + R^5B + 3R^4B^2 + 4R^3B^3 + 3R^2B^4 + RB5 + B^6.$$

Thus we have one all-red orbit, one all-blue orbit, one orbit with five reds and a blue, one with five blues and a red, three orbits with four reds and two blues as well as three with two reds and four blues, and four orbits with three reds and three blues.

(b) Find the generating function (in variables R, B) for colorings of the faces of a cube with two colors (red and blue). What does the generating function tell you about the number of ways to color the cube (up to spatial movement) with various combinations of the two colors.

Solution.   We want to think of the rotation group of the cube acting on the faces of the cube, in order to see what kinds of colorings are left fixed. For this purpose we note that each element of the rotation group gives us a permutation of the faces of the cube, and if two faces are in the same cycle of this permutation, they must have the same color, but if they are in different cycles, they may get different colors. The 90 and 270 degree rotations around an axis through two faces have a four-cycle and two one-cycles. The 180 degree rotation around an axis through two faces has two two-cycles and two one-cycles. The 120 and 240 degree rotations around an axis connecting diagonally opposite vertices have two three-cycles. The 180 degree rotations around an axis connecting two opposite edges have three two-cycles, and the identity has six one-cycles. A coloring is fixed by a permutation if and only if it is constant on (i.e. assigns the same color to all elements of) each cycle of the permutation. The picture enumerator for (constant) colorings of an i-cycle is $R^i + B^i$. Thus, using Problem 308, if σ is a 90 or 270 degree rotation, its fixed point enumerator is $(R^4 + B^4)(R + B)^2$, if it is a 180 degree rotation around an axis connecting two opposite faces, its fixed point enumerator is $(R^2 + B^2)^2(R + B)^2$, if it is a 180 degree rotation around an axis connecting two opposite edges, its fixed point enumerator is $(R^2 + B^2)^3$, if it is a 120 or 240 degree rotation its fixed point enumerator is $(R^3 + B^3)^2$, and if it is the identity, then its fixed point enumerator is $(R + B)^6$. Therefore the generating function is

$$\frac{1}{24}\left((R + B)^6 + 8(R^3 + B^3)^2 + 3(R^2 + B^2)^2(R + B)^2 + \right.$$
$$\left. 6(R^2 + B^2)^3 + 6(R^4 + B^4)(R + B)^2\right),$$

which expands to

$$RB^5 + 2R^4B^2 + R^5B + 2R^3B^3 + 2R^2B^4 + R^6 + B^6.$$

There is one way to color the cube all red or all blue, one way to color it with exactly five red or exactly five blue faces, there are two ways to color it with exactly four red or four blue faces and two ways to color it with exactly three red (and three blue) faces.

The Pólya-Redfield Theorem

Pólya's (and Redfield's) famed enumeration theorem deals with situations such as those in Problems 310 and ⟨⟨Unresolved xref, reference "polya2"; check spelling or use "provisional" attribute⟩⟩ in which we want a generating function for the set of all functions from a set S to a set T on which a picture function is defined, and the picture of a function is the product of the pictures of its multiset of values. The point of the next series of problems is to analyze the solution to Problems 310 and ⟨⟨Unresolved xref, reference "polya2"; check spelling or use "provisional" attribute⟩⟩ in order to see what Pólya and Redfield saw (though they didn't see it in this notation or using this terminology).

Activity 311. In Problem 310 we have four kinds of group elements: the identity (which fixes every coloring), the rotations through 60 or 300 degrees, the rotations through 120 and 240 degrees, and the rotation through 180 degrees. The fixed point enumerator for the rotation group acting on the functions is by definition the sum of the fixed point enumerators of colorings fixed by the identity, of colorings fixed by 60 or 300 degree rotations, of colorings fixed by 120 or 240 degree rotations, and of colorings fixed by the 180 degree rotation. Write down each of these enumerators (one for each kind of permutation) individually and factor each one (over the integers) as completely as you can.

Solution. In the solution to Problem 310 we wrote: "The picture enumerator of the two constant functions is $R^6 + B^6$. If the vertices are numbered one through six clockwise, then ρ^2 is (1 3 5)(2 4 6). Thus for a function f to be fixed by ρ^2, $f(1) = f(3) = f(5)$ and $f(2) = f(4) = f(6)$. Thus the color of vertex 1 is repeated on vertex 3 and 5, and the color of vertex 2 is repeated on vertices 4 and 6. Therefore the picture enumerator for functions fixed by ρ^2 and ρ^4 is $R^3R^3 + R^3B^3 + B^3R^3 + B^3B^3 = (R^3 + B^3)^2$. We can also think of this as the picture enumerator for functions defined on the cycles of the permutation. There are two possible pictures of a colored cycle, R^3 and B^3, so the picture enumerator for one cycle is $R^3 + B^3$. Thus the picture of assigning red to a cycle is R^3 and the picture of assigning blue to a cycle is B^3. By Problem 308, the picture enumerator for functions defined on the two cycles is then $(R^3 + B^3)^2$. If a function f is fixed by $\rho^3 = $ (1 4)(2 5)(3 6), then $f(1) = f(4)$, $f(2) = f(5)$, and $f(3) = f(6)$, so f is determined by the three values $f(1)$, $f(2)$, and $f(3)$. The picture enumerator for these functions is, by Problem 308 applied to colorings of the cycles, $(R^2 + B^2)^3$." In factored form, these enumerators are $R^6 + B^6$, $(R + B)^6$, $(R^3 + B^3)^2$, and $R^2 + B^2)^3$.

Activity 312. In ⟨⟨Unresolved xref, reference "polya2"; check spelling or use "provisional" attribute⟩⟩ we have five different kinds of group elements, and the fixed point enumerator is the sum of the fixed point enumerators of each of these kinds of group elements. For each kind of element, write down the fixed point enumerator for the elements of that kind. Factor the enumerators as completely as you can.

Solution.   In the solution to ⟨⟨Unresolved xref, reference "polya2"; check spelling or use "provisional" attribute⟩⟩, we wrote "Thus, using Problem 308, if σ is a 90 or 270 degree rotation, its fixed point enumerator is $(R^4+B^4)(R+B)^2$, if it is a 180 degree rotation around an axis connecting two opposite faces, its fixed point enumerator is $(R^2 + B^2)^2(R + B)^2$, if it is a 180 degree rotation around an axis connecting two opposite edges, its fixed point enumerator is $(R^2 + B^2)^3$, if it is a 120 or 240 degree rotation its fixed point enumerator is $(R^3 + B^3)^2$, and if it is the identity, then its fixed point enumerator is $(R + B)^6$." We just wrote the enumerators out in factored form.

Activity 313. In Problem 311, each "kind" of group element has a "kind" of cycle structure. For example, a rotation through 180 degrees has three cycles of size two. What kind of cycle structure does a rotation through 60 or 300 degrees have? What kind of cycle structure does a rotation through 120 or 240 degrees have? Discuss the relationship between the cycle structures and the factored enumerators of fixed points of the permutations in Problem 311.

Solution.   A rotation through 60 or 300 degrees is a five-cycle; a rotation through 120 or 240 degrees is a product of two three-cycles. The cycle structure determines the factored enumerator; a cycle of size i gives a factor of $(R^i + B^i)$. That is because a coloring fixed by a group element has to be constant on the cycles of that group element. If a cycle has size i, it contributes a summand of P^i to the picture enumerator for colorings of that cycle for each picture P of a possible value of the function. Problem 308 tells us to multiply these individual picture enumerators together.

Recall that we said that a group of permutations acts on a set if, for each member σ of G there is a bijection β_σ of S such that

$$\beta_{\sigma \circ \varphi} = \beta_\sigma \circ \beta_\varphi$$

for every member σ and φ of G. Since β_σ is a bijection of S to itself, it is in fact a permutation of S. Thus β_σ has a cycle structure (that is, it is a product of disjoint cycles) as a permutation of S (in addition to whatever its cycle structure is in the original permutation group G).

Activity 314.

Activity 315. The usual way of describing the Pólya-Redfield enumeration theorem involves the "cycle indicator" or "cycle index" of a group acting on a set. Suppose we have a group G acting on a finite set S. Since each group element σ gives us a permutation β_σ of S, as such it has a decomposition into disjoint cycles as a permutation of S. Suppose σ has c_1 cycles of size 1, c_2 cycles of size 2, ..., c_n cycles of size n. Then the *cycle monomial* of σ is

$$z(\sigma) = z_1^{c_1} z_2^{c_2} \cdots z_n^{c_n}.$$

The *cycle indicator* or *cycle index* of G acting on S is

$$Z(G, S) = \frac{1}{|G|} \sum_{\sigma : \sigma \in G} z(\sigma).$$

What is the cycle index for the group D_6 acting on the vertices of a hexagon? What is the cycle index for the group of rotations of the cube acting on the faces of the cube?

Solution. For D_6, we get

$$\frac{1}{12}\left(z_1^6 + 2z_6 + 2z_3^2 + z_2^3 + 3z_2^3 + 3z_2^2z_1^2\right) =$$

$$\frac{1}{12}\left(z_1^6 + 2z_6 + 2z_3^2 + 4z_2^3 + 3z_1^2z_2^2\right)$$

For the rotation group of the cube, we get

$$\frac{1}{24}\left(3z_1^2z_2^2 + 6z_1^2z_4 + 8z_3^2 + 6z_3^2 + z_1^6\right) =$$

$$\frac{1}{24}\left(3z_1^2z_2^2 + 6z_1^2z_4 + 14z_3^2 + z_1^6\right)$$

Activity 316. How can you compute the Orbit Enumerator of G acting on functions from S to a finite set T from the cycle index of G acting on S? (Use $P(t)$ as the notation for the picture of an element t of T.) State and prove the relevant theorem! This is Pólya's and Redfield's famous enumeration theorem.

Solution. To compute the orbit enumerator of G acting on functions from S to a finite set T, we substitute $\sum_{t:t\in T} P(t)^i$ for z_i in the cycle index for G acting on S. (This is the Pólya-Redfield Theorem.) By the Orbit-Fixed Point Theorem, we need to sum the fixed point enumerators of the permutations in G. However if σ fixes f, then f is constant on the cycles of σ. Thus we may think of our function f as a function defined on the cycles of σ. If an i-cycle is sent to the member t of T, then the picture of the restriction of our function to that cycle is $P(t)^i$. The picture enumerator for all functions defined and constant on that cycle is then $\sum_{t:t\in T} P(t)^i$. By [Problem 308](#) the picture enumerator for all functions constant on the cycles of σ, then is the result of substituting $\sum_{t:t\in T} P(t)^i$ for z_i in $z(\sigma)$. This proves the Pólya-Redfield Theorem.

Activity 317. Suppose we make a necklace by stringing 12 pieces of brightly colored plastic tubing onto a string and fastening the ends of the string together. We have ample supplies blue, green, red, and yellow tubing available. Give a generating function in which the coefficient of $B^iG^jR^kY^h$ is the number of necklaces we can make with i blues, j greens, k reds, and h yellows. How many terms would this generating function have if you expanded it in terms of powers of B, G, R, and Y? Does it make sense to do this expansion? How many of these necklaces have 3 blues, 3 greens, 2 reds, and 4 yellows?

Solution. We are asking for a generating function for the orbits of a colored 12-gon under the action of D_{12}. To apply the Pólya-Redfield theorem we need the cycle index for D_{12}. If ρ is a 30 degree rotation, then ρ, ρ^5, ρ^7 and ρ^{11} are 12-cycles. The elements ρ^2 and ρ^{10} are products of two six-cycles. The elements ρ^3 and ρ^9 are products of three four-cycles. The elements ρ^4 and

ρ^8 are products of four three-cycles. The element ρ^6 is a product of six two-cycles. The element $\iota = \rho^0$ is a product of 12 one-cycles. There are six flips around axes through opposite vertices; each is a product of five two-cycles and two one-cycles. There are six flips around axes perpendicular to two opposite sides; each is a product of six two-cycles. Summarizing this in the cycle index, we write

$$Z(G, S) = \frac{1}{24} \left(z_1^{12} + z_2^6 + 2z_3^4 + 2z_4^3 + 2z_6^2 + 4z_{12} + 6z_2^5 z_1^2 + 6z_2^6 \right)$$

$$= \frac{1}{24} \left(z_1^{12} + 7z_2^6 + 2z_3^4 + 2z_4^3 + 2z_6^2 + 4z_{12} + 6z_2^5 z_1^2 \right).$$

When we substitute $B^i + G^i + R^i + Y^i$ for z_i and expand, we would get get 12^4 terms, one for each possible term $B^i G^j R^k Y^h$. Thus it does not make sense to expand the polynomial. The unexpanded form is

$\frac{1}{24} \Big((B+G+R+Y)^{12} + 7(B^2+G^2+R^2+Y^2)^6 + 2(B^3+G^3+R^3+Y^3)^4 +$
$2(B^4 + G^4 + R^4 + Y^4)^3 + 2(B^6 + G^6 + R^6 + Y^6)^2 +$
$4(B^{12} + G^{12} + R^{12} + Y^{12}) + 6((B^2 + G^2 + R^2 + Y^2)^5)(B + G + R + Y)^2 \Big)$

We can compute the coefficient of $B^3 G^3 R^2 Y^4$ by computing the contribution of each term of the sum to the coefficient. We get

$$\frac{1}{24} \left(\binom{12}{3, 3, 2, 4} + 6 \binom{5}{1, 1, 1, 2} \binom{2}{1, 1} \right) = \frac{12!}{24 \cdot 3!3!2!4!} + \frac{6 \cdot 5!2!}{24 \cdot 2!}$$

$$= \frac{11!}{2 \cdot 3!3!2!4!} + 5 \cdot 3 \cdot 2$$

$$= 11 \cdot 5 \cdot 3 \cdot 2 \cdot 7 \cdot 5 + 30$$

$$= 11,550 + 30 = 11,580$$

Activity 318. What should we substitute for the pictures of each of the elements of T in the orbit enumerator of G acting on the set of functions from S to T in order to compute the total number of orbits of G acting on the set of functions? What should we substitute into the variables in the cycle index of a group G acting on a set S in order to compute the total number of orbits of G acting on the functions from S to a set T? Find the number of ways to color the faces of a cube with four colors.

Solution.   Substitute the number one for each picture in the picture enumerator. Substitute $|T|$ for each variable in the cycle index. The cycle index for the rotation group of the cube acting on the faces is

$$\frac{1}{24} \left(3z_1^2 z_2^2 + 6z_1^2 z_4 + 14z_3^2 + z_1^6 \right)$$

Substituting 4 for each variable gives us $\frac{1}{24} \left(3 \cdot 4^4 + 6 \cdot 4^3 + 14 \cdot 4^2 + 4^6 \right)$
$= 228$.

Activity 319. We have red, green, and blue sticks all of the same length, with a dozen sticks of each color. We are going to make the skeleton of a cube

by taking eight identical lumps of modeling clay and pushing three sticks into each lump so that the lumps become the vertices of the cube. (Clearly we won't need all the sticks!) In how many different ways could we make our cube? How many cubes have four edges of each color? How many have two red, four green, and six blue edges? For this problem we are interested in the action of the rotation group of the cube on the edges. Now we think of the group elements as permutations of the edges and analyze their cycle structure.

(a) The identity is a product of 12 one-cycles.

(b) A 90 or 270 degree rotation around an axis perpendicular to two opposite faces is a product of three four-cycles.

(c) A 180 degree rotation around an axis perpendicular to two opposite faces is a product of six two-cycles.

(d) A 180 degree rotation around an axis perpendicular to two opposite edges is a product of five two-cycles and two one-cycles.

(e) A 120 degree rotation around an axis connecting two diagonally opposite vertices is a product of four three-cycles.

Thus the cycle index is

$$\frac{1}{24}\left(z_1^{12} + 6z_4^3 + 3z_2^6 + 6z_2^5 z_1^2 + 8z_3^4\right).$$

We substitute the number three for each of the variables to get

$$\frac{1}{24}\left(3^{12} + 6 \cdot 3^3 + 3 \cdot 3^6 + 6 \cdot 3^7 + 8 \cdot 3^4\right) = 22815$$

ways to make the cube. To compute the number of ways with four sticks of each color, we need to apply the Pólya-Redfield theorem. Substituting $R^i + B^i + G^i$ for z_i in the cycle index gives us

$$\frac{1}{24}\Big((R+B+G)^{12} + 6((R^4 + B^4 + G^4)^3 + 3(R^2 + B^2 + G^2)^6 +$$

$$6(R^2 + B^2 + B^2)^5 (R+B+G)^2 + 8(R^3 + B^3 + G^3)^4\Big).$$

The coefficient of $R^4 B^4 G^4$ is

$$\frac{1}{24}\left(\binom{12}{4,4,4} + 6\binom{3}{1,1,1} + 3\binom{6}{2,2,2} + 6\binom{3}{1}\binom{5}{2,2,1}\binom{2}{2,0,0}\right) =$$

$$\frac{1}{24}\left(\frac{12!}{4!4!4!} + 6 \cdot 3! + 3\frac{6!}{2!2!2!} + 6 \cdot 3\frac{5!}{2!2!1!}\right) = 1479$$

The coefficient of $R^2 B^4 G^6$ is

$$\frac{1}{24}\left(\binom{12}{2,4,6} + 3\binom{6}{1,2,3} + 6\left(\binom{5}{0,2,3} + \binom{5}{1,1,3} + \binom{5}{1,2,2}\right)\right) =$$
$$\frac{1}{24}\left(\frac{12!}{2!4!6!} + 3\frac{6!}{1!2!3!} + 6\left(\frac{5!}{2!3!} + \frac{5!}{1!1!3!} + \frac{5!}{1!2!2!}\right)\right) = 600.$$

Activity 320. How many cubes can we make in Problem 319 if the lumps of modelling clay can be any of four colors? We can first consider all ways of coloring the vertices of the cube with the four colors; once those vertices are in place, the number of ways to put the sticks in is the result of Problem 319, so by the product principle then number of ways to choose the colors of the vertices and edges is the product of the number of ways to choose each. Thus we just need to consider the action of the rotation group of the cube on the vertices:

-+ identity is the product of eight one-cycles.

-+ 90 or 270 degree rotation around an axis perpendicular to two opposite faces is a product of two four-cycles.

-+ 180 degree rotation around an axis perpendicular to two opposite faces is a product of four two-cycles.

-+ 180 degree rotation around an axis joining two opposite edges is a product of four two-cycles.

-+ 120 or 240 degree rotation around an axis through two diagonally opposite vertices is a product of two three-cycles and two one-cycles.

Thus the cycle index for this action is

$$\frac{1}{24}\left(z_1^8 + 6z_4^2 + 9z_2^4 + 8z_3^2 z_1^2\right).$$

Substituting 4 for each variable gives us

$$\frac{1}{24}\left(4^8 + 6 \cdot 4^2 + 9 \cdot 4^4 + 8 \cdot 4^2 \cdot 4^2\right) = 2916.$$

Thus we have $22815 \cdot 2916 = 66,528,540$ possible colored cubes.
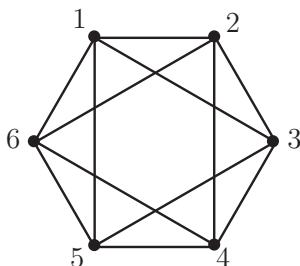
Figure 6.3.1: A possible computer network.

Activity 321. In Figure 6.3.1 we see a graph with six vertices. Suppose we have three different kinds of computers that can be placed at the six vertices of

the graph to form a network. In how many different ways may the computers be placed? (Two graphs are not different if we can redraw one of the graphs so that it is identical to the other one.) This is equivalent to permuting the vertices in some way so that when we apply the permutation to the endpoints of the edges to get a new edge set, the new edge set is equal to the old one. Such a permutation is called an *automorphism* of the graph. Then two computer placements are the same if there is an automorphism of the graph that carries one to the other.

Solution. The computer placements are functions from the vertices of the graph to the set of three kinds of computers, say $\{A, B, C\}$. Thus we are asking for the number of orbits of the automorphism group of the graph on functions from the vertices of $\{A, B, C\}$. To find this number we need to compute the cycle index of the automorphism group. An automorphism will send vertices 1, 2, and 3 to three vertices that are mutually connected; i.e. a triangle. Further, each of vertices 4, 5, and 6 is adjacent to exactly two of vertices of 1, 2, and 3. In fact, for any of the eight triangles in the graph, each vertex not in the triangle is adjacent to exactly two vertices of the triangle (and a different two for each vertex). Therefore we can send the vertices 1, 2, and 3, to any of the eight triangles in any of six orders, and this completely determines an automorphism. Thus there are $6 \cdot 8 = 48$ elements in the group of automorphisms. The dihedral group D_6 is a subgroup of the group of automorphisms. The two-cycles (1 4), (2 5), and (3 6) are also in the group of automorphisms. (For example, 1 is adjacent to everything but 4, and 4 is adjacent to everything but 1, so interchanging them leaves us with exactly the same edges.) We will write out D_6 in disjoint cycle notation, and then the coset (1 4)D_6, and from those we will be able to get the cycle index of the automorphism group acting on the vertices. In cycle notation, D_6 is

$$
\begin{array}{cccc}
(1\ 2\ 3\ 4\ 5\ 6) & (1\ 3\ 5)(2\ 4\ 6) & (1\ 4)(2\ 5)(3\ 6) & (1\ 5\ 3)(2\ 6\ 4) \\
(1\ 6\ 5\ 4\ 3\ 2) & (1)(2)(3)(4)(5)(6) & (1)(4)(2\ 6)(3\ 5) & (2)(5)(1\ 3)(4\ 6) \\
(3)(6)(1\ 5)(2\ 4) & (1\ 2)(3\ 6)(4\ 5) & (1\ 6)(2\ 5)(3\ 4) & (1\ 4)(2\ 3)(5\ 6)
\end{array}
$$

and (1 4)D_6 is

$$
\begin{array}{cccc}
(1\ 2\ 3)(4\ 5\ 6) & (1\ 3\ 5\ 4\ 6\ 2) & (1)(4)(2\ 5)(3\ 6) & (1\ 5\ 3\ 4\ 2\ 6) \\
(1\ 6\ 5)(2\ 4\ 3) & (1\ 4)(2)(3)(5)(6) & (1\ 4)(2\ 6)(3\ 5) & (2)(5)(1\ 3\ 4\ 6) \\
(3)(6)(1\ 5\ 4\ 2) & (1\ 2\ 4\ 5)(3\ 6) & (1\ 6\ 4\ 3)(2\ 5) & (1)(4)(2\ 3)(5\ 6).
\end{array}
$$

Notice that neither (2 5) nor (3 6) is in the coset, so our group is the union

$$D_6 \cup (1\ 4)D_6 \cup (2\ 5)D_6 \cup (3\ 6)D_6.$$

But by symmetry, the cycle structure of (1 4)D_6, (2 5)D_6 and (3 6)D_6 will be the same, so the cycle index for our group is.

$$\frac{1}{48}\left(Z_1^6 + 8Z_3^2 + 8z_6^1 + 7z_2^3 + 9z_1^2 z_2^2 + 3z_1^4 z_2 + 6z_2 z_4 + 6z_1^2 z_4\right).$$

Since we have three kinds of computers, we substitute 3 for each variable to get

$$\frac{1}{48} \left(3^6 + 8 \cdot 3^2 + 8 \cdot 3^1 + 7 \cdot 3^3 + 9 \cdot 3^4 + 3 \cdot 3^5 + 6 \cdot 3^2 + 6 \cdot 3^3 \right) = 56.$$

Activity 322. Two simple graphs on the set $[n] = \{1, 2, \ldots, n\}$ with edge sets E and E' (which we think of a sets of two-element sets for this problem) are said to be *isomorphic* if there is a permutation σ of $[n]$ which, in its action of two-element sets, carries E to E'. We say two graphs are different if they are not isomorphic. Thus the number of different graphs is the number of orbits of the set of all two-element subsets of $[n]$ under the action of the group S_n. We can represent an edge set by its characteristic function (as in problem 33). That is we define

$$\chi_E(\{u, v\}) = \begin{cases} 1 & \text{if } \{u, v\} \in E \\ 0 & \text{otherwise.} \end{cases}$$

Thus we can think of the set of graphs as a set of functions defined on the set of all two-element subsets of $[n]$. The number of different graphs with vertex set $[n]$ is thus the number of orbits of this set of characteristic functions under the action of the symmetric group S_n on the set of two-element subsets of $[n]$. Use this to find the number of different graphs on five vertices.

Solution.   For this problem we need the cycle index for the action of the symmetric group on five letters acting on two-element subsets of those five letters. Each way of partitioning the number five describes the cycle structure of an element of S_5 acting on $[5]$. The cycle structure of a permutation σ on the two-element subsets of $[5]$ will be determined by its cycle structure on $[5]$. The partitions of five and the cycle structures they give on two-element subsets are:

- $(1, 1, 1, 1, 1) = 1^5$ is the cycle structure of the identity; acting on two-element subsets the identity is a product of $\binom{5}{2} = 10$ one-cycles. Of course there is only one identity permutation.

- A permutation with cycle structure $2^1 1^3$ will fix the two element set in the two cycle and each pair of the three elements outside for a total of four one-cycles; each pair of an element in the two cycle and an element not will be in a two cycle. There are 6 such pairs and thus three two-cycles, so such a permutation has 3 two-cycles and four one-cycles. There are $\binom{5}{2} = 10$ such permutations.

- A permutation with cycle structure $2^2 1^1$ will have a one-cycle of two-sets for each two cycle, any other two-set will be in a two cycle. There are $4 + 2 + 2 = 8$ such two-sets and thus four two-cycles, so such a permutation has four two cycles and two one cycles. There are $\binom{5}{2}\binom{3}{2}/2 = 15$ such permutations.

- A permutation with cycle structure $3^1 1^2$ will have one cycle of size 1 on two-sets from the two one-element cycles of the original action; each other two-element subset will be in a three cycle. There are $\binom{3}{2} + 3 \cdot 2 = 9$ such pairs, and so there are three three-cycles of two-element subsets. Thus such a permutation has one one-cycle and three three-cycles in its action of two-sets. There are $2\binom{5}{3} = 20$ such permutations.

- A permutation with cycle structure $3^1 2^1$ will have one one-cycle on two-sets from the two-cycle of the original action, it will have $\binom{3}{2}$ two-sets in three cycles, and will have six two-sets in six-cycles, and so will have just one six-cycle of two-sets. Thus such a permutation has one one-cycle, one three-cycle and one six-cycle when it acts on two-sets. There are $2\binom{5}{3} = 20$ such permutations.

- A permutation with cycle structure $4^1 1^1$ has a four-cycle of two-subsets and one two-cycle of two-subsets from its own four-cycle and another four cycle of pairs of members of the four-cycle and one cycle. thus it has two four-cycles and two one-cycles There are $6\binom{5}{4} = 30$ such permutations.

- A permutation with cycle structure 5^1 will have two five-cycles in its action on two-sets. There are $4! = 24$ such permutations.

Thus for the group S_5 acting on pairs from $[5]$, the cycle index is

$$\frac{1}{120}\left(z_1^{10} + 10z_2^3 z_1^4 + 15z_2^4 z_2 + 20z_3^3 z_1 + 20z_6 z_3 z_1 + 30z_4^2 z_2 + 24z_5^2\right).$$

Substituting 2 for each variable gives us that there are

$$\frac{1}{120}\left(2^{10} + 10 \cdot 2^7 + 15 \cdot 2^5 + 20 \cdot 2^4 + 20 \cdot 2^3 + 30 \cdot 2^3 + 24 \cdot 2^2\right) = 30$$

graphs on five vertices.

6.4 Supplementary Problems

1. Show that a function from S to T has an inverse (defined on T) if and only if it is a bijection.

2. How many elements are in the dihedral group D_3? The symmetric group S_3? What can you conclude about D_3 and S_3?

3. A tetrahedron is a thee dimensional geometric figure with four vertices, six edges, and four triangular faces. Suppose we start with a tetrahedron in space and consider the set of all permutations of the vertices of the tetrahedron that correspond to moving the tetrahedron in space and returning it to its original location, perhaps with the vertices in different places. Explain why these permutations form a group. What is the size of this group? Write down in two-row notation a permutation that is *not* in this group.

4. Find a three-element subgroup of the group S_3. Can you find a different three-element subgroup of S_3?

5. Prove true or demonstrate false with a counterexample: "In a permutation group, $(\sigma\varphi)^n = \sigma^n\varphi^n$."

6. Describe a permutation group with 60 elements.

7. If a group G acts on a set S, and if $\sigma(x) = y$, is there anything interesting we can say about the subgroups $Fix(x)$ and $Fix(y)$?

8. Find the number of ways to color the faces of a tetrahedron with two colors.

9. Find the number of ways to color the faces of a tetrahedron with four colors so that each color is used.

10. Find the cycle index of the group of spatial symmetries of the tetrahedron acting on the vertices. Find the cycle index for the same group acting on the faces.

11. Find the generating function for the number of ways to color the faces of the tetrahedron with red, blue, green and yellow.

12. (interesting) Find the generating function for the number of ways to color the faces of a cube with four colors so that all four colors are used.

13. (interesting) How many different graphs are there on six vertices with seven edges?

Appendix A

Relations

A.1 Relations as sets of Ordered Pairs

The relation of a function

Activity 323. Consider the functions from $S = \{-2, -1, 0, 1, 2\}$ to $T = \{1, 2, 3, 4, 5\}$ defined by $f(x) = x + 3$, and $g(x) = x^5 - 5x^3 + 5x + 3$. Write down the set of ordered pairs $(x, f(x))$ for $x \in S$ and the set of ordered pairs $(x, g(x))$ for $x \in S$. Are the two functions the same or different?

Solution. We get $\{(-2, 1), (-1, 2), (0, 3), (1, 4), (2, 5)\}$ in both cases, so the functions are the same.

Problem 323 points out how two functions which appear to be different are actually the same on some domain of interest to us. Most of the time when we are thinking about functions it is fine to think of a function casually as a relationship between two sets. In Problem 323 the set of ordered pairs you wrote down for each function is called the *relation* of the function. When we want to distinguish between the casual and the careful in talking about relationships, our casual term will be "relationship" and our careful term will be "relation." So *relation* is a technical word in mathematics, and as such it has a technical definition. A *relation* from a set S to a set T is a set of ordered pairs whose first elements are in S and whose second elements are in T. Another way to say this is that a *relation* from S to T is a subset of $S \times T$.

A typical way to define a *function* f from a set S to a set T is that f is a relationship between S to T that relates one and only one member of T to each element of X. We use $f(x)$ to stand for the element of T that is related to the element x of S. If we wanted to make our definition more precise, we could substitute the word "relation" for the word "relationship" and we would have a more precise definition. For our purposes, you can choose whichever definition you prefer. However, in any case, there is a relation associated with each function. As we said above, the relation of a function $f : S \to T$ (which is the standard shorthand for "f is a function from S to T" and is usually

read as *f maps S to T*) is the set of all ordered pairs $(x, f(x))$ such that x is in S.

Activity 324.

(a) Here are some questions that will help you get used to the formal idea of a relation and the related formal idea of a function. S will stand for a set of size s and T will stand for a set of size t.

   (a) What is the size of the largest relation from S to T?

   Solution. st because that is the size of the relation that has all the ordered pairs (x, y) with $x \in S$ and $y \in T$.

   (b) What is the size of the smallest relation from S to T?

   Solution.   0, because the empty set of ordered pairs is a relation.

   (c) The relation of a function $f : S \to T$ is the set of all ordered pairs $(x, f(x))$ with $x \in S$. What is the size of the relation of a function from S to T? That is, how many ordered pairs are in the relation of a function from S to T?

   Solution.   The size of the relation of $f : S \to T$ is s.

   (d) We say f is a *one-to-one* function or *injection* from S to T if each member of S is related to a *different* element of T. How many different elements must appear as second elements of the ordered pairs in the relation of a one-to-one function (injection) from S to T?

   Solution. s different elements must appear, one for each element of S.

   (e) A function $f : S \to T$ is called an *onto function* or *surjection* if each element of T is $f(x)$ for some $x \in S$ What is the minimum size that S can have if there is a surjection from S to T?

   Solution.   In order to have a surjection from S to T, the size of S must be at least t.

(b) When f is a function from S to T, the sets S and T play a big role in determining whether a function is one-to-one or onto (as defined in Problem 324). For example, if S and T are both the nonnegative real numbers, and $f : S \to T$ is given by $f(x) = x^2$, is f one-to-one? Is f onto? Now assume S' is the set of all real numbers and $g : S' \to T$ is given by $g(x) = x^2$. Is g one-to-one? Is g onto? Assume that T' is

the set of all real numbers and $h : S \to T'$ is given by $h(x) = x^2$. Is h one-to-one? Is h onto? And if the function $j : S' \to T'$ is given by $j(x) = x^2$, is j one-to-one? Is j onto?

Solution. With S as domain, f is both one-to-one and onto. The function g is not one-to-one, but it is onto. The function h is one-to-one but not onto. The function j is neither one-to-one nor onto.

(c) If $f : S \to T$ is a function, we say that f *maps* x to y as another way to say that $f(x) = y$. Suppose $S = T = \{1, 2, 3\}$. Give a function from S to T that is not onto. Notice that two different members of S have mapped to the same element of T. Thus when we say that f associates one and only one element of T to each element of S, it is quite possible that the one and only one element $f(1)$ that f maps 1 to is exactly the same as the one and only one element $f(2)$ that f maps 2 to.

Solution. The function given by $f(x) = 1$ for all x in S is not onto.

Directed graphs

We visualize numerical functions like $f(x) = x^2$ with their graphs in Cartesian coordinate systems. We will call these kinds of graphs *coordinate graphs* to distinguish them from other kinds of graphs used to visualize relations that are non-numerical.
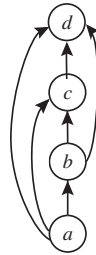
Figure A.1.1: The alphabet digraph.

In Figure A.1.1 we illustrate another kind of graph, a "directed graph" or "digraph" of the "comes before in alphabetical order" relation on the letters a, b, c, and d. To draw a *directed graph* of a relation on a set S, we draw a circle (or dot, if we prefer), which we call a *vertex*, for each element of the set, we usually label the vertex with the set element it corresponds to, and we draw an arrow from the vertex for a to that for b if a is related to b, that is, if the ordered pair (a, b) is in our relation. We call such an arrow an *edge* or a *directed edge*. We draw the arrow from a to b, for example, because a comes before b in alphabetical order. We try to choose the locations where we draw our vertices so that the arrows capture what we are trying to illustrate as

well as possible. Sometimes this entails redrawing our directed graph several times until we think the arrows capture the relationship well.

   We also draw digraphs for relations from a set S to a set T; we simply draw vertices for the elements of S (usually in a row) and vertices for the elements of T (usually in a parallel row) draw an arrow from x in S to y in T if x is related to y. Notice that instead of referring to the vertex representing x, we simply referred to x. This is a common shorthand. Here are some exercises just to practice drawing digraphs.

Activity 325.

   (a) Draw the digraph of the "is a proper subset of" relation on the set of subsets of a two element set. How many arrows would you have had to draw if this problem asked you to draw the digraph for the subsets of a three-element set?
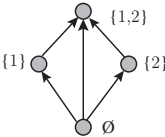
Solution.   We would need to draw 19 arrows, some of them curved, for the subsets of a three-element set.

   (b) Draw the digraph of the relation from the set {A, M, P, S} to the set {Sam, Mary, Pat, Ann, Polly, Sarah} given by "is the first letter of."
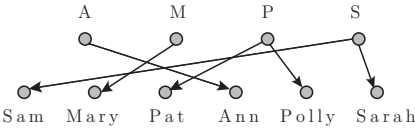
Solution.

Digraphs of Functions

Need to crosscheck some references back in 1.2.2, since they likely aren't to the right thing at this point.

Equivalence relations

So far we've used relations primarily to talk about functions. There is another kind of relation, called an equivalence relation, that comes up in the counting problems with which we began. In Problem 8 with three distinct flavors, it was probably tempting to say there are 12 flavors for the first pint, 11 for the second, and 10 for the third, so there are $12 \cdot 11 \cdot 10$ ways to choose the pints of ice cream. However, once the pints have been chosen, bought, and put into a bag, there is no way to tell which is first, which is second and which is third. What we just counted is lists of three distinct flavors—one to one functions from the set $\{1, 2, 3\}$ in to the set of ice cream flavors. Two of those lists

become equivalent once the ice cream purchase is made if they list the same ice cream. In other words, two of those lists become equivalent (are related) if they list same subset of the set of ice cream flavors. To visualize this relation with a digraph, we would need one vertex for each of the $12 \cdot 11 \cdot 10$ lists. Even with five flavors of ice cream, we would need one vertex for each of $5 \cdot 4 \cdot 3 = 60$ lists. So for now we will work with the easier to draw question of choosing three pints of ice cream of different flavors from four flavors of ice cream.

Activity 326. Suppose we have four flavors of ice cream, V(anilla), C(hocolate), S(trawberry) and P(each). Draw the directed graph whose vertices consist of all lists of three distinct flavors of the ice cream, and whose edges connect two lists if they list the same three flavors. This graph makes it pretty clear in how many ways we may choose 3 flavors out of four. How many is it?

Solution.   We used double-headed arrows in place of one arrow going in each direction to reduce the clutter in the picture. Note that there is an arrow from each vertex to itself. We may choose 3 flavors in four ways.
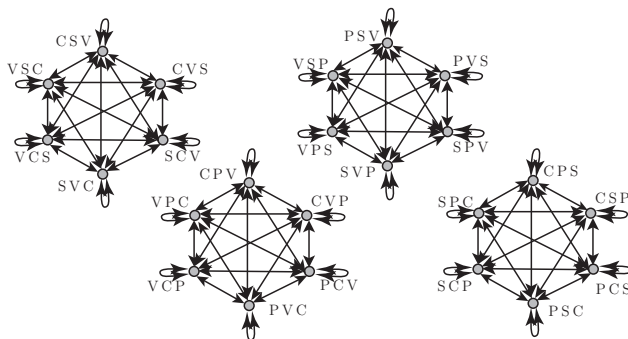
Activity 327. Now suppose again we are choosing three distinct flavors of ice cream out of four, but instead of putting scoops in a cone or choosing pints, we are going to have the three scoops arranged symmetrically in a circular dish. Similarly to choosing three pints, we can describe a selection of ice cream in terms of which one goes in the dish first, which one goes in second (say to the right of the first), and which one goes in third (say to the right of the second scoop, which makes it to the left of the first scoop). But again, two of these lists will sometimes be equivalent. Once they are in the dish, we can't tell which one went in first. However, there is a subtle difference between putting each flavor in its own small dish and putting all three flavors in a circle in a larger dish. Think about what makes the lists of flavors equivalent, and draw the directed graph whose vertices consist of all lists of three of the flavors of ice cream and whose edges connect two lists that we cannot tell the difference between as dishes of ice cream. How many dishes of ice cream can we distinguish from one another?

Solution.   We can distinguish eight different dishes of ice cream.

CSV   CVS   CPV   CVP

SVC   VCS   VSC   SCV   PCV   VCP   PVC   VPC

PSV   PVS   CPS   CSP

SVP   VPS   VSP   SPV   PSC   SCP   SPC   PCS
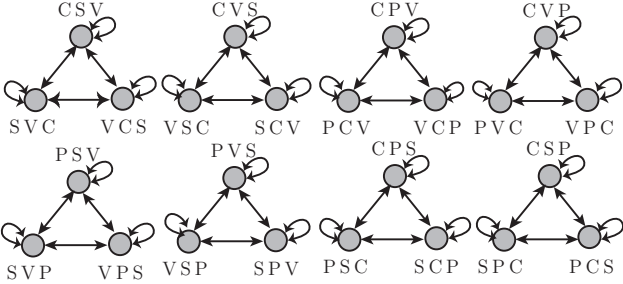
Activity 328. Draw the digraph for Problem 38 in the special case where we have four people sitting around the table.
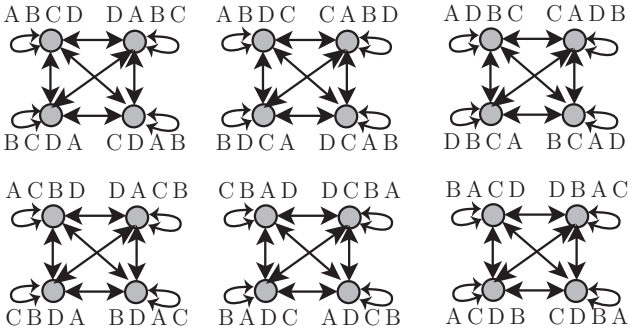
ABCD   DABC   ABDC   CABD   ADBC   CADB

BCDA   CDAB   BDCA   DCAB   DBCA   BCAD

ACBD   DACB   CBAD   DCBA   BACD   DBAC

Solution.   CBDA   BDAC   BADC   ADCB   ACDB   CDBA

In Problems 326, Activity 327, and ⟨⟨Unresolved xref, reference "roundtable-four"; check spelling or use "provisional" attribute⟩⟩ (as well as Problems 34, Activity 38, and Activity 39) we can begin with a set of lists, and say when two lists are equivalent as representations of the objects we are trying to count. In particular, in Problems 326, Activity 327, and ⟨⟨Unresolved xref, reference "roundtablefour"; check spelling or use "provisional" attribute⟩⟩ you drew the directed graph for this relation of equivalence. Technically, you should have had an arrow from each vertex (list) to itself. This is what we mean when we say a relation is *reflexive*. Whenever you had an arrow from one vertex to a second, you had an arrow back to the first. This is what we mean when we say a relation is *symmetric*.

When people sit around a round table, each list is equivalent to itself: if List1 and List 2 are identical, then everyone has the same person to the right in both lists (including the first person in the list being to the right of the last person). To see the symmetric property of the equivalence of seating arrangements, if List1 and List2 are different, but everyone has the same person to the right when they sit according to List2 as when they sit according to List1, then everybody better have the same person to the right when they sit according to List1 as when they sit according to List2.

In Problems 326, Activity 327 and ⟨⟨Unresolved xref, reference "roundtable-four"; check spelling or use "provisional" attribute⟩⟩ there is another property of those relations you may have noticed from the directed graph. Whenever you had an arrow from L_1 to L_2 and an arrow from L_2 to L_3, then there was an arrow from L_1 to L_3. This is what we mean when we say a relation is

transitive. You also undoubtedly noticed how the directed graph divides up into clumps of mutually connected vertices. This is what equivalence relations are all about. Let's be a bit more precise in our description of what it means for a relation to be reflexive, symmetric or transitive.

- If R is a relation on a set X, we say R is *reflexive* if $(x, x) \in R$ for every $x \in X$.

- If R is a relation on a set X, we say R is *symmetric* if (x, y) is in R whenever (y, x) is in R.

- If R is a relation on a set X, we say R is *transitive* if whenever (x, y) is in R and (y, z) is in R, then (x, z) is in R as well.

Each of the relations of equivalence you worked with in the Problem 326, Activity 327 and ⟨⟨Unresolved xref, reference "roundtablefour"; check spelling or use "provisional" attribute⟩⟩ had these three properties. Can you visualize the same three properties in the relations of equivalence that you would use in problems 34, Activity 38, and Activity 39? We call a relation an *equivalence relation* if it is reflexive, symmetric and transitive.

After some more examples, we will see how to show that equivalence relations have the kind of clumping property you saw in the directed graphs. In our first example, using the notation $(a, b) \in R$ to say that a is related to B is going to get in the way. It is really more common to write aRb to mean that a is related to b. For example, if our relation is the less than relation on $\{1, 2, 3\}$, you are much more likely to use $x < y$ than you are $(x, y) \in <$, aren't you? The reflexive law then says xRx for every x in X, the symmetric law says that if xRy, then yRx, and the transitive law says that if xRy and yRz, then xRz.

Activity 329.

(a) For the necklace problem, Problem 43, our lists are lists of beads. What makes two lists equivalent for the purpose of describing a necklace? Verify explicitly that this relationship of equivalence is reflexive, symmetric, and transitive.

Solution.   Two lists are equivalent if I can get one from the other by some combination of cyclic permutations (putting the last thing in the list at the front and moving everything else one place right) and reversals. The combination could include no operation at all. Since it is possible to have no operation, the relation is reflexive. (Even without the opportunity to do no operation, if we do two reversals to a list we get the original list back so it is equivalent to itself.) Suppose we have n beads. Then if I can get from list A to list B with a cyclic permutations, then $n - 1$ more cyclic permutations give us the original list. Also if I get from a list A to a list B by a reversal, then another reversal takes B to A. Thus any sequence of cyclic permutations and reversals can be undone. Therefore if list A is equivalent to list B, then list B is equivalent to list

A. Following one combination of operations with another one still gives a combination of operations, so our relation is transitive.

(b) Which of the reflexive, symmetric and transitive properties does the $<$ relation on the integers have?

Solution. It is transitive, but not reflexive or symmetric.

(c) A relation R on the set of ordered pairs of positive integers that you learned about in grade school in another notation is the relation that says (m, n) is related to (h, k) if $mk = hn$. Show that this relation is an equivalence relation. In what context did you learn about this relation in grade school?

Solution. $mn = mn$ so the relation is reflexive. If $mk = hn$, then $hn = mk$, so if (m, n) is related to (h, k), then (h, k) is related to (m, n). If (m, n) is related to (h, k) and (h, k) is related to (p, q), then $mk = hn$ and $hq = pk$, which gives us $mkhq = hnpk$, and cancelling h and k gives us $mq = np = pn$, so (m, n) is related to (p, q). Therefore, the relation is transitive. This is the relation of equality of the fractions $\frac{m}{n}$ and $\frac{h}{k}$.

(d) Another relation that you may have learned about in school, perhaps in the guise of "clock arithmetic," is the relation of equivalence modulo n. For integers (positive, negative, or zero) a and b, we write $a \equiv b \pmod{n}$ to mean that $a - b$ is an integer multiple of n, and in this case, we say that a is *congruent to b modulo n* and write $a \equiv b \pmod{n}$.. Show that the relation of congruence modulo n is an equivalence relation.

Solution. $a - a = 0 = 0 \cdot n$, so $a \equiv a \pmod{n}$. Thus the relation is reflexive. If $a - b = kn$ for some integer k, then $b - a = -kn$, and -k is an integer, so if $a \equiv b \pmod{n}$, then $b \equiv a \pmod{n}$. If $a - b = kn$ and $b - c = jn$, then $a - b + b - c = kn + jn$, so $a - c = (k + j)n$ and since $k + j$ is an integer this means that $a \equiv c \pmod{n}$. Therefore the relation of congruence mod n is an equivalence relation.

(e) Define a relation on the set of all lists of n distinct integers chosen from $\{1, 2, \ldots, n\}$, by saying two lists are related if they have the same elements (though perhaps in a different order) in the first k places, and the same elements (though perhaps in a different order) in the last $n - k$ places. Show this relation is an equivalence relation.

Solution. The relation is reflexive, for a list L has the same elements as the list L in the first k places and the last $n - k$ places. If L_1 and L_2

have the same elements in the first k places and have the same elements in the last k places, then L_2 and L_1 have the same elements in the first k places and have the same elements in the last $n - k$ places, so our relation is symmetric. If L_1 and L_2 have the same elements in the first k places and L_2 and L_3 have the same elements in the first k places, then L_1 and L_3 have the same elements in the first k places. Similarly with the last $n - k$ places. Therefore our relation is transitive, and so it is an equivalence relation.

(f) Suppose that R is an equivalence relation on a set X and for each $x \in X$, let $C_x = \{y | y \in X \text{ and } yRx\}$. If C_x and C_z have an element y in common, what can you conclude about C_x and C_z (besides the fact that they have an element in common!)? Be explicit about what property(ies) of equivalence relations justify your answer. Why is every element of X in some set C_x? Be explicit about what property(ies) of equivalence relations you are using to answer this question. Notice that we might simultaneously denote a set by C_x and C_y. Explain why the union of the sets C_x is X. Explain why two distinct sets C_x and C_z are disjoint. What do these sets have to do with the "clumping" you saw in the digraph of Problem 326 and Activity 327?

Solution. If C_x and C_y have the element z in common, then by symmetry and transitivity, all elements in C_x are related to z and by symmetry and transitivity, all elements in C_y are related to z. Then by symmetry and transitivity again, all elements of C_y are related to x, so $C_y \subseteq C_x$. By the same kind of reasoning, $C_x \subseteq C_y$. Therefore, $C_x = C_y$. Every element x is in the set C_x by reflexivity Thus the union of the sets C_x is X. The sets C_x and C_y are disjoint if they are different, because if they have a common element z then they are equal. By definition, the sets C_x form a partition of X. The clumps that we saw in those problems are the blocks of the partition.

In ⟨⟨Unresolved xref, reference "equivalenceclasses"; check spelling or use "provisional" attribute⟩⟩ the sets C_x are called *equivalence classes* of the equivalence relation R. You have just proved that if R is an equivalence relation of the set X, then each element of X is in exactly one equivalence class of R. Recall that a *partition* of a set X is a set of disjoint sets whose union is X. For example, $\{1,3\}$, $\{2,4,6\}$, $\{5\}$ is a partition of the set $\{1,2,3,4,5,6\}$. Thus another way to describe what you proved in ⟨⟨Unresolved xref, reference "equivalenceclasses"; check spelling or use "provisional" attribute⟩⟩ is the following:

Theorem A.1.2. *If R is an equivalence relation on X, then the set of equivalence classes of R is a partition of X.*

Since a partition of S is a set of subsets of S, it is common to call the subsets into which we partition S the *blocks* of the partition so that we don't

find ourselves in the uncomfortable position of referring to a set and not being sure whether it is the set being partitioned or one of the blocks of the partition.

Activity 330.

(a) In each of Problems 38, Activity 39, Activity 43, Activity 326, and Activity 327, what does an equivalence class correspond to? (Five answers are expected here.)

Solution.   In Problem 38 the equivalence classes correspond to seating arrangements. In Problem 39 the equivalence classes correspond to the k-element subsets of our n-element set S. In Problem 43, the equivalence classes correspond to necklaces. In Problem 326 the equivalence classes correspond to choices of three flavors of ice cream out of a possible four flavors. In Problem 327 the equivalence classes correspond to the ways we can choose scoops of ice cream of three different flavors out of four and put them into a dish in a symmetric fashion.

(b) Given the partition $\{1, 3\}$, $\{2, 4, 6\}$, $\{5\}$ of the set $\{1, 2, 3, 4, 5, 6\}$, define two elements of $\{1, 2, 3, 4, 5, 6\}$ to be related if they are in the same part of the partition. That is, define 1 to be related to 3 (and 1 and 3 each related to itself), define 2 and 4, 2 and 6, and 4 and 6 to be related (and each of 2, 4, and 6 to be related to itself), and define 5 to be related to itself. Show that this relation is an equivalence relation.

Solution.   We have said for each element of our set that it is related to itself, so the relation is reflexive. If x and y are in a given one of those sets, then y and x are in that same given set. If x and y are in the same set, and if y and z are in the same set, then x and z must be in the same set because there is one and only one set that y is in. Thus the relation is an equivalence relation.

(c) Suppose $P = \{S_1, S_2, S_3, \ldots, S_k\}$ is a partition of S. Define two elements of S to be related if they are in the same set S_i, and otherwise not to be related. Show that this relation is an equivalence relation. Show that the equivalence classes of the equivalence relation are the sets S_i.

Solution.   Each element is in a set S_i with itself, so the relation is reflexive. If x and y are in a given one of those sets S_i, then y and x are in that same set S_i. If x and y are in the same set S_i, and if y and z are in the same set S_j, then S_i must equal S_j because y is in one and only one block of the partition. Therefore, x and z must be in the same set S_i. Thus the relation is an equivalence relation. If $x \in S_i$, then by definition S_i consists of all elements related to x, so it is the equivalence class containing x.

In Problem 330 you just proved that each partition of a set gives rise to an equivalence relation whose classes are just the parts of the partition. Thus in ⟨⟨Unresolved xref, reference "equivalenceclasses"; check spelling or use "provisional" attribute⟩⟩ and Problem 330 you proved the following Theorem.

Theorem A.1.3. *A relation R is an equivalence relation on a set S if and only if S may be partitioned into sets S_1, S_2, ..., S_n in such a way that x and y are related by R if and only if they are in the same block S_i of the partition.*

In Problems 326, Activity 327, Activity 38 and Activity 43 what we were doing in each case was counting equivalence classes of an equivalence relation. There was a special structure to the problems that made this somewhat easier to do. For example, in Activity 326, we had $4 \cdot 3 \cdot 2 = 24$ lists of three distinct flavors chosen from V, C, S, and P. Each list was equivalent to $3 \cdot 2 \cdot 1 = 3! = 6$ lists, including itself, from the point of view of serving 3 small dishes of ice cream. The order in which we selected the three flavors was unimportant. Thus the set of all $4 \cdot 3 \cdot 2$ lists was a union of some number n of equivalence classes, each of size 6. By the product principle, if we have a union of n disjoint sets, each of size 6, the union has $6n$ elements. But we already knew that the union was the set of all 24 lists of three distinct letters chosen from our four letters. Thus we have $6n = 24$, or $n = 4$ equivalence classes.

In Problem 327 there is a subtle change. In the language we adopted for seating people around a round table, if we choose the flavors V, C, and S, and arrange them in the dish with C to the right of V and S to the right of C, then the scoops are in different relative positions than if we arrange them instead with S to the right of V and C to the right of S. Thus the order in which the scoops go into the dish is somewhat important—somewhat, because putting in V first, then C to its right and S to its right is the same as putting in S first, then V to its right and C to its right. In this case, each list of three flavors is equivalent to only three lists, including itself, and so if there are n equivalence classes, we have $3n = 24$, so there are $24/3 = 8$ equivalence classes.

Activity 331.

(a) If we have an equivalence relation that divides a set with k elements up into equivalence classes each of size m, what is the number n of equivalence classes? Explain why.

Solution. The number of equivalence classes is k/m, because by the product principle, $mn = k$.

(b) In Problem 329, what is the number of equivalence classes? Explain in words the relationship between this problem and the Problem 39.

Solution. There are $n!$ lists, and each is in an equivalence class of size $k!(n-k)!$, so the number of equivalence classes is $\frac{n!}{k!(n-k)!}$ by Problem 331. This is a way of computing the number of k-element subsets

that shows why the final answer we got in Problem 39 is symmetric in k and $n-k$.

(c) Describe explicitly what makes two lists of beads equivalent in Problem 43 and how Problem 331 can be used to compute the number of different necklaces.

Solution. Two lists are equivalent if I can get one from the other by some combination of cyclic shifts and reversals. A cyclic shift on the list $a_1, a_2, \ldots, a_{n-1}, a_n$ gives either the list $a_n, a_1, a_2, \ldots, a_{n-1}$ or the list $a_2, \ldots, a_{n-1}, a_n, a_1$. There are n possible results of repeated cyclic shifts, and each of them may be reversed to give a new list if $n \geq 3$. Further, these are the only lists we can get from shifts and reversals. (a_1 must go to one of n places, and that leaves two choices for where a_2 goes. Then the rest of the list is determined.) Thus we can get exactly $2n$ lists from combinations of cyclic shifts and reversals. We define two lists to be equivalent if they give the same necklace; we've seen that this is an equivalence relation and that it has $2n$ elements per equivalence class. Since there are $n!$ lists, this gives us $(n-1)!/2$ equivalence classes, or necklaces.

(d) What are the equivalence classes (write them out as sets of lists) in Problem 45, and why can't we use Problem 331 to compute the number of equivalence classes?

Solution. The equivalence classes are

$$\{RRBB, BRRB, BBRR, RBBR\} \text{ and } \{RBRB, BRBR\}.$$

We can't use Problem 331 to compute the number of equivalence classes because the equivalence classes don't have the same size.

In Problem 331 you proved our next theorem. In Chapter 1 (Problem 42) we discovered and stated this theorem in the context of partitions and called it the *Quotient Principle*

Theorem A.1.4. *If an equivalence relation on a set S size k has n equivalence classes each of size m, then the number of equivalence classes is k/m.*

Appendix B

Mathematical Induction

B.1   The Principle of Mathematical Induction

The ideas behind mathematical induction

There is a variant of one of the bijections we used to prove the Pascal Equation that comes up in counting the subsets of a set. In the next problem it will help us compute the total number of subsets of a set, regardless of their size. Our main goal in this problem, however, is to introduce some ideas that will lead us to one of the most powerful proof techniques in combinatorics (and many other branches of mathematics), the principle of mathematical induction.

Activity 332.

(a) Write down a list of the subsets of $\{1, 2\}$. Don't forget the empty set! Group the sets containing containing 2 separately from the others.

   Solution. $\emptyset$, $\{1\}$, $\{2\}$, $\{1, 2\}$.

(b) Write down a list of the subsets of $\{1, 2, 3\}$. Group the sets containing 3 separately from the others.

   Solution. $\emptyset$, $\{1\}$, $\{2\}$, $\{1, 2\}$, $\{3\}$, $\{1, 3\}$, $\{2, 3\}$, $\{1, 2, 3\}$.

(c) Look for a natural way to match up the subsets containing 2 in Part (a) with those not containing 2. Look for a way to match up the subsets containing 3 in Part (b) containing 3 with those not containing 3.

   Solution.   Adjoin 2 to each subset not containing 2 and you get each set containing 2. Adjoin 3 to each subset not containing 3, and you get each subset containing 3.

(d) On the basis of the previous part, you should be able to find a bijection between the collection of subsets of $\{1, 2, \ldots, n\}$ containing n and those not containing n. (If you are having difficulty figuring out the bijection, try rethinking Parts (a) and (b), perhaps by doing a similar exercise with the set $\{1, 2, 3, 4\}$.) Describe the bijection (unless you are very familiar with the notation of sets, it is probably easier to describe to describe the function in words rather than symbols) and explain why it is a bijection. Explain why the number of subsets of $\{1, 2, \ldots, n\}$ containing n equals the number of subsets of $\{1, 2, \ldots, n-1\}$.

Solution. If we adjoin n to the subsets not containing n we get the subsets containing n. This is a bijection because if we start with two different sets, adjoining n to them can't make them the same, and every subset S containing n must arise in this way from the set $S - \{n\}$ not containing n.

(e) Parts (a) and (b) suggest strongly that the number of subsets of a n-element set is 2^n. In particular, the empty set has 2^0 subsets, a one-element set has 2^1 subsets, itself and the empty set, and in Parts a and b we saw that two-element and three-element sets have 2^2 and 2^3 subsets respectively. So there are certainly some values of n for which an n-element set has 2^n subsets. One way to prove that an n-element set has 2^n subsets for all values of n is to argue by contradiction. For this purpose, suppose there is a nonnegative integer n such that an n-element set doesn't have exactly 2^n subsets. In that case there may be more than one such n. Choose k to be the smallest such n. Notice that $k - 1$ is still a positive integer, because k can't be 0, 1, 2, or 3. Since k was the smallest value of n we could choose to make the statement "An n-element set has 2^n subsets" false, what do you know about the number of subsets of a $(k-1)$-element set? What do you know about the number of subsets of the k-element set $\{1, 2, \ldots, k\}$ that don't contain k? What do you know about the number of subsets of $\{1, 2, \ldots, k\}$ that do contain k? What does the sum principle tell you about the number of subsets of $\{1, 2, \ldots, k\}$? Notice that this contradicts the way in which we chose k, and the only assumption that went into our choice of k was that "there is a nonnegative integer n such that an n-element set doesn't have exactly 2^n subsets." Since this assumption has led us to a contradiction, it must be false. What can you now conclude about the statement "for every nonnegative integer n, an n-element set has exactly 2^n subsets?"

Solution. We know that the number of subsets of a $(k-1)$-element set is 2^{k-1}. The number of subsets of $\{1, 2, \ldots, k\}$ that do not contain k is the number of subsets of the $k-1$-element set $\{1, 2, \ldots, k-1\}$, so we know this number is 2^{k-1}. We know that the number of subsets that do contain k equals the number that don't, so the number that do

contain k is also 2^{k-1}. The sum principle tells us that the number os subsets of $\{1, 2, \ldots, k\}$ is $2^{k-1} + 2^{k-1} = 2^k$. We can conclude that the statement "for every nonnegative integer n, an n-element set has exactly 2^n subsets" is true.

Activity 333. The expression

$$1 + 3 + 5 + \cdots + 2n - 1$$

is the sum of the first n odd integers. Experiment a bit with the sum for the first few positive integers and guess its value in terms of n. Now apply the technique of Problem 332 to prove that you are right.

Solution. We guess that $1 + 3 + 5 + \cdots + 2n - 1 = n^2$. Clearly this is true when n is 1, 2, or 3. Suppose there is an n for which this formula is not true, and let k be the smallest such n. Then $1 + 3 + 5 + \cdots + 2(k-1) - 1 = (k-1)^2$. Simplifying, $1 + 3 + 5 + \cdots + 2k - 3 = (k-1)^2$. Now suppose we add $2k - 1$ to both sides of this equation. Then we get

$$1 + 3 + 5 + \cdots + 2k - 3 + 2k - 1 = (k-1)^2 + 2k - 1 = k^2 - 2k + 1 + 2k - 1 = k^2.$$

But this is a contradiction, because we assumed that k was the smallest value of n for which the sum on the left is not n^2. Therefore the assumption that there is an n for which $1 + 3 + 5 + \cdots + 2n - 1 \neq n^2$ must be false, so the equation $1 + 3 + 5 + \cdots + 2n - 1 = n^2$ must be true for all positive integers n.

In Problems 332 and $\langle\langle$Unresolved xref, reference "sumodd"; check spelling or use "provisional" attribute$\rangle\rangle$ our proofs had several distinct elements. We had a statement involving an integer n. We knew the statement was true for the first few nonnegative integers in Problem 332 and for the first few positive integers in $\langle\langle$Unresolved xref, reference "sumodd"; check spelling or use "provisional" attribute$\rangle\rangle$. We wanted to prove that the statement was true for all nonnegative integers in Problem 332 and for all positive integers in $\langle\langle$Unresolved xref, reference "sumodd"; check spelling or use "provisional" attribute$\rangle\rangle$. In both cases we used the method of proof by contradiction; for that purpose we assumed that there was a value of n for which our formula wasn't true. We then chose k to be the smallest value of n for which our formula wasn't true. This meant that when n was $k - 1$, our formula was true, (or else that $k - 1$ wasn't a nonnegative integer in Problem 332 or that $k - 1$ wasn't a positive integer in $\langle\langle$Unresolved xref, reference "sumodd"; check spelling or use "provisional" attribute$\rangle\rangle$). What we did next was the crux of the proof. We showed that the truth of our statement for $n = k - 1$ implied the truth of our statement for $n = k$. This gave us a contradiction to the assumption that there was an n that made the statement false. In fact, we will see that we can bypass entirely the use of proof by contradiction. We used it to help you discover the central ideas of the technique of proof by mathematical induction.

The central core of mathematical induction is the proof that the truth of a statement about the integer n for $n = k - 1$ implies the truth of the statement

for $n = k$. For example, once we know that a set of size 0 has 2^0 subsets, if we have proved our implication, we can then conclude that a set of size 1 has 2^1 subsets, from which we can conclude that a set of size 2 has 2^2 subsets, from which we can conclude that a set of size 3 has 2^3 subsets, and so on up to a set of size n having 2^n subsets for any nonnegative integer n we choose. In other words, although it was the idea of proof by contradiction that led us to think about such an implication, we can now do without the contradiction at all. What we need to prove a statement about n by this method is a place to start, that is a value b of n for which we know the statement to be true, and then a proof that the truth of our statement for $n = k - 1$ implies the truth of the statement for $n = k$ whenever $k > b$.

Mathematical induction

The *principle of mathematical induction* states that

> In order to prove a statement about an integer n, if we can
>
> 1. Prove the statement when $n = b$, for some fixed integer b
>
> 2. Show that the truth of the statement for $n = k - 1$ implies the truth of the statement for $n = k$ whenever $k > b$,
>
> then we can conclude the statement is true for all integers $n \geq b$.

As an example, let us return to Problem 332. The statement we wish to prove is the statement that "A set of size n has 2^n subsets."

> Our statement is true when $n = 0$, because a set of size 0 is the empty set and the empty set has $1 = 2^0$ subsets. (This step of our proof is called a *base step*.) Now suppose that $k > 0$ and every set with $k - 1$ elements has 2^{k-1} subsets. Suppose $S = \{a_1, a_2, \ldots a_k\}$ is a set with k elements. We partition the subsets of S into two blocks. Block B_1 consists of the subsets that do not contain a_n and block B_2 consists of the subsets that do contain a_n. Each set in B_1 is a subset of $\{a_1, a_2, \ldots a_{k-1}\}$, and each subset of $\{a_1, a_2, \ldots a_{k-1}\}$ is in B_1. Thus B_1 is the set of all subsets of $\{a_1, a_2, \ldots a_{k-1}\}$. Therefore by our assumption in the first sentence of this paragraph, the size of B_1 is 2^{k-1}. Consider the function from B_2 to B_1 which takes a subset of S including a_n and removes a_n from it. This function is defined on B_2, because every set in B_2 contains a_n. This function is onto, because if T is a set in B_1, then $T \cup \{a_k\}$ is a set in B_2 which the function sends to T. This function is one-to-one because if V and W are two different sets in B_2, then removing a_k from them gives two different sets in B_1. Thus we have a bijection between B_1 and B_2, so B_1 and B_2 have the same size. Therefore by the sum principle the size of $B_1 \cup B_2$ is $2^{k-1} + 2^{k-1} = 2^k$. Therefore S has 2^k subsets. This shows that if a set of size $k - 1$ has 2^{k-1} subsets, then a set of size k has 2^k

subsets. Therefore by the principle of mathematical induction, a set of size n has 2^n subsets for every nonnegative integer n.

The first sentence of the last paragraph is called the *inductive hypothesis*. In an inductive proof we always make an inductive hypothesis as part of proving that the truth of our statement when $n = k - 1$ implies the truth of our statement when $n = k$. The last paragraph itself is called the *inductive step* of our proof. In an inductive step we derive the statement for $n = k$ from the statement for $n = k - 1$, thus proving that the truth of our statement when $n = k - 1$ implies the truth of our statement when $n = k$. The last sentence in the last paragraph is called the *inductive conclusion*. All inductive proofs should have a base step, an inductive hypothesis, an inductive step, and an inductive conclusion.

There are a couple details worth noticing. First, in this problem, our base step was the case $n = 0$, or in other words, we had $b = 0$. However, in other proofs, b could be any integer, positive, negative, or 0. Second, our proof that the truth of our statement for $n = k - 1$ implies the truth of our statement for $n = k$ required that k be at least 1, so that there would be an element a_k we could take away in order to describe our bijection. However, condition (2) of the principle of mathematical induction only requires that we be able to prove the implication for $k > 0$, so we were allowed to assume $k > 0$.

Activity 334. Use mathematical induction to prove your formula from ⟨⟨Unresolved xref, reference "sumodd"; check spelling or use "provisional" attribute⟩⟩.

Proving algebraic statements by induction

Activity 335.

(a) Use mathematical induction to prove the well-known formula that for all positive integers n,

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2}.$$

Solution. When $n = 0$, $0 = 0(0 + 1)/2$, so our formula holds. Now suppose that $k > 0$ and that our formula holds when $n = k - 1$, so that $1 + 2 + \cdots + k - 1 = (k - 1)k/2$. Add k to both sides of this equation to get

$$
\begin{aligned}
1 + 2 + \cdots + (k - 1) + k &= (k - 1)k/2 + k \\
&= k^2/2 - k/2 + k \\
&= k^2/2 + k/2 \\
&= k(k + 1)/2.
\end{aligned}
$$

Thus the truth of our formula for $n = k - 1$ implies its truth for $n = k$. Therefore by the principle of mathematical induction, our formula holds for all nonnegative integers n.

(b) Experiment with various values of n in the sum

$$\frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \frac{1}{3\cdot 4} + \cdots + \frac{1}{n\cdot(n+1)} = \sum_{i=1}^{n} \frac{1}{i\cdot(i+1)}.$$

Guess a formula for this sum and prove your guess is correct by induction.

Solution.   We guess the formula

$$\sum_{i=1}^{n} \frac{1}{i(i+1)} = \frac{n}{n+1}.$$

When $n = 1$ this formula says that $\frac{1}{1\cdot 2} = \frac{1}{1\cdot 2}$, so our formula holds when $n = 1$. Now assume that $k > 1$ and that our formula holds when $n = k - 1$. Then

$$\sum_{i=1}^{k-1} \frac{1}{i(i+1)} = \frac{k-1}{k}.$$

Adding $\frac{1}{k(k+1)}$ to both sides of this equation gives us

$$\sum_{i=1}^{k-1} \frac{1}{i(i+1)} + \frac{1}{k(k+1)} = \frac{k-1}{k} + \frac{1}{k(k+1)}$$

$$\sum_{i=1}^{k} \frac{1}{i(i+1)} = \frac{(k-1)(k+1)}{k(k+1)} + \frac{1}{k(k+1)}$$

$$= \frac{k^2 - 1 + 1}{k(k+1)}$$

$$= \frac{k}{k+1}.$$

Thus whenever our formula is true with $n = k-1$, it is true with $n = k$ as well. Therefore by the principle of mathematical induction, our formula is true for all positive integers.

(c) For large values of n, which is larger, n^2 or 2^n? Use mathematical induction to prove that you are correct.

Solution.   We note that $0^2 = 0$, while $2^0 = 1$, that $1^2 = 1$, while $2^1 = 2$, that $2^2 = 4$, while $2^2 = 4$, that $3^2 = 9$ while $2^3 = 8$, that $4^2 = 16$ while $2^4 = 16$, and $5^2 = 25$ while $2^5 = 32$. We suspect that $2^n > n^2$ for $n \geq 5$, so we try to prove this by induction. We have already shown that $2^5 > 5^2$. Now suppose that $k > 5$ and $2^{k-1} > (k-1)^2$. Then $2^k = 2 \cdot 2^{k-1} > 2(k-1)^2 = 2k^2 - 4k + 1$. Now since $k > 5$, $k^2 > 5k$, so

that $k^2 - 4k + 1 = k^2 + k^2 - 4k + 1 > k^2 + 5k - 4k + 1 = k^2 + k + 1 > k^2$. Thus for $k > 5$, the statement $2^{k-1} > (k-1)^2$ implies the statement $2^k > k^2$. Therefore, by the principle of mathematical induction, $2^n > n^2$ for all $n \geq 5$.

(d) What is wrong with the following attempt at an inductive proof that all integers in any consecutive set of n integers are equal for every positive integer n? For an arbitrary integer i, all integers from i to i are equal, so our statement is true when $n = 1$. Now suppose $k > 1$ and all integers in any consecutive set of $k - 1$ integers are equal. Let S be a set of k consecutive integers. By the inductive hypothesis, the first $k - 1$ elements of S are equal and the last $k - 1$ elements of S are equal. Therefore all the elements in the set S are equal. Thus by the principle of mathematical induction, for every positive n, every n consecutive integers are equal.

Solution.   One possible value of k that is greater than 1 is 2. When we have a set S of two elements and we argue that the first $k - 1$ elements are equal and the last $k - 1$ elements are equal, we cannot conclude from those equalities that all elements of S are equal, because there is no overlap among the first $k - 1 = 1$ elements of S and the last $k - 1 = 1$ elements of S. Thus our inductive step does not cover the possibility that $k = 2$. Therefore our inductive step does not show that the truth of our statement for $n = k - 1$ implies the truth of our statement for $n = k$ for *all* integers $n > 1$. Therefore the principle of mathematical induction does not apply.

Strong Induction

One way of looking at the principle of mathematical induction is that it tells us that if we know the "first" case of a theorem and we can derive each other case of the theorem from a smaller case, then the theorem is true in all cases. However the particular way in which we stated the theorem is rather restrictive in that it requires us to derive each case from the immediately preceding case. This restriction is not necessary, and removing it leads us to a more general statement of the principal of mathematical induction which people often call the *strong principle of mathematical induction*. It states:

In order to prove a statement about an integer n if we can

1. prove our statement when $n = b$ and

2. prove that the statements we get with $n = b$, $n = b + 1$, ... $n = k - 1$ imply the statement with $n = k$,

then our statement is true for all integers $n \geq b$.

Activity 336. What postage do you think we can make with five and six cent stamps? Is there a number N such that if $n \geq N$, then we can make n cents worth of postage?

Solution. We can make 10, 11, and 12 cents in postage, but not 13 cents. We can also make 15, 16, 17, and 18, but not 19 cents. However when we try starting with 20 cents, we can make 20, 21, 22, 23, 24, 25, 26, 27,...cents, and so it seems for all $n \geq 20$, we can make n cents in stamps. Once we know we can make 20 cents through 24 cents, by adding 5 cents to each of these we can get 25 through 29 cents, and so we expect to be able to keep going. However making 29 cents does not depend on our ability to make 28 cents; rather we know we can make 29 cents because we know we can make 24 cents and $24 + 5 = 29$ or we know we can make 23 cents and $23 + 6 = 29$. Thus it certainly seems as if for all $n \geq 20$ we can make n cents in postage.

You probably see that we can make n cents worth of postage as long as n is at least 20. However you didn't try to make 26 cents in postage by working with 25 cents; rather you saw that you could get 20 cents and then add six cents to that to get 26 cents. Thus if we want to prove by induction that we are right that if $n \geq 20$, then we can make n cents worth of postage, we are going to have to use the strong version of the principle of mathematical induction.

We know that we can make 20 cents with four five-cent stamps. Now we let k be a number greater than 20, and assume that it is possible to make any amount between 20 and $k - 1$ cents in postage with five and six cent stamps. Now if k is less than 25, it is 21, 22, 23, or 24. We can make 21 with three fives and one six. We can make 22 with two fives and two sixes, 23 with one five and three sixes, and 24 with four sixes. Otherwise $k - 5$ is between 20 and $k - 1$ (inclusive) and so by our inductive hypothesis, we know that $k - 5$ cents can be made with five and six cent stamps, so with one more five cent stamp, so can k cents. Thus by the (strong) principle of mathematical induction, we can make n cents in stamps with five and six cent stamps for each $n \geq 20$.

Some people might say that we really had five base cases, $n = 20, 21, 22, 23$, and 24, in the proof above and once we had proved those five consecutive base cases, then we could reduce any other case to one of these base cases by successively subtracting 5. That is an appropriate way to look at the proof. A logician would say that it is also the case that, for example, by proving we could make 22 cents, we also proved that if we can make 20 cents and 21 cents in stamps, then we could also make 22 cents. We just didn't bother to use the assumption that we could make 20 cents and 21 cents! So long as one point of view or the other satisfies you, you are ready to use this kind of argument in proofs.

Activity 337. A number greater than one is called prime if it has no factors other than itself and one. Show that each positive number is either a prime or a power of a prime or a product of powers of prime numbers.

Solution. We note that $1 = 2^0$, so 1 is a power of a prime. Now suppose that all positive numbers less than n are primes, powers of primes, or products

of powers of primes. If n has no proper factors, it is a prime. If it does have proper factors, say $n = mk$, both factors are less than n and greater than 1. Therefore each factor is a prime, a power of a prime, or a product of powers of primes. When we multiply m and k together, the result will still be a power of a prime or a product of powers of primes. Thus the statement that all positive numbers less than n are primes, powers of primes, or products of powers of primes implies the statement that n is a prime, a power of a prime, or a product of powers of primes. Therefore by the strong principle of mathematical induction, all positive numbers are either primes, powers of primes, or products of powers of primes.

Activity 338. Show that the number of prime factors of a positive number $n \geq 2$ is less than or equal to $\log_2 n$. (If a prime occurs to the kth power in a factorization of n, you can consider that power as k prime factors.) (There is a way to do this by induction and a way to do it without induction. It would be ideal to find both ways.)

Solution.   First, we will prove this by induction.   The number of prime factors of 2 is 1, which is less than or equal to $\log_2 2 = 1$. Now assume that the number of prime factors of any number k greater than 1 and less than n is no more than $\log_2 k$. If n is prime, then its number of prime factors is less than or equal to $\log_2 n$. Otherwise n is a product of two factors, $n = mk$. Then by our inductive hypothesis, the number of prime factors of m is less than or equal to $\log_2 m$ and the number of prime factors of k is less than or equal to $\log_2 k$. But the number of prime factors of the product is the sum of the number of prime factors of each factor, so the number of prime factors of n is no more than $\log_2 m + \log_2 k = \log_2 mk = \log_2 n$. Thus statement that the number of prime factors of any number k between 2 and $n-1$ inclusive is no more than $\log_2 k$, implies the statement that the number of prime factors of n is no more than $\log_2 n$. Therefore, by the principle of mathematical induction, the number of prime factors of n is less than or equal to $\log_2 n$ for all integers $n \geq 2$.

Solution.   For a noninductive proof, note that all factors of n are at least 2. If n is a power of two, then the number of times 2 is a factor of n is exactly $\log_2 n$. But if n is not a power of 2, we still have that $2^{\log_2 n} = n$, so the product of $\lceil \log_2 n \rceil$ numbers including some greater than 2 must be greater than n. Therefore, the number of prime factors of n is no more than $\log_2 n$. Thus for all $n \geq 2$, the number of prime factors of n must be less than or equal to $\log_2 n$.

Exponential Generating Functions

B.2 Indicator Functions

When we introduced the idea of a generating function, we said that the formal sum

$$\sum_{i=0}^{n} a_i x^i$$

may be thought of as a convenient way to keep track of the sequence a_i. We then did quite a few examples that showed how combinatorial properties of arrangements counted by the coefficients in a generating function could be mirrored by algebraic properties of the generating functions themselves. The monomials x^i are called *indicator polynomials*. (They indicate the position of the coefficient a_i.) One example of a generating function is given by

$$(1+x)^n = \sum_{i=0}^{\infty} \binom{n}{i} x^i.$$

Thus we say that $(1+x)^n$ is the generating function for the binomial coefficients $\binom{n}{i}$. The notation tells us that we are assuming that only i varies in the sum on the right, but that the equation holds for each fixed integer n. This is implicit when we say that $(1+x)^n$ is the generating function for $\binom{n}{i}$, because we haven't written i anywhere in $(1+x)^n$, so it is free to vary.

Another example of a generating function is given by

$$x^{\underline{n}} = \sum_{i=0}^{\infty} s(n,i) x^i.$$

Thus we say that $x^{\underline{n}}$ is the generating function for the Stirling numbers of the first kind, $s(n,i)$. There is a similar equation for Stirling numbers of the second kind, namely

$$x^n = \sum_{i=0}^{\infty} S(n,i) x^{\underline{i}}.$$

However with our previous definition of generating functions, this equation would not give a generating function for the Stirling numbers of the second kind, because $S(n, i)$ is not the coefficient of x^i. If we were willing to consider the falling factorial powers $x^{\underline{i}}$ as indicator polynomials, then we could say that x^n is the generating function for the numbers $S(n, i)$ relative to these indicator polynomials. This suggests that perhaps different sorts of indicator polynomials go naturally with different sequences of numbers.

The binomial theorem gives us yet another example.

Activity 339. Write $(1 + x)^n$ as a sum of multiples of $\frac{x^i}{i!}$ rather than as a sum of multiples of x^i.

Solution.

$$(1 + x)^n = \sum_{i=0}^{\infty} \binom{n}{i} x^i = \sum_{i=0}^{\infty} \frac{n!}{i!(n - i)!} x^i = \sum_{i=0}^{\infty} \frac{n!}{(n - i)!} \frac{x^i}{i!} = \sum_{i=0}^{\infty} n^{\underline{i}} \frac{x^i}{i!}.$$

This example suggests that we could say that $(1 + x)^n$ is the generating function for the falling factorial powers $n^{\underline{i}}$ relative to the indicator polynomials $\frac{x^i}{i!}$. In general, a sequence of polynomials is called a family of *indicator polynomials* if there is one polynomial of each nonnegative integer degree in the sequence. Those familiar with linear algebra will recognize that this says that a family of indicator polynomials form a basis for the vector space of polynomials. This means that each polynomial way can be expressed as a sum of numerical multiples of indicator polynomials in one and only one way. One could use the language of linear algebra to define indicator polynomials in an even more general way, but a definition in such generality would not be useful to us at this point.

B.3   Exponential Generating Functions

We say that the expression $\sum_{i=0}^{\infty} a_i \frac{x^i}{i!}$ is the *exponential generating function* for the sequence a_i. It is standard to use *EGF* as a shorthand for exponential generating function. In this context we call the generating function $\sum_{i=0}^{n} a_i x^i$ that we originally studied the *ordinary generating function* for the sequence a_i. You can see why we use the term exponential generating function by thinking about the exponential generating function (EGF) for the all ones sequence,

$$\sum_{i=0}^{\infty} 1 \frac{x^i}{i!} = \sum_{i=0}^{\infty} \frac{x^i}{i!} = e^x,$$

which we also denote by $\exp(x)$. Recall from calculus that the usual definition of e^x or $\exp(x)$ involves limits at least implicitly. We work our way around that by defining e^x to be the power series $\sum_{i=0}^{\infty} \frac{x^i}{i!}$.

Activity 340. Find the EGF (exponential generating function) for the sequence $a_n = 2^n$. What does this say about the EGF for the number of subsets of an n-element set?

Solution. $\sum_{i=0}^{\infty} 2^i \frac{x^i}{i!} = e^{2x}$. It says that the EGF for subsets of an n-element set is e^{2x}.

Activity 341.

(a) Find the EGF (exponential generating function) for the number of ways to paint the n streetlight poles that run along the north side of Main Street in Anytown, USA using four colors.

   Solution. The number of ways to paint the streetlights is 4^n, so the EGF is $\sum_{n=0}^{\infty} 4^n \frac{x^n}{n!} = e^{4x}$.

(b) For what sequence is $\frac{e^x - e^{-x}}{2} = \cosh x$ the EGF (exponential generating function)?

   Solution. For the sequence $\frac{1-(-1)^n}{2}$ which, starting with $n = 0$ is the alternating sequence 0,1,0,1... of zeros and ones.

Activity 342. For what sequence is $\ln(\frac{1}{1-x})$ the EGF? ($\ln(y)$ stands for the natural logarithm of y. People often write $\log(y)$ instead.) Hint: Think of the definition of the logarithm as an integral, and don't worry at this stage whether or not the usual laws of calculus apply, just use them as if they do! We will then define $\ln(1 - x)$ to be the power series you get. [1]

Solution.

$$\ln(\frac{1}{1-x}) = -\ln(1-x) = \int_0^x \frac{1}{1-t} dt$$

$$= \int_0^x (1 + t + t^2 + \cdots) dt$$

$$= \sum_{i=1}^{\infty} \frac{x^i}{i}$$

$$= \sum_{i=1}^{\infty} (i-1)! \frac{x^i}{i!}$$

so the sequence is $a_n = (n - 1)!$.

---

[1] It is possible to define the derivatives and integrals of power series by the formulas

$$\frac{d}{dx} \sum_{i=0}^{\infty} b_i x^i = \sum_{i=1}^{\infty} i b_i x^{i-1}$$

and

$$\int_0^x \sum_{i=0}^{\infty} b_i x^i = \sum_{i=0}^{\infty} \frac{b_i}{i+1} x^{i+1}$$

rather than by using the limit definitions from calculus. It is then possible to prove that the sum rule, product rule, etc. apply. (There is a little technicality involving the meaning of composition for power series that turns into a technicality involving the chain rule, but it needn't concern us at this time.)

Activity 343. What is the EGF for the number of permutations of an n-element set?

Solution. $\frac{1}{1-x}$.

Activity 344. What is the EGF for the number of ways to arrange n people around a round table? Notice that we may think of this as the EGF for the number of permutations on n elements that are cycles.

Solution.   By Problem 342 the EGF is $\ln\frac{1}{1-x}$.

Activity 345. What is the EGF $\sum_{n=0}^{\infty} p_{2n}\frac{x^{2n}}{(2n)!}$ for the number of ways p_{2n} to pair up $2n$ people to play a total of n tennis matches (as in Problems 12 and Activity 44)?

Solution.   Recall that $p_{2n} = (2n-1)(2n-3)\cdots 1 = \frac{(2n)!}{2^n n!}$. Thus

$$\sum_{n=0}^{\infty} p_{2n}\frac{x^{2n}}{(2n)!} = \sum_{n=0}^{\infty}\frac{x^{2n}}{2^n n!} = \sum_{n=0}^{\infty}\frac{(x^2/2)^n}{n!} = e^{x^2/2}.$$

Activity 346. What is the EGF for the sequence $0, 1, 2, 3, \ldots$? You may think of this as the EFG for the number of ways to select one element from an n element set. What is the EGF for the number of ways to select two elements from an n-element set?

Solution. $\sum_{n=0}^{\infty}\frac{nx^n}{n!} = \sum_{n=1}^{\infty}\frac{x^n}{(n-1)!} = x\sum_{n=1}^{\infty}\frac{x^{n-1}}{(n-1)!} = xe^x$.
$\sum_{n=0}^{\infty}\frac{n(n-1)x^n}{2n!} = \sum_{i=2}^{\infty}\frac{x^n}{2(n-2)!} = x^2 e^x/2$.

Activity 347. What is the EGF for the sequence $1, 1, \cdots, 1, \cdots$? Notice that we may think of this as the EGF for the number of identity permutations on an n-element set, which is the same as the number of permutations of n elements that are products of 1-cycles, or as the EGF for the number of ways to select an n-element set (or, if you prefer, an empty set) from an n-element set.

Solution. e^x.

Activity 348. What is the EGF for the number of ways to select n elements from a one-element set? What is the EGF for the number of ways to select a positive number n of elements from a one element set?

Solution. $1 + x$, x.

Activity 349. What is the EGF for the number of partitions of a k-element set into exactly one block? (Hint: is there a partition of the empty set into exactly one block?)

Solution. $e^x - 1$.

Activity 350. What is the EGF for the number of ways to arrange k books on one shelf (assuming they all fit)? What is the EGF for the number of ways to arrange k books on a fixed number n of shelves, assuming that all the books can fit on any one shelf?

Solution. $\frac{1}{1-x}$, $\sum_{k=0}^{\infty}\binom{n+k-1}{k}k!\frac{x^k}{k!} = (1-x)^{-n}$.

B.4 Applications to recurrences.

We saw that ordinary generating functions often play a role in solving recurrence relations. We found them most useful in the constant coefficient case. Exponential generating functions are useful in solving recurrence relations where the coefficients involve simple functions of n, because the $n!$ in the denominator can cancel out factors of n in the numerator.

Activity 351. Consider the recurrence $a_n = na_{n-1} + n(n-1)$. Multiply both sides by $\frac{x^n}{n!}$, and sum from $n = 2$ to ∞. (Why do we sum from $n = 2$ to infinity instead of $n = 1$ or $n = 0$?) Letting $y = \sum_{i=0}^{\infty} a_i x^i$, show that the left-hand side of the equation is $y - a_0 - a_1 x$. Express the right hand side in terms of y, x, and e^x. Solve the resulting equation for y and use the result to get an equation for a_n. (A finite summation is acceptable in your answer for a_n.)

Solution.

$$\sum_{n=2}^{\infty} a_n \frac{x^n}{n!} = \sum_{n=2}^{\infty} a_{n-1} \frac{x^n}{(n-1)!} + \sum_{n=2}^{\infty} \frac{x^n}{(n-2)!}$$

$$= x \left(\sum_{n=0}^{\infty} a_n \frac{x^n}{n!} - a_0 \right) + x^2 \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

$$= x \left(\sum_{n=0}^{\infty} a_n \frac{x^n}{n!} \right) - a_0 x + x^2 e^x$$

We sum from $n = 2$ because otherwise we would have a factorial of a negative number in the denominator. Thus $\sum_{n=0}^{\infty} a_n \frac{x^n}{n!} - a_0 - a_1 x = x \left(\sum_{n=0}^{\infty} a_n \frac{x^n}{n!} \right) - a_0 x + x^2 e^x$, or

$$(1 - x) \sum_{n=0}^{\infty} a_n \frac{x^n}{n!} = a_0 + a_1 x - a_0 x + x^2 e^x.$$

This gives us

$$\sum_{n=0}^{\infty} a_n \frac{x^n}{n!} = \frac{1}{1 - x} (a_0 + a_1 x - a_0 x + x^2 e^x).$$

computing the coefficient of a_n gives us $a_n = a_1 + \sum_{i=0}^{n-2} \frac{1}{i!}$.

Activity 352. The telephone company in a city has n subscribers. Assume a telephone call involves exactly two subscribers (that is, there are no calls to outside the network and no conference calls), and that the configuration of the telephone network is determined by which pairs of subscribers are talking. Give a recurrence for the number c_n of configurations of the network.)Hint: Person n is either on the phone or not.) What are c_0 and c_1? What are c_2 through c_6? Notice that we may think of a configuration of the telephone

network as a permutation that is a product of disjoint two-cycles (and one-cycles[2]), that is, we may think of a configuration as an involution in the symmetric group S_n.

Solution. $c_n = (n-1)c_{n-2} + c_{n-1}$. (The first term counts the number of network configurations in which person n is in a phone call with someone else, and the second term counts the number of network configurations in which person n is not in a phone call.) $c_0 = 1$ and $c_1 = 1$, because there is only one configuration of a network with 0 or one phones.) Then $c_2 = 2$, $c_3 = 2 \cdot 1 + 2 = 4$, $c_4 = 3 \cdot 2 + 4 = 10$, $c_5 = 4 \cdot 4 + 10 = 26$, and $c_6 = 5 \cdot 10 + 26 = 76$.

Activity 353. Recall that a *derangement* of $[n]$ is a permutation of $[n]$ that has no fixed points, or equivalently is a way to pass out n hats to their n different owners so that nobody gets the correct hat. Use d_n to stand for the number of derangements of $[n]$. We can think of derangement of $[n]$ as a list of 1 through n so that i is not in the ith place for any n. Thus in a derangement, some number k different from n is in position n. Consider two cases: either n is in position k or it is not. Notice that in the second case, if we erase position n and replace n by k, we get a derangement of $[n-1]$. Based on these two cases, find a recurrence for d_n. What is d_1? What is d_2? What is d_0? What are d_3 through d_6?

Solution. $d_n = (n-1)d_{n-1} + (n-1)d_{n-2}$. $d_1 = 0$ and $d_2 = 1$. Thus d_0 must be 1 for our recurrence to be valid. (For those familiar with functions as sets of ordered pairs, the empty function is not only a permutation, but it does not map i to i for any integer i, so it is a derangement as well! Thus the definition of a derangement also gives us that $d_0 = 1$.) $d_3 = 2$, $d_4 = 3 \cdot 1 + 3 \cdot 2 = 9$, $d_5 = 4 \cdot 2 + 4 \cdot 9 = 44$, and $d_6 = 5 \cdot 9 + 5 \cdot 44 = 256$.

Using calculus with exponential generating functions

Activity 354. Your recurrence in Problem 352 should be a second order recurrence.

(a) Assuming that the left hand side is c_n and the right hand side involves c_{n-1} and c_{n-2}, decide on an appropriate power of x divided by an appropriate factorial by which to multiply both sides of the recurrence. Using the fact that the derivative of $\frac{x^n}{n!}$ is $\frac{x^{n-1}}{(n-1)!}$, write down a differential equation for the EGF $T(x) = \sum_{i=0}^{\infty} c_i \frac{x^i}{i!}$. Note that it makes sense to substitute 0 for x in $T(x)$. What is $T(0)$? Solve your differential equation to find an equation for $T(x)$.

Solution.

$$\sum_{n=2}^{\infty} c_n \frac{x^{n-1}}{(n-1)!} = \sum_{n=2}^{\infty} (n-1)c_{n-2} \frac{x^{n-1}}{(n-1)!} + \sum_{n=2}^{\infty} c_{n-1} \frac{x^{n-1}}{(n-1)!}$$

---

[2]When we think of writing a permutation as a product of disjoint cycles, we often do not include the one-cycles in our notation because a one-cycle is an identity permutation. However any element not moved by the cycles we do write down is in a one cycle, and those one cycles are implicit in the product of cycles we do write down.

$$\sum_{n=1}^{\infty} c_n \frac{x^{n-1}}{(n-1)!} - c_1 = x \sum_{n=2}^{\infty} c_{n-2} \frac{x^{n-2}}{(n-2)!} + \sum_{n=0}^{\infty} c_n \frac{x^n}{n!} - c_0$$

$$T'(x) = xT(x) + T(x)$$

$T(0) = c_0 = 1$. Then $\frac{T'(x)}{T(x)} = x + 1$, giving $\ln T(x) = x^2/2 + x + k$, and $T(x) = e^k e^{x + x^2/2} = e^{x + x^2/2}$, since $T(0) = 1$.

(b) Use your generating function to compute a formula for c_n.

Solution. $T(x) = \sum_{i=0}^{\infty} (x + x^2/2)^i / i!$. By the binomial theorem, this gives $T(x) = \sum_{i=0}^{\infty} \frac{\sum_{j=0}^{i} \binom{i}{j} x^j \left(\frac{x^2}{2}\right)^{i-j}}{i!} = \sum_{i=0}^{\infty} \frac{\sum_{j=0}^{i} \binom{i}{j} x^{2i-j} 2^{j-i}}{i!}$. Then the coefficient c_n of x^n is the sum over all i and j with $2i - j = n$ and $j \le i$ of $\binom{i}{j} \frac{n!}{i!} 2^{j-i}$. But if $2i - j = n$, then $j = 2i - n$, and if $2i - n \le i$, then $i \le n$, so that $c_n = \frac{n!}{2^n} \sum_{i=0}^{n} \binom{i}{2i-n} \frac{2^i}{i!}$. Note that $\binom{i}{2i-n}$ is the same as $\binom{i}{n-i}$, which is 0 unless $i \ge n/2$, which reduces our sum to $c_n = \frac{n!}{2^n} \sum_{i=\lceil n/2 \rceil}^{n} \binom{i}{n-i} \frac{2^i}{i!}$.

Activity 355. Your recurrence in Problem 353 should be a second order recurrence.

(a) Assuming that the left-hand side is d_n and the right hand side involves d_{n-1} and d_{n-2}, decide on an appropriate power of x divided by an appropriate factorial by which to multiply both sides of the recurrence. Using the fact that the derivative of $\frac{x^n}{n!}$ is $\frac{x^{n-1}}{(n-1)!}$, write down a differential equation for the EGF $D(x) = \sum_{i=0}^{\infty} d_i \frac{x^i}{i!}$. What is $D(0)$? Solve your differential equation to find an equation for $D(x)$.

(b) Use the equation you found for $D(x)$ to find an equation for d_n. Compare this result with the one you computed by inclusion and exclusion.

Solution.

$$\sum_{n=2}^{\infty} d_n \frac{x^{n-1}}{(n-1)!} = \sum_{n=2}^{\infty} d_{n-1} \frac{x^{n-1}}{(n-2)!} + \sum_{n=2}^{\infty} d_{n-2} \frac{x^{n-1}}{(n-2)!}$$

$$\sum_{n=1}^{\infty} d_n \frac{x^{n-1}}{(n-1)!} - d_1 = x \sum_{n=2}^{\infty} d_{n-1} \frac{x^{n-2}}{(n-2)!} + xD(x)$$

$$D'(x) - d_1 = xD'(x) + xD(x)$$

$$D'(x)(1 - x) = xD(x)$$

$$\frac{D'(x)}{D(x)} = \frac{x}{1 - x}$$

This gives us $\ln D(x) = -\ln(1 - x) - x + c$, so that $D(x) = \frac{1}{1-x}e^{-x}e^c$. Since $d_0 = 1$, we have $d(0) = 1$, so $c = 0$ and

$$D(x) = \frac{e^{-x}}{1 - x}$$

$$= \sum_{i=0}^{\infty}(-1)^i\frac{x^i}{i!} \cdot \sum_{j=0}^{\infty}x^j = \sum_{i=0}^{\infty}\left(\sum_{j=0}^{i}\frac{(-1)^j}{j!}\right)x^i.$$

Thus $d_n = n!\sum_{j=0}^{n}\frac{(-1)^j}{j!}$, as we computed by inclusion and exclusion.

B.5   The product principle

One of our major tools for ordinary generating functions was the product principle. It is thus natural to ask if there is a product principle for exponential generating functions. In Problem 350 you likely found that the EGF for the number of ways of arranging n books on one shelf was exactly the same as the EGF for the number of permutations of $[n]$, namely $\frac{1}{1-x}$ or $(1 - x)^{-1}$. Then using our formula from Problem 112 and the generating function for multisets, you probably found that the EGF for number of ways of arranging n books on some fixed number m of bookshelves was $(1 - x)^{-m}$. Thus the EGF for m shelves is a product of m copies of the EGF for one shelf.

Activity 356. In Problem 341 what would the exponential generating function have been if we had asked for the number of ways to paint the poles with just one color of paint? With two colors of paint? What is the relationship between the EGF for painting the n poles with one color of paint and the EGF for painting the n poles with four colors of paint? What is the relationship between the EGF for painting the n poles with two colors of paint and the EGF for painting the poles with four colors of paint?

Solution.   With one color of paint, there would have been one way to paint each pole so our EGF would be $\sum_{n=0}^{\infty}\frac{x^n}{n!}$, or e^x. With two colors of paint, it would be e^{2x} by analogy with the solution to Problem 341. Thus the EGF for two colors of paint would be the square of the EGF for one color of paint. The EGF for four colors of paint is the fourth power of the EGF for one color of paint and the square of the EGF for two colors of paint.

In Problem 352 you likely found that the EGF for the number of network configurations with n customers was $e^{x+x^2/2} = e^x \cdot e^{x^2/2}$. In Problem 347 you saw that the generating function for the number of permutations on n elements that are products of one cycles was e^x, and in Problem 345 you likely found that the EGF for the number of tennis pairings of $2n$ people, or equivalently, the number of permutations of $2n$ objects that are products of n two-cycles is $e^{x^2/2}$.

Activity 357. What can you say about the relationship among the EGF for the number of permutations that are products of disjoint two-cycles and

one-cycles, i.e., are involutions, the exponential generating function for the number of permutations that are the product of disjoint two-cycles only and the generating function for the number of permutations that are the product of disjoint one cycles only?

Solution. The EGF for involutions in S_n is the product of the EGF for the permutations in S_n that are products of $n/2$ two-cycles, and the EGF for permutations in S_n that are products of n one-cycles.

In Problem 355 you likely found that the EGF for the number of permutations of $[n]$ that are derangements is $\frac{e^{-x}}{1-x}$. But every permutation is a product of derangements and one cycles, because the permutation that sends i to i is a one-cycle, so that when you factor a permutation as a product of disjoint cycles, the cycles of size greater than one multiply together to give a derangement, and the elements not moved by the permutation are one-cycles.

Activity 358. If we multiply the generating function for derangements times the generating function for the number of permutations that are products of one-cycles only, what EGF for what set of permutations do we get? (Notice that there are two questions here.)

Solution. We get the EGF $\frac{1}{1-x}$ for all permutations of $[n]$. Notice that any permutation is a product of a derangement of the elements not fixed by the permutation times the one cycles of elements that are fixed by the permutation.

We now have four examples in which the EGF for a sequence or a pair of objects is the product of the EGFs for the individual objects making up the sequence or pair.

Activity 359. What is the coefficient of $\frac{x^n}{n!}$ in the product of two EGFs $\sum_{i=0}^{\infty} a_i \frac{x^i}{i!}$ and $\sum_{j=0}^{\infty} b_j \frac{x^j}{j!}$? (A summation sign is appropriate in your answer.)

Solution.
$$\sum_{i,j:\ i+j=n} n! \frac{a_i}{i!} \frac{b_j}{j!},$$
which can be better written as
$$\sum_{i=0}^{n} \frac{n!}{i!(n-i)!} a_i b_{n-i} = \sum_{i=0}^{n} \binom{n}{i} a_i b_{n-i}.$$

Our product principle for ordinary generating functions involved the idea of a value function. In particular, if we have a set S of objects we call a function v from S into the nonnegative integers a *value function*. Our combinatorial interpretation of the product of ordinary generating functions in Problem 167 is the following theorem.

Theorem B.5.1. *Suppose we have a set S with a value function v from S into the nonnegative integers and a set T with a value function u from T into*

the nonnegative integers. If a_i is the number of objects s in S with value $v(s) = i$ and b_j is the number of objects t in T with value $u(t) = j$, and c_n is the number of ordered pairs (s,t) of objects in $S \times T$ with total value $v(s) + u(t) = n$, then the ordinary generating function for c_n is the product of the ordinary generating function for a_i and the ordinary generating function for b_j.

We ask if there is a similar interpretation for the products of exponential generating functions we have just seen. In the case of painting streetlight poles in Problem 356, what is important is not just how many light poles are painted with each color, but which set of poles is painted with each color. Let us examine the relationship between the EGF for painting poles with two colors, e^{2x} and the EGF for painting poles with four colors, e^{4x}. To be specific, the EGF for painting poles red and white is e^{2x} and the EGF for painting poles blue and yellow is e^{2x}. To decide how to paint poles with red, white, blue, and yellow, we can decide which set of poles is to be painted with red and white, and which set of poles is to be painted with blue and yellow. Notice that the number of ways to paint a set of poles with red and white depends only on the size of that set, and the number of ways to paint a set of poles with blue and yellow depends only on the size of that set. (It is a coincidence that the number of ways to paint a set of poles with red and white equals the number of ways to paint the same set of poles with blue and yellow. The coincidence is the result of trying to keep our example simple!)

Activity 360. Suppose that a_i is the number of ways to paint a set of i poles with red and white, and b_j is the number of ways to paint a set of j poles with blue and yellow. In how many ways may we take a set N of n poles, divide it up into two sets I and J (using i to stand for the size of I and j to stand for the size of the set J, and allowing i and j to vary) and paint the poles in I red and white and the poles in J blue and yellow? (Don't figure out formulas for a_i and b_j to use in your answer; that will make it harder to get the point of the problem!) How does this relate to Problem 359?

Solution. $\binom{n}{i}a_i b_j$. This shows that the coefficient of $\frac{x^n}{n!}$ in the EGF for painting poles with four colors is the coefficient of $\frac{x^n}{n!}$ in the product of the EGF for painting poles with two colors and the EGF for painting poles with two colors.

Problem 360 shows that the formula you got for the coefficient of $\frac{x^n}{n!}$ in the product of two EGFs is the formula we get by splitting a set N of poles into two parts and painting the poles in the first part with red and white and the poles in the second part with blue and yellow. More generally, you could interpret your result in Problem 359 to say that the coefficient of $\frac{x^n}{n!}$ in the product $\sum_{i=0}^{\infty} a_i \frac{x^i}{i!} \sum_{j=0}^{\infty} b_j \frac{x^j}{j!}$ of two EGFs is the sum, over all ways of splitting a set N of size n into an ordered pair of disjoint sets I and J, of the product $a_{|I|} b_{|J|}$. In contrast, when we multiplied ordinary generating functions, the coefficient of x^n in $\sum_{i=0}^{\infty} a_i x^i \sum_{j=0}^{\infty} b_j x^j$ is the sum of all $a_i b_j$ over ordered pairs of integers i and j with $i + j = n$. In our combinatorial

interpretation of the product of two ordinary generating functions, we had two sets S and T of objects, a_i was the number of objects in S of value i, b_j was the number of objects in T of value j, and $i + j$ was the total value of an ordered pair of objects, one from S and one from T. In painting poles for streetlights, what was important was not only the number of poles we selected to paint red and white, but the actual *set* of poles we selected to paint red and white. This suggests that the value of a selection in this context ought to be a set instead of simply an integer, even though the size of the set will still play a role. For example, the number of ways of paint the poles in a set red and white depends only on the size of the set, and not which set we choose, but in forming the product of the exponential generating functions, we have to analyze all ways of choosing a set of that size. This suggests that to get a combinatorial interpretation of the product of EGFs we should consider a different kind of value function, one whose values are sets and not integers. Thus what we want to consider is a "set-valued value function." Since this is a mouthful to say, we will shorten it in the definition that follows. We define a *set-value function* from a set S to a set Y to be a function V from S to the set of subsets of Y such that for all subsets I of Y of the same size, the number of elements s of S whose value $V(s)$ is I is the same. For example if S is all ways of painting some of the street light poles on the north side on Main Street using red and white, and if $V(s)$ is the set of poles actually painted red or white, then the number of ways to paint some of the poles red and white depends on the size of the set of poles being painted, so for each set of poles of a given size, the number of ways to paint that set of poles using red and white is the same.

Activity 361. In Problem 350, why is the set of books that we actually put onto a shelf a set-value function on the set of all ways to put some of the books on that shelf?

Solution. Because the number of ways to put a set S of books onto a shelf is the same (namely $|S|!$) for all sets S of the same size.

Activity 362. In Problem 352, why is the set of people actually using their phones a set-value function (assume nobody is calling outside the telephone network)? Equivalently, given an involution, that is, a permutation that is a product of two cycles and one cycles, why is the set of elements that are actually in two-cycles a set-value function? Why is the set of people who are not using their phones (or, equivalently, the set of elements in a product of two-cycles and one-cycles that are in one-cycles) a set-value function?

Solution. Because the number of ways to break a given set of $2n$ people into two-cycles depends only on n and not the particular set of $2n$ people we choose. The number of ways to break up a set of size n into one-cycles is one, so it doesn't depend on which set of size n we are breaking up. (In fact it doesn't depend on n either, but that is irrelevant here.)

Activity 363. If S is a set of objects with V a set-value function from S to some set Y and T is a set of objects with U a set-value function from T to the

same set Y, then what is the relationship among the EGF for the number a_i of elements of S whose set-value is any one particular set of size i, the EGF for the number b_j of elements of T whose set-value is any one particular set of size j and the EGF for the number c_n of ordered pairs (s, t) in $S \times T$ such that the set values of s and t are disjoint sets whose union is any one particular set of size n?

Solution.   The EGF for a_i times the EGF for b_j is the EGF for c_n. Stated as a theorem,

Solution-Theorem   If S is a set of objects with V a set-value function from S to some set Y and the EGF for the number a_i of elements of S whose V-value is any particular set N of size n is $f(x) = \sum_{n=0}^{\infty} a_n \frac{x^n}{n!}$ and T is a set of objects with U a set-value function from T to the same set Y and the EGF for the number of elements b_n of T whose U-value is any particular set N of size n is $g(x) = \sum_{i=0}^{\infty} b_n \frac{x^n}{n!}$, then the EGF for the number c_n of ordered pairs (s, t) in $S \times T$ such that the set values of s and t are disjoint sets whose union is any one particular set of size n is $f(x)g(x)$. That is, the number of ordered pairs of the type described is the coefficient of $\frac{x^n}{n!}$ in $f(x)g(x)$.

   The proof consists of interpreting Problem 359 in this context.

   The theorem you proved in Problem 363 is called the *product principle for exponential generating functions*.

Activity 364. Use the product principle for EGFs to explain the results of Problems 357 and Activity 358.

Solution.   Every involution is a unique product of disjoint two-cycles and one-cycles, so we can think of it as an ordered pair of whose first entry is a set of disjoint two-cycles and whose second element is a set of disjoint one-cycles. The value of a set of two-cycles is the union of their support sets and the value of a set of one-cycles is the union of their support sets. Then the EGF for involutions is the product of the EGF for permutations that are the product of only disjoint two-cycles and the EGF for permutations that are the product of only disjoint one-cycles (i.e. identity permutations).

   We noted in the solution to Problem 358 that every permutation is the product of a derangement (on the elements that are not fixed by the permutation) and a product of one cycles (on the elements that are fixed by the permutation). Thus we can think of every permutation as an ordered pair consisting of a derangement and a product of one-cycles, and the product principle tells us that the EGF for all permutations is the product of the EGF for derangements and the EGF for identity permutations.

   The product principle for EGFs has a natural extension to a product of some arbitrary number k of exponential generating functions. Instead of dealing with ordered pairs, it deals with (ordered) k-tuples. Since it is inconvenient to state unless one is careful with notation, we will state it here. The proof would be quite similar to your proof in Problem 363.

Theorem B.5.2. *If $f_1(x)$, $f_2(x)$, ..., $f_k(x)$ are the exponential generating functions for sets S_1, S_2, ..., S_k according to the set-value functions V_1, V_2, ..., V_k, then $f_1(x)f_2(x)\cdots f_k(x)$ is the exponential generating function for the number c_n of (ordered) k-tuples $(s_1, s_2, \ldots, s_k)$ with $s_i \in S_i$ such that $V_1(s_1)$, $V_2(s_2)$, ..., $V_k(s_k)$ are disjoint sets whose union is any one particular set N of size n.*

Corollary B.5.3. *If $f(x)$ is the exponential generating function for a set S according to the set-value function V, then $f(x)^k$ is the exponential generating function in which a_n is the number of k-tuples of elements of S whose values are disjoint sets whose union is any particular set of size n.*

Activity 365. Use the general product principle for EGFs or its corollary to explain the relationship between the EGF for painting streetlight poles in only one color and the EGF for painting streetlight poles in 4 colors in Problems 341 and Activity 356. What is the generating function for the number p_n of ways to paint n streetlight poles with some fixed number k of colors of paint.

Solution. We can think of a painting of a set of street poles as a four-tuple of sets, the sets painted each of the four colors. Then Corollary B.5.3 tells us that the EGF for such four-tuples is the fourth power of the EGF for the number of ways to paint streetlight poles with one color. The EGF for painting streetlight poles with k colors of paint is e^{kx}.

Activity 366. Use the general product principle for EGFs or its corollary to explain the relationship between the EGF for arranging books on one shelf and the EGF for arranging books on n shelves in Problem 350.

Solution. An arrangement of books on n shelves may be thought of as a n-tuple of arrangements of books on one shelf, and the value of an arrangement may be taken to be the set of books actually on the shelf (or shelves). The Corollary B.5.3 tells us that the EGF is the nth power of the EGF for arranging books on one shelf, which is the EGF for permutations. Thus the EGF for arranging books on n shelves is $(1 - x)^{-n}$.

Activity 367.

(a) (Optional) Our very first example of exponential generating functions used the binomial theorem to show that the EGF for k-element permutations of an n element set is $(1 + x)^n$. Use the EGF for k-element permutations of a one-element set and the product principle to prove the same thing. Hint: Review the alternate definition of a function in Section .

Solution. In Section we remarked that an alternate definition of a function from S to T is that it is an assignment of disjoint subsets of S to elements of T so that the union of the subsets is S. Thus a function from $[k]$ to $[n]$ may be thought of as an n-tuple of disjoint subsets of S whose union is $[k]$. In particular, an injection from $[k]$ to $[n]$ (which is a k-element permutation of $[n]$) can be thought of as a n-tuple of disjoint

singleton sets and empty sets whose union is n. The number of such n-tuples is therefore the number of k-element permutations of $[n]$. If $n = 1$, the possible n-tuples are $(\emptyset)$ and $\{1\}$, and so the EGF for such n-tuples is $1 + x$. Then by Corollary B.5.3, if we let the value of an n-tuple of disjoint singletons and empty sets be the union of the sets in the n-tuple, the EGF for the number of n-tuples whose value is $[k]$ is $(1 + x)^n$. Thus this is the EGF for k-element permutations of $[n]$.

(b) What is the EGF for the number of ways to paint n streetlight poles red, white blue, green and yellow, assuming an even number of poles must be painted green and an even number of poles must be painted yellow?

Solution. By the product principle for exponential generating functions it is

$$e^{2x}\left(1 + \frac{(2x)^2}{2!} + \frac{(2x)^4}{4!} + \cdots\right) = e^{2x}\frac{e^{2x} - e^{-2x}}{2} = e^{2x}\cosh(2x).$$

Activity 368. What is the EGF for the number of functions from an n-element set onto a one-element set? (Can there be any functions from the empty set onto a one-element set?) What is the EGF for the number c_n of functions from an n-element set onto a k element set (where k is fixed)? Use this EGF to find an explicit expression for the number of functions from a k-element set onto an n-element set and compare the result with what you got by inclusion and exclusion.

Solution. There are no onto functions from a 0-element set to a 1-element set; otherwise there is exactly one function from an n-element set onto a one element set so the EGF for functions from an n-element set onto a one element set is $e^x - 1$. A function from an n-element set onto a k-element set may be thought of as a $k - tuple$ of functions from disjoint subsets whose union is the n-element set onto the one element subsets of the k-element set. Therefore by the corollary to the product principle for EGFs the EGF for functions from an n-element set onto a k-element set is $(e^x - 1)^k$. By the binomial theorem, this is

$$\sum_{i=0}^{k}\binom{k}{i}(-1)^{k-i}e^{ix} = \sum_{i=0}^{k}\binom{k}{i}(-1)^{k-i}\sum_{j=0}^{\infty}\frac{(ix)^j}{j!} = \sum_{i=0}^{k}(-1)^{k-i}\binom{k}{i}\sum_{j=0}^{\infty}i^j\frac{(x)^j}{j!}.$$

Thus $c_n = \sum_{i=0}^{k}(-1)^{k-i}\binom{k}{i}n^i$, which is consistent with the formula we got by inclusion and exclusion.

Activity 369. In Problem 127 You showed that the Bell Numbers B_n satisfy the equation $B_{n+1} = \sum_{k=0}^{n}\binom{n}{k}B_{n-k}$ (or a similar equation for B_n.) Multiply both sides of this equation by $\frac{x^n}{n!}$ and sum from $n = 0$ to infinity. On the left hand side you have a derivative of a certain EGF we might call $B(x)$. On

the right hand side, you have a product of two EGFs, one of which is $B(x)$. What is the other one? What differential equation involving $B(x)$ does this give you. Solve the differential equation for $B(x)$. This is the EGF for the Bell numbers!.

Solution.

$$\sum_{n=0}^{\infty} B_{n+1} \frac{x^n}{n!} = \sum_{n=0}^{\infty} \binom{n}{k} B_{n-k} \frac{x^n}{n!} = \sum_{i=0}^{\infty} B_i \frac{x^i}{(i)!} \sum_{j=0}^{\infty} \frac{x^j}{j!}.$$

Thus $B'(x) = B(x)e^x$, which gives us $\ln B(x) = e^x + c$, or $B(x) = e^{ex+c} = e^c e^{(e^x)}$. Since $B_0 = B(0)$ and $B_0 = 1$, we have $c = -1$ and $B(x) = \exp(e^x - 1)$.

Activity 370. Prove that $n2^{n-1} = \sum_{k=1}^{n} \binom{n}{k} k$ by using EGFs.

Solution. By the product principle for EGFs, the EGF for the right hand side is

$$e^x \cdot xe^x = xe^{2x} = x \sum_{i=o}^{\infty} \frac{(2x)^i}{i!} = \sum_{j=1}^{\infty} 2^{j-1} \frac{x^j}{(j-1)!} = \sum_{j=1}^{\infty} j2^{j-1} \frac{x^j}{j!}.$$

Thus the coefficient of $\frac{x^n}{n!}$ is $n2^{n-1}$, as well as $\sum_{k=1}^{n} \binom{n}{k} k$.

Activity 371. In light of Problem 349, why is the EGF for the Stirling numbers $S(n,k)$ of the second kind not $(e^x - 1)^n$? What is it equal to instead?

Solution. Notice that a one block partition is the same thing as a function from that block onto a one element set. However a partition with n blocks is not an n-tuple of blocks, but rather a set of n blocks. An n-tuple of blocks corresponds to a function from the union of the blocks onto an n-element set, and $n!$ different onto functions correspond to the same partition into n blocks. Thus the EGF for partitions of an n-element set into k parts (where n is fixed but k varies) is $\frac{1}{n!}(e^x - 1)^n$.

The idea of the set-value function in the product principle for exponential generating functions helps to resolve a mystery that you may have been consciously or unconsciously wondering about. How do we know when ordinary generating functions will be most useful and how do we know when exponential generating functions will be most useful? When we are in a situation—such as distributing fruit to children, or partitioning an integer—where a combinatorial structure is determined by the number of objects of a certain type (for example a partition of an integer is determined by the number of ones, the number of twos and so on), ordinary generating functions are most likely to be useful. However when what determines a combinatorial structure is the set of objects, or some structure (such as a permutation) on a set of objects, then exponential generating functions are most likely to be useful. In particular, in situations where our structure comes with labels, then exponential generating functions are likely to be useful. However there is a grey area. When we are distributing identical candy to children, the children are labelled (they have

names), but the candy is not. Experience tells us, though that since the candy is not labelled, ordinary generating functions are useful. So while the question of whether the most natural value function seems to be an integer value or a set-value is a good guideline, in the end experience helps tremendously!

B.6   The Exponential Formula

Exponential generating functions turn out to be quite useful in advanced work in combinatorics. One reason why is that it is often possible to give a combinatorial interpretation to the composition of two exponential generating functions. In particular, if $f(x) = \sum_{i=0}^{n} a_i \frac{x^i}{i!}$ and $g(x) = \sum_{j=1}^{\infty} b_j \frac{x^j}{j!}$, it makes sense to form the composition $f(g(x))$ because in so doing we need add together only finitely many terms in order to find the coefficient of $\frac{x^n}{n!}$ in $f(g(x))$ since in the EGF $g(x)$ the dummy variable j starts at 1. Since our study of combinatorial structures has not been advanced enough to give us applications of a general formula for the composition of the EGF, we will not give here the combinatorial interpretation of this composition. However we have seen some examples where one particular composition can be applied. Namely, if $f(x) = e^x = \exp(x)$, then $f(g(x)) = exp(g(x))$ is well defined when $b_0 = 0$. We have seen three examples in which an EGF is $e^{f(x)}$ where $f(x)$ is another EGF. There is a fourth example in which the exponential function is slightly hidden.

Activity 372. If $f(x)$ is the EGF for the number of partitions of an n-set into one block, and $g(x)$ is the EGF for the total number of partitions of an n-element set, that is, for the Bell numbers B_n, how are the two generating functions related?

Solution. The EGF for one-block partitions is $e^x - 1$ and for the Bell numbers is $\exp(e^x - 1)$, and so the EGF for the Bell numbers is composition of the exponential function with the EGF for one block partitions.

Activity 373. Let $f(x)$ be the EGF for the number of permutations of an n-element set with one cycle of size one or two. What is $f(x)$? What is the EGF $g(x)$ for the number of permutations of an n-element set all of whose cycles have size one or two, that is, the number of involutions in S_n? How are these two exponential generating functions related?

Solution. There is one permutation with one cycle of size 1, and one permutation with one cycle of size 2. Therefore the EGF for such permutations is $x + \frac{x^2}{2!} = x + x^2/2$. The EGF for involutions is $e^{x+x^2/2}$. Thus $g(x) = \exp(f(x))$.

Activity 374. Let $f(x)$ be the EGF for the number of permutations of an n-element set that have exactly one two-cycle and no other cycles. Let $g(x)$ be the EGF for the number of permutations which are products of two-cycles only, that is, for tennis pairings. (That is, they are not a product of two-cycles and a nonzero number of one-cycles). What is $f(x)$? What is $g(x)$? How are these to exponential generating functions related?

Solution. The EGF $f(x)$ for permutations of an n element set that have exactly one two-cycle (and no other cycles) is $\frac{x^2}{2!}$ or $x^2/2$. By Problem 345, the EGF to permutations that are products of two-cycles only is $\exp(x^2/2)$. Thus $g(x) = \exp(f(x)) = e^{x^2/2}$.

Activity 375. Let $f(x)$ be the EGF for the number of permutations of an n-element set that have exactly one cycle. (This is the same as the EGF for the number of ways to arrange n people around a round table.) Let $g(x)$ be the EGF for the total number of permutations of an n-element set. What is $f(x)$? What is $g(x)$? How are $f(x)$ and $g(x)$ related?

Solution. In Problem 344 we showed that $f(x) = \ln\left(\frac{1}{1-x}\right)$. In Problem 343 we showed that $g(x) = \frac{1}{1-x}$. Therefore $g(x) = \exp(f(x))$.

There was one element that our last four problems had in common. In each case our EGF $f(x)$ involved the number of structures of a certain type (partitions, telephone networks, tennis pairings, permutations) that used only one set of an appropriate kind. (That is, we had a partition with one part, a telephone network consisting either of one person or two people connected to each other, a tennis pairing of one set of two people, or a permutation with one cycle.) Our EGF $g(x)$ was the number of structures of the same "type" (we put type in quotation marks here because we don't plan to define it formally) that could consist of any number of sets of the appropriate kind. Notice that the order of these sets was irrelevant. For example we don't order the blocks of a partition and a product of disjoint cycles is the same no matter what order we use to write down the product. Thus we were relating the EGF for structures which were somehow "building blocks" to the EGF for structures which were sets of building blocks. For a reason that you will see later, it is common to call the building blocks *connected* structures. Notice that our connected structures were all based on nonempty sets, so we had no connected structures whose value was the empty set. Thus in each case, if $f(x) = \sum_{i=0}^{\infty} a_i \frac{x^i}{i!}$, we would have $a_0 = 0$. The relationship between the EGFs was always $g(x) = e^{f(x)}$. We now give a combinatorial explanation for this relationship.

Activity 376. Suppose that S is a set with a set-value function V from S to a set Y such that no element of S has the empty set as a value. Let $f(x)$ be the generating function for S according to the value function V.

(a) In the power series

$$e^{f(x)} = 1 + f(x) + \frac{f(x)^2}{2!} + \cdots + \frac{f(x)^k}{k!} + \cdots = \sum_{k=0}^{\infty} \frac{f(x)^k}{k!},$$

the product principle tells us that the coefficient of $\frac{x^n}{n!}$ in $f(x)^k$ is the number of k-tuples of elements of S whose values are disjoint sets whose union is any one particular subset N of Y of size n. How do you know that all of the elements of the k-tuple are different?

> **Solution.** Because the coefficient of x^0 is 0, there are no elements in S whose value is the empty set.

(b) When you divide the coefficient of $\frac{x^n}{n!}$ in $f(x)^k$ by $k!$, it no longer counts k-tuples whose values are disjoint sets whose union is that set N. What does it count instead? (Hint: how many k-tuples can you form with a set of k distinct elements?)

> **Solution.** It counts k-element subsets since the entries of the k-tuples are distinct.

(c) Why does this prove that the coefficient of $\frac{x^n}{n!}$ in $e^{f(x)}$ is the number of subsets of S such that the values of th elements of the subset partition any one particular set $N \subseteq T$ of size n?

> **Solution.** We know that the coefficient of $\frac{x^n}{n!}$ in $e^{f(x)}$ is the sum over all k of the number of k-element subsets of S whose values partition a particular set N. (The values must partition N because none of them are empty, they are disjoint, and their union is N.) This sum is the number of subsets of S such that the values of the elements of the subset partition a particular set N of size n.

In Problem 376 we proved the following theorem, which is called the *exponential formula*.

Theorem B.6.1. *Suppose that S is a set with a set-value function V from it to a set Y such that no element of S has the empty set as a value. Let $f(x)$ be the generating function for S according to the value function V. Then the coefficient of $\frac{x^n}{n!}$ in $e^{f(x)}$ is the number of subsets of S such that the values of the elements of the subset partition any one particular set $N \subseteq T$ of size n.*

Since the statement of the theorem is rather abstract, let us see how it applies to the examples in Problems 372, Activity 373, Activity 374 and Activity 375. In Problem 349 our set S should consist of one-block partitions of sets; that is, it should be a set of nonempty sets. The value of a one-block partition will be the set it partitions, and we want every subset (of a given size) of our set Y to be the value of the same number of partitions. Thus we can take S to be the set of all nonempty finite subsets of the positive integers and take Y to be the set of positive integers. Since a partition of a set is a set of blocks whose union is S, a one block partition whose block is B is the set $\{B\}$. We take $V(\{B\}) = B$. Then any nonempty finite subset of of the positive integers is the value of exactly one member of S, and the empty set and all infinite subsets of the positive integers are the value of exactly zero members of S. Thus we have a value function. As you showed in Problem 349 the generating function for partitions with just one block is $e^x - 1$. Thus by

the exponential formula, $\exp(e^x - 1)$ is the generating function for sets of sub-sets of the positive integers whose values are disjoint sets whose union is any particular set N of size n. Since the values are nonempty sets, this means they partition the set N. Thus $\exp(e^x - 1)$ is the generating function for partitions of sets of size n. (Technically, it is the generating function for partitions of subsets of the integers of size n, but any two n-element sets have the same number of partitions.)

Activity 377. Explain how the exponential formula proves the relationship we saw in Problem 375.

Solution. We take S to be the set of permutations of sets of positive integers that have just one-cycle. We take the value of a one-cycle permutation to be the support set of the cycle. The number of one-cycle permutations whose value is any given n-element set of positive integers is $(n-1)!$, so we have defined a value function. We take the set Y to be the set of positive integers. Then by the Exponential Formula, if $f(x)$ is the EGF for permutations with one cycle, $\exp(f(x)) = g(x)$ is the EGF in which the coefficient of $\frac{x^n}{n!}$ is the number of sets of permutations with one cycle whose support sets partition any given set N of size n. That is, the coefficient of $\frac{x^n}{n!}$ in $g(x)$ is the the number of products of disjoint cycles which partition any given set N; that is, it is the EGF for permutations of N.

Activity 378. Explain how the exponential formula proves the relationship we saw in Problem 374.

Solution. We let S be the set of permutations of sets of positive integers that consist of exactly one two-cycle. We take the set Y to be the set of positive integers. We take the value of a two-cycle to be its support set. Since the number of two-cycles with a given support set is one, this is a value function. We saw in Problem 374 that the EGF for S with this value function is $x^2/2$. By the exponential formula, the EGF for sets of disjoint two-cycles is $\exp(x^2/2)$. But a sets of disjoint two-cycles correspond bijectively with permutations of subsets of the positive integers which are products of disjoint two-cycles (only), and this confirms the result of Problem 345.

Activity 379. Explain how the exponential formula proves the relationship we saw in Problem 373.

Solution. We let the set S be the set of permutations of subsets of the positive integers that consist of a one-cycle or consist of a two-cycle. We let the set Y be the set of positive integers. We let the value of a permutation in S be the support set of its cycle. We saw in problem 373 that the EGF for S is $x + x^2/2$. By the exponential formula, the EGF for sets of disjoint one and two-cycles is $e^x + x^2/2$. But there is a bijection between the sets of disjoint one and two-cycles and permutations that are a product of disjoint one and two-cycles. This confirms the result of Problem .

Activity 380. In Problem 341 we saw that the generating function for the number of ways to use four colors of paint to paint n light poles along the north

side of Main Street in Anytown was e^{4x}. We should expect an explanation of this generating function using the exponential formula. Let S be the set of all ordered pairs consisting of a light pole and a color. Thus a given light pole occurs in four ordered pairs. What is a natural set-value function on S? What is the exponential generating function $f(x)$ for S according to that value? Assuming that there is no upper limit on the number of light poles, what subsets of S does the exponential formula tell us are counted by the coefficient of x^n in $e^{f(x)}$? How do the sets being counted relate to ways to paint light poles?

Solution. A natural set-value function on S assigns to each ordered pair the singleton set consisting of the light pole in the pair. The EGF for S according to that value is $4x$, because there are four ordered pairs with any given value. Note that a set of ordered pairs whose values partition a set N of light poles N is exactly a function from the set N of light poles to the set of colors. Then by the exponential formula, the generating function for the number of ways to paint n light poles with four colors is e^{4x}.

One of the most spectacular applications of the exponential formula is also the reason why, when we regard a combinatorial structure as a set of building block structures, we call the building block structures *connected*. In Chapter 2 we introduced the idea of a connected graph and in Problem 96 we saw examples of graphs which were connected and were not connected. A subset C of the vertex set of a graph is called a *connected component* of the graph if

- every vertex in C is connected to every other vertex in that set by a walk whose vertices lie in C, and

- no other vertex in the graph is connected by a walk to any vertex in C.

In Problem 224 we showed that each connected component of a graph consists of a vertex and all vertices connected to it by walks in the graph.

Activity 381.

Activity 382. Explain why no edge of the graph connects two vertices in different connected components.

Solution. If an edge connected two vertices in different connected components, that edge would give a walk from a vertex in one of the connected components to a vertex in the other connected component, and thus not in the first component, violating the second part of the definition of a connected component.

Activity 383. Explain why it is that if C is a connected component of a graph and E' is the set of all edges of the graph that connect vertices in C, then the graph with vertex set C and edge set E' is a connected graph. We call this graph a *connected component graph* of the original graph.

Solution. If there is a walk between two vertices in a connected component, all edges of the walk must connect two vertices in the connected component,

because if there were an edge in the walk that did not do so, it would violate the second part of the definition of a connected component.

The last sequence of problems shows that we may think of any graph as the set of its connected component graphs. (Once we know them, we know all the vertices and all the edges of the graph). Notice that a graph is connected if and only if it has exactly one connected component. Since the connected components form a partition of the vertex set of a graph, the exponential formula will relate the generating function for the number of connected graphs on n vertices with the generating function for the number of graphs (connected or not) on n vertices. However because we can draw as many edges as we want between two vertices of a graph, there are infinitely many graphs on n vertices, and so the problem of counting them is uninteresting. We can make it interesting by considering *simple graphs*, namely graphs in which each edge has two distinct endpoints and no two edges connect the same two vertices. It is because connected graphs form the building blocks for viewing all graphs as sets of connected components that we refer to the building blocks for structures counted by the generating functions in the exponential formula as connected structures.

Activity 384. Suppose that $f(x) = \sum_{n=0}^{\infty} c_n \frac{x^n}{n!}$ is the exponential generating function for the number of simple connected graphs on n vertices and $g(x) = \sum_{i=0}^{\infty} a_i \frac{x^i}{i!}$ is the exponential generating function for the number of simple graphs on i vertices.

(a) Is $f(x) = e^{g(x)}$, is $f(x) = e^{g(x)-1}$, is $g(x) = e^{f(x)-1}$ or is $g(x) = e^{f(x)}$?

    Solution.   To apply the exponential formula, we must take the exponential function of an EGF whose constant term is zero, or in other words for a set of elements whose values are nonempty sets. We can let S be the set of nonempty connected graphs. (Technically, the graph with the empty set of vertices and the empty set of edges is connected.) Therefore $f(x) - 1$ is the EGF for S. By the exponential formula, $g(x) = e^{f(x)-1}$ because a simple graph may be thought of as a set of simple connected graphs, namely its connected component graphs. (Note that $g(x)$ has 1 for its constant term, which corresponds to thinking of the empty graph as having an empty set of nonempty connected components.)

(b) One of a_i and c_n can be computed by recognizing that a simple graph on a vertex set V is completely determined by its edge set and its edge set is a subset of the set of two element subsets of V. Figure out which it is and compute it.

    Solution.   To specify a simple graph on a vertex set V, we have to specify its set of edges. The possible sets of edges thus correspond bijectively to sets of two-element subsets of V. But if V has size i the

set of all two element subsets of V has $\binom{i}{2}$ elements. Thus the number of sets of two element subsets of V is $2^{\binom{i}{2}}$. Therefore $a_i = 2^{\binom{i}{2}}$.

(c) Write $g(x)$ in terms of the natural logarithm of $f(x)$ or $f(x)$ in terms of the natural logarithm of $g(x)$.

Solution. Since $g(x) = e^{f(x)-1}$, $f(x) = 1 + \ln g(x)$.

(d) Write $\log(1 + y)$ as a power series in y.

Solution.

$$\log(1 + y) = \int_0^y \frac{1}{1+x} dx = \int_0^y \sum_{i=0}^{\infty} (-1)^i x^i = \sum_{j=1}^{\infty} (-1)^{j-1} \frac{y^j}{j}.$$

(e) Why is the coefficient of $\frac{x^0}{0!}$ in $g(x)$ equal to one? Write $f(x)$ as a power series in $g(x) - 1$.

Solution.   The coefficient of $\frac{x^0}{0!}$ is 1 because there is one graph on the empty set; the one with no edges.

$$f(x) = 1 + \ln(1 + (g(x) - 1)) = 1 + \sum_{j=1}^{\infty} (-1)^{j-1} \frac{(g(x) - 1)^j}{j}.$$

(f) You can now use the previous parts of the Problem to find a formula for c_n that involves summing over all partitions of the integer n. (It isn't the simplest formula in the world, and it isn't the easiest formula in the world to figure out, but it is nonetheless a formula with which one could actually make computations!) Find such a formula.

Solution.

$$f(x) = 1 + \sum_{j=1}^{\infty} (-1)^{j-1} \frac{(g(x) - 1)^j}{j} = 1 + \sum_{j=1}^{\infty} (-1)^{j-1} \frac{(\sum_{i=1}^{\infty} 2^{\binom{i}{2}} \frac{x^i}{i!})^j}{j}.$$

From the right-hand expression, we get a term involving x^n whenever we have an x^n term in the jth power of $\sum_{i=1}^{\infty} 2^{\binom{i}{2}} \frac{x^i}{i!}$. So the coefficient of $\frac{x^n}{n!}$ is the sum over all j and all sequences $i_1, i_2, \ldots, i_j$ with $i_1 + i_2 + \cdots + i_j = n$ of terms of the form

$$\frac{n!}{j} (-1)^j \prod_{k=1}^{j} \frac{2^{\binom{i_k}{2}}}{i_k!},$$

where each $i_k > 0$. Notice that reordering the numbers i_1, i_2, ... i_k does not change the value of the expression. The sequence of i_ks is a composition of n into positive parts. If we knew how many compositions of n into j parts correspond to one partition of n into j parts, we could sum over a much smaller set of terms. If we use the type vector notation for a partition, namely that it has p_1 parts of size 1, p_2 parts of size 2, ..., p_n parts of size n, then the number of compositions corresponding to that partition, i.e. the number of compositions with the type vector $(p_1, p_2, \ldots, p_n)$ is the number of ways to take j places in a vector and label p_1 of them with 1, p_2 of them with 2, and so on until we label p_n of them with n. This number is the multinomial coefficient $\binom{j}{p_1, p_2, \ldots, p_n}$. Thus our sum over all j and all compositions of n into j parts becomes

$$\sum_{j=1}^{n} \frac{n!}{j}(-1)^j \sum_{\substack{p_1, p_2, \ldots, p_n : \sum_{r=1}^{n} r p_r = n \\ \text{and} \sum_{r=1}^{n} p_r = j}} \binom{j}{p_1, p_2, \ldots, p_n} \prod_{r=1}^{n} \frac{2^{\binom{r}{2}p_r}}{(r!)^{p_r}}.$$

We can remove one of the summation signs and the condition that $\sum_{r=1}^{n} p_r = j$ by substituting $\sum_{r=1}^{n} p_r$ for j, and we get

$$\sum_{\substack{p_1, p_2, \ldots, p_n : \sum_{r=1}^{n} r p_r = n}} n!(-1)^{\sum_{r=1}^{n} p_r} \left(\sum_{r=1}^{n} p_r - 1\right)! \prod_{r=1}^{n} \frac{2^{\binom{r}{2}p_r}}{(r!)^{p_r} p_r!}$$

for the number of connected graphs on n vertices. If we want to shorten the appearance of the formula we can keep j in our sum and explain its value afterwards, as in

$$\sum_{\substack{p_1, p_2, \ldots, p_n : \sum_{r=1}^{n} r p_r = n}} n!(-1)^j (j-1)! \prod_{r=1}^{n} \frac{2^{\binom{r}{2}p_r}}{(r!)^{p_r} p_r!},$$

where $j = \sum_{r=1}^{n} p_r$.

The point to the last problem is that we can use the exponential formula in reverse to say that if $g(x)$ is the generating function for the number of (nonempty) connected structures of size n in a given family of combinatorial structures and $f(x)$ is the generating function for all the structures of size n in that family, then not only is $f(x) = e^{g(x)}$, but $g(x) = \ln(f(x))$ as well. Further, if we happen to have a formula for either the coefficients of $f(x)$ or the coefficients of $g(x)$, we can get a formula for the coefficients of the other one!

B.7 Supplementary Problems

1. Use product principle for EGFs and the idea of coloring a set in two colors to prove the formula $e^x \cdot e^x = e^{2x}$.

2. Find the EGF for the number of ordered functions from a k-element set to an n-element set.

3. Find the EGF for the number of ways to string n distinct beads onto a necklace.

4. Find the exponential generating function for the number of broken permutations of a k-element set into n parts.

5. Find the EGF for the total number of broken permutations of a k-element set.

6. Find the EGF for the number of graphs on n vertices in which every vertex has degree 2.

7. Recall that a cycle of a permutation cannot be empty.

    (a) What is the generating function for the number of cycles on an even number of elements (i.e. permutations of an even number n of elements that form an n-cycle)? Your answer should not have a summation sign in it. Hint: If $y = \sum_{i=0}^{\infty} \frac{x^{2i}}{2i}$, what is the derivative of y?

    (b) What is the generating function for the number of permutations on n elements that are a product of even cycles?

    (c) What is the generating function for the number of cycles on an odd number of elements?

    (d) What is the generating function for the number of permutations on n elements that are a product of odd cycles?

    (e) How do the generating functions in parts (b) and (d) of this problem related to the generating function for all permutations on n elements?

Bibliography

Index